

JAPAN INTELLECTUAL PROPERTY ASSOCIATION

Nihombashi 3-Chome Square Bldg.
6F, 9-1 Nihonbashi 3-chome,
Chuo-ku, Tokyo 103-0027, JAPAN



TEL: +81-3-6262-1570
FAX: +81-3-6262-5655
URL: <http://www.jipa.or.jp/>

Oct 30th, 2023

Dr Shira Perlmutter

Register of Copyrights and Director, U.S. Copyright Office
101 Independence Ave. S.E.
Washington, D.C.
20559-6000

Dear Dr Shira Perlmutter,

Re: Request for comment on Artificial Intelligence and Copyright by U.S. Copyright Office

We, the Japan Intellectual Property Association “JIPA”, are a private user organization with about 990 major Japanese companies as members. When appropriate opportunities arise, we offer our opinions on the intellectual property system of other countries and make recommendations for more effective implementation of the systems.

Having learned a request for comment on Artificial Intelligence and Copyright on your website, we would like to offer our opinions as follows.

Your consideration on our opinions would be greatly appreciated.

Sincerely yours,

Toru Ito
Managing Director
Japan Intellectual Property Association

JIPA Comment:

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?

In Japan, issues regarding copyright and AI are sorted out into "(1) learning phase" and "(2) use phase", in order to elaborate the discussion.

As for (1) the learning phase, when the use of the work is deemed to be not for purpose of enjoyment, the use of copyrighted works is organized to be subject to restrictions on rights to the extent necessary, unless it does not unreasonably prejudice the benefits of rightsholders.

In addition, (2) in the use phase, "reproduction" applicability (judgement regarding "basis" that is a requirement for the authorization of reproduction) is particularly problematic when copyrighted works similar to those in learning datasets are output in a way unintended by a developer or a user.

We believe that sorting the issues into those phases is also effective in the United States.

5. Is new legislation warranted to address copyright or related issues with generative AI? If so, what should it entail?

(Although this not only can be said for this question but also for multiple questions below relating to whether laws and regulations need to be developed,)

As AI is a rapidly changing area, it is desirable to strike a balance between protection and use. by focusing more on soft law (for example, measures by guidelines) that can be flexibly implemented, or by architecture (implementation of opt out based on the rightsholder's will, and mechanisms to prevent output of content similar to that of others, and the like using technology) rather than hard law that requires time for enactment and enforcement.

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?

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?

(To answer 9.2 and 9.3 collectively), For websites, it is possible to disallow crawling on one's own site for certain or all crawlers by utilizing a disallow tag of robot.txt, and this is conceived to function to a certain extent as an opt out by the rightsholder. However, it should be noted that (1) the web site installer are not necessarily the same as the rightsholder, (2) robot.txt is originally created to manage crawling for the purpose of search and basically

cannot distinguish crawling for the purpose of search and crawling for the purpose of machine learning from each other,(3) disallow for all crawlers would also disqualify the site from search,(4)the legal assessment without regard to robot.txt is unclear, and (5) the introduction of new opt-out mechanisms specific to machine learning is subject to a major global standardization effort.

10.1. Is direct voluntary licensing feasible in some or all creative sectors?

Since there is no one-size-fits-all approach to generative AI, and various legal risks may be taken into account in some cases depending on the specific content of the AI learning model and interpretation of copyright law, we believe that legal safety for the user and an appropriate return of compensation to the rightsholder can be balanced by a licensing scheme of the rightsholder.

Where the rightsholder voluntarily provides a licensing scheme under reasonable conditions, no rights limitations are applied (Fair Use is not applied) to use that may harm the market of the license, and the user needs to get permission from the rightsholder. especially in the case of fine tuning, where there may be inherent high value in the work being learned, the license scheme may provide the rights holder with an appropriate return of consideration.

As this issue is an area where the interpretation of copyright law is involved in the effectiveness of the licensing scheme, accumulation of judicial precedents and discussions regarding legal interpretation are awaited first.

In any case, whether the rightsholder wants such license needs to be examined for each type of content.

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?

While it is important to respect the rights of creators in the first place, a company that performs fine tuning of a foundation model, for example, may input content that falls under trade secrets. In addition, the type of data used for learning itself may have commercial value, and hence it is important to appropriately protect such values that should be kept safe by companies from the standpoint of appropriately facilitating corporate activities.

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?

We believe that at least those generated with AI by a user intentionally trying to create something similar to certain copyrighted works can conflict with the right of reproduction, the derivative work right, and the like and

infringe rights unless rights limitation provisions are applied. Meanwhile, it is debatable whether those works conflict with such rights, when copyrighted works similar to those in datasets for learning are output in a way unintended by a developer or a user.

In Japan, in order to conflict with the right of reproduction and the derivative work right, the copyrighted work of concern requires to be created on the "basis" of the original copyrighted work, but in what cases are understood to be the "basis" (whether the internal intention of use of the user is a requirement or whether a conflict with the right of reproduction and the like occurs on the "basis" of datasets for learning when there are similar works in the datasets for learning and those works are used, and what "use" means in that case) is under consideration by the government.

End of Document