Comments

Of

THE CENTER FOR AI AND DIGITAL POLICY (CAIDP)

To the

U.S. COPYRIGHT OFFICE, LIBRARY OF CONGRESS

On

ARTIFICIAL INTELLIGENCE AND COPYRIGHT

October 30, 2023

We write to you on behalf of the Center for AI and Digital Policy (CAIDP), in response to the notice of inquiry and request for comments (RFC) issued by the U.S. Copyright Office on Artificial Intelligence and Copyright.[1] In the Comments below we provide (1) general recommendations on AI and Copyright and (2) responses to specific questions set out by the Copyright Office. We have also promoted this opportunity for public comment on AI policy through our Public Voice project.[2]

In summary, we recommend:

1. The U.S. Copyright Office should ensure that the development of AI systems is human-centered and trustworthy. The value of copyright is recognized in the U.S. Constitution[3] and is integral to the livelihood of millions of creators, actors, artists, designers, musicians, photographers, poets, writers and many others in the United States.

2. Accountability and transparency mechanisms to ensure governance of AI techniques should incorporate best practices set out in the Universal Guidelines for Artificial

[1] Notice of inquiry and request for comments, *Artificial Intelligence and Copyright*, U.S. Copyright Office, Library of Congress, 88 FR 59942, https://www.govinfo.gov/content/pkg/FR-2023-08-30/pdf/2023-18624.pdf.
[2] CAIDP, Public Voice, *Copyright Right Office, Copyright and Artificial Intelligence*, https://www.caidp.org/public-voice/copyright-office-us-2023/
[3] Constitution of the United States, Art.1, Section 7: "To promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries", https://constitution.congress.gov/constitution/

Intelligence,[4] the OECD AI Principles,[5] and the UNESCO Recommendations on AI Ethics.[6]

3. The allocation of rights concerning those who are impacted by AI systems should be based on pre-deployment assessments, audits, certifications throughout the AI lifecycle to ensure transparency, explainability, accountability, and traceability.

4. Legal standards should also be established to ensure AI transparency and accountability. Accountability mechanisms or practices will have no meaningful impact in the absence of clearly defined legal standards and enforceable remedies.

About CAIDP

The CAIDP is an independent non-profit organization that advises national governments and international organizations on artificial intelligence and digital policy based in Washington, DC. We are a global network of AI policy experts and advocates. We publish annually the ***Artificial Intelligence and Democratic Values Index***, a comprehensive review of AI policies and practices around the world.[7] We advise national governments and international organizations regarding artificial intelligence and digital policy.[8] We set out below our general recommendations and specific comments to the questions presented in your RFC.

CAIDP General Recommendations on AI and Copyright

1. **Ensure that federal legislation is consistent with human-centered values, fairness and transparency as set out in the OECD Principles on Artificial Intelligence.**

In 2019, the Organization for Economic Cooperation and Development (OECD) promulgated the OECD Principles on Artificial Intelligence which was adopted by both member and non-member countries. The U.S. endorsed the OECD AI Principles.[9] According to OECD AI Principle on Human-centered values and fairness (Principle 1.2): "*AI actors should respect the rule of law, human rights and democratic values, throughout the AI system lifecycle. These include*

[4] Universal Guidelines for Artificial Intelligence, https://thepublicvoice.org/ai-universal-guidelines/
[5] Recommendation of the Council on Artificial Intelligence, OECD (May 21, 2019), https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449
[6] UNESCO Recommendation on the Ethics of AI 2021, https://unesdoc.unesco.org/ark:/48223/pf0000377897
[7] CAIDP, *Artificial Intelligence and Democratic Values* (2023), https://www.caidp.org/reports/aidv-2022/
[8] CAIDP, https://www.caidp.org
[9] *U.S. Joins with OECD in Adopting Global AI Principles*, NTIA (May 22, 2019), https://www.ntia.doc.gov/blog/2019/us-joins-oecd-adopting-global-ai-principles

freedom, dignity and autonomy, privacy and data protection, non-discrimination and equality, diversity, fairness, social justice, and internationally recognized labor rights…AI actors should implement mechanisms and safeguards, such as capacity for human determination, that are appropriate to the context and consistent with the state of art" (emphasis added).[10]

OECD AI Principle on Transparency and explainability (Principle 1.3) states: "*AI Actors should commit to transparency and responsible disclosure regarding AI systems…they should…foster a general understanding of AI systems…make stakeholders aware of their interactions with AI systems, including in the workplace…enable those affected by an AI system to understand the outcome, and…enable those adversely affected by an AI system to challenge its outcome based on plain and easy-to-understand information on the factors, and the logic that served as the basis for the prediction, recommendation or decision"* (emphasis added).[11]

The principles of human-centered values, fairness and transparency should lie at the core of any legislation enacted on AI. These principles can help ensure that stakeholders impacted by AI, particularly creators and creative industries, can participate in when and how artificial intelligence is used at work. The principles can also help to ensure that AI is used to benefit creators, while also mitigating the direct risk that workers are displaced by AI and excessive automation.[12]

2. Regarding Generative AI, follow the direction set out by President Biden and implement accountability provisions of the AI Bill of Rights.[13]

In remarks made before the President's Council of Advisors on Science and Technology last month, President Biden noted that his administration will "continue to work with bipartisan legislation so America leads the way toward responsible AI innovation" by ensuring that "AI technology is safe, secure, and trustworthy before it's released to the public."[14] President Biden said that he is keen to "harness the power of artificial intelligence for good while protecting people

[10] OECD, "Human-centred values and fairness (Principle 1.2), https://oecd.ai/en/dashboards/ai-principles/P6
[11] OECD, "Transparency and explainability (Principle 1.3), https://oecd.ai/en/dashboards/ai-principles/P7
[12] *See* European Commission and the U.S. Council of Economic Advisors, *The Impact of Artificial Intelligence on the Future of Workforces in the European Union and the United States of America,* December 5, 2022, https://www.whitehouse.gov/wp-content/uploads/2022/12/TTC-EC-CEA-AI-Report-12052022-1.pdf
[13] The White House, *Blueprint for An AI Bill of Rights*, https://www.whitehouse.gov/ostp/ai-bill-of-rights/
[14] The White House, *Remarks by President Biden Before Meeting with the President's Council of Advisors on Science and Technology*, September 27, 2023 https://www.whitehouse.gov/briefing-room/speeches-remarks/2023/09/27/remarks-by-president-biden-before-meeting-with-the-presidents-council-of-advisors-on-science-and-technology-san-francisco-ca/

from the profound risk it also presents."[15] He applauded the resolution reached in the writers' strike, "including insurances on how the use of AI will occur."[16] President Biden has stated that the purpose of the proposed AI Bill of Rights is "to ensure that important protections are built into the AI systems from the very start."[17] During her meeting with consumer protection, labor and civil rights leaders on AI, Vice President Harris added that the U.S. does not need to make a "false choice" between AI innovation and improving the condition of people's lives.[18]

President Biden has announced a new Executive Order today on "Safe, Secure, and Trustworthy Artificial Intelligence".[19] In this Order, the President directs federal agencies to "Develop principles and best practices to mitigate the harms and maximize the benefits of AI for workers by addressing job displacement; labor standards; workplace equity, health, and safety; and data collection."

CAIDP strongly supports a human-centric approach to the governance of AI, and Generative AI in particular. CAIDP has already endorsed the AI Bill of Rights. CAIDP has also made specific recommendations for this initiative.[20] Given the urgency of this moment, it is our view that the U.S. Copyright Office can play a role in advising Congress on national and international issues relating to AI.

CAIDP Specific Responses to Questions in RFC

General Questions (Responses to Question 1)

1. As described above, generative AI systems have the ability to produce material that would be copyrightable if it were created by a human author. What are your views on the potential

[15] *Id.*
[16] *Id.*
[17] The White House, *Remarks by President Biden on Seizing the Opportunities and Managing the Risks of Artificial Intelligence*, June 20, 2023 https://www.whitehouse.gov/briefing-room/speeches-remarks/2023/06/20/remarks-by-president-biden-on-seizing-the-opportunities-and-managing-the-risks-of-artificial-intelligence/
[18] The White House, *Readout of Vice President Harris's Meeting with Consumer Protection, Labor, and Civil Rights Leaders on AI,* July 13, 2023, https://www.whitehouse.gov/briefing-room/statements-releases/2023/07/13/readout-of-vice-president-harriss-meeting-with-consumer-protection-labor-and-civil-rights-leaders-on-ai/
[19] The White House, *FACT SHEET: President Biden Issues Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence,* Statements and Releases, Oct. 30, 2023, https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/
[20] Lorraine Kisselburgh and Marc Rotenberg, *Next Steps on the AI Bill Of Rights,* Washington Spectator (Nov. 2021), https://washingtonspectator.org/author/lorraine-marc/; CAIDP, Public Voice, https://www.caidp.org/public-voice/

benefits and risks of this technology? How is the use of this technology currently affecting or likely to affect creators, copyright owners, technology developers, researchers, and the public?

The UNESCO Recommendation on the Ethics of AI address the benefits and risks of AI on culture. The Recommendation states:

- AI education and digital training for artists and creative professionals should be promoted to assess the sustainability of AI technologies for use in their profession, and contribute to the design and implementation of suitable AI technologies (Rec #96)
- Awareness and evaluation of AI tools should be encouraged among local cultural industries and small and medium enterprises working in the cultural field, to avoid the risk of concentration in the cultural market (Rec #97)
- Technology companies and other stakeholders should engage with government to promote a diverse supply of and plural access to cultural expressions, and in particular to ensure that algorithmic recommendation enhances the visibility and discoverability of local content (Rec #98)
- Research at the intersection of AI and intellectual property (IP) should be fostered, for example to determine whether or how to protect the IP rights of works created by means of AI technologies and how AI affects the rights and interests of IP owners, whose works are used to research, develop, train or implement AI applications (Rec #99)
- Museums, galleries, libraries and archives at the national level should be encouraged to use AI systems to highlight their collections and enhance their libraries, databases, and knowledge base, while also providing access to their users (Rec #100)

In addition, CAIDP applauds the agreement reached between the Writers Guild of America and the Alliance of Motion Picture and Television Producers as an example of how the benefits of generative AI can be harnessed without replacing workers.[21] The agreement states that studios cannot use AI in place of a credited and paid guild member, and that studios cannot mandate guild members to use AI. Overall, the agreement "guarantees a contractually mandated context in which A.I. can be utilized – one that benefits, rather than impedes or replaces, the workers."[22] The

[21] Memorandum of Agreement for the 2023 WGA Theatrical and Television Basic Agreement, https://www.wgacontract2023.org/wgacontract/files/memorandum-of-agreement-for-the-2023-wga-theatrical-and-television-basic-agreement.pdf

[22] Adam Seth Litwin, "Hollywood's Deal With Screenwriters Just Rewrote the Rules Around A.I." *New York Times*, September 29, 2023, https://www.nytimes.com/2023/09/29/opinion/wga-strike-deal-ai-jobs.html

agreement is one that other industries can look to as a model of how to negotiate around the use of a new technology to benefit the greatest number of workers without simply resisting technological change.[23]

Training (Responses to Questions 7.4, 8.3, 8.4)

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

There is consensus among AI experts that "pre-training" processes of generative AI models use copyrighted data. In his written testimony[24] for Congress,[25] Christopher Callison-Burch testified that "AI systems use large amounts of training data in the process called pre-training (described above). The process of gathering pre-training data for AI systems is similar to the "web crawling" process that Google and other companies use in order to create a searchable index of the web. Because pre-training data is largely gathered through web crawling, a very large fraction of the data consists of copyrighted sources. This is a result of the fact that nearly all content posted online is protected by U.S. copyright laws, since copyright protection arises automatically when an author creates an original work and fixes it in a tangible medium."

In October 2023, *Scientific American* compiled views of several AI experts including Emily Bender, Meredith Broussard, and Ben Zhao in a report on how personal data is used to train AI Models. [26] The *Scientific American* report on generative AI found that "[M]achine learning models are only capable of pumping out images and text because they've been trained on mountains of real people's creative work, much of it copyrighted.

[23] *Id.*
[24] Written Testimony of Christopher Callison-Burch, Ph.D., *Understanding Generative Artificial Intelligence and Its Relationship to Copyright,* Before The U.S. House of Representatives Judiciary Committee Subcommittee on Courts, Intellectual Property, and the Internet Hearing on Artificial Intelligence and Intellectual Property: Part I – Interoperability of AI and Copyright Law, https://www.cis.upenn.edu/~ccb/publications/understanding-generative-AI-and-its-relationship-to-copyright.pdf
[25] House of Representatives, Judiciary Committee, *Hearing on Artificial Intelligence and Intellectual Property: Part I — Interoperability of AI and Copyright Law*, May 17, 2023, https://judiciary.house.gov/committee-activity/hearings/artificial-intelligence-and-intellectual-property-part-i
[26] Lauren Leffer, *Your Personal Information Is Probably Being Used to Train Generative AI Models, Scientific American,* Oct. 19. 2023, https://www.scientificamerican.com/article/your-personal-information-is-probably-being-used-to-train-generative-ai-models/

Several AI developers including OpenAI, Meta and Stability AI now face multiple lawsuits because they use Other People's Data.[27] Such legal claims are supported by independent analyses; in August, for instance, *The Atlantic* reported that Meta trained its large language model (LLM) in part on a data set called Books3, which contained more than 170,000 pirated and copyrighted books. In the rush to build and train ever-larger AI models, developers have swept up much of the searchable Internet. This not only has the potential to violate copyrights but also threatens the privacy of the billions of people who share information online. It also means that supposedly neutral models could be trained on biased data. A lack of corporate transparency makes it difficult to figure out exactly where companies are getting their training data."[28]

AI expert Meredith Whittaker has spoken on the lack of transparency on training of AI systems and their datasets. "Popular AI programs such as DALL-E2 and ChatGPT remain "gated to public" access, and are therefore not open to public scrutiny."[29]

Therefore, access to pre-training datasets, based upon the principle of 'transparency' and 'explainability' of generative AI systems is critical without which opacity would persist in the data used to train AI models.

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. How should the fair use analysis apply if AI models or datasets are later adapted for use of a commercial nature? Does it make a difference if funding for these noncommercial or research uses is provided by for-profit developers of AI systems? And 8.4. What quantity of training materials do developers of generative AI models use for training? Does the volume of material used to train an AI model affect the fair use analysis? If so, how?

There is consensus among experts that current large language models (LLMs) are a black-box and suffer from severe opacity. Researchers at the Center for Research on Foundation Models (CRFM) have published their evaluation of transparency of a variety of LLMs.[30] "The Index results were sobering: No major foundation model developer was close to providing adequate

[27] See Louis Brandeis*, Other People's Money And How The Bankers Use* It (2014).
[28] Lauren Leffer, *Your Personal Information Is Probably Being Used to Train Generative AI Models, Scientific American,* Oct. 19. 2023, https://www.scientificamerican.com/article/your-personal-information-is-probably-being-used-to-train-generative-ai-models/
[29] Damien Black, *Signal boss: there's nothing open about AI,* Cybernews, Aug. 18, 2023, https://cybernews.com/tech/meredith-whittaker-open-ai-corporate-control/
[30] CRFM, *The Foundation Model Transparency Index,* Oct. 19, 2023, https://crfm.stanford.edu/fmti/

transparency, according to the researchers — the highest overall score was 54% — revealing a fundamental lack of transparency in the AI industry." [31]

Without transparency of the workings of these systems and their datasets a conclusion on fair use would be premature and likely to suffer infirmities of being over or under inclusive. There is a divergence of views among experts as to whether training of AI models would constitute fair use.

AI companies have made voluntary commitments to the President and administration "to develop and deploy advanced AI systems to help address society's greatest challenges."[32] Research purposes to address issues of mitigating bias and create fairer systems could possibly constitute fair use.

Amanda Levendowski, Associate Professor Law and founding Director of the Intellectual Property and Information policy (IPIP) Clinic at Georgetown Law School claims that copyright law can positively shape how AI systems are developed and who is empowered to develop them.[33] Specifically, she argues the fair use doctrine and the normative values embedded within traditional fair use align with the goals of mitigating AI bias and can create fairer AI systems.[34] Levendowski states: "If we hope to create less biased commercial AI systems, using copyright-protected works as AI training data will be key."[35] She notes that recent studies have shown that AI systems trained using larger datasets perform more accurately than those trained with smaller ones. In addition, she says the quantity of data needed for AI training is too large for many would-be AI creators to obtain without relying on others' copyrighted works.[36] Levandowski adds that AI training data does not harm the commercial market for copyrighted works, and that "while there may be a licensing market, the value of datasets is derived from a diversity of sources and content."[37]

---

[31] VentureBeat, *How transparent are AI models? Stanford researchers found out,* Oct.18, 2023, https://venturebeat.com/ai/how-transparent-are-ai-models-stanford-researchers-found-out/
[32] The White House, *FACT SHEET: Biden-Harris Administration Secures Voluntary Commitments from Leading Artificial Intelligence Companies to Manage the Risks Posed by AI,* Statements and Releases, Jul.21, 2023, https://www.whitehouse.gov/briefing-room/statements-releases/2023/07/21/fact-sheet-biden-harris-administration-secures-voluntary-commitments-from-leading-artificial-intelligence-companies-to-manage-the-risks-posed-by-ai/
[33] Amanda Levendowski, "How Copyright Law Can Fix Artificial Intelligence's Implicit Bias Problem," 93 Wash. L. Rev. 579, 587 (2018).
[34] *Id.* at 621.
[35] *Id.*
[36] *Id.*
[37] *Id.* at 630.

The Congressional Research Services report on Generative Artificial Intelligence and Copyright Law[38] has stated that "Whether or not copying constitutes fair use depends on four statutory factors under 17 U.S.C. § 107:

1. the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes;
2. the nature of the copyrighted work;
3. the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and
4. the effect of the use upon the potential market for or value of the copyrighted work."

The CRS report discusses the positions taken by AI companies in their public comments and prior decided cases like *The Authors Guild, Inc. v. Google, Inc.,* as well as the positions taken by experts and states, "Regarding the fourth fair use factor, some generative AI applications have raised concern that training AI programs on copyrighted works allows them to generate similar works that compete with the originals."[39]

Machine learning systems mimic and manipulate human behavior.[40] Understanding this fundamental nature of the technology is also key to contextualizing the "fair use" doctrine both for inputs and outputs of AI systems.

Considering the above, a conclusion on whether the use of copyrighted material to train generative AI models, even for non-commercial purposes, would constitute "fair use" merits further examination of the systems after establishing assessment procedures to ensure transparency and explainability of AI systems.

Transparency & Recordkeeping (Responses to Questions 15, 15.1, 15.2, 15.3, 15.4, 16)

15. In order to allow copyright owners to determine whether their works have been used, should developers of AI models be required to collect, retain, and disclose records regarding the materials used to train their models? Should creators of training datasets have a similar obligation?

15.1. What level of specificity should be required?

15.2. To whom should disclosures be made?

[38] Congressional Research Service, *Generative Artificial Intelligence and Copyright Law,* Sept. 29, 2023, pg. 3, 4, https://crsreports.congress.gov/product/pdf/LSB/LSB10922
[39] *Id* at pg. 4
[40] CAIDP, *In the Matter of OpenAI* (FTC 2023), at pg. 19, https://www.caidp.org/cases/openai/; CAIDP, *In the Matter of Zoom* (FTC 2023), at pg. 21, https://www.caidp.org/cases/zoom/

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

15.4. What would be the cost or other impact of such a recordkeeping system for developers of AI models or systems, creators, consumers, or other relevant parties?

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

As we stated in our complaint to the Federal Trade Commission on OpenAI, it is important that companies comply with recommended business practices established by the FTC for commercial products.[41] The failure of OpenAI and other companies to provide basic information on details about the architecture, model size, hardware, computing resources, training techniques, dataset construction and training methods is concerning.[42]

In our complaint we discuss the need for documentation to ensure transparency of AI models.[43] Leading AI experts and researchers Timnit Gebru, Emily Bender and others made clear the importance of documentation and level of specificity required to help ensure transparency that enables evaluation. As the experts explain:

"As a part of careful data collection practices, researchers must adopt frameworks to describe the uses for which their models are suited and benchmark evaluations for a variety of conditions. This involves providing thorough documentation on the data used in model building, including the motivations underlying data selection and collection processes. This documentation should reflect and indicate researchers' goals, values, and motivations in assembling data and creating a given model."[44]

They further state "While documentation allows for potential accountability, undocumented training data perpetuates harm without recourse. Without documentation, one cannot try to understand training data characteristics in order to mitigate some of these attested

[41] CAIDP, *In the Matter of OpenAI* (FTC 2023), https://www.caidp.org/cases/openai/
[42] *Id.* at pg. 26.
[43] *Id* at pg. 25
[44] Emily M. Bender, Timnit Gebru, Angelina McMillan-Major, Margaret Mitchell, *On the Dangers of Stochastic Parrots: Can Language Models Be Too Big*, FAccT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (March 2022), pg. 618, https://doi.org/10.1145/3442188.3445922

issues or even unknown ones."[45] The experts recommend that "all stakeholders of potential use cases" should be considered in the training and pre-deployment stage of generative AI models.[46]

As Dr. Kate Crawford, founder and former director of research at the AI Now Institute at NYU and author of *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence* (2021) has stated:

> There is a real problem here. Scientists and researchers like me have no way to know what Bard, GPT4, or Sydney are trained on. Companies refuse to say. This matters, because training data is part of the core foundation on which models are built. Science relies on transparency. Without knowing how these systems are built, there is no reproducibility. You can't test or develop mitigations, predict harms, or understand when and where they should not be deployed or trusted. The tools are black boxed.[47]

Therefore, without extensive documentation and record-keeping requirements it will be impossible to assess whether or not copyrighted material has been used in training data, to what extent it has been used, and any potential infringement of existing copyrights by AI-generated outputs. Transparency and accountability mechanisms for large language models are not a novel feature unique to copyright considerations. The EU AI Act prescribes extensive documentation, disclosure, and audit requirements of such systems which could be multi-purpose or multi-modal.

The Federal Trade Commission held a roundtable on "Creative Economy and Generative AI" in an effort "to better understand how these tools may impact open and fair competition or enable unlawful business practices across markets, including in creative industries."[48] Chair Khan in her address stated, "We've already heard significant concern about how these technologies could virtually overnight significantly disempower creators and artists who may watch their life's creation be appropriated to models over which they have no control."[49]

Speaking at the event, Duncan Cabtree-Ireland, National Executive Director and Chief Negotiator at SAG-AFTRA stated "When used ethically and in a manner that recognizes intellectual property rights, AI can help people in their careers and can further opportunities. It can

---

[45] *Id* at pg. 610, 615; *See also,* CAIDP, *In the Matter of OpenAI* (FTC 2023), pg. 11, https://www.caidp.org/cases/openai/
[46] CAIDP, *In the Matter of OpenAI* (FTC 2023), pg. 25, https://www.caidp.org/cases/openai/
[47] Kate Crawford, Twitter, March 22 2023, https://x.com/katecrawford/status/1638524013432516610
[48] The Federal Trade Commission Event, *Creative Economy and Generative AI,* Oct. 4, 2023, https://www.ftc.gov/news-events/events/2023/10/creative-economy-generative-ai
[49] *Id.,* Transcript-Files, at pg. 2, https://www.ftc.gov/system/files/ftc_gov/pdf/creative-economy-and-generative-ai-transcript-october-4-2023.pdf

create access to employment for people with disabilities and those who would otherwise be prevented from pursuing work in the entertainment industry. What SAG-AFTRA is eager to do is to channel the benefits of AI into a future that's beneficial to our members, to workers in other industries, and to the public in general. The key is that the companies using AI technology must be required to get the informed consent of any individuals whose voice, likeness, performance, persona, or intellectual property is being used to generate content and companies need to compensate these individuals fairly."[50] *(emphasis added)*

The SAG-AFTRA list of demands and key issues for negotiations state "Performers need the protection of our images and performances to prevent replacement of human performances by artificial intelligence technology."[51]

There is a clear demand from copyright owners to be notified, consent to, and be compensated for their creative works that are used in AI models or systems. The Copyright Office would be well advised to take note of these current concerns and the "identification" obligation[52] set out in established AI governance frameworks like the Universal Guidelines on AI.

Labelling or Identification (Responses to Questions 28, 28.1, 28.2, 28.3, 29)

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?

The White House AI Executive Order[53] directs federal agencies to "*Protect Americans from AI-enabled fraud and deception by establishing standards and best practices for detecting AI-generated content and authenticating official content.*" Building on this principle for the creative economy and to ascertain the allocation of authorship and/or attribution or distribution rights, AI-generated material should be labeled or otherwise publicly identified as being generated by AI. The effects of such labels would differ depending on what kind of content is included in the definition of "material." In regards to AI-generated reviews or testimonials that present false statements could unduly influence a consumer's purchasing decisions, potentially contravening

[50] *Id,* at pg. 6
[51] SAG-AFTRA, *We're Fighting for the Survival of our Profession,* Jul. 17, 2023, https://www.sagaftra.org/files/sa_documents/SAG-AFTRA_Negotiations_Status_7_13_23.pdf
[52] CAIDP, *Universal Guidelines for AI,* https://www.caidp.org/universal-guidelines-for-ai/
[53] [53] The White House, *FACT SHEET: President Biden Issues Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence,* Statements and Releases, Oct. 30, 2023, https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/

fair trade and advertising standards that safeguard against misleading consumer practices.[54] AI-generated images and highly realistic but entirely synthetic deep fakes pose a threat to individual rights to privacy and reputation.[55] Without a legal mandate for clear labeling, it becomes nearly impossible for creators and consumers to seek redress when harmed by AI-generated content. Furthermore, numerous governance bodies, including the European Union, have advocated for similar measures as outlined in the EU AI Act.[56] Given that technology companies often operate beyond traditional national boundaries, it is highly probable that prominent U.S. firms will seek standardization and implement these measures in their U.S. operations.

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

To ensure transparency and clarity in the domain of artificial intelligence image generation, it is recommended that each image produced via such mechanisms be adorned with a discernible watermark, designating its status as "AI-Generated." The obligation to incorporate this watermark should squarely reside with the originating entity or corporation responsible for facilitating the AI-driven image creation, specifically the primary platform or portal from which users procure the aforementioned imagery. Many stakeholders such as Google have already implemented such a practice. Google Cloud launched "SynthID, a tool for watermarking and identifying AI-generated images." This tool works by embedding "a digital watermark directly into the pixels of an image, making it imperceptible to the human eye but detectable for identification."[57] This serves as evidence that corporations possess the capability to institute such measures.

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

Such measures may take some time before they are freely made available to the public. Google's SynthID is only the beta version and has only been released to a limited number of Vertex AI customers.[58] Furthermore, should enterprises commence the development of proprietary digital

[54] Federal Trade Commission PRESS RELEASE, *Federal Trade Commission Announces Proposed Rule Banning Fake Reviews and Testimonials*, https://www.ftc.gov/news-events/news/press-releases/2023/06/federal-trade-commission-announces-proposed-rule-banning-fake-reviews-testimonials.
[55] Michael Atleson, *Chatbots, deepfakes, and voice clones: AI deception for sale*, https://www.ftc.gov/business-guidance/blog/2023/03/chatbots-deepfakes-voice-clones-ai-deception-sale.
[56] European Union, Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL
LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS, https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52021PC0206.
[57] Google DeepMind, *Identifying AI-generated images with SynthID*, https://www.deepmind.com/blog/identifying-ai-generated-images-with-synthid.
[58] Google DeepMind, *Identifying AI-generated images with SynthID*, https://www.deepmind.com/blog/identifying-ai-generated-images-with-synthid.

markers, this fragmented approach to marking may introduce complexities in standardization and tracking. Additionally, in the event of a malfunction or downtime in a company's digital watermark detection system, many individuals may find themselves unable to verify the authenticity of images upon which they depend, potentially undermining the primary purpose of distinguishing AI-generated content.

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?

Possible consequences could include an imposition of monetary fines, especially if the failure leads to harm to the consumer. Continued failures could lead to mandatory audits. In all instances of failure, the organization should be required to make the appropriate remedies or corrections within a reasonable and effectively communicated timeframe.

A. Additional Questions About Issues Related to Copyright (Responses to Questions 30, 34)

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?

To ensure appropriate regulation and the imposition of fines, the Copyright Office must take into account the specific legal entitlements provided to individuals when their personal data is utilized without authorization. The determination of legal rights pertaining to AI-generated creations hinges on numerous factors, notably the degree of human participation in the material's inception. Some protections exist for privacy torts when there is appropriation of the other's name or likeness or when there is publicity that unreasonably places the other in a false light before the public.[59] However, none of these protections are specifically for harms caused by AI-generated content.

Duncan Cabtree-Ireland speaking at the FTC roundtable mentioned above stated, "It's important to understand that all AI-generated content originates from a human creative source. No AI algorithm is able to make something out of nothing. And that human generated content that's used in the training data reflects real and substantial work and its intellectual property. It deserves legal protection".[60]

[59] Restatement (2nd) of Torts (1977).
[60] The Register, *Acting union calls out Hollywood studios for 'double standard' on AI use,* Oct. 5, 2023, https://www.theregister.com/2023/10/05/sagaftra_ftx_ai_ip/

Dan Navarro in his statement[61] to the House Judiciary Committee stated "Training AI to mimic professional performers or "generate" new works based on millions of copies of published songs and recordings presents a host of legal implications, from copyright infringement to violations of rights of publicity and trademark, to name, voice, and likeness abuse"

We strongly support the decision of the Judge Howell to affirm the U.S. Copyright Office's decision to withhold copyright protections for authorship by an AI system.[62] While there may be future cases where generative AI is recognized as tool to augment human creativity, like photography, that allows for copyright protection, we would oppose any decision to assign authorship to an AI system.

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?

Dan Navarro in his Congressional testimony stated, "The Copyright clause of the Constitution exists to incentivize humans to create –machines don't need incentives."[63]

Ashley Irwin, President, Society of Composers & Lyricists in his Congressional testimony[64] stated "The market will be diluted due to AI generated works and as a result, copyright protection should not be granted to AI-generated works."

There is a clear demand from various creative groups for protections against an AI system generating outputs that imitate the artistic style of a human creator. From an intellectual property perspective, an artist's unique style can be seen as a proprietary form of expression developed over years of training, experience, and personal introspection. Allowing unrestricted imitation by AI could dilute the distinctiveness of such intellectual assets, potentially leading to market confusion and unfair competition. Secondly, on economic grounds, artists depend on the exclusivity of their work for livelihood. If AI can replicate their signature style en masse, it might undermine the

[61] Written Statement of Dan Navarro before the Committee on the Judiciary, United States House of Representatives Subcommittee on Courts, Intellectual Property and the Internet on *Artificial Intelligence and Intellectual Property: Part I — Interoperability of AI and Copyright Law*, pg. 1, https://judiciary.house.gov/sites/evo-subsites/republicans-judiciary.house.gov/files/evo-media-document/navarro-testimony.pdf ("Dan Navarrro Testimony")
[62] *Thaler v. Perlmutter*, 2023 U.S. Dist. LEXIS 145823.
[63] Dan Navarro Testimony, pg. 2
[64] Statement of Ashley Irwin, President Society of Composers & Lyricists Before the Subcommittee on Courts, Intellectual Property and the Internet of the House Committee of the Judiciary Hearing on "*Artificial Intelligence and Intellectual Property: Part I — Interoperability of AI and Copyright Law",* pg. 3, https://judiciary.house.gov/sites/evo-subsites/republicans-judiciary.house.gov/files/evo-media-document/irwin-testimony.pdf

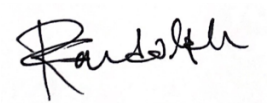
market value of their creations, unjustly depriving them of economic benefits. Copyright law protects original works of authorship,[65] and while individual techniques or ideas are not copyrightable, an artist's unique style and skill should be protected.

We support the U.S. Copyright Office's effort to undertake a study of copyright law and policy issues raised by AI systems. Thank you for your consideration of our views.

We welcome the opportunity to speak with you further about these recommendations.

Marc Rotenberg
CAIDP Executive Director

Merve Hickok
CAIDP President



Christabel Randolph
Law Fellow

Madeleine Matsui
Law Fellow

Janhvi Patel
CAIDP Research Assistant

[65] U.S. Copyright Office, What Does Copyright Protect?, Copyright.gov, https://www.copyright.gov/help/faq/faq-protect.html#:~:text=Copyright%2C%20a%20form%20of%20intellectual,%2C%20computer%20software%2C%20and%20architecture..