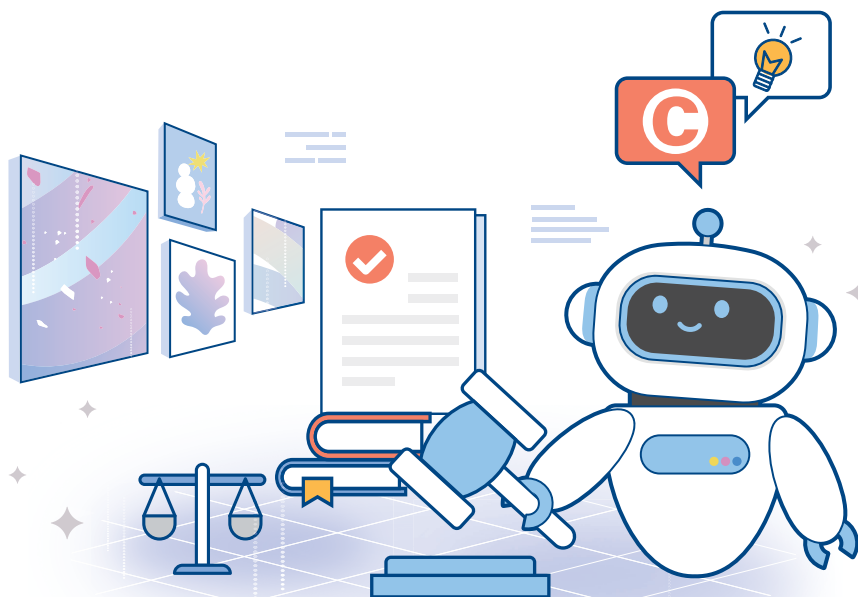# A Guide on Generative AI and Copyright

Ministry of Culture, Sports and Tourism
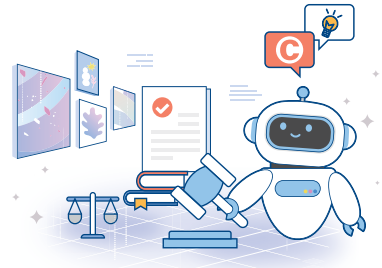
KOREA COPYRIGHT COMMISSION

Note to Readers

- The purpose of *A Guide to Generative AI and Copyright* ("Guide") is to inform artificial intelligence (AI) businesses, copyright holders, and users, who are the major stakeholders in generative AI, of what they need to be aware of in relation to copyright.

- The rapid development of AI technology has created an unpredictable environment for the copyright industry specifically and creative activities more generally. The use of AI technology in various fields has contributed to economic and social benefits, and in the creative sector, many are deploying generative AI as a tool to assist creative work. However, it's also true that some are concerned about the financial loss and threat to jobs that generative AI may bring.

- In response, the Ministry of Culture, Sports and Tourism (MCST) and the Korea Copyright Commission (KCC) brought together experts and stakeholders from the academia, legal profession, and various industrial technology sectors into a working group to review copyright issues triggered by the emergence of generative AI and to seek responses. This *AI-Copyright Law Working Group (ACL WG)* was established in February 2023 and has been in operation since. The discussions of the ACL WG led to the publishing of this Guide.

- The Guide focuses on the process of using data to train AI and then generating output, during which legal issues may arise - specifically during the process from building a foundation model for generative AI to deriving an AI output. It provides information, from the perspective of the current copyright law, for AI businesses, rights holders, and users of AI services. Readers will also benefit from explanations of recent generative AI use cases and related legislative trends in and outside Korea.

- This Guide is not intended to direct, in any way, the ongoing discussions on amendment bills to the Korea Copyright Act (KCA) currently pending in the National Assembly. However, after additional discussion, research, and opinion gathering, the ACL WG plans to come up with well-informed proposals for legislative and institutional reform.

- Going forward, the MCST and KCC will continue to exert effort through copyright-related laws and institutions to forge an enabling copyright environment where human creative endeavors are respected, and AI technological and industrial development is promoted.

* This Guide is not a legal interpretation of copyright issues related to generative AI. The final determination on whether copyright has been infringed shall be made by a court of law, considering all facts and details regarding an individual case.

* This Guide was written based on the law and legal precedents as of December 2023 and may be revised depending on future legislation, revisions, new rulings, or technological advancement.

Contents

A Guide on
Generative
AI and
Copyright
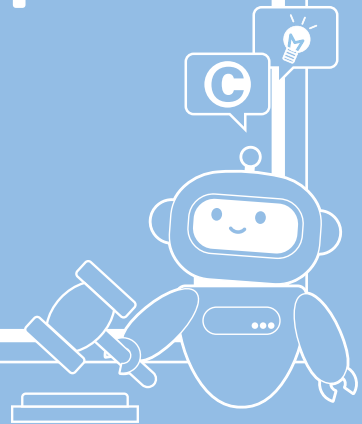
A Guide on Generative AI and Copyright

Chapter 1

Generative AI Technology and Copyright

01

Significance of generative AI technology and use cases

A Significance of generative AI

■ Generative AI refers to artificial intelligence technology that leverages deep learning of large datasets to present output, at the instruction of a user, in the form of new data or content.

> 🔍 **Deep Learning**
>
> Deep learning is an AI technology that mimics the human brain's neural network to make inferences from patterns of layered data it had trained on.

■ Generative AI is essentially a computer producing new output, particularly an output resembling pre-existing data, by using artificial neural networks as learning models.
  ※ An artificial neural network refers to a training algorithm that mimics the way a human brain works.

■ The rise of generative AI is revolutionizing how content is created across various sectors.
  − ChatGPT[OpenAI] can develop storyline ideas for a screenplay and even propose camera angles, actors' expressions and positions, lighting, and props, just like a film director.
  − RadioGPT[Futuri] is an artificial intelligence DJ that automatically collects information from the internet to generate a broadcast script. It also makes an actual broadcast.

B How generative AI technology is being used

■ Natural language generation: Industries are developing and applying various generative AI models to fit their service. Such models are mainly used in natural language generation (chatbot service), including writing news articles, novels, or reviews.
  − A large language model (LLM) is a typical AI model that generates natural language. HyperCLOVA X[Naver], ChatGPT[OpenAI], and Bard[Google] provide services in the form of chatbots for easier use by non−experts and are used for productivity enhancement, such as in Q&As, summaries, and translations.

> 🔍 **Large language model (LLM)**
>
> A LLM is a generative AI model that learns large−scale language data (text) to sequentially predict the most probable words in a given context, outputting sentences with high relevance.

- Image generation: Some services are based on a diffusion model, which is a technology that generates high-quality images in response to text or image input. Examples include DALL-E3[OpenAI], Make-A-Scene[Meta], Imagen[Google], and Generative-Fill[Adobe].
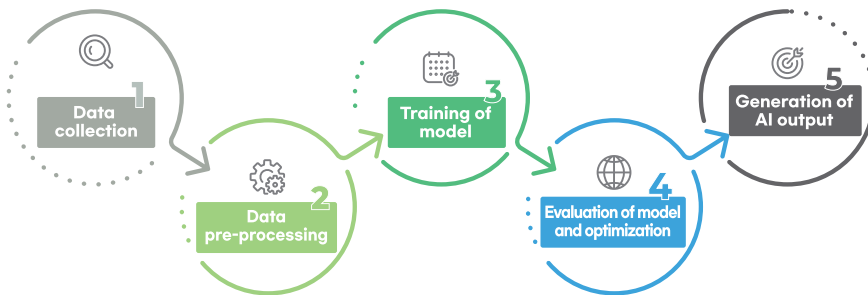
> **🔍 Diffusion model**
>
> Diffusion model is a generative AI model that makes new data (image) through the process of adding noise and then deleting the original data (image) from the noise.

- Voice and music generation: There are services (MusicLM[Google], MusicGen[Meta], Stable Audio[Stability AI], among others) that generate new music based on various inputs, such as another but similar type of music, a music score, MIDI music, chords, and a specific tempo, in addition to linguistic input regarding mood, genre or length.
  - Some services even use written text and human voice samples to generate narrations or songs in the voice of AI.

- Video generation: Although there are still technical issues to be resolved due to the temporal nature of video data (thematic consistency, flickering, etc.), advances in deep learning technology are enabling the transfer of styles and real-time image synthesis, allowing intuitive and high-performance video editing (Make-A-Video[Meta AI Research]).

- Other applications: Generative AI is also being used to write code, such as in optimizing or automating the coding process or in assisting code-writing (Copilot[GitHub], CodeWhisperer[Amazon]), to provide metaverse services (in areas like tourism, education, manufacturing, defense, entertainment, etc.), such as in 3D modeling, constructions of spatial environments, and creation of digital humans, by generating the appearance and texture of objects based on user input (text, image, GUI, etc.). It is also being used in CAD services.

Chapter 01
Chapter 02
Chapter 03
Chapter 04
Chapter 05
Chapter 06

02

Generative AI technology from the perspective of copyright

- Below is a diagram expressing various copyright issues that may emerge along the path from collecting training data to creating generative AI output.

Sequence	Process	Explanation	Copyright issue
1	Data collection	■ Original data (text, images, music, videos, etc.) are collected to train AI.	■ Including a copyrighted work in the data used to train AI involves the act of copying a work.
2	Data pre-processing	■ Unnecessary data are deleted (cleaned), divided (tokenized), given consistency (normalized), etc.. ■ Quality of data is enhanced and transformed into formats appropriate for training.	
3	Training of model	■ Pre-processed data is used to train the AI model.	
4	Evaluation of model and optimization	■ Performance is evaluated using a verified dataset. ■ The model's structure, training completion rate and parameters are adjusted to optimize the model.	
5	Generation of AI output	■ AI output is generated when a user inputs a prompt.	■ Output generated by AI may infringe copyright.

A Copyright issues during AI training stage

> AI training data can include not only works, but also databases, performances, phonograms, and broadcasts, leading to relevant rights being infringed. However, this Guide will focus on works since copyright issues that arise are essentially not different from those related to the latter.

- Training AI involves several steps to collect and process the data, and to form a dataset, which is then passed through an artificial neural network. These steps inevitably involve the act of copying a work.

  - Suppose a person uses publicly available data on the internet to train an AI model without permission from the rights holder. In that case, the right to reproduce a particular work, protected by KCA, may be violated during the training.

Copyright

Creative works that express human thoughts or emotions are protected by copyright, which, as a rule, is given to the author. Copyright is comprised of moral rights (right to make public, right to attribution, right to integrity) and economic rights (rights to reproduce, perform, publicly transmit, exhibit, distribute or lend a work, or to create a derivative work).

Database

KCA also protects, under certain conditions, databases that may not be creative in the selection and arrangement of their content, such as telephone directories, personnel directories, or address books. According to the KCA, databases must be systematically arranged or constructed and searchable to be protected.

Right to reproduce a work

Reproducing a work is the act of making another physical form of a work by printing, photographing, duplicating, making audio or visual recording, or making digital copies. (KCA Article 2.22) The author holds the right to make copies of his or her work — i.e., right to reproduction (KCA Article 16) — meaning that others who intend to copy that work have to get permission from the holder of this right.

B   Copyright infringement issues related to AI output

Diverse terms are used to refer to the output generated by AI. However, since the KCA only recognizes humans as creative authors, the Guide recommends terms such as "AI output" or "AI result" to be used rather than "AI creations" or "AI works." This Guide uses "AI output."

- Generative AI output is stochastically generated from an AI model that has completed its training based on input such as prompts of users.
  - If an AI output is identical or similar to parts of an original copyrighted work included in the training data, copyright issues may arise regarding whether the original work's copyright was infringed, and if so, who is responsible.

Prompt

Traditionally, the term "prompt" is used as "a sign on a computer screen that shows that the computer is ready to receive your instructions." (Source: Cambridge Dictionary) However, in the context of generative AI, the term is generally accepted to refer to a certain input value (text, image, video, etc.) to instruct or order the AI model to perform a function.

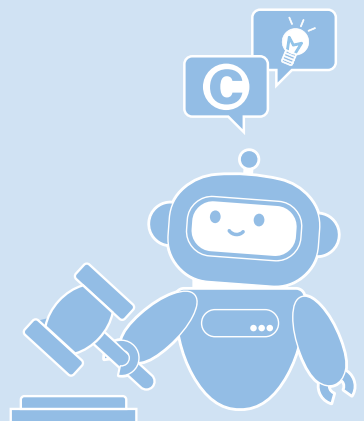# Rights granted by KCA to holders of copyright and neighboring rights

Copyright	Moral rights	Right to make a work public
		Right to attribution
		Right to integrity
	Economic rights	Right to reproduce a work
		Right to perform a work
		Right to publicly transmit a work
		Right to exhibit a work
		Right to distribute a work
		Right to lend a work
		Right to produce a derivative work
Neighboring rights[1]	**Performers** — Moral rights	Right to attribution
		Right to integrity
	Performers — Economic rights	Right to reproduce a work
		Right to distribute a work
		Right to lend a work
		Right to perform a work (excluding broadcasted performances)
		Right to broadcast a work (excluding audio–recorded performances)
		Right to interactively transmit a work
	Performers — Right to claim remuneration	Right to claim remuneration for broadcasts
		Right to claim remuneration for digital sound transmissions
		Right to claim remuneration for performances
	Phonogram producers — Economic rights	Right to reproduce a work
		Right to distribute a work
		Right to lend a work
		Right to interactively transmit a work
	Phonogram producers — Right to claim remuneration	Right to claim remuneration for broadcasts
		Right to claim remuneration for digital sound transmissions
		Right to claim remuneration for performances
	Broadcasters — Economic rights	Right to reproduce a work
		Right to simultaneously broadcast live events
		Right to perform a work (only when payment is received for admission to watch a broadcast)

---

1) For an explanation of this term, see the section explaining "neighboring rights" in Chapter 4.

A Guide on Generative AI and Copyright

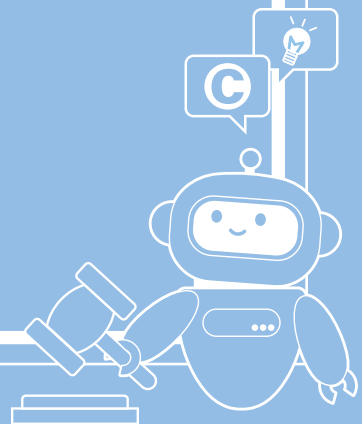# A Guide on Generative AI and Copyright

Chapter 2

Guide for AI Businesses

01

Generative AI training stage

A Use of works during training of AI

- The training process for building an AI model requires large amounts of different types of data and may involve works that are protected by the KCA (literary works, music, art, etc.).

- Under KCA, using someone else's work without permission may lead to copyright infringement.
- However, if a work is exploited under permission from the rights holder, or the use of the work falls under limitations to economic rights, it won't be considered an infringement.

> ### 🔍 Limitations to economic rights
>
> Economic rights are essentially rights that guarantee exclusive and monopolistic use of a work. However, considering that a work is created with direct and indirect support from society at large, giving unrestricted monopoly to the rights holder is not in line with public interest nor is it conducive to cultural development. This is why the KCA stipulates certain circumstances where economic rights are limited so that under those circumstances, using a work without the rights holder's permission won't be considered copyright infringement.

- Since the current KCA does not limit economic rights when the purpose of using a work is for training AI, an alternative argument may be whether the fair use provision (KCA Article 35-5), a general clause limiting economic rights, applies to AI.
  - There is an ongoing debate within academia on the applicability of the fair use rule, and there is not yet any legal precedent that directly applies the rule to AI training in Korea or elsewhere.
- ※ Regarding the use of works in training AI, an amendment to the KCA is currently pending in the National Assembly to introduce the so-called text and data mining (TDM), which allows copying of works without permission from the rights holder if specific requirements are fulfilled.

> ### 🔍 Fair use provision
>
> The fair use provision in the KCA was introduced to cover diverse situations where works were being used but regulation was difficult to be imposed through the existing specified exceptions and limitations provisions against the backdrop of digitalization and the changing distribution environment. Unlike provisions on limitations that specify individual acts, the fair use provision is general and comprehensive and does not specify a particular act.
> Whether or not an exploitation is fair use will depend on (1) the purpose and characteristics of the use, (2) the purpose and type of work, (3) the proportion and significance of the part used out of the whole, and (4) the effect the use will have on the existing or potential market or value of that work. Ultimately, a court of law shall make the final decision based on circumstances and facts pertaining to each individual case.

Chapter2

Guide for AI Businesses

Chapter 01
Chapter 02
Chapter 03
Chapter 04
Chapter 05
Chapter 06

■ Until several related court precedents accumulate, the applicability of the fair use defense will remain unclear, meaning there will always be the possibility that using a work for AI training without permission from the copyright holder[2] may infringe copyright.

B   How to legitimately use works

■ An AI business, even when using a work solely for AI training, should try to prevent possible disputes by securing authorization from the rights holder prior to use, whether by paying adequate remuneration or through other means.[3]

– An AI business should remain vigilant when using copyrighted work because the fact that a work is publicly available through a website, blog, or social media does not automatically mean it can be exploited without permission from the copyright holder.

※ If the rights holder of a work to be used for AI training is unclear or unknown, the statutory licensing system in the KCA may help secure the right to use the work legally.

■ If an AI business continues to store the data that had been copied for AI training or uses it for another purpose, it may be infringing copyright. Therefore, the company needs to detail the purpose, scope, and duration of its use when entering a contract with the copyright holder.

📢 Related examples

- In July 2023, company "O" signed a license agreement with a telecom company "A" and a stock image company "S" to attain data necessary for training its large language model.
- The National Institute of Korean Language launched its corpus project in 2018 to support development of Korean language-based AI technology, linguistic studies, and establishment of literary policies – the results of which are accessible by the general public. The institute had signed a license agreement regarding the data it had used to train the AI.
- Company "A" used proprietary and copyright-expired content (photos, videos, illustrations, etc.) to train its image generation AI model.
- Company "P" inputted, in formats such as midi files, music that had been written directly by its employees or those whose term of protection had expired, so as to be able to use the data to train its music generating AI.
- In August 2023, company "N" announced that it was developing an AI-powered creative tool that could assist a particular artist by training AI solely with that artist's images.

■ In addition to signing a license agreement directly with a rights holder, an AI business can also consider using works in the public domain, such as the National Institute of Korean Language's *Corpus for All* or KCC's *Artificial Intelligence Seed Project*, as a way to secure data for AI training.

---

2) The fair use provision is one of several limitations to economic rights. However, some cases may also involve moral rights, which is why the Guide refers to "copyright holder" as the licenser.

3) The KCA has separate clauses (KCA Article 25, etc.) specifying cases where a user, having used a work without the copyright holder's permission, only has to pay remuneration rather than be liable for infringement. However, "remuneration" used in this Guide is not limited just to the KCA concept of remuneration but encompasses a broader meaning.

🔍 Works in the public domain

Works in the public domain are those whose copyright has expired, or where the author had given up his or her copyright and are thus available to be used freely by anyone.

National Institute of Korean Language's *Corpus for All*	Korea Copyright Commission's *Artificial Intelligence Seed Project*	National Information Society Agency's *AI Hub*
The National Institute of Korean Language's *Corpus for All* project collects language-related, licensed material from various fields. It converts the material into a computer-readable format to accumulate into a corpus that anyone can use.[4]	KCC's *Artificial Intelligence Seed Project* cleans and processes public domain works provided through its *Gong-u-madang* website to allow those works to be used as source data for AI. It is essentially a project to set up and share source data that anyone can use.[5]	National Information Society Agency's *AI Hub* is an integrated AI platform providing infrastructural support necessary for developing AI technology. It offers free access to datasets set up through the AI training data project.[6]

- Further discussion is needed on whether to introduce measures to improve the transparency of AI services, such as labeling the source of the training data and works that had been used.

  − Going forward, the type and scale of data used for training, as well as technological feasibility, all need to be considered, and consensus among the stakeholders will be crucial.

📢 Related examples

- European Union's AI Act (parliament draft) mandates that a sufficiently detailed summary of the data protected by copyright law and used in training AI be disclosed in the form of a document.
- In September 2023, company "I" announced that it will fully disclose the datasets it had used in AI training in order to alleviate any fear from its customers regarding legal risks related to its AI service.

---

4) 〈https://corpus.korean.go.kr/〉
5) 〈https://gongu.copyright.or.kr/gongu/bbs/B0000026/list.do?menuNo=200311〉
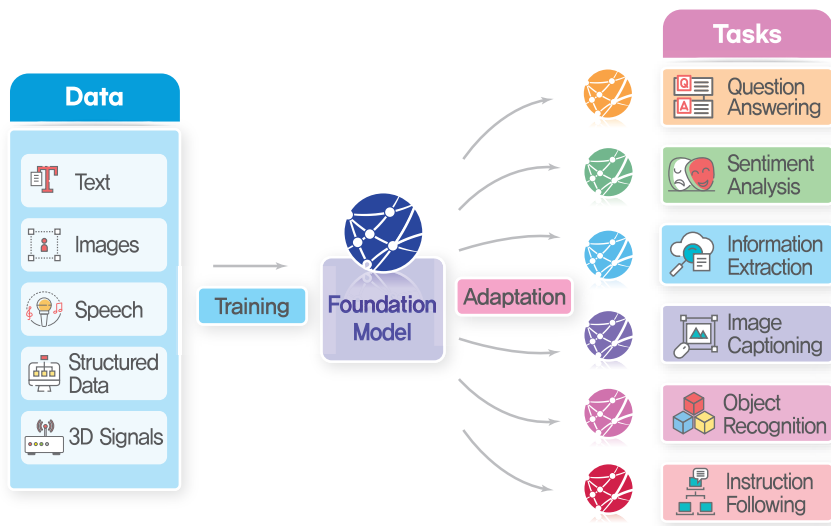6) 〈https://www.aihub.or.kr/〉

02

AI output generation stage

A   When does an AI output infringe copyright?

- If an AI output is deemed to be the same as or similar to an existing work, a copyright infringement claim can be raised.
    - ①Whether the AI output was generated with the pre-existing work in mind, i.e., had access to the original work, and ②whether the output is the same as or similar to a pre-existing work, i.e., has substantial similarities to the original work, will determine whether copyright was infringed.

- Ultimately, a court of law will determine whether an AI output infringed the copyright of an existing work based on specific facts relevant to the case.

B   Measures to prevent copyright infringement

- AI businesses are encouraged to prevent copyright infringement by ensuring that, when they provide their services, their AI output is not identical or similar to existing works (e.g., by adopting filtering).
    - In particular, a business that provides application services using an existing foundation model may not be able to track all the data the AI had been trained on, requiring a separate technology to be adopted to prevent its AI output from infringing copyright.

- A business that fine-tunes its AI model to focus and train on a particular artist's work or a specific set of images faces a higher likelihood that its generated AI output will infringe copyright.
    - If a business fine-tunes a model directly or provides related services to users, it must take particular precaution to avoid becoming liable for copyright infringement.

Chapter 01
Chapter 02
Chapter 03
Chapter 04
Chapter 05
Chapter 06

[For reference] Foundation model conceptual diagram [7]

🔍 **Foundation model**

A foundation model is trained on a huge amount data through self-supervision. It's a general model – meaning it can be fine-tuned according to the user's needs. GPT, BERT or CLIP are some examples.

🔍 **Fine-tuning**

Fine-tuning is the process of additionally training and optimizing an AI model that has already completed pre-training. The aim is to enable it to perform a specified task.

■ When a copyright dispute arises regarding an AI output, related businesses may also additionally face the issue of which company is liable for what. Therefore, a foundation model provider and a customer company using that model to provide application services are recommended to clarify the obligations and liability of each party when signing a contract.

📢 **Related examples**

■ Companies "A" (June 2023), "G" (October 2023), "M" (September 2023) and "O" (November 2023) each announced its User Protection Policy where they state that they will bear all legal responsibility including pay litigation expenses, should any of their customers face copyright litigation as a result of using their service.

---

7) R. Bommasani et al., "On the Opportunities and Risks of Foundation Models," arXiv:2108.07258v3 [cs.LG], 2022.

Chapter2

Guide for AI Businesses

Chapter 01
Chapter 02
Chapter 03
Chapter 04
Chapter 05
Chapter 06

03

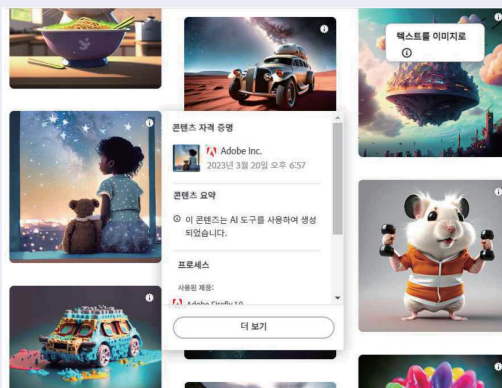# Distinguishing between AI output and human-created works

- Rapid development of generative AI technology is making it increasingly difficult to distinguish between a human creative work and an AI output. With human creative activities under threat, copyright regimes now have to be able to protect the rights of human authors while at the same time encouraging the use of works.

- Recently, there have been discussions both within Korea and internationally on ways to label generative AI output* to disclose the use of AI technology and distinguish it from human-created works.

  - The AI business community is encouraged to take on the task of researching and developing related technology while the government institutionalizes support for such technology development.

  * For example, a label can inform the user whether something is an AI output or a human-created work, or differentiate between parts generated by an AI and those created by a human.

📢 Related examples

- An amendment[8] to the Content Industry Promotion Act, proposed in May 2023, stipulates that if a content creator creates something using AI, they have to label that AI was used.
- In July 2023, the U.S. government reached an agreement with seven companies, including companies "A", "G", "M", and "O", on introducing safety measures such as watermarking AI-generated content, and in October, issued an executive order on the development of guidelines for the use of watermarks on such content.
- Image and video editing tools of company "A" include a feature called Content Credentials that allow users to track any production or editing that involved AI.[9]

※ The ACL WG, in 2023, discussed the issue of whether to mandate labeling for the use of AI but concluded that further discussion was needed on who has the obligation, what the label should contain, and who the target should be. ACL WG plans to discuss these topics at future meetings and, if necessary, propose revisions to the law.

---

8) Initiation led by Sang-Heon Lee, member of the National Assembly. Motion #2122180.
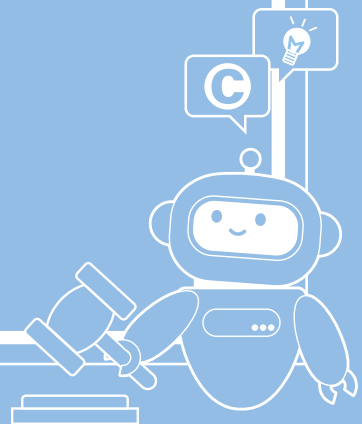9) 〈https://helpx.adobe.com/kr/creative-cloud/help/content-credentials.html〉

A Guide on Generative AI and Copyright

Chapter 3

Guide for Rights Holders

> AI businesses and rights holders, as stakeholders in the use of data to train AI, face similar copyright issues and thus need to take similar precautions. Therefore, the general explanation regarding copyright issues in Chapter 2 of this Guide also applies to rights holders.

01

AI training stage

- Many AI businesses crawl the internet to collect publicly available works that are accessible to anyone and create a dataset to use during their training of AI models.
  - If a rights holder does not want their work to be used for AI training, it is recommendable that they express objection in a reasonable manner or adopt technical measures to prevent such exploitation (e.g., by specifying as such in the Terms of Use or applying the robot exclusion standard, etc.).
  - For works that are already publicly available on the internet, failure on the part of the rights holder to take precautionary measures may result in a dispute over whether implicit permission had, de facto, been given to the user.

🔍 Crawling

Crawling refers to collecting and saving internet data using automatized means. Programs developed to crawl are called crawlers.
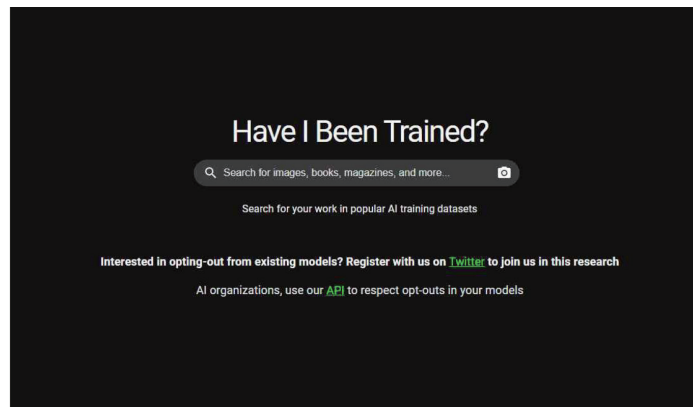
🔍 Robot exclusion standard (robots.txt)

This standard prevents crawlers and other forms of robots from accessing an internet website. However, it is not legally binding.

- Even after belatedly learning that a work had been used in AI training, a rights holder can still consider taking the above measures to prevent further use of their work in AI training.[10]

---

10) Recently, many researchers are looking into "machine unlearning" – deleting parts of the data that AI had already trained on. Ronen Eldan and Mark Russinovich, "Who's Harry Potter? Approximate Unlearning in LLMs", arXiv:2310.02238v2 [cs.CL], 2023.

[For reference] An internet site that tracks if a work had been used for AI training[11]

> 📢 **Related examples**
>
> ▪ Article 4 of the *EU Directive on Copyright in the Digital Single Market* regulates exceptions and limitations for data mining. However, if a rights holder explicitly opts out, then the provision will not be applied.
>
> ▪ Many internet websites, including broadcasters and media outlets, include in their Terms of Use a provision requiring prior discussion when AI developers want to use news reports and other content for training an AI model. Sometimes they also use the robot exclusion standard to prevent crawlers accessing their content.

---

11) 〈https://haveibeentrained.com/〉

Chapter 01
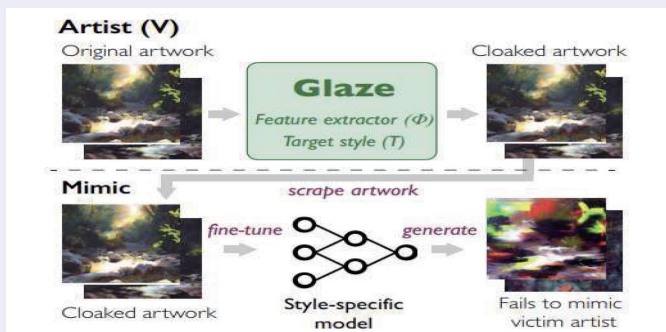Chapter 02
Chapter 03
Chapter 04
Chapter 05
Chapter 06

02

Recommendations for preventing an AI output from infringing copyright

- Rights holders generally have little leeway to intervene when someone uses their work for AI training and their copyright is infringed.

  - However, in recent years, technologies have emerged that prevent similar outputs from being generated by AI, even if a work had been exploited in training AI models. Thus, rights holders are encouraged to adopt such technological measures to prevent infringement of their rights (e.g., *Glaze* by University "C"'s research team and *Photo Guard* by University "M"'s research team).
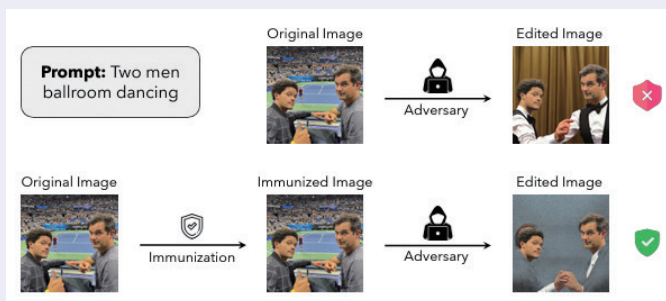
📢 Related examples

- A research team at University "C" developed a technology (*Glaze*) that prevents AI models from copying the style of an original work by making the AI to recognize the work differently, even if the work had already been used in training the AI. The university distributed this application online so that anyone could use it.[12]

- A research team at University "M" developed and made public a technology (*Photo Guard*) that manipulates an original image in ways unrecognizable to the human eye to prevent AI from editing or altering the image.[13]
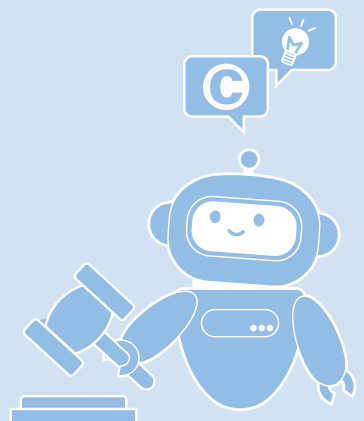

---

12) Shawn Shan et al., "Glaze: Protecting Artists from Style Mimicry by Text-to-Image Models," arXiv:2302.04222 [cs.CR], 2023.
13) Hadi Salman et al., "Raising the Cost of Malicious AI-Powered Image Editing," arXiv:2302.06588 [cs.LG], 2023.
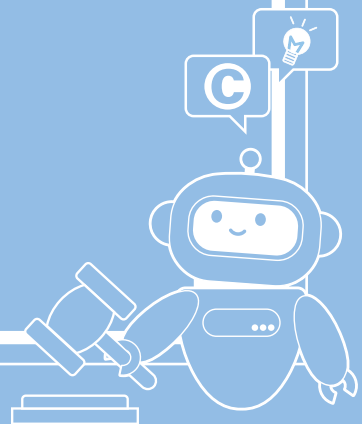
A Guide on Generative AI and Copyright

A Guide on Generative AI and Copyright

Chapter 4

Guide for Users of AI

01

Being aware of potential copyright infringement when using generative AI

A   Possibility of copyright infringement

> Depending on how one uses generative AI, infringement of not only copyright but also of rights to portrait, sound or name may also occur. However, this Guide will focus only on precautions related to copyright infringement.

- Users of generative AI can potentially infringe the copyright of others by getting AI to generate an output that is identical or similar to an existing work.
  - In particular, building a generative AI model requires training using a large amount of data, and that data may contain significant amounts of existing copyrighted works and could result in identical or similar outputs, thus increasing the risk of copyright infringement.

Related examples

- In a lawsuit involving a code-generating AI[14], the plaintiffs argued that the program generated codes that were the same as ones already existing.[15] Depending on how a prompt is articulated, an image-generating AI can produce characters that are exactly the same as those that are already well-known.[16]

"Realistic 3D rendering of Mickey Mouse working on a vintage computer doing his taxes"
on DALL · E 2 (left) vs. Stable Diffusion (right)

---

14) DOE 1 v. GitHub, Inc., 4:22-cv-06823, (N.D. Cal.). For details, see ⟨https://githubcopilotlitigation.com/⟩.
    The complaint is available at ⟨https://githubcopilotlitigation.com/pdf/06823/1-0-github_complaint.pdf⟩.
15) For more information about the lawsuit, see ⟨https://www.courtlistener.com/docket/65669506/doe-1-v-github-inc/⟩.
16) ⟨https://waxy.org/2022/08/exploring-12-million-of-the-images-used-to-train-stable-diffusions-image-generator/⟩

■ When copyright has been infringed, the infringer may be subject to civil and criminal penalties – i.e., be liable for civil damages for tortious acts or criminal penalties under the penalty provisions of the KCA. (Imprisonment of no more than five years or a fine of no more than KRW 50 million for infringement of economic rights, and imprisonment of no more than three years or a fine of no more than KRW 30 million for defamation of an author by infringing their moral rights).

B   Precautions when using generative AI

■ Users of generative AI should ensure that data, such as text, images, or audio, that they input to produce the desired AI output does not infringe or induce infringement of someone else's copyright.

– In particular, a user should take special care not to infringe copyright by using AI output that is the same as or similar to an existing work by performing, exhibiting, distributing or publicly transmitting that output, or otherwise making it publicly available.

🔍 Performing

Performing refers to presenting to the general public a work by acting, playing an instrument, singing, reciting, screening, replaying or other means. It can also include replaying a cinematographic work in front of many people.

🔍 Exhibiting

Exhibiting refers to displaying or hanging an artistic, photographic, or architectural work, in either in its original or reproduced form – i.e., in physical form – to allow the general public to view the work. The act of exhibiting does not apply to a work that is not artistic, photographic, or architectural.

🔍 Distributing

Distributing is transferring a work or its copy by assigning it or lending it to the public, with or without payment. Major examples include selling or lending CDs or DVDs.

🔍 Publicly transmitting

Publicly transmitting an output or a work means transmitting it to the public through wired or wireless means. Broadcasting (e.g., transmitting terrestrial broadcasts through television), interactive transmissions (e.g., posting on website, OTT services) and digital sound transmissions (e.g., music service through podcasts) are all public transmissions.

Chapter 01
Chapter 02
Chapter 03
Chapter 04
Chapter 05
Chapter 06

1) Text

- If a user generates an AI output that is identical or similar to an existing work by entering the work into the prompt window in its original form or by entering a text inducing the generation of an identical or similar output and then goes on to post the result on the internet, the user may be infringing various rights such as the rights to reproduction, public transmission, and integrity.[17)]

> 📢 **Examples of common forms of copyright infringement**
>
> - A person inputs a prompt inducing an image-generating AI to produce popular characters like Pororo and then uses the generated image as the profile picture on one's social media account.
> - A person inputs a detailed story of an existing novel as a prompt to induce a similar story to be generated, and then posts the output on one's blog.
>
>
>
> [For reference] These images were generated by inputting prompts such as "Winnie-the-Pooh" and "Snoopy." The outputs show the features of the original characters sufficiently enough to be able to recognize them.[18)]

2) Images and videos

- If a person inputs a specific image or video to create an AI image or video without permission and then uses the generated AI output, they may be infringing copyright, such as the right to reproduce a work.

> 📢 **Examples of common forms of copyright infringement**
>
> - A person produces an advertisement poster by inputting a copyrighted image into an image-generating AI program, changes the text to express one's own company products and then uploads the poster onto the company website.
> - A person uses AI to merge an image of a popular character with a video of an existing musician performing and then uploads the result onto YouTube.

---

17) What rights are actually infringed will depend on the act involved.
18) Images generated using Microsoft Bing's *Image Creator*, a generative AI service (21st November, 2023). 〈https://www.bing.com/images/create〉

Chapter4

Guide for Users of AI

Chapter 01
Chapter 02
Chapter 03
Chapter 04
Chapter 05
Chapter 06

3) Music

■ If a user generates an AI output (video, music, etc.) by inputting someone else's music without authorization from the owner and then uses that output, they may be infringing copyright, including the right to reproduce a work.

- In particular, if a person produces a cover song – where AI trains on a musician's music and then "sings" the song in another musician's voice – but in the process, uses someone else's music without permission or decides to upload the AI cover song onto an online platform, this person may be infringing copyright as well as neighboring rights because the entire process involves copying someone else's music or album.

- In addition, in some cases, these actions may be violating the Unfair Competition Prevention Act or infringing general publicity rights, among other rights.

> 🔍 Neighboring rights
>
> Neighboring rights are rights granted to those who had contributed financially or creatively so that a work can be delivered to the public, such as performers, phonogram producers or broadcasters. KCA protects these rights as those neighboring to copyright.

※ In Korea, there was a discussion on a general level about unauthorized use of portraits of famous people and their publicity rights, which are rights to commercially use, control, or exclusively dominate the economic benefits or value of a person's portrait or name. Then, an amendment, effective since June 2022, was made to the Unfair Competition Prevention Act, stipulating unauthorized use of a celebrity's personal identification mark (Article 2-1-(l)) to be an act of unfair competition and thus regulating such use. On the other hand, a proposal to comprehensively amend the KCA, including incorporating a provision directly regulating publicity rights, is currently pending in the National Assembly. (Proposal initiated by member Jong-Hwan Doh on 15th January 2021. Motion No. 2107440.)

> 📢 Examples of common forms of copyright infringement
>
> ■ A person inputs the melody of a song sung by a certain singer (by humming, whistling or using a sound source) as parameters along with a text and a melody similar to someone else's to produce music.
> ■ A person uses AI to generate music by making Michael Jackson's voice sing to a BTS song.

4) Other cases

■ Some generative AI models (e.g., Stable Diffusion of company "S") allow the user to fine-tune the model directly to reflect one's style of choice.

- Fine-tuning additionally involves the user training the AI using specific selective data, which, again, means that copyright infringement of the training data can become an issue.

- Suppose a person trains the AI to focus on a certain artist's work or a specific image to induce the AI to render a particular style. In that case, the output will more likely be identical or similar to the inputted works. Thus, the user needs to take extra precautions due to the higher likelihood of copyright infringement.

02

Ethical and policy considerations in areas such as research, education, or creativity

A   Discussion regarding generative AI in areas such as research, education, or creativity

- Ethical issues related to the use of generative AI are being discussed mainly by academic societies and universities in areas of research or education. Such institutions have established AI-related policies for publishing papers in scholarly journals and using AI in class.

> 📢 **Related examples**
>
> - University "A" in Korea established a guideline for its professors, who are required to choose one out of three options regarding the use of generative AI (①Use of generative AI forbidden, ②Use of generative AI allowed only after prior approval from professor or after labeling the source, ③Use of generative AI permitted), which is then stated in the course curriculum for the students to follow.

- There is also an ongoing discussion on the use of generative AI in creative fields, such as art and webtoons, while art competition organizers in different genres are exploring the possibility of adopting an AI policy.

> 📢 **Related examples**
>
> - In May 2023, Korean company "N" was swept in controversy when it hosted a webtoon competition without setting a policy banning AI. It adopted a ban only later, forbidding the use of AI from the second round of submissions. Company "K" also held a webtoon competition where an explicit condition for submitting a work was that "the work must be drawn by a human hand."

B   Considerations when submitting a work to an academic journal or competition

1) Check AI-related policies and guidelines of journals and competitions

- A person seeking to make a submission should check and comply with the policies and guidelines of journals and competitions regarding the use of generative AI.
  - An incompliant submission may result in its publication or award being canceled while failing to disclose the use of AI may be considered a false submission and lead to a legal claim.

> 📣 **Related examples**
>
> - Academic journal N does not recognize a LLM tool as a CRediT author of a research paper that it publishes. It explicitly requires that if a LLM tool was used, it be adequately explained in the research methodology of the paper while the use of a generative AI image is disallowed altogether.
> - In September 2023, Korean company "K" did not allow AI-made messenger emoticons from entering its store and announced that it will continue to look into whether or not it will accept AI technology.

2) Label the source of generative AI output and use of AI-assisted tools

- An author is encouraged to verify the source of generative AI-powered text before citing it in academic papers and to disclose the source where possible.[19]

  - Articles written by generative AI may be inaccurate due to biases in the training data and its technical nature of generating output based on probability.
  - If an AI generates an output similar to an article contained in the training data and the user quotes the output verbatim, the user can be accused of plagiarism in addition to copyright infringement.

> 🔍 **Plagiarism**
>
> Plagiarism refers to the act of pretending to have created a work when in fact someone else had. It is a moral or ethical concept, not a legal one. Therefore, the scope of plagiarism may not coincide with that of copyright infringement.

- When citing an article produced by generative AI, the author must distinguish the citation from the parts they had written directly and clearly label it as such.

  - The author should choose a citation method appropriate to the situation and purpose. In the case of an academic paper, they should follow the citation style recommended by the relevant discipline (APA, MLA, Chicago Style, etc.).[20]
  - Also, they should indicate, in as much detail as possible, what generative AI-powered creative tools they had used and how.

> 📣 **Related example (of a Korean university "A"'s policy on how to make citations)**
>
> - If a text generative AI was used: e.g., ChatGPT3.5 (20th April 2023). "Content of prompt." ChatGPT3.5 of OpenAI was used to generate the text. https://chat.openai.com/
> - If an image generative AI was used: e.g., Stable Diffusion (20th April 2023). "Content of prompt." Stable Diffusion web version was used to generate the image. https://stabledifusionweb.com/

---

19) Many language models now provide its users with internet links to resources a model had used in the generation process when they render output. These links can be used as a reference when making citations. However, in other areas including images, it may be realistically difficult to identify the source of the original data used to generate the AI output.

20) These citation styles are the most common ways of citing sources when writing academic papers, reports and other forms of written text. For citation methods for generative AI, refer to the relevant websites.
APA Style: ⟨https://apastyle.apa.org/blog/how-to-cite-chatgpt⟩, MLA Style: ⟨https://style.mla.org/citing-generative-ai/⟩, Chicago Style: ⟨https://www.chicagomanualofstyle.org/qanda/data/faq/topics/ Documentation/faq0422.html⟩

Chapter 01
Chapter 02
Chapter 03
Chapter 04
Chapter 05
Chapter 06

03

Other recommendations

A Check regulations such as those in AI service providers' Terms of Use

- In a separate provision in their Terms of Use, generative AI providers usually stipulate how and to what extent the user is allowed to use the output.[21] Therefore, a person using the service must check all regulations in advance to prevent problems arising from violating the Terms of Use.[22]

B Label the use of generative AI

- Even if a user does not intend to submit an article or a work to an academic journal or competition, it is recommended that the user clearly label, in a reasonable manner, that they have used AI to generate texts, images, or videos.

- Labeling will minimize potential confusion between AI output and human-created works, thereby increasing public trust in how content is created, and rights are attributed.
  - It will also contribute to specifying the owners of various rights to content, thereby clarifying the liability of each party should a legal conflict arise due to copyright infringement.

📢 Related example

- In November 2023, video platform "Y" announced its AI policy, under which a creator is required to disclose whether they had used AI to generate a video or to change parts of an existing content. If the uploader does not comply, the content could be deleted, profit-sharing scheme stopped or the uploader imposed with other forms of penalty.

---

21) For an AI output, whether or not the output is copyrightable and who copyright belongs to, will depend on whether creativity of that output is enough to be recognized under KCA.

22) Example: OpenAI's Terms of Use (updated on 14th November 2023). 〈https://openai.com/policies/terms-of-use〉

"Your Content. You may provide input to the Services ("Input"), and receive output from the Services based on the Input ("Output"). Input and Output are collectively "Content." You are responsible for Content, including ensuring that it does not violate any applicable law or these Terms. You represent and warrant that you have all rights, licenses, and permissions needed to provide Input to our Services.

Ownership of Content. As between you and OpenAI, and to the extent permitted by applicable law, you (a) retain your ownership rights in Input and (b) own the Output. We hereby assign to you all our right, title, and interest, if any, in and to Output."

For reference

Key guidance for AI businesses, rights holders, and users of AI

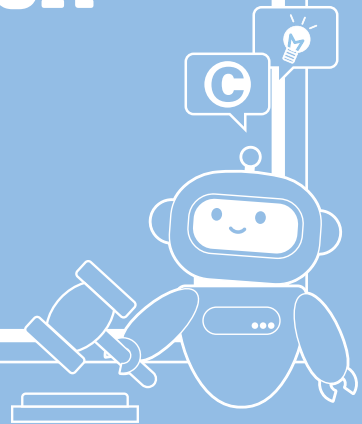| Who	Key guidance
AI businesses	■ Legitimately obtain the right to use a copyrighted work prior to using it for AI training by, for example, paying adequate remuneration. ■ State in detail the purpose, scope, and duration of the exploitation when signing a contract.
	■ Prevent copyright infringement by ensuring an AI output is not identical or similar to an existing work.
	■ Consider adopting specific labeling for AI output to distinguish it from human creative works.
Rights holders	■ To prevent a work from being used for AI training, explicitly opt-out or consider adopting specific technical measures.
AI users	■ Ensure that prompt inputs and the use of AI output do not infringe or induce infringement of someone else's copyright.
	■ Check policies related to the use of AI in articles or works when submitting them to academic journals or competitions. ■ Cite the source and indicate whether AI-assisted creative tools were used to produce generative AI output.
	■ Check relevant regulations, such as the Terms of Use, applicable to the AI service. ■ Label the use of generative AI.

A Guide on Generative AI and Copyright

Chapter 5

AI Output and Copyright Registration

01

Copyright issues related to AI output

A Copyrightability of an AI output

- Under KCA, a *work* is "a creative production that expresses human thoughts and emotions." (KCA Article 2.1)
  - According to the current interpretation of the law, output rendered by an AI, and not a human, is not copyrightable.

> 📢 **Related example**
>
> - In July 2022, Korea Music Copyright Association (KOMCA) decided to stop distributing royalties for 6 songs that had been composed by an AI program "E" because KOMCA does not recognize AI output as a copyrighted work.
> - A U.S. district court, in August 2023, denied copyrightability of an alleged work in a litigation raised against the decision of the U.S. Copyright Office (USCO), which had rejected copyright registration of an output made by a generative AI program developer entirely without human intervention.

- However, if a human had added his or her creativity to the AI output by modifying, making additions or deletions, editing, or arranging the production, then those aspects would be considered copyrightable.

B Authorship of an AI output

- KCA defines a work as "a creative production that expresses human thoughts and emotions" and an author as "a person who creates a work." (Articles 2.1 and 2.2)
  - In other words, only a natural person can become an author while a legal entity or an organization may be recognized as an author only if the law specifically provides as thus, such as works made for hire. Therefore, under the current law, AI itself cannot become an author.

- AI output itself cannot be protected as a work. Only when a human has contributed his or her creativity to the output can it be copyrightable and eligible for any consideration on authorship or granting of copyright.
  - In this case, whether or not a human contributed creativity and how much will be determined by general copyright jurisprudence while a court of law makes the final judgment.

Chapter5

AI Output and Copyright
Registration

Chapter 01
Chapter 02
Chapter 03
Chapter 04
Chapter 05
Chapter 06

02.

Is copyright registration possible for an AI output?

- Copyright registration for an AI output is impossible if a human did not contribute creatively to the expressive form.
  - Copyright registration is only available for creative works that express human ideas or emotions.

※ Copyright arises the moment a work is created and does not require any procedure or formality, such as registration, submission, or depositing, to become effective. (KCA Article 10.2 on the principle of non-formality) Furthermore, copyright registration does not automatically mean that a work is copyrightable nor will the registration guarantee protection of copyright for something that is not a work.

Copyright registration

Registering copyright refers to disclosing information about a work (name of author, date of creation, date it was first made public, etc.) and any changes made to rights to the work, such as the assignment of economic rights, restrictions on its disposal, or creation of a pledge, by inscribing it in a public ledger called a copyright register, available for public perusal.

A registered author can benefit from legal presumption of authorship, which means that the person will be presumed to be the author of the registered work. When changes to a right is registered, the rights holder gain opposing power, which means they gain the power to argue that the change is valid vis-à-vis a person who has legal incompatibility. Registration also allows a rights holder to benefit from statutory damages, whereby the plaintiff will be able to receive a certain amount of legally set compensation without providing proof of the actual value of the damages.

- However, if a human had performed *additional work* on the AI output, such as modifying, or making additions or deletions, only the part that had undergone such change is copyrightable and thus eligible for registration. (The effects of the copyright registration will also be limited to that part.)
  - Also, even if the AI output, as it is, is not eligible for registration, if a person selects and re-arranges the AI output creatively, the author can register the work as a *compilation*.

Compilation

A compilation refers to a collection of edited elements such as symbols, letters, sounds or videos, whereby the selection, arrangement or composition of those elements show creativity. In a compilation, it's creativity of the "selection, arrangement or composition of the elements" that is being protected, regardless of whether an individual element has copyright. Examples include a newspaper, a magazine, an encyclopedia, a court decisions report, or a journal of academic theses.

- An author is a person who creates a work, so AI itself, as opposed to a human (or a legal entity), cannot be an author. Therefore, AI cannot be registered as an author.
  - AI can be registered as neither a sole author nor a co-author, nor can AI output be registered in the name of the AI developer, the CEO, or legal entity as a work made for hire.

03

Examples from Korea and other countries regarding copyright registration of AI output

A Example from Korea

- A person had used a famous writer's poem as a text prompt to get AI to generate a video, for which the person then attempted to register the copyright, claiming the video was a derivative work.

  - The KCC rejected registration of this AI-powered video (registration application filed on 4th October 2022 and rejection notice sent on 3rd November 2022), on the grounds that under the current KCA, a work is a creative production that expresses *human* thoughts and emotions. (KCA Article 2.1)

⋯▸ Even if a human had inputted a prompt, the video was generated entirely by an AI, which means that the resulting video is not eligible to be registered.

⋯▸ According to the KCA Article 2.1, only a *human* creation is considered a work, meaning copyright registration is also limited to a human creation. Therefore, an AI output cannot be registered.

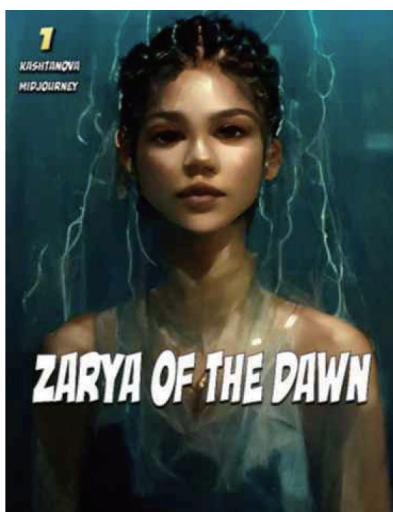## B Examples from the U.S.

1) A Recent Entrance to Paradise

⟨A Recent Entrance to Paradise[23]⟩

Chapter5
AI Output and Copyright
Registration

Chapter 01
Chapter 02
Chapter 03
Chapter 04
Chapter 05
Chapter 06

- Stephen Thaler, an applicant, attempted to register the copyright to an image "A Recent Entrance to Paradise" as a work made for hire. He claimed that the author was an AI program called Creativity Machine and that he owned the machine.

    - In response, the USCO rejected the application for registration on the grounds that the image lacked the "human authorship requirement."

⋯▶ The USCO had determined that an AI output was not a work created by a human and thus not eligible for copyright registration.

2) Zarya of the Dawn

⟨Zarya of the Dawn[24]⟩

- Applicant Kristina Kashtanova registered copyright to an 18-page comic book titled *Zarya of the Dawn*, made using an image-generating AI, Midjourney.

    - The USCO subsequently canceled the registration and issued a new registration certificate. The new registration only recognized the applicant's actual contribution to (1) the creation of the text (literary work) and (2) the selection, arrangement, and adjustment of the AI-powered illustrations (compilation), but not the illustrations themselves, which were produced entirely by an AI.

⋯▶ The USCO had rejected copyright registration of the AI output on the grounds that it lacked the "human authorship requirement" — i.e., that there was no human creative intervention.

⋯▶ However, if a human had modified the AI output by selecting or arranging it, only that partial, creative contribution was copyrightable.

24) ⟨https://www.copyright.gov/docs/zarya-of-the-dawn.pdf⟩

04

Precautions when registering copyright

A   Registering copyright only for a part of the AI output that is modified, added or deleted

■ An AI output must have undergone additional touch by a human, such as modification, addition or deletion, and the altered amount should be sufficient for it to be copyrightable and thus eligible to be registered independently.

  − If a certain part of the AI output had been modified or had parts added or deleted but that change had only been *minor*, that part won't be sufficient enough to be recognized as a human's creative endeavor and thus not eligible for copyright registration.

  − An AI output is not a copyrighted work, which means that a *derivative work* whose *underlying work* is the AI output will also not be eligible for copyright registration.

  ※ If an AI output had been modified or edited to create another work, then that work is not a derivative work of the AI output but an original work on its own right.

> ### 🔍 Derivative work
>
> A derivative work is a work that had been created by translating, arranging, modifying, dramatizing, cinematizing, or using the work in other ways. Examples of a derivative work are a TV series or a movie based on a novel, or a novel that had been translated into Korean from another language. The TV series, the movie and the translated novel are derivative works.

■ To register copyright, an applicant must fill in the "Contents of the Work" section of the registration application, detailing which part was generated entirely by AI and what the applicant had creatively modified.

  − Copyright registration is only valid for the additional touch given, and the registration may be used as evidence in the event of a dispute.

■ The KCC, as the copyright registration authority, decides whether or not a part of an AI output that has undergone human modification, addition or deletion is eligible for copyright registration after going through a minimum copyrightability review based on the *Copyright Registration Examination Manual*.

  − The KCC examiners exclude any AI output from the scope of review and only include parts created by humans.

Chapter 01
Chapter 02
Chapter 03
Chapter 04
Chapter 05
Chapter 06

B **When registering a compilation after having selected and arranged elements of an AI output**

■ An AI output can be registered as a *compilation* if there is creativity in the selection, arrangement, or composition of the AI output.

- Each element, such as text, image or videos, may have individually been created by AI and thus not copyrightable. However, if the selection and arrangement of the different elements were done creatively by a human, then those aspects can be registered.
- When registering a compilation, it's not the individual element – i.e., AI output – that is registered, which means that an element does not enjoy the legal effects of registration (such as legal presumption).

■ To register the copyright of a compilation, the applicant must describe in detail the creative contribution they had made when selecting and arranging the different elements in the "Contents of the Work" section of the registration application.

- Copyright registration is only valid for the selection and arrangement, and the registration may be used as evidence in the event of a dispute.

C **Penalty for false registration**

■ If a person applies for registration of an AI output, which is not eligible for registration, by intentionally faking it as if it were indeed a work of human creation, they may be penalized for false registration. (KCA Article 136.2.2)

- Even if registration has passed through, it can later be canceled ex officio if found to be an AI output. (KCA Article 55-4.1)

■ The KCC has inserted a warning in its registration application form to remind applicants that an AI output is not eligible for registration.

[Reference] Warning in the copyright application form
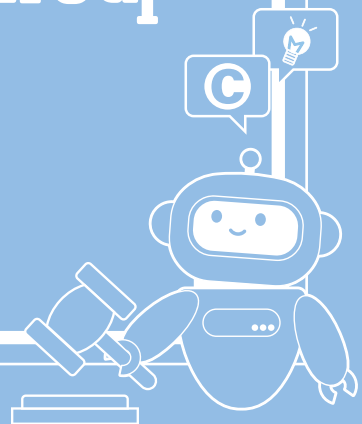
A Guide on Generative AI and Copyright

Chapter 6

Progress and Future Plans of AI-Copyright Law Working Group

01

Progress-to-date of the ACL WG

- In February 2023, the Ministry of Culture, Sports, and Tourism (MCST) and the Korea Copyright Commission (KCC) formed the AI−Copyright Law Working Group (ACL WG), which has been in operation since. The ACL WG comprises copyright academics, legal experts, industrial technology experts, and stakeholders to examine copyright issues related to the emergence and spread of generative AI and to discuss institutional reforms in response to the evolution of AI technology.
  - The ACL WG discussed recent trends in generative AI technology and related copyright issues. These discussions led to the issuing of this Guide.
- The members and progress-to-date of the ACL WG are as follows:
  - ＊ ACL WG members

Sector and Name		Affiliation and position
Academia	Dae−Hee Lee	Professor, Korea University School of Law
	Chul−Nam Lee	Professor, Chungnam National University School of Law
	Byung−Pil Kim	Professor, Dept. of Technology Management, Korea Advanced Institute of Science and Technology (KAIST)
Law	Kyu−Hong Lee	Presiding Judge, Seoul High Court
	Kwang−Nam Kim	Presiding Judge, Seoul High Court
	Min−Jung Kim	Prosecutor, Seoul Central District Prosecutors'Office
	Yong−Min Lee	Attorney−at−law, Yulchon Law Firm
Industrial Technology	Won−Young Yoo	Principal Researcher, Content Research Division, Electronics and Telecommunications Research Institute (ETRI)
	Han−Kyu Koh	Principal Researcher, AI Lab, LG Electronics
	Sung−Won Ahn	Director, AI Policy Research, Software Policy Research Institute (SPRi)
	Kyung−Il Lee	CEO, Saltlux
Rights holders	Dong−Hyun Kim	Secretary General, Korean Literature, Academic Works and Art Copyright Association
	Hyun−Sook Kim	Director, Policy and Legal Research Institute, Korea Music Content Association
	Min−Jae Choi	Director, Media Research Centre, Korea Press Foundation
Government/ Public sector	Gyung−Geun Chang	Director, Copyright Policy Division, Ministry of Culture, Sports and Tourism (MCST)
	Gyu−Ho Kang	Deputy Director, Copyright Policy Division, MCST
	Jin−Sook Jang	Assistant Director, Copyright Policy Division, MCST
	Hye−Chang Kim	Director General, Policy Research Department, Korea Copyright Commission (KCC)
	Chan−Dong Kim	Director, Legal & Policy Research Team, KCC
	Jung−Hoon Park	Senior Researcher, Legal & Policy Research Team, KCC
	Tae−Hwa Gye	Assistant Director, Legal Support Department, Korean Copyright Protection Agency (KCOPA)

Chapter6

Progress and Future Plans of
AI−Copyright Law Working Group

Chapter 01
Chapter 02
Chapter 03
Chapter 04
Chapter 05
Chapter 06

∗ ACL WG progress−to−date

Date of meeting	Name of meeting	Discussion topics
24th February 2023	Kickoff meeting (1st meeting)	▪ Terms of reference of ACL WG and introduction to discussion topics, etc.
20th March 2023	2nd meeting	▪ AI model training and ChatGPT ▪ Classification of AI−related copyright issues
24th April 2023	3rd meeting	▪ Generative AI and related copyright issues (Fair use, TDM, etc.)
30th May 2023	4th meeting	▪ AI technology and its application in the music industry, etc.
20th June 2023	5th meeting	▪ Mandatory labeling of AI output; AI and copyright registration
11th July 2023	6th meeting	▪ Improving AI−related laws and institutions from rights holders' perspective
8th August 2023	7th meeting	▪ Listening to the AI industry perspective and discussion on the Guide
12th September 2023	8th meeting	▪ Sharing of the draft Guide and first round of feedback
31st October 2023	9th meeting	▪ Sharing of the draft Guide and second round of feedback

▪ In addition to holding meetings, the ACL WG invited the public to comment on the Guide. At the same time, the Guide itself underwent separate reviews (written and in−person) by stakeholders, such as AI businesses and rights holders, as well as copyright academics and legal experts.

Sector		Method of feedback
Feedback from public		Seoul Copyright Forum (26th October), Culture, Sports and Tourism Digital Innovation Forum (29th November), Korea−China Copyright Forum (7th December), etc.
Feedback from individuals	Rights holders	(Nine) Rights holders' organizations such as the Korean Association of Newspapers and the Korea Music Copyright Association, (seven) large AI development corporations in Korea and overseas, and start−ups (18th October, 14th November, 24th November)
	AI businesses	

02

Future plans

- This Guide is based on the interpretation of the current KCA. It is intended to provide AI businesses, rights holders, and users of AI services with reference material regarding the use of generative AI. For further development, there is a need for continuous discussion on how to improve copyright-related laws and institutions concerning generative AI.

  - In particular, measures to secure legitimate rights through a system providing adequate remuneration when using a work for AI training, ways to facilitate the transaction of works, issue of whether to protect AI output, or whether AI output should require labeling, are some of the topics that need further consideration to better protect the rights of creators as well as develop the AI industry. The ACL WG will discuss these topics as part of its future research.

- Currently, the MCST is preparing a plan to support the development of technologies* to protect copyright and manage distribution across the entire process from training generative AI models to generating and distributing output.

  * Such technologies include those that compare the similarity of generative AI output (text, video, music, etc.), track original work, or insert watermarking into AI-based 3D output, among others.

- If a remuneration scheme is to be adopted for works used in AI training, a transparent and reliable communication channel will need to be established to ensure that a virtuous cycle that protects creators while advancing AI technology and related industries is created.
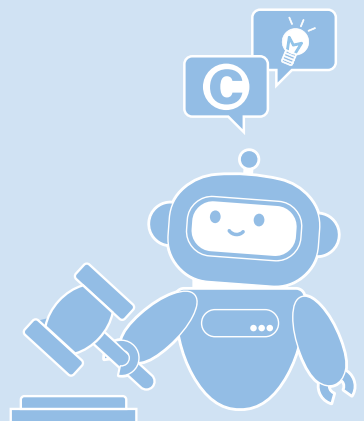
  - If the cost of building an AI training model is excessively high, it may result in the shrinking of AI-related industries and less use of copyrighted works. On the other hand, excessively low cost would lead to human creativity being devalued to a mere means to generate AI output.
  - Finding a win-win solution that can facilitate the necessary exploitation of works for AI training while fully recognizing the value of fruits of human creativity will increasingly become critical.

- More consideration will need to be made not only to AI-related legislative trends overseas but also to the technological development of AI and the relevant industrial structure in Korea and abroad to enable a more fine-tuned alternative. Also, the overall copyright protection regime will need to be revisited internationally.

A Guide on Generative AI and Copyright

A Guide on
Generative
AI and
Copyright

A Guide on
Generative AI and
Copyright