

NO&T IP Law Update

知的財産法ニュースレター

July, 2023 No.3

Japan a Paradise for Machine Learning, Not Generative AI

Kenji Tosaki
Takahiro Hatori
Tomoki Abe

I. Introduction

Generative artificial intelligence (generative AI) has captured the attention of the world. We have recently seen the launch of a wide variety of generative AI services, some of which have even become familiar in our daily lives and at work. For example, a new image or portrait can be easily generated by using services such as “Stable Diffusion” or “Imagen”. Also, “Chat GPT” (Chat Generative Pre-training Transformer), which has already had an enormous impact, enables us to obtain response to almost any question posed to an artificial intelligence in natural language.

In tandem with the development of these generative AI services and technologies, there has been worldwide discussion about various related legal issues. Japan has had its own lively discussions on the subject, including consideration of the relation between generative AI and copyright law.

In light of Article 30-4 of the Copyright Act of Japan, Japan is sometimes described as a “Machine Learning Paradise”. But can the same Article support the description of Japan as a “Generative AI Paradise”?

Copyright infringement in relation to generative AI may take place in two situations because acts of exploitation defined in the Copyright Act may take place in the two situations. One is when developing generative AI. In making generative AI learn a work, datasets for training are created and the datasets are input into a training program. Acts of reproduction of copyrighted work, which are designated as acts of exploitation in the Copyright Act, are likely to take place in these processes. The other is when using generative AI. In generating works by using generative AI, acts of reproduction, acts of adaptation and/or acts of public transmission of copyrighted work, which are designated as acts of exploitation in the Copyright Act, may take place. For this reason, we would like to explain (x) whether acts of exploitation when developing generative AI constitute copyright infringement in **Part II.** below, and (y) whether acts in the course of producing works by using a generative AI fall under acts of exploitation of existing copyrighted works in **Part III.** below.

II. Possibility of Copyright Infringement in the Course of the Development of Generative AI**1. Overview**

When developing generative AI, generative AI developers (1) copy, collect and preprocess large amounts of copyright-protected and non-copyright-protected data, (2) create datasets for training, and (3) input these datasets into a training program (See, Fig. 1 below). In the course of the processes (1) and (2) above, it is likely that acts of reproduction of copyrighted works take place. An issue that presents itself at this stage is whether the acts of reproduction of a copyrighted work for use in generative AI development constitutes copyright infringement of such work.

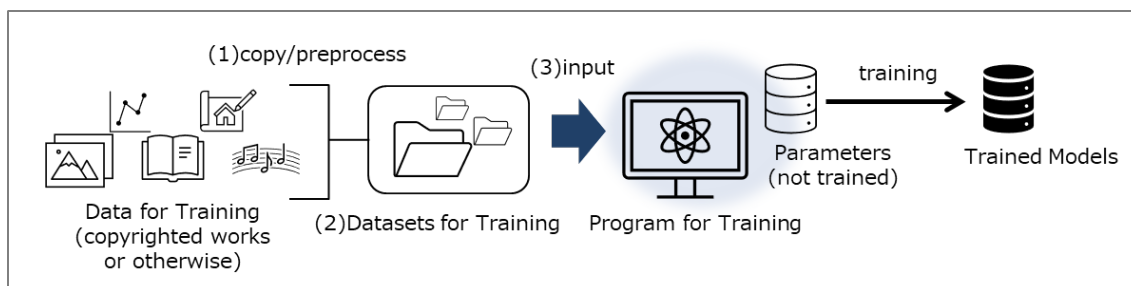


Fig. 1

Under the Copyright Act, an act of making a given work that is identical or similar to an existing copyrighted work in terms of original expression in reliance on the said existing copyrighted work is considered to fall under “reproduction” under the Copyright Act and therefore, constitutes, in principle, copyright infringement. When developing generative AI, the data from a large volume of copyrighted work is duplicated for use in such development and it is considered that an act of such duplication falls under “reproduction” under the Copyright Act. The key question concerning copyright infringement at this stage becomes whether Article 30-4 of the Copyright Act applies.

2. Applicability of Article 30-4 of the Copyright Act

In light of the above, it is necessary to consider the applicability of Article 30-4 of the Copyright Act, **which is just the article that caused Japan to be described as a “Machine Learning Paradise.”**

Under this Article, a work may be exploited in any way and to the extent considered necessary, (I) in cases in which it is not a person’s purpose to personally enjoy or cause another person to enjoy the thoughts or sentiments expressed in that work. However, this exception does not apply (II) if the action would unreasonably prejudice the interests of the copyright owner in light of the nature or purpose of the work or the circumstances of its exploitation. Please be advised that Article 30-4 lists three examples of cases in which it is not a person’s purpose to personally enjoy or cause another person to enjoy the thoughts or sentiments expressed in a given work, one of which is the exploitation of a work for use in data analysis (meaning the extraction, comparison, classification, or other statistical analysis of the constituent language, sounds, images, or other elemental data from a large number of works or a large volume of other such data).

In determining whether Article 30-4 of the Copyright Act applies, the most important considerations are (1) whether, as set out in (I) above, the purpose of the exploitation of a work is “to enjoy the thoughts or sentiments expressed in that work” (hereinafter referred to as the “**Non-Enjoyment Purpose Requirement**”) and, (2) if the answer to (1) is negative, whether such exploitation falls under (II) above, concerning unreasonable prejudice to the interests of the copyright owner.

(1) The Non-Enjoyment Purpose Requirement

Whether the Non-Enjoyment Purpose Requirement is satisfied is determined with regard to whether the exploitation of the copyrighted work at issue has, as its intent, the satisfaction of one’s intellectual desire or spiritual needs, through an action such as viewing. The following three examples illustrate the Non-Enjoyment Purpose Requirement.

Example 1: A person collects and preprocesses various copyrighted landscape pictures (including copying these pictures) for the purpose of developing AI for creating 3D CGI movies, which movies have the “essential features of expression” of such copyrighted works.¹

In this case, it is highly likely that a Japanese court would determine that **the Non-Enjoyment Purpose Requirement has NOT been satisfied, and that a copyright infringement has been established.** This is because,

¹ See, Copyright Division, Agency of Cultural Affairs “*AI to Chosakuken no Kankeitou ni tsuite*” (Relationship between AI and copyrights, etc.).

while the purpose of the developer in this example is to use files containing these pictures for the purpose of data analysis (i.e., machine learning to develop AI), the developer also intends to create 3D CGI movies, and to watch these movies or show them to other humans. The court thus would be highly likely to find that this intent constitutes the purpose to personally enjoy or cause another person to enjoy the thoughts or sentiments expressed in those works.

Example 2: A person uses copyrighted works (e.g., pictures, music, and literary works) only for training a generative AI, and configures this generative AI only to produce works that are not identical or similar to any copyrighted elements of such copyrighted works.

While this hypothetical is a bit less certain than the previous, it is more likely than not that the court would determine that **the Non-Enjoyment Purpose Requirement is satisfied**, although we do not know if one can configure a generative AI only to produce works that are not identical or similar to any copyrighted work that the AI learned in terms of original expression. This is because in this case, the works that the generative AI produce do not have the “essential features of expression” of the copyrighted works that the AI learned and one cannot enjoy the thoughts or sentiments expressed in the copyrighted works which he uses for AI training.

Example 3: A person copies copyrighted works (e.g., pictures, music, and literary works) in order to create datasets to be used in the development of AI, after which he sells these datasets to third parties. He does not carry out this activity with the purpose of allowing others to use these datasets for the development of generative AI which will generate thoughts or sentiments expressed in the copyrighted works.²

In this case, the purpose of the dataset creator is not to enjoy thoughts or sentiments expressed in the collected copyrighted works, but only to profit by selling datasets built on data analysis thereof. However, unless the dataset creator confirms that the third party’s purpose is a non-enjoyment purpose, the court will likely find that the dataset creator’s purpose includes a purpose to cause another person to enjoy the thoughts or sentiments expressed in a work. Therefore, if (i) each of the third party discloses his/her specific non-enjoyment purpose of using the dataset to the dataset creator, (ii) there are reasonable grounds for the dataset creator to believe that each of the third party’s purpose is non-enjoyment purpose and (iii) all the third parties, users of the datasets, also specifically agree with using the datasets only for the specifically described non-enjoyment purpose and not transferring the datasets to others, it is likely that **the Non-Enjoyment Purpose Requirement is satisfied in this case, irrespective of his purpose for gaining profits**.

Please be advised that, even where the main purpose of using copyrighted work is *not* enjoyment, if enjoyment is found to be a subsidiary purpose in the use thereof, such use does **NOT** satisfy the Non-Enjoyment Purpose Requirement.

Based on the foregoing, the use of data from copyrighted work for developing generative AI causes the probability of not being found that the person does not have a purpose to enjoy or cause another person to enjoy the thoughts or sentiments expressed in the copyrighted work, as in Example 1. It is important to carefully consider whether or not the exploitation of the copyrighted work at issue actually satisfies the Non-Enjoyment Purpose Requirement in developing generative AI.

(2) Exceptions to the Applicability of Article 30-4 of the Copyright Act

The following factors should be considered in determining whether the exploitation of copyrighted work has the effect of “unreasonable prejudice to the interests of the copyright owner,” in which case such exploitation would constitute copyright infringement regardless of whether the Non-Enjoyment Purpose Requirement is met:

- whether such exploitation would interfere with an existing market for the copyrighted work; and

² This example is referred to in Copyright Division, Agency of Cultural Affairs “*Chosakuken no ichibu wo kaiseisuru houritsu (Heisei 30-nen kaisei) ni tsuite*” (Explanation regarding the act to amend a part of the Copyright Act of Japan in 2018), Copyright no. 692 vol. 58, p. 35.

- whether such exploitation would hinder the development of a potential market for the copyrighted work.

For example, where a person copies an AI training dataset available on a market, conducts data analysis of this dataset as a whole, and uses this data analysis to sell a competing AI training dataset³, this interference with the existing market could likely constitute unreasonable prejudice to the interests of the copyright owner.

III. Possibility of Copyright Infringement in the Use of Generative AI

When generating a work by using a generative AI, the generated work may be identical or similar to the existing copyrighted work that was included in the dataset used for the development of the AI, or to the existing copyrighted work that was included in the instruction by the user (See, Fig. 2 below). A copyright infringement issue that presents itself at this stage is whether (1) the copying and saving an existing copyrighted work on a server when such work is input into the AI by the user as part of the instruction to the AI, and (2) the sale, copying or public transmission of the work generated by the AI constitute copyright infringement.

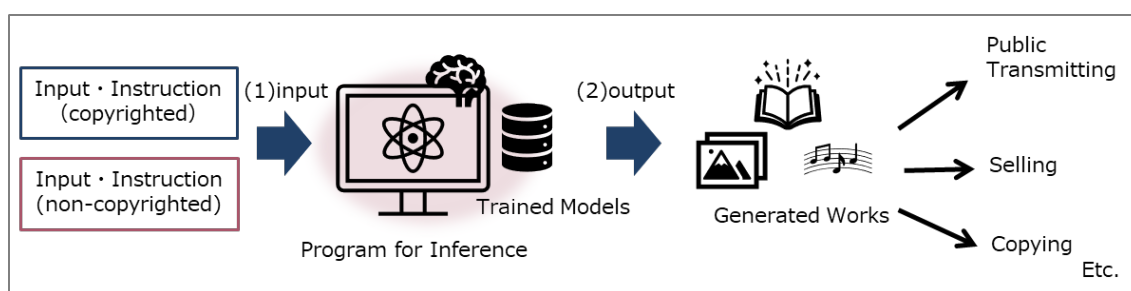


Fig. 2

As noted above, an act of making a given work that is identical or similar to an existing copyrighted work in terms of original expression in reliance on the said existing copyrighted work (the “**Reliance**”) is considered to fall under “reproduction” under the Copyright Act and therefore, constitutes, in principle, copyright infringement. With respect to “adaptation” and “making a transmission to the public” under the Copyright Act as well, an act of exploitation of the given work in reliance on the existing copyrighted work (the **Reliance**) is also required.

With respect to the Reliance, some argue that the Reliance can be established when the producer of a new work already had knowledge of an existing copyrighted work and made use of such copyrighted work, e.g., the form of expression thereof in the creation of the new work. Some argue that the reliance is found when a person who produced a new work had recognized the content of the expression in the existing copyrighted work and had an intent to make use of it in the new work. Whether a new work is produced in reliance on an existing copyrighted work is generally determined by examining (a) whether the producer had knowledge of the existence and the content of the existing copyrighted work, or had a reasonable opportunity to access the existing copyrighted work, (b) the extent of identity between the existing copyrighted work and the new work, including deliberate errors and shared mistakes. However, as the users of generative AI will not likely know specifically which copyrighted works were included in the dataset used to train a given AI, this conventional reliance standard may not be a good fit when it comes to works produced by generative AI. Some argue that, as long as the existing copyrighted work was included in the dataset, the user of the generative AI is deemed to have had access to the existing copyrighted work and therefore should be found to have relied upon same. Some argue that the reliance should not be found in the case where the existing copyrighted work was fragmented into parameters in the training process because in this case the expression of the work is not really used. In any event, it is highly likely that a court will find copyright infringement where existing copyrighted work was included in the dataset used to train generative AI. It should be noted that further discussion between the government, practitioners and scholars is expected with respect to this issue.

Whether a produced work is identical or similar to the copyrighted work is determined with reference to the same

³ This example is also referred to in Copyright Division, Agency of Cultural Affairs “AI to Chosakuken” (AI and Copyright), p.39.

criteria by which a court determines whether, in a non-AI generated work, such work shares the same “essential features of expression” as the copyrighted work.

As a side note, Article 30-4 of the Copyright Act does not apply to the production of a new work by using generative AI because it is obvious that the production of a new work does not meet the Non-Enjoyment Purpose Requirement.

On the basis of the foregoing, the copying, sale or public transmission of a work that shares the same “essential features of expression” as an existing copyrighted work produced via generative AI that learned the dataset including the existing copyrighted work is likely to fall under “reproduction,” “adaptation” or “making a transmission to the public” under the Copyright Act and therefore constitute copyright infringement. Therefore, when making generative AI produce a new work, it would be advisable to check with the developer of the generative AI if any copyrighted work is included in the dataset, and if the answer is yes, check with the developer if licenses are obtained from the copyright holder, or if there is any other legal reason to prevent copyright infringement (e.g., the expiry of the copyright term).

IV. Future Prospects

Generative AI technology is advancing daily, and, along with this advance, generative AI is becoming more and more important in the business world. Under the current copyright law, both (i) use of an existing copyrighted work as data to be input into a training program of a generative AI to produce a new work and (ii) producing a new work that shares the same “essential features of expression” as an existing copyrighted work by using a generative AI are likely to constitute copyright infringement. It should be noted that the law may change in the future in accordance with further discussion between the government, practitioners and scholars. We will follow the discussion and keep you updated.

[Authors]

**Kenji Tosaki** (Nagashima Ohno & Tsunematsu Partner)

kenji_tosaki@noandt.com

Kenji Tosaki specializes in intellectual property litigation and complex commercial litigation, and he also covers the area of TMT. He has been handling both IP infringement litigations and IP invalidation litigations before the IP High Court, the Supreme Court, District Courts and Japan Patent Office. His IP expertise includes a wide variety of IP matters (patents, copyrights, trademarks, design rights, unfair competition and trade secrets) in many areas, such as pharmaceuticals, telecommunications, electronics, social games, medical devices and chemicals. He also provides pre-litigation counseling, including infringement/invalidity analysis.

He is a member of AIPPI (International Association for the Protection of Intellectual Property), AIPLA (American Intellectual Property Law Association), LES (Licensing Executives Society) and INTA (International Trademark Association).

**Takahiro Hatori** (Nagashima Ohno & Tsunematsu)

takahiro_hatori@noandt.com

Takahiro Hatori's practice focuses on intellectual property and litigation. He has experienced a variety of IP matters, especially, disputes and legal advice in relation to patent infringement, patent invalidation, trademark, unfair competition prevention act, copyright, IP-related agreement.

Graduated from the University of Tokyo Faculty of Law (LL.B.) in 2011 and Waseda Law School (J.D.) in 2013, registered as a lawyer (Dai-ichi Tokyo Bar Association) in 2014, joined Nagashima Ohno and Tsunematsu in 2014, graduated from Munich Intellectual Property Law Center (LL.M.) in 2020, worked at Gleiss Lutz (Munich) in 2020-2021, member of Intellectual Property Law Center of Japan Federation of Bar Association (2022-)

**Tomoki Abe** (Nagashima Ohno & Tsunematsu)

tomoki_abe@noandt.com

Graduated from the University of Waseda Faculty of Law (LL.B.) in 2019 and Waseda Law School (J.D.) in 2021, registered as a lawyer (Dai-ichi Tokyo Bar Association) in 2022.

Nagashima Ohno & Tsunematsu (2022-)

This newsletter is given as general information for reference purposes only and therefore does not constitute our firm's legal advice. Any opinion stated in this newsletter is a personal view of the author(s) and not our firm's official view. For any specific matter or legal issue, please do not rely on this newsletter but make sure to consult a legal adviser. We would be delighted to answer your questions, if any.

NAGASHIMA OHNO & TSUNEMATSU

www.noandt.com

JP Tower, 2-7-2 Marunouchi, Chiyoda-ku, Tokyo 100-7036, Japan

Tel: +81-3-6889-7000 (general) Fax: +81-3-6889-8000 (general) Email: info@noandt.com



Nagashima Ohno & Tsunematsu is the first integrated full-service law firm in Japan and one of the foremost providers of international and commercial legal services based in Tokyo. The firm's overseas network includes locations in New York, Singapore, Bangkok, Ho Chi Minh City, Hanoi, Jakarta and Shanghai, and collaborative relationships with prominent local law firms throughout Asia and other regions. The over 500 lawyers of the firm, including about 40 experienced attorneys from various jurisdictions outside Japan, work together in customized teams to provide clients with the expertise and experience specifically required for each client matter.

If you would like to receive future editions of the NO&T IP Law Update by email directly to your Inbox, please fill out our newsletter subscription form at the following link: https://www.noandt.com/en/newsletters/nl_ip/.

Should you have any questions about this newsletter, please contact us at [<newsletter-ip@noandt.com>](mailto:newsletter-ip@noandt.com).

Please note that other information related to our firm may be also sent to the email address provided by you when subscribing to the NO&T IP Law Update.