Labor, 200 Constitution Ave. NW, Room S–3323, Washington, DC 20210.

Electronic submission: You may submit comments and attachments electronically at http://www.regulations.gov. Follow the online instructions for submitting comments.

FOR FURTHER INFORMATION CONTACT: Anjanette Suggs, Office of Workers’ Compensation Programs, Division of Federal Employees Longshore, and Harbor Workers’ Compensation, OWCP/DFELHWC, at suggs.anjanette@dol.gov (email); (202) 354–9660.

SUPPLEMENTARY INFORMATION:

I. Background

The Office of Workers’ Compensation Programs (OWCP) is the agency responsible for administration of the Longshore and Harbor Workers’ Compensation Act (LHWCA), and the Federal Employees’ Compensation Act (FECA). 33 U.S.C. 939 (LHWCA) and 5 U.S.C. 8104 and 8111 (FECA) authorizes OWCP to pay for approved vocational rehabilitation services to eligible workers with work-related disabilities. In order to decide whether to approve a rehabilitation plan, OWCP must receive a copy of the plan, supporting vocational testing materials and the estimated cost to implement the plan, broken down to the fees, supplies, tuition and worker maintenance payments that are contemplated. OWCP also must receive the signatures of the worker and the rehabilitation counselor to show that the worker agrees to follow the proposed plan, and that the proposed plan is appropriate. Form OWCP–16 is the standard format for the collection of this information. The regulations implementing these statutes allow for the collection of information needed for OWCP to determine if a rehabilitation plan should be approved and payment of any related expenses should be authorized. (LHWCA, 702.506 and 702.507, (FECA, 20 CFR 10.518, 10.519).

II. Desired Focus of Comments

OWCP is soliciting comments concerning the proposed information collection (ICR) titled, “Rehabilitation Plan and Award”, OWCP–16. OWCP/DFELHWC is particularly interested in comments that:

- Evaluate whether the collection of information is necessary for the proper performance of the functions of the Agency, including whether the information has practical utility;
- Evaluate the accuracy of OWCP/DFELHWC’s estimate of the burden related to the information collection, including the validity of the methodology and assumptions used in the estimate;
- Suggest methods to enhance the quality, utility, and clarity of the information to be collected; and
- Minimize the burden of the information collection on those who are to respond, including through the use of appropriate automated, electronic, mechanical, or other technological collection techniques or other forms of information technology, e.g., permitting electronic submission of responses.

Background documents related to this information collection request are available at https://regulations.gov and at DOL–OWCP/DFELHWC located at 200 Constitution Avenue NW, Room S–3323, Washington, DC 20210. Questions about the information collection requirements may be directed to the person listed in the FOR FURTHER INFORMATION section of this notice.

III. Current Actions

This information collection request concerns the Rehabilitation Plan and Award, OWCP–16. OWCP/DFELHWC has updated the data with respect to the number of respondents, responses, burden hours, and burden costs supporting this information collection request from the previous information collection request.

Type of Review: Extension, without change, of a currently approved collection.

Agency: Office of Workers’ Compensation Programs, Division of Federal Employees’ Longshore, and Harbor Workers’ Compensation, OWCP/DFELHWC.

OMB Number: 1240–0045.

Affected Public: Not-for-profit institutions, Businesses or other for-profits.

Number of Respondents: 3,413.

Frequency: On occasion.

Number of Responses: 3,413.

Annual Burden Hours: 1,707 hours.

Total Respondent or Recordkeeper Cost: $0.

OWCP Form 16, Rehabilitation Plan and Award.

Comments submitted in response to this notice will be summarized in the request for Office of Management and Budget approval of the proposed information collection request; they will become a matter of public record and will be available at https://www.reginfo.gov.

Anjanette Suggs,
Certifying Officer.

BILLING CODE 4510–CH–P

LIBRARY OF CONGRESS

Copyright Office

[DOCKET No. 2023–6]

Artificial Intelligence and Copyright

AGENCY: U.S. Copyright Office, Library of Congress.

ACTION: Notice of inquiry and request for comments.

SUMMARY: The United States Copyright Office is undertaking a study of the copyright law and policy issues raised by artificial intelligence (“AI”) systems. To inform the Office’s study and help assess whether legislative or regulatory steps in this area are warranted, the Office seeks comment on these issues, including those involved in the use of copyrighted works to train AI models, the appropriate levels of transparency and disclosure with respect to the use of copyrighted works, and the legal status of AI-generated outputs.

DATES: Written comments are due no later than 11:59 p.m. Eastern Time on Wednesday, October 18, 2023. Written reply comments are due no later than 11:59 p.m. Eastern Time on Wednesday, November 15, 2023.

ADDRESSES: For reasons of governmental efficiency, the Copyright Office is using the regulations.gov system for the submission and posting of public comments in this proceeding. All comments should be submitted electronically through regulations.gov. Specific instructions for submitting comments are available on the Copyright Office website at https://copyright.gov/policy/artificial-intelligence. If electronic submission is not feasible, please contact the Office using the contact information below for special instructions.

FOR FURTHER INFORMATION CONTACT: Rhea Efthimiadis, Assistant to the General Counsel, by email at meft@copyright.gov or telephone at 202–707–8350.

SUPPLEMENTARY INFORMATION:

I. Introduction

Over the last year, artificial intelligence (“AI”) systems and the rapid growth of their capabilities have attracted significant media and public attention. One type of AI, “generative AI” technology, is capable of producing outputs such as text, images, video, or audio (including emulating a human voice) that would be considered copyrightable if created by a human author.\footnote{Generative AI technologies produce outputs based on “learning” statistical patterns in existing}

\footnote{Generative AI technologies produce outputs based on “learning” statistical patterns in existing}
generative AI systems by millions of Americans—and the resulting volume of AI-generated material—have sparked widespread public debate about what these systems may mean for the future of creative industries and raise significant questions for the copyright system.

Some of these questions relate to the scope and level of human authorship, if any, in copyright claims for material produced in whole or in part by generative AI. Over the past several years, the Office has begun to receive applications to register works containing AI-generated material, some of which name AI systems as an author or co-author. At the same time, copyright owners have brought infringement claims against AI data, which may include copyrighted works. Kim Martinneau, What is generative AI?, IBM Research Blog Apr. 20, 2023, https://research.ibm.com/blog/what-is-generative-ai (“At a high level, generative models encode a simplified representation of their training data and draw from it to create a new work that’s similar, but not identical, to the original data.”). The Office has defined “generative AI” and other key terms in a glossary of this Notice.


A History of Generative AI: From GAN to GPT-4, Marketchpost (Mar. 21, 2023), https://www.marketchpost.com/2023/03/21/a-history-of-generative-ai-from-gan-to-gpt-4/. Generative AI is also a point of contention in the labor disputes between the Alliance of Motion Picture and Television Producers and both the Writers Guild of America and SAG–AFTRA (the guild representing actors and other media professionals). See Andrew Webster, Actors say Hollywood studios want their AI replicas—for free, forever, The Verge (July 13, 2023), https://www.theverge.com/2023/7/13/ 23794224/ai-sag-actors-ai-image-rights.

4 See U.S. Copyright Office Review Board, Decision Affirming Refusal of Registration of A Recent Entrance to Paradise at 2 (Feb. 14, 2022), https://www.copyright.gov/ rulings-filings/review-board/docs/a-recent-entrance-to-paradise.pdf (noting visual work was submitted listing the author as the “Creativity Machine”).


7 Id.


9 Id.

The Copyright Office has long been engaged in questions involving machine learning and copyright. In 1965, the Office’s annual report noted that developments in computer technology had begun to raise “difficult questions of authorship”—namely the question of the authorship of works “‘written’ by computers.” As the then-Registrar of Copyrights observed:

The crucial question appears to be whether the “work” is basically one of human authorship, with the computer merely being an assisting instrument, or whether the traditional elements of authorship in the work (literary, artistic, or musical expression or elements of selection, arrangement, etc.) were actually conceived and executed not by man but by a machine.

Because the answer depends on the circumstances of a work’s creation, the head of the Office’s Examining Division (and future Register) Barbara Ringer warned that the Office could not “take the categorical position that registration will be denied merely because a computer may have been used in some manner in creating the work.”

As she noted, “a typewriter is a machine that is used in the creation of a manuscript[,] but this does not result in the manuscript being uncopyrightable.”

This view was echoed a decade later by the National Commission on New Technological Uses of Copyrighted Works (“CONTU”), which agreed with the Office but declined to discuss the issue in depth because “[t]he development of this capacity for ‘artificial intelligence’ has not yet come to pass, and, indeed, it has been suggested to this Commission that such a development is too speculative to consider at this time.” In the intervening years, as AI moved out of the realm of speculation, the Office continued to participate in discussions on AI issues, from a 1991 conference hosted by the World Intellectual Property Organization (“WIPO”) to more recent events the Office co-hosted with WIPO and with the U.S. Patent and Trademark Office.

Last year, in two separate copyright registration matters, the Office publicly addressed the question of copyright in AI-generated material. In the first instance, the Office refused to register a claim for two-dimensional artwork described as “autonomously created by a computer algorithm running on a machine.” The Office’s Review
Board 17 explained that the work could not be registered because it was made “without any creative input or intervention from a human author,” and that “statutory text, judicial precedent, and longstanding Copyright Office practice” all require human authorship as a condition of copyrightability.18 The Office’s registration denial, as well as the supporting legal analysis, was recently affirmed in federal district court.19

A second registration application, submitted in 2022, involved a work containing both human authorship and generative AI material. The work was a graphic novel with text written by the human applicant and illustrations created through the use of Midjourney, a generative AI system. After soliciting information from the applicant about the process of the work’s creation, the Office determined that copyright protected both the human-authored text and human selection and arrangement of the text and images, but not the AI-generated images themselves.20 The Office explained that where a human author lacks sufficient creative control over the AI-generated components of a work, the human is not the “author” of those components for copyright purposes. The Office continues to receive applications to register works incorporating AI-generated material, involving different levels of human contributions.22

III. The Office’s AI Initiative

In response to growing Congressional21 and public interest,24 the Office launched a comprehensive AI Initiative in early 2023. The Initiative identified a number of questions that the Office would take to further explore the copyright policy questions surrounding AI, including hosting public listening sessions and publishing a notice of inquiry.25 At the same time, the Office created a website, www.copyright.gov/ai, to provide information about the Initiative, including planned events and opportunities for public engagement.

a. March 2023 Registration Guidance

At the outset of the Initiative, the Office issued a statement of policy providing registration guidance on works containing AI-generated material (“AI Registration Guidance”).26 The AI Registration Guidance reiterated the principle that copyright protection in the United States requires human authorship. Under well-established case law, the Guidance explained, “the term ‘author,’ used in both the Constitution and the Copyright Act, excludes non-humans.”27 In the context of generative AI, this means that “[i]f a work’s traditional elements of authorship were produced by a machine, the work lacks human authorship and the Office will not register it.”28 The Guidance instructed applicants seeking to register works containing more than de minimis AI-generated material to disclose that the work contains such material and provide a brief explanation of the human author’s contributions.29

b. Public Listening Sessions

In April and May 2023, the Office held four public listening sessions to gather input on the copyright issues raised by generative AI. Each session focused on a different category of creative work: literary works, including print journalism and software; works of visual art; audiovisual works, including video games; and musical works and sound recordings. Over the four listening sessions, nearly 90 participants representing individual artists, academic experts, legal practitioners, technology companies, and industry associations shared their views with the Office. Transcripts, videos recordings, and agendas for each session are available on the Office’s website.30

c. Educational Webinars

In June and July 2023, the Office held two public webinars on generative AI, each of which drew an audience of nearly 2,000. The first webinar focused on registration of works containing AI-generated material. It included an overview of the Office’s general rules on how to register works containing material created or owned by someone other than the applicant, followed by examples illustrating how those rules apply to works that incorporate AI-generated material.33 The second webinar convened experts on different regions of the world to discuss international developments in generative AI and copyright law. These experts discussed how other countries are addressing copyright issues, including authorship, training, and exceptions and limitations. They provided an overview of legislative
developments and highlighted possible areas of convergence and divergence.32

d. Engagement With Stakeholders

In addition to the public events described above, the Office has spoken with a broad spectrum of stakeholders, participating in dozens of meetings with academics, trade groups, individual creators, technology companies, and creative industries.33 These meetings have provided valuable information on the technical aspects of generative AI models and systems, how creators are using generative AI, and the continuing questions copyright applicants have about registering works that include AI-generated material.

IV. The Current Inquiry

Drawing on our prior AI Initiative work, including discussions with stakeholders, the Office has identified a wide range of copyright policy issues arising from the development and use of AI. These relate to: (1) the use of copyrighted works to train AI models; (2) the copyrightability of material generated using AI systems; (3) potential liability for infringing works generated using AI systems; and (4) the treatment of generative AI outputs that imitate the identity or style of human artists. The Office seeks public comments on these and related issues.

As to the first issue, the Office is aware that there is disagreement about whether or when the use of copyrighted works to develop datasets for training AI models (in both generative and non-generative systems) is infringing.34 This Notice seeks information about the collection and curation of AI datasets, how those datasets are used to train AI models, the sources of materials ingested into training, and whether permission by and/or compensation for copyright owners is or should be required when their works are included.

32 The transcript and recording of the international webinar are available at https://www.copyright.gov/events/international-ai-copyright-webinar/.

33 Additionally, the Office has offered guidance to The Mechanical Licensing Collective (“The MLC”), explaining that AI-generated music is not eligible for the statutory mechanical blanket license in section 115 of the Copyright Act and that The MLC should not disburse royalties for such musical works. See Letter from Suzanne V. Wilson, General Counsel and Associate Register of Copyrights, U.S. Copyright Office, to Kris Ahrend, Chief Exec. Officer, The MLC, at 2–3 (Apr. 20, 2023), https://www.copyright.gov/USCO-Guidance-Letter-to-The-MLC-Letter-on-AI-Created-Works.pdf.

34 In some cases, a non-generative AI model may be trained on copyrighted material. In other cases, the same AI model may be capable of being deployed in both a generative AI system and a non-generative one. The Office’s consideration of training is framed broadly in order to encompass these and other situations.


36 For example, a system that facilitates the use of AI to generate content that is substantially similar to that of a pre-existing work without human intervention is infringing Human v. Meta Platforms, Inc., No. 4:22–cv–6823 (N.D. Cal.) (affirming the Office’s registration denial of AI-generated work).


38 The purpose of this Notice is to collect factual information and views relevant to the copyright law and policy issues raised by recent advances in generative AI. The Office undertakes this study pursuant to its statutory mandate in title 17 to “[c]onduct studies and “[a]dvise Congress on national and international issues relating to copyright, other matters arising under this title, and related matters.”39 It intends to use this information to advise Congress by providing analyses of the current state of the law, identifying unresolved issues, and evaluating potential areas for congressional action. The Office will also use this record to inform its regulatory work and to offer information and resources to the public, courts, and other government entities considering these issues.

The questions are grouped into several categories. This Notice begins with several general high-level questions and then inquires about AI training, including questions of transparency and accountability; generative AI outputs, including questions of copyrightability, infringement, and labeling or identification of such works; and other issues related to copyright. Because of the importance of using shared language in discussing these issues, the questions are followed by a glossary of key terms for the purposes of this Notice. The Office welcomes input from commenters on the definitions.

VI. Instructions and Questions

The Office does not expect that every party choosing to respond to this Notice will address every question raised below. The questions are designed to gather views from a broad range of parties. The Office does request that, when responding to a question, commenters clearly identify each


question for which they submit a response, address questions separately, and provide the factual, legal, or policy basis for their responses. Commenters should make clear whether they are submitting in a personal capacity or on behalf of an organization or entity they are authorized to represent. Commenters are particularly encouraged to explain any technical understandings that inform their legal and policy viewpoints, as well as whether their answers are applicable only to certain industries, technologies, or types of copyrighted works. Although some questions seek technical information about generative AI systems, commenters do not need to be affiliated with a technical entity to answer these questions.

General Questions

The Office has several general questions about generative AI in addition to the specific topics listed below. Commenters are encouraged to raise any positions or views that are not elicited by the more detailed questions further below.

1. As described above, generative AI systems have the ability to produce material that would be copyrightable if it were created by a human author. What are your views on the potential benefits and risks of this technology? How is the use of this technology currently affecting or likely to affect creators, copyright owners, technology developers, researchers, and the public?

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

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.

4. Are there any statutory or regulatory approaches that have been adopted or are under consideration in other countries that relate to copyright and AI that should be considered or avoided in the United States? How important a factor is international consistency in this area across borders?

5. Is new legislation warranted to address copyright or related issues with generative AI? If so, what should it entail? Specific proposals and legislative text are not necessary, but the Office welcomes any proposals or text for review.

Training

If your comment applies only to a specific subset of AI technologies, please make that clear.

6. What kinds of copyright-protected training materials are used to train AI models, and how are those materials collected and curated?

6.1. How or where do developers of AI models acquire the materials or datasets that their models are trained on? To what extent is training material first collected by third-party entities (such as academic researchers or private companies)?

6.2. To what extent are copyright works licensed from copyright owners for use as training materials? To your knowledge, what licensing models are currently being offered and used?

6.3. To what extent is non-copyrighted material (such as public domain works) used for AI training? Alternatively, to what extent is training material created or commissioned by developers of AI models?

6.4. Are some or all training materials retained by developers of AI models after training is complete, and for what purpose(s)? Please describe any relevant storage and retention practices.

7. To the extent that it informs your views, please briefly describe your personal knowledge of the process by which AI models are trained. The Office is particularly interested in:

7.1. How are training materials used and/or reproduced when training an AI model? Please include your understanding of the nature and duration of any reproduction of works that occur during the training process, as well as your views on the extent to which these activities implicate the exclusive rights of copyright owners.

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

7.3. Is it possible for an AI model to “unlearn” inferences it gained from training on a particular piece of training material? If so, is it economically feasible? In addition to retraining a model, are there other ways to “unlearn” inferences from training?

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

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

8.1. In light of the Supreme Court’s recent decisions in Google v. Oracle America and Andy Warhol Foundation v. Goldsmith, how should the “purpose and character” of the use of copyrighted works to train an AI model be evaluated? What is the relevant use to be analyzed? Do different stages of training, such as pre-training and fine-tuning, raise different considerations under the first fair use factor?

8.2. How should the analysis apply to entities that collect and distribute copyrighted material for training but may not themselves engage in the training?

8.3. The use of copyrighted materials in a training dataset or to train generative AI models may be done for noncommercial or research purposes. 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?

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

8.5. Under the fourth factor of the fair use analysis, how should the effect on the potential market for or value of a copyrighted work used to train an AI model be evaluated? What is the market for the output of AI models? What is the effect on or value of the output of AI models? What is the protected, copyrighted work to which the output of AI models may be in infringement?

8.6. Does the use of training data protected under copyright law in the training of an AI model constitute fair use under 17 U.S.C. 107? How important is the purpose of the use to the fair use determination? How important is the amount of use to the fair use determination?

8.7. Should fair use be considered in the copyright enforcement process, and if so, what role should it play?

8.8. How do you think the fair use doctrine should apply to the use of training data protected under copyright law in the training of an AI model?

8.9. Should the fair use doctrine be modified or expanded to address the unique circumstances of training AI models?

8.10. What is the potential market for or value of a copyrighted work used to train an AI model? Should fair use be considered in the copyright enforcement process, and if so, what role should it play?

8.11. How do you think the fair use doctrine should apply to the use of training data protected under copyright law in the training of an AI model?

8.12. Should the fair use doctrine be modified or expanded to address the unique circumstances of training AI models?
model be measured? 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?

9. Should copyright owners have to affirmatively consent (opt in) to the use of their works for training materials, or should they be provided with the means to object (opt out)?

10.1. Should consent of the copyright owner be required for all uses of copyrighted works to train AI models or only commercial uses? 47

10.2. Is a voluntary collective licensing scheme feasible or desirable approach? Are there existing collective management organizations that are well-suited to provide those licenses, and are there legal or other impediments that would prevent those organizations from performing this role? Should Congress consider statutory or other changes, such as an antitrust exception, to facilitate negotiation of collective licenses?

10.3. Should Congress consider establishing a compulsory licensing regime? If so, what should such a regime look like? What activities should the license cover, what works would be subject to the license, and would copyright owners have the ability to opt out? How should royalty rates and terms be set, allocated, reported and distributed?

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

10.5. Should licensing regimes vary based on the type of work at issue? What legal, technical or practical issues might there be with respect to obtaining appropriate licenses for training? Who, if anyone, should be responsible for securing them (for example when the curator of a training dataset, the developer who trains an AI model, and the company employing that model in an AI system are different entities and may have different commercial or noncommercial roles)?

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.

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

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

15.1. What level of specificity should be required?

15.2. To whom should disclosures be made?

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

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

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

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 used to train their models? Should creators of training datasets have a similar obligation?

Generative AI Outputs

If your comment applies only to a particular subset of generative AI technologies, please make that clear.

Copyrightability

18. Under copyright law, are there circumstances when a human using a generative AI system should be considered the “author” of material produced by the system? If so, what factors are relevant to that determination? For example, is selecting what material an AI model is trained on and/or providing an iterative series of

46 Collective licensing is one alternative to a direct licensing regime, in which copyright owners negotiate and enter into private agreements on an individual basis. Under a collective licensing arrangement, rights are aggregated and administered by a management organization. The management organization negotiates the terms of use and distributes payment to participating copyright owners. See WIPO, WIPO Good Practice Toolkit for CMRs at 6 (2021), https://www.wipo.int/publications/en/details.jsp?id=4561.

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/regstal071205.html.


49 Collective licensing is one alternative to a direct licensing regime, in which copyright owners negotiate and enter into private agreements on an individual basis. Under a collective licensing arrangement, rights are aggregated and administered by a management organization. The management organization negotiates the terms of use and distributes payment to participating copyright owners. See WIPO, WIPO Good Practice Toolkit for CMRs at 6 (2021), https://www.wipo.int/publications/en/details.jsp?id=4561.

50 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/regstal071205.html.


26. If a generative AI system is trained on copyrighted works containing copyright management information, how does 17 U.S.C. 1202(b) apply to the treatment of that information in outputs of the system?
27. Please describe any other issues that you believe policymakers should consider with respect to potential copyright liability based on AI-generated output.

Labeling or Identification
28. Should the law require AI-generated material to be labeled or otherwise publicly identified as being generated by AI? If so, in what context should the requirement apply and how should it work?
29. Who should be responsible for identifying a work as AI-generated?
30. Are there technical or practical barriers to labeling or identification requirements?
31. 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?
32. What tools exist or are in development to identify AI-generated material, including by standard-setting bodies? How accurate are these tools?
33. What are their limitations?

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?
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?
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?
33. With respect to sound recordings, how does section 114(b) of the Copyright Act relate to state law, such as state right of publicity laws? Does AI (July 18, 2023), https://ai.meta.com/llama/license/ (requiring users of Llama 2 AI model to include an attribution notice and excluding use in services with greater than 700 million monthly active users)

VII. Glossary of Key Terms

The Office has included definitions of key terms as they are used in this Notice to clarify the technical processes involved in generative AI systems. The following definitions are used for purposes of this Notice only; they do not necessarily reflect the government’s legal position with respect to any particular term.

Artificial Intelligence (AI): A general classification of automated systems designed to perform tasks typically associated with human intelligence or cognitive functions. Generally, AI technologies may use different techniques to accomplish such tasks. This Notice uses the term “AI” in a more limited sense to refer to technologies that employ machine learning, a technique further defined below.

AI Model: A combination of computer code and numerical values (or “weights,” which is defined below) that may be used to produce outputs from inputs. In this Notice, the term “AI model” generally refers to a software product or service that substantially incorporates one or more AI models and is designed for use by an end-user. An AI system may be created by a developer of an AI model, or it may incorporate one or more AI models developed by third parties.

Generative AI: An application of AI used to generate outputs in the form of expressive material such as text, images, audio, or video. Generative AI systems may take commands or instructions

independent fixation of other sounds, even though such sounds imitate or simulate those in the copyrighted sound recording.

54 Under 17 U.S.C. 114(b), the reproduction and derivative works rights for sound recordings “do not extend to the making or duplication of another sound recording that consists entirely of an

53 Some AI models are released by their developers for download and use by members of the general public. Such so-called “open-source” models may restrict how those models can be used through the terms of a licensing agreement. See, e.g., Llama 2 Community License Agreement, Meta


51 17 U.S.C. 117–263, Sec. 7223(4)(A), 136 Stat. 2395, 3669 (2022) (defining “artificial intelligence system” as “any data system, software, application, tool, or utility that operates in whole or in part using dynamic or static machine learning algorithms or other forms of artificial intelligence”).
from a human user, which are sometimes called “prompts.” Examples of generative AI systems include Midjourney, OpenAI’s ChatGPT, and Google’s Bard.

Machine Learning: A technique for building AI systems that is characterized by the ability to automatically learn and improve on the basis of data or experience, without relying on explicitly programmed rules.57 Machine learning involves ingesting and analyzing materials such as quantitative data or text and obtaining inferences about qualities of those materials and using those inferences to accomplish a specific task. These inferences are represented within an AI model’s weights.

Training Datasets: A collection of training material (as defined below) that is compiled and curated for use in machine learning. Examples of training datasets include BookCorpus, ImageNet, and LAION.

Training Material: Individual units of material that are used for purposes of training an AI model. They may include a combination of text, images, audio, or other categories of expressive material, as well as annotations describing the material. An example of training material would be an individual image and an associated text “label” that describes the image.

Weights: A collection of numerical values that define the behavior of an AI model. Weights are stored within an AI model and reflect inferences from the training process.


Suzanne V. Wilson,
General Counsel and Associate Register of Copyrights.

Maria Strong,
Associate Register of Copyrights and Director of Policy and International Affairs.

[FR Doc. 2023–18624 Filed 8–29–23; 8:45 am]