

## Supplementary Material for Video 22

### “AI Learns and AI Creates? Generative AI and Copyright”

# AI and Copyright

After the release of the “Learning about Copyright from the Basics” series, several documents relating to AI and copyright have been published.

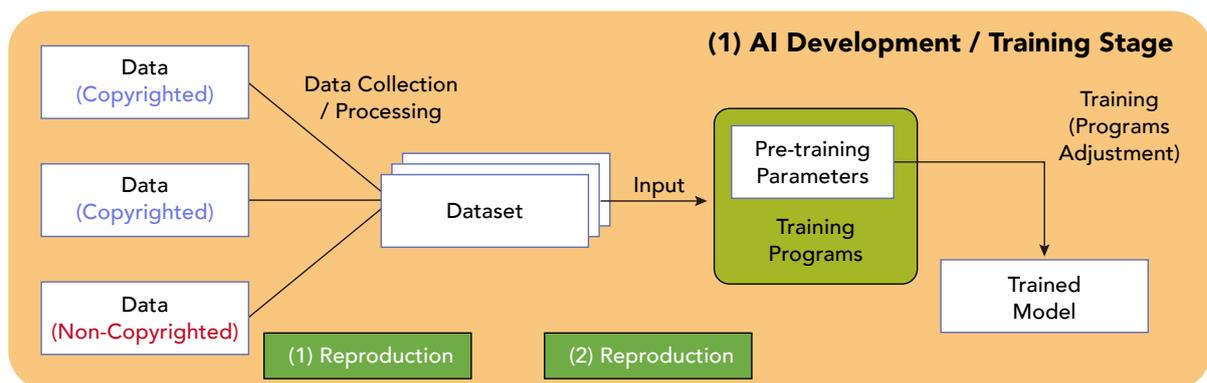
- Japan Copyright Office, General Understanding on AI and Copyright in Japan-Overview, May 2024 (the “General Understanding Overview”)

[https://www.bunka.go.jp/english/policy/copyright/pdf/94055801\\_01.pdf](https://www.bunka.go.jp/english/policy/copyright/pdf/94055801_01.pdf)

We provide supplementary explanations on the three basic topics related to AI and copyright: 1) the application of copyright exceptions for non-enjoyment use during the training phase, 2) copyright infringement caused by AI-generated works during the utilization phase, and 3) the copyrightability of AI-generated works.

## 1. Provisions on the limitation of rights for non-enjoyment use during the training stage (Article 30-4 of the Copyright Act)

When considering AI and copyright, the process is organized into two stages: AI development/training and AI utilization.



Source: Copyright Division, Agency for Cultural Affairs, Japan, AI and Copyright II (Japanese), August 2024, p.29

Article 30-4 of the Copyright Act refers to the limitation of rights provision that allows the use of copyrighted works without the copyright owner's permission when such use is not for the “enjoyment” of the expression of the copyrighted work. “Enjoyment” means deriving intellectual or mental satisfaction from viewing or hearing the work.

When using a work as training data for AI development, such use typically falls under this non-

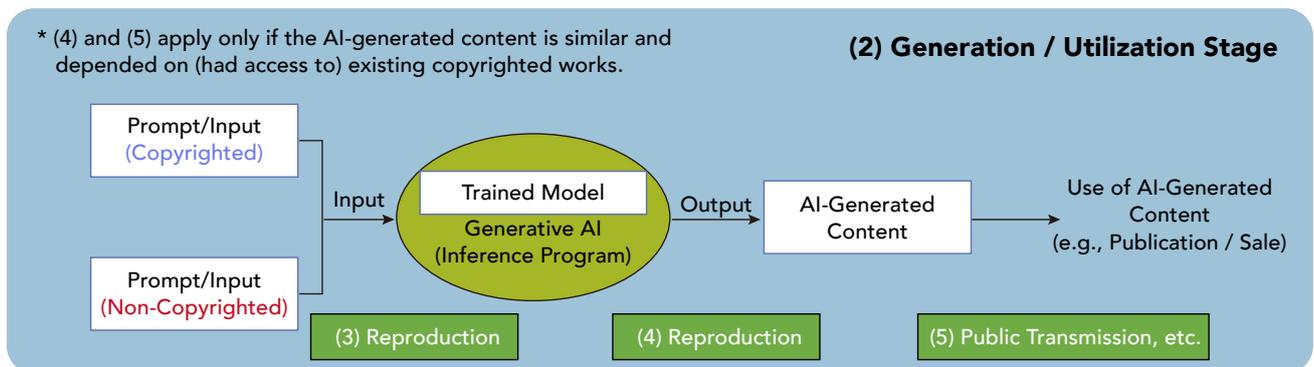
enjoyment purpose and may be used without the copyright owner's permission. However, in the following cases, the “enjoyment purpose” coexists, and this limitation does not apply (p.8 of the General Understanding Overview):

- Use of copyrighted works for additional training (such as intentional overtraining) for generating similar works to the training data.
- Using copyrighted works as input data for generating similar works of existing copyrighted works through retrieval-augmented generation (RAG) or similar methods.

Using works of a particular creator as training data to train AI for generating the creative expressions of that creator.

Additionally, when using databases of works sold for AI training purposes without permission, or in other cases where the use “unreasonably prejudices the interests of the copyright owner”, this copyright limitation provision does not apply (Article 30-4, proviso. See p.10 of the General Understanding Overview).

## 2. Copyright infringement by AI-generated works at the utilization stage



Source: Copyright Division, Agency for Cultural Affairs, Japan, AI and Copyright II (Japanese), August 2024, p.46

Whether AI-generated works infringe the copyright of existing works is determined based on the following two criteria, as explained in “The Difficulty of Similarity” in Video 16:

- Similarity: The essential characteristics of the expression of existing work can be directly perceived.
- Access/Dependency (Ikkyo in Japanese): The AI-generated work was created based on the existing work.

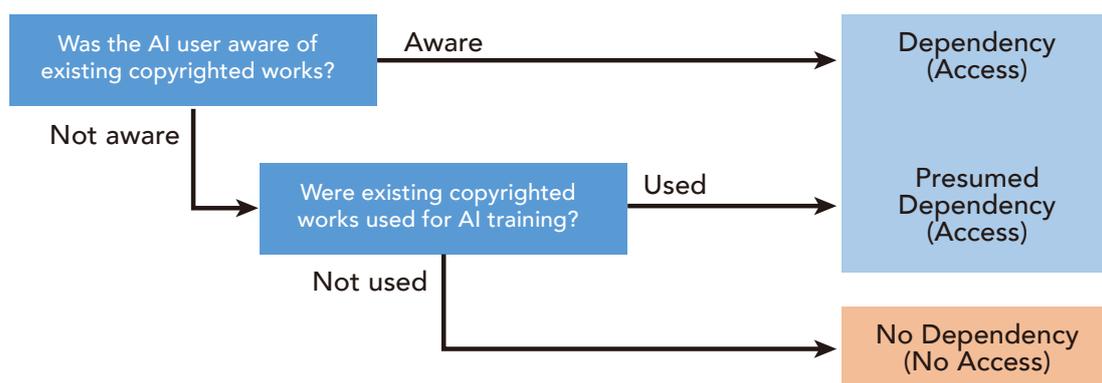
In the case of AI-generated works, the following scenarios may be found in access/dependency, potentially constituting copyright infringement (p.13 of the General Understanding Overview):

- When It is Unknown whether the Existing Copyrighted Work is Used for AI Training

If it is uncertain whether a particular copyrighted material is used in the AI training data, dependency will be presumed if the copyright owner can prove that “the AI user had access to the existing copyrighted work” or “the AI-generated material has a high degree of similarity with the work”. This means that it is possible for the copyright owner to establish dependency even though it is unknown whether the AI has used the work during AI training.

b) When the Copyright Owner Can Prove the Existing Copyrighted Work is Used for AI Training

It is generally assumed that there was dependency on a preexisting copyrighted work, even if the user of an AI was not aware of it, if the work was used for AI training during the development stage of that AI.



Source: Copyright Division, Agency for Cultural Affairs, Japan, *AI and Copyright II (Japanese)*, August 2024, p.52

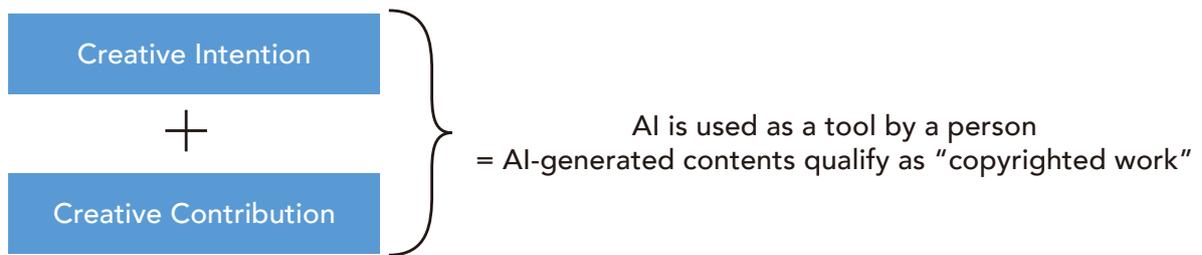
### 3. Copyrightability of AI-generated works

---

For AI-generated works to be protected as copyrighted work, they must meet the definition of a copyrighted work, “a creatively produced expression of thoughts or sentiments that falls within the literary, academic, artistic, or musical domain” (Article 2 (1)(i) of the Copyright Act) outlined in “What is Copyright?” in Video 1.

Additionally, the “author” is defined as “a person who creates a work” (Article 2 (1)(ii)), referring either to a natural person or a legal entity. Therefore, AI itself does not qualify as an author, and a person who uses AI as a tool to create the work is the author of the AI-generated work.

Furthermore, whether a person uses AI as a tool depends on whether the AI user has “creative intention” and whether there are “creative contributions” (p.16 of the General Understanding Overview).



Source: Copyright Division, Agency for Cultural Affairs, Japan, *AI and Copyright II (Japanese)*, August 2024, p.60

Whether "creative contributions" exist is determined by comprehensively considering the following elements (p.17 of the General Understanding Overview):

- Amount of instructions/input: "Detailed instructions that specifically indicate what constitutes creative expression" are more likely to be considered as creative contributions. However, lengthy instructions (i.e., prompts) that merely suggest an idea do not influence the assessment of creative contribution.
- Number of generation attempts: A large number of attempts alone does not affect the assessment of creative contribution. Repeated attempts, while checking the generated materials and correcting the instructions/input, may be recognized as a creative contribution.
- Selection from multiple output materials: The mere act of selection itself does not influence the determination of creative contribution. However, certain elements of choice may be involved, which may be considered as creative.

AI-generated content that is created autonomously cannot be a creative expression of human thoughts or sentiments. Therefore, such content does not qualify as copyrighted work. However, if a person uses AI as a "tool" to creatively express their own thoughts or sentiments, the portion created in this manner may be considered copyrighted work.

Additionally, parts of AI-generated content to which the user has added creative expressions through additions or corrections are generally considered copyrighted works (p.17 of the General Understanding Overview).