The Copyright Licensing Agency Ltd

Madeleine Pow-Jones, Policy and Contracts Manager

1 St Katharine's Way

London

E1W 1UN

Date: 30 October 2023

Response to US Copyright Office Notice of inquiry and request for comments concerning Artificial Intelligence and Copyright made by a collective management organisation, The Copyright Licensing Agency Ltd (www.cla.co.uk)

Introduction

CLA is the recognised UK collective management organisation (CMO) in the UK and is a regulated not-for-profit organisation which licenses organisations, in the UK and, in certain cases overseas, to lawfully use, copy, and share text and image-based content owned by authors, publishers, and visual artists. Revenues are distributed to CLA's member-owners (more below), ensuring fair compensation for rights holders and support for the UK's creative economy.

CLA has been providing licences, as well as a growing range of related services, that simplify copyright and make it easier to access content for over 40 years.

Collective licensing is a cost-effective blanket licensing solution and offers a practical alternative where it is not easy to license on an individual basis for specific uses due to the volume of rightsholders/users and the scale of use. Since it is not possible to take account of the exact rights ownership of each extract which may be copied or used, the licence fees are shared between all the relevant rightsholders.

CLA has four members: Authors Licensing and Collecting Society Ltd (ALCS), Design and Artists Copyright Society (DACS), PICSEL Ltd (Picture Industry Collecting Society for Effective Licensing) and Publishers' Licensing Services Ltd (PLS) and distributes the revenue it collects to its members, who in turn distribute to authors, publishers and visual artists.

CLA is a member of the International Federation of Reproduction Rights Organisation (IFRRO) and works closely with our counterpart in the USA, Copyright Clearance Centre (CCC).

General Questions

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

Whilst we recognise the opportunities for efficiencies and innovation with the expanding development and use of generative AI systems, the technology has created a number of new societal, legal and compliance risks and is having a detrimental impact upon the creators and rightsholders (such as the authors, artists and publishers) that CLA represents. These are current, ongoing issues and present significant risks. We detail these risks, and their mitigation as follows:

a) Unauthorised use of copyright protected works

In order to train a generative AI system, an AI developer makes a digital copy of a piece of data, often scraped from the internet, which is then fed into the system for ingestion. A number of companies developing AI systems across the world have made copies of copyright-protected works for use as training data without permission from creators and rightsholders, and in many cases using pirated works. In the UK, and to the best of our knowledge in the USA, this is an infringement of copyright and has caused economic loss and damage.

To comply with the law, AI developers must secure permission for the use of copyright protected works and enter into appropriate licence agreements with rightsholders prior to ingestion and use. Such agreements must provide adequate compensation to creators and rightsholders for the use of their material as well as the right for creators to be credited for such use.

The exceptions and limitations within the UK Copyright, Designs and Patents Act (CDPA) are generally based on a 'fair dealing' approach, rather than the US approach of 'fair use'. In 2014 an exception to the CDPA was introduced for text and data mining for non-commercial research[1].

b) Lack of transparency

The lack of transparency around the development of generative AI systems undermines the creative sector and the livelihoods of creators and rightsholders.

It is essential for the responsible development of generative AI systems that the provenance of training data, together with an explanation of how the AI system has

[1] S29(a) of the UK CDPA https://www.legislation.gov.uk/ukpga/1988/48/section/29A

been developed, is made available in an accessible and understandable form. It is vital that rightsholders can establish where and how their content is used and to seek fair compensation and remuneration accordingly.

Creators and rightsholders' works must be accurately attributed when included in generative AI outputs.

c) Weakening of creator and rightsholder freedom

Creators and rightsholders must have the choice as to whether their works are included in generative AI systems or not, prior to ingestion taking place. It must be simple for the creators and rightsholders to be able to communicate whether they want their works included or not. Consent is essential to developing copyright compliant AI systems. Licensing should be undertaken on a voluntary basis and licences may be collective (such as those offered by CLA in the UK and CCC in the USA) or direct.

d) Lack of trustworthiness

The quality of the information used to train generative AI systems will determine the quality of the system's output. Those AI systems trained on unauthorised and unreliable content will produce untrustworthy and unreliable outputs, with the risk of the creation and dissemination of misinformation as well as the inclusion of various forms of bias and potential discrimination.

Through greater transparency, AI developers can help build trust and show that the data being used is accurate and well sourced. In addition, AI developers must also provide accessible information as to how they mitigate and correct for elements such as bias and false information within their systems.

e) Labelling of AI-generated works

The outputs from generative AI systems must be clearly labelled as AI-generated outputs to avoid causing public confusion. There is a risk to creators and rightsholders, of losing work, both in terms of losing the opportunity to produce new works and losing control over maintaining the integrity of their works, as well as a risk of imitation – this may have the effect of causing serious and irreparable reputational damage and economic loss. There are examples of generative AI systems being used to create works in the style and name of a known author but which an author has no knowledge or involvement of. Without clear identification as to whether a work is human-created, AI-created or hybrid, this will continue to proliferate. There are numerous examples of the risks posed by AI-generated works, for example:
https://www.theguardian.com/technology/2023/sep/01/mushroom-pickers-urged-to-avoid-foraging-books-on-amazon-that-appear-to-be-written-by-ai

f) Risk to licensing and creator livelihoods

From a UK perspective, there is a risk that the continued unauthorised use of copyright works in generative AI systems and a failure to uphold compliance with the copyright legal framework in the UK, means existing and potential new markets for copyright licensing, whether individual or collective, will be undermined. The consequences of this will be the loss of important income for creators and rightsholders. The authors, visual artists and publishers that CLA represent rely on the revenue generated from the licensing of their copyright works to not only make a living but to support the investment of time and resources needed to create new original works.

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?

   Please see our responses to question 1 above.

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.

   CLA has recently published a set of principles for Copyright and Generative AI: https://cla.co.uk/principles-for-copyright-and-generative-ai/

   CLA is a member of the International Federation of Reproduction Rights Organisations (IFRRO) and supports its position statement on Artificial Intelligence: https://ifrro.org/resources/documents/General/V2_IFRRO_AI_Position_Statement_29_March_2023.pdf

   Collective Licensing and AI: A Memo by Prof. Daniel Gervais for CCC (https://www.copyright.com/landing/ccc-collective-licensing-for-ai-memo-danielgervais/)

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?

There are divergent approaches emerging to regulating AI globally. In the UK, the Government consulted on its AI Regulation White Paper[2] in June 2023, with the intention of delivering a "pro-innovation regulatory framework". The UK Government has, to date, adopted a context specific approach rather than establishing a central regulator for AI. There is therefore a question as to how each regulator will consider intellectual property and how issues such as transparency and labelling may ultimately diverge between regulators. We are concerned therefore with the approach of the UK Government as the rights of 'content-producers' have been expressly excluded from the White Paper and thus the scope of the following consultation. CLA believes that this approach is flawed. The rights of creators and rightsholders are integral to the development, and subsequent success, of generative AI systems and must therefore be considered in any regulation.

In parallel to the pro-innovation regulatory framework consultation, the UK Intellectual Property Office is working to develop a voluntary Code of Practice for copyright and generative AI which will seek to set out certain good-practice principles. This follows the UK government's proposal in June 2022 to introduce an exception for text and data mining which covered all uses, which the UK Government subsequently withdrew in February 2023. Whilst the discussions relating to the Code of Practice between copyright holders and AI companies are beneficial, the nature of a voluntary code means that it is likely that further action will be needed particularly in regard to regulating generative AI systems in relation to their ingestion and use of copyright-protected works.

The UK Government's Department for Science, Innovation and Technology published an interim report in August 2023, which recognises that "policy must establish the rights of the originators of this content, and these rights must be enforced."[3] Amongst other areas of regulation of AI, national governments must ensure compliance with intellectual property law and that developers of AI models have the appropriate licences in place for the use of copyright-protected works, and that the creators and rightsholders whose works are used by generative AI systems are fairly remunerated.

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.

   For the development and adoption of generative AI to be safe, ethical, legal and successful, as noted above, there needs to be regulation of AI systems. The UK

[2] https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper
[3] https://committees.parliament.uk/publications/41130/documents/205611/default/

Government's pro-innovation approach to regulating AI sets out five key principles to guide the responsible development of AI. The five principles are as follows:

- Safety, security and robustness
- Appropriate transparency and explainability
- Fairness
- Accountability and governance
- Contestability and redress

Currently, the UK Government does not intend to cover the rights of 'content creators' in its regulation of AI which, as noted in our response to Q.4 is flawed. The works of rightsholders and the use of those works are integral to the successful development of AI and any legislation must be applied and upheld, and regulation must therefore take the rights of creators and rightsholders into account. Copyright law in the UK (and to the best of our knowledge in the USA) is flexible and able to deal with the challenges posed by emerging technologies, as it has done so in the past. Voluntary licensing, whether direct or collective, has proven to be able to evolve to deal with technological developments.

Training

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

There is a lack of transparency as to what materials have been used to train AI models to date, and how these materials have been collected and curated. It is well understood that data and materials used to train AI models have been scraped from the internet, without appropriate permission. Based on the information that is available CLA has undertaken some initial research (based on the data available in the article published in The Atlantic.[4]) We have identified that at least 25% of the Books3 database mentioned in the article are currently represented by CLA.

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)?
N/A

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?

[4] https://www.theatlantic.com/technology/archive/2023/08/books3-ai-meta-llama-pirated-books/675063/

With regards to generative AI, CLA is in the process of discussing and consulting with its members (authors, publishers and visual artists) to develop and launch voluntary collective licences for generative AI. It may be that the rightsholders that CLA represents licence their works individually for this purpose on a direct basis, and it is critical that creators and rightsholders continue to have a choice - whether to licence or not, and on a collective or individual, direct basis.

CLA is also in the process of developing a stand-alone licence for text and data mining for non-generative AI purposes.

To date, no AI firm has approached CLA to seek a licence.

Whilst it does not cover generative AI, CLA offers a text and data mining licence to Media Monitoring Organisations (MMOs). CLA licenses MMOs operating in the UK and overseas for indexing and TDM of magazine and news website content. The licence permits MMOs to carry out various restricted acts protected by copyright on relevant website content, including 'accessing and copying web pages … using so-called "web-crawler", "spider" or "robot" software or other automated and/or manual review methods, processes or means in order to derive and store searchable index entries' (i.e. TDM) and permitting MMOs to 'store the indexed material … in order to search, retrieve and evaluate (by automated and/or manual review methods, processes or means) [the material]'.

The media monitoring industry has grown significantly in the last fifteen years and is now worth £3.8 billion. MMOs generate revenue from a suite of products and services all of which are based, to a varying degree, on published content. The MMOs licensed by CLA are generating significant and growing revenue from the analysis and insight provided by mining published copyright-protected content. The licence offered by CLA, and the revenue collected from MMOs ensures that MMOs have the permissions needed to develop their business and serve their customers, and rightsholders whose works are used are fairly remunerated for the use of their works.

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?

The current lack of transparency from the developers and deployers of generative AI systems means it is very difficult to establish exactly what information a model has been trained on and the sources of that information. However, it has been established that copyright-protected works have been widely used for training purposes.

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.

N/A

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?

For the responsible development of AI systems, and to ensure that the process by which AI models are trained is safe, ethical and legal (and therefore responsibly developed), the developers of AI systems should be required to adopt principles similar to those required by data protection law in the UK and Europe – namely, privacy by design. Those training AI systems should evaluate at the outset the provenance of training data, if permission from rightsholders to use the data for the intended purpose has been secured and ensure rightsholders have the option to choose whether their works are included or not, and include clear attribution where permission is granted.

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 non-commercial 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 non-commercial or research uses is provided by for-profit developers of AI systems?

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

8.5. Under the fourth factor of the fair use analysis, how should the effect on the potential market for or value of a copyrighted work used to train an AI model be measured? 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?

N/A

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)?

Rightsholders should have the right to choose whether their works are used or not. The choice may be dependent on the AI model, and it can be either opt-in or opt-out and both should be possible. CLA offers collective licences for internal corporate or educational uses that operate predominantly an opt-out basis, but on an opt-in basis where this is a necessary condition for rightsholders. One of the benefits of collective licensing is that it can offer such flexibility.

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

Consent must be required for all uses of copyright-protected works to train AI models.

The high commercial value of AI models means that introducing an exception to copyright permitting training for (initially) non-commercial purposes without consent is not appropriate, given that it would almost certainly contravene the three-step test as set out in the Berne Convention and TRIPS. The non-commercial research nature of the initial training would be beside the point, as future uses of any AI tool using this research would most likely 'reasonably prejudice the legitimate interests of the rights holder'.

9.2 If an "opt out" approach were adopted, how would that process work for a copyright owner who objected to the use of their works for training? Are there technical tools that might facilitate this process, such as a technical

flag or metadata indicating that an automated service should not collect and store a work for AI training uses?

The majority of CLA's voluntary blanket licences operate on an opt-out basis where rightsholders communicate to CLA, via its member organisations, if they wish their works to be excluded from a particular licence, and this can be as granular as an identified text work. What is covered by the licence is managed by CLA's repertoire look-up tool, Check Permissions, (https://cla.co.uk/resources/tools/check-permissions/) which conveys whether a work is included or not to rightsholders.

9.3. What legal, technical, or practical obstacles are there to establishing or using such a process? Given the volume of works used in training, is it feasible to get consent in advance from copyright owners?

CLA believes it is feasible to get consent in advance from rightsholders. Generative AI models must take responsibility for this. By way of example, collective licensing, as operated by CLA, is a cost-effective blanket licensing solution and offers a practical alternative where it is not easy to license on an individual basis for specific uses due to the volume of rightsholders/users and the scale of use. Collective management organisations are therefore well placed to offer voluntary licences for generative AI.

9.4. If an objection is not honored, what remedies should be available? Are existing remedies for infringement appropriate or should there be a separate cause of action?

9.5. In cases where the human creator does not own the copyright—for example, because they have assigned it or because the work was made for hire—should they have a right to object to an AI model being trained on their work? If so, how would such a system work?

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?

CLA cannot speak for all creative sectors but direct voluntary licensing is feasible within the text and image sector, by way of example:

BRIA AI: https://bria.ai/

Heni Editions and Damien Hirst: https://heni.com/nft/more-info/the-beautiful-paintings

vAIsual: https://vaisual.com/datasets/

10.2. Is a voluntary collective licensing scheme a 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?

Voluntary collective licensing schemes are both feasible and desirable in certain circumstances where rightsholders believe it is an appropriate solution to allow them to control the use of their works and obtain appropriate compensation. As referenced in our response to 6.2, CLA already offers a voluntary collective licence which grants text and data mining (TDM) rights to Media Monitoring Organisations, and in consultation with rightsholders is in the process of developing licences for TDM (non-generative AI and generative AI). In the UK, CLA is well placed to develop and offer such licences, and the current copyright framework in the UK supports this approach, as well as direct licensing.

CLA works closely with its counterpart in the USA, Copyright Clearance Centre (CCC). CLA and CCC have been in partnership for almost 40 years and CCC is well placed to offer a similar collective licensing solution in the USA.

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?

N/A

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

Extended collective licensing may be a feasible and desirable approach, but it is dependent on whether rightsholders believe this to be an appropriate approach, and appropriate safeguards, such as transparency and a robust authorisation process being in place. Regulations for extended collective licensing exist in the UK (although, due to the complexities of the UK's exit from the European Union, no scheme has yet been approved). It is possible that it may be utilised to develop a licensing scheme for generative AI. For extended collective licensing to be feasible, a collective management organisation must be able to demonstrate, at least, that it is representative of the relevant class of rightsholders, and rightsholders must be able to exercise their right to opt-out of the licensing scheme.

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

In order for a licensing regime to be successful it must be flexible by design to maintain rightsholder choice and provide the most appropriate solution to users. Regimes may therefore vary based on the type of work, if such an approach is what works best. It is

unlikely that there will be a 'one-size-fits-all' solution for generative AI models, and rightsholders, users and regulators need to be mindful of this, ensuring flexibility.

11. What legal, technical or practical issues might there be with respect to obtaining appropriate licenses for training? Who, if anyone, should be responsible for securing them (for example when the curator of a training dataset, the developer who trains an AI model, and the company employing that model in an AI system are different entities and may have different commercial or non-commercial roles)?

The responsibility for obtaining appropriate licences for training should be with the developers of the AI system and, at all points in the chain – whether the developers, curators or a company employing an AI model – they must have due regard to whether the system they are using is compliant with all relevant laws.

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.

N/A

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

CLA are unable to comment on what the economic impacts of a licensing requirement will be directly, but it has to be noted that obtaining appropriate copyright permissions is a cost of doing business, and AI models should not be benefitting economically to the economic loss of the creative sectors.

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

N/A

Transparency & Recordkeeping

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

15.1. What level of specificity should be required?

15.2. To whom should disclosures be made?

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

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

As stated in CLA's principles for Copyright and Generative AI, it is essential that developers of generative AI systems are transparent so as to clearly show what works have been ingested and used for training, as well as transparent as to how the AI system has been developed. Developers of AI models and creators of training datasets must be required to collect, retain and disclose their records. From a rightsholders perspective, the level of specificity should be sufficient to ensure that individual works can be identified (for text works, this is typically by ISBN / ISSN). Disclosures of works ingested and used to train should be made available, and this helps ensure that rightsholders have knowledge of where their works are used, and a subsequent right to object or accept this.

Developers of AI systems that incorporate models from third parties should have similar obligations and should undertake due diligence to ensure that the model they are incorporating maintains high standards of record-keeping.

Transparency regarding the materials used to train AI models ensures compliance with copyright law, that rightsholders can be fairly remunerated for the use of their works and helps to mitigate the risk of bias within AI models and systems. Transparency helps uses to make informed decisions as to the trustworthiness and ethics of such systems. The ability to scrutinise the systems benefits not only rightsholders, but also users of the systems.

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

For models that have already been trained, there must be an obligation to notify all rightsholders if their work has been used with an option provided by the developer for the rightsholders to opt out of the training of future systems. For the future development of AI models, permission must be sought from rightsholders before ingestion and training takes place and there must the opportunity for rightsholders to opt-out in in an accessible and simple way. Permission may be sought from rightsholders directly, or may be available via a collective licence, as discussed above. The model is then trained only on works where permission has been granted.

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?

N/A

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?

N/A

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?

N/A

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?

N/A

> 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?

N/A

21. Does the Copyright Clause in the U.S. Constitution permit copyright protection for AI-generated material? Would such protection "promote the progress of science and useful arts"? If so, how?

N/A

Infringement

22. Can AI-generated outputs implicate the exclusive rights of preexisting copyrighted works, such as the right of reproduction or the derivative work right? If so, in what circumstances?

N/A

23.Is the substantial similarity test adequate to address claims of infringement based on outputs from a generative AI system, or is some other standard appropriate or necessary?

N/A

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

N/A

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

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

N/A

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?

N/A

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

N/A

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?

As detailed in our response to Q.1, fully AI-generated material should be clearly labelled as such. Not doing so presents serious risks to creators and rightsholders.

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

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

28.3. If a notification or labeling requirement is adopted, what should be the consequences of the failure to label a particular work or the removal of a label?

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?

N/A

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?

N/A

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?

N/A

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?

N/A

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 this issue require legislative attention in the context of generative AI?

N/A

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