A Study on International Jurisdiction and Governing Law Regarding Generative AI Models - From the perspective of Korean private international law[1]

EungJun Jeon*

Table of Contents

I . Introduction - How Generative AI Works and the Types of Disputes Involved

1. How Generative AI Works

Generative AI is an artificial intelligence technology that enables AI to generate content such as text, images, audio, and video with varying degrees of autonomy in response to specific requests (prompts) from users.[2] It is believed that the important requirements for defining this concept are that the purpose of AI is to express and generate content, that AI has a significant degree of autonomy when generating content, and that content generation is directed and guided by user prompts. Generative AI, such as Generative Adversarial Networks (GANs), ChatGPT, and Stable Diffusion, is distinguished from Discriminative AI, which is designed to

[1] See the link below for the English version of the Korean Private International Law.
https://elaw.klri.re.kr/kor_service/lawView.do?hseq=60196&lang=ENG
Refer to the link below for the original version of the Korean Private International Law.
https://www.law.go.kr/%EB%B2%95%EB%A0%B9/%EA%B5%AD%EC%A0%9C%EC%82%AC%EB%B2%95

* Korean Attorney (LIN Law Firm).

[2] An amendment to the EU Artificial Intelligence Bill published by the Internal Market Committee and Civil Liberties Committee of the European Parliament on May 11, 2023 defines generative AI as "AI systems specifically intended to generate, with varying levels of autonomy, content such as complex text, images, audio, or video" (Art. 28b 4). The above EU Parliamentary proposal, with some amendments, was passed by the Parliament on June 14, 2023.

classify or categorize data, by focusing on the expressive power of content generation.[3] In statistical terms, a generative model such as Stable Diffusion and VAE is an AI model that learns from a given source data and generates similar (fake) data that follows the probability distribution of the source data, so learning from the distribution of the source data is of utmost importance. [4]

Due to the technical characteristics of trying to imitate the distribution of training data, the output of generative AI may be similar to other people's works or personal information, which causes problems such as copyright infringement and personal information leakage (OUTPUT phase). However, compared to the model that adopts the conventional DB search method (Retrieval Based Model), the retrieval-based model selects (searches) works or personal data stored in the DB and discloses them as they are, but in the generative model, information is newly generated and disclosed by the AI algorithm, so the nature of the problem that occurs in the above two models is quite different. Since generative AI has a significant degree of autonomy in the process of generating representations, the degree of equivalence (substantial similarity) between the trained source data and the generated representations becomes an issue. From the point of view of 'access', generative AI is guided by user prompts, but the AI itself has considerable autonomy and generates the output according to a black-box path. Even if the work or personal information contained in the training data is generated identically in the output, it is difficult to conclude that this phenomenon is an intentional infringement or to admit that the AI generated a copy by referring only to the training data in question. Furthermore, the common problems that occur in artificial intelligence models that perform machine learning, such as the collection and use of personal information and the reproduction of works that occur in the artificial intelligence learning stage (INPUT phase), are also tasks that generative AI must solve first.

[3] Background: What is a Generative Model? https://developers.google.com/machine-learning/gan/generative?hl=en

[4] Ian Goodfellow, 'NIPS 2016 Tutorial: Generative Adversarial Networks', 2016, p. 2, Ian Goodfellow, creator of GAN, describes a generative model as follows. "The term "generative model" is used in many different ways. In this tutorial, the term refers to any model that takes a training set, consisting of samples drawn from a distribution, and learns to represent an estimate of that distribution somehow."

This study first examines the problematic aspects of copyright infringement based on the operating principles of AI machine learning, dividing it into the stages of collecting and using training data (INPUT phase), generating outputs from AI models (OUTPUT phase), and providing AI systems as a service (SERVICE phase). This research methodology seems to be universally adopted in AI-related research. In each of these three stages, we will discuss the issues of international jurisdiction and governing law from the perspective of Korean private international law.

2. Types of Disputes Involving Generative AI

Since the class action lawsuit against GitHub Copilot's code generation service in the U.S. in early November 2022, which was primarily based on open source license violations, copyright infringement lawsuits against image generation AIs such as Stable Diffusion and Midjourney, copyright and privacy infringement lawsuits against OpenAI's ChatGPT, copyright infringement lawsuits against Meta's LLaMA, and copyright and privacy infringement lawsuits against AI products such as Google's Bard have been filed since early 2023. This has led to a situation where almost every AI service known to the public has been sued. In the Copilot case, the defendants' motion to dismiss has been largely denied and the case has moved to discovery, with a full trial expected.[5]

It is worth noting that most of the lawsuits have been filed in the United States, where class action lawsuits have developed. This is not surprising since the major AI developers are based in the U.S. However, Getty Images filed a lawsuit against Stability AI, which provides Stable Diffusion, not only in the U.S. but also in London, England. Getty Images' decision to sue in the UK may have been based on the fact that Stability AI's operations are based in London, but it may also have been based on the fact that the scope of fair dealing, a defense under UK copyright law, is narrower than the scope of fair use in the US. This suggests that the legality of AI model development and service provision may vary depending on what each

[5] Among other things, claims for unauthorized removal of CMI (DMCA §1202(b)) and claims for violation of open source licenses will be heard.

country's copyright and privacy laws say. In addition, the fact that all of the lawsuits, except for Getty Images, are in the form of class action lawsuits reflects the characteristics of AI machine learning. Since machine learning utilizes large-scale works and data, a small number of individual rights holders cannot qualify as plaintiffs. For AI that learns large amounts of data, a copyright trust organization can file a lawsuit or right holders must take the form of a U.S.-style class action lawsuit.

The main claim in these lawsuits is that works and personal information were used in the machine learning process of AI (INPUT phase) without the consent of the right holder. Furthermore, in the case of generative AI, it has been pointed out that the output of AI is a secondary work or derivative of training data (OUTPUT phase. Since generative AI is essentially a technique that tries to mimic the distribution of training data, there is a potential for the output to be similar to the works or personal information of others used as training data. However, the percentage of AI outputs that resemble the original training data appears to be extremely small. The plaintiffs in the Copilot lawsuit rely on a study that found that approximately 1% of Copilot's code output was similar to the training data code.

II . International Jurisdiction and Choice of Law Issues in Copyright Infringement

1. AI training(learning) phase (= INPUT phase)

1.1. Substantive law issues

Currently, copyright infringement lawsuits related to AI models in the United Kingdom and the United States are mainly concerned with copyright infringement in the AI training phase.

During the A.I. training stage or TDM process, the training data is duplicated (data scraping) from the Internet, etc. and entered into the software for A.I. learning, and the duplication of the training data occurs again.

Since AI learning often uses all or most of the data that has been preprocessed as training data, it is inevitable that large-scale copyright infringement will occur if there is no legal right to use the data.

From a TDM (Text/Data Mining) perspective, these input steps can be categorized as accessing data (step 1), extracting and replicating data (step 2), and text/data mining and knowledge discovery (step 3). Among them, legal issues are likely to arise especially in the second stage.[6]

In the second step, you need to verify whether the data to be extracted and reproduced is a copyrighted work, whether it is contained in a database (database protection regulations exist in Korea and the EU), and whether there are other legal restrictions (privacy, personal information, contract, etc.).

Reproduction of the work may also occur in the third stage. The third stage involves data pre-processing and structured data extraction. [7]

The third stage includes tokenization, cleaning, normalization, stemming, lemmatization, and stopword removal. This process is sometimes done by human labor, which raises privacy concerns for sensitive data.

The current AI technology that adopts machine learning methods requires the input of training data, so we believe that such a breach is an urgent problem that needs to be solved.

The Korean Copyright Act has a general fair use provision, but the Korean Parliament is currently discussing a bill to amend the Copyright Act to introduce a TDM exception. In this regard, there is a strong view that it is preferable to utilize the general fair use provision first in legal systems that have a general fair use provision rather than having a separate TDM exemption. However, considering the speed of artificial intelligence development, it is believed that case judgments and legal interpretations are insufficient and that specific TDM exceptions and legislation are necessary, such as in the EU and Japan.

---

[6] . Eleonora Rosati, Copyright as an obstacle or an enabler? A European perspective on text and data mining and its role in the development of AI creativity, Asia Pacific Law Review, 27:2, p. 204 (2019); Gyooho Lee, Protection of datasets for AI training under copyright law and unfair competition law and its limitations, Human Rights and Justice Vol. 494, 2020, p. 91.

[7] Eleonora Rosati, Ibid., p. 209

If the TDM exemption or the doctrine of fair use is applied to the AI learning process (INPUT phase) and copyright infringement liability is exempted, the question is whether the AI output process (OUTPUT phase) should be completely exempted even if the AI output is substantially similar to any prior work. From a principle point of view, the TDM exception applies only to the AI training phase and not to the AI output phase where a new output is created.

1.2. International Jurisdiction

1.2.1. General Jurisdiction

KPIL(hereinafter referred to as 'KPIL') has a "general jurisdiction" rule that applies to claims against legal entities or organizations whose principal office or place of incorporation or place of central administration is located in Korea and against legal entities or organizations established under Korean law (Article 3(3)). If an AI service company has its place of central administration, such as its principal office, in Korea or is established under Korean law, the Korean courts have jurisdiction under the 'general jurisdiction' provision. Where jurisdiction is granted to a Korean court under this 'general jurisdiction' provision, unlike the special jurisdiction for intellectual property rights described later, jurisdiction is not limited to results arising in Korea.

1.2.2. Special Jurisdiction

KPIL, which was fully revised in 2022, provides two articles related to intellectual property rights: (1) Special Jurisdiction of Lawsuits for Intellectual Property Contracts (Article 38) and (2) Special Jurisdiction of Lawsuits for Infringement of Intellectual Property Rights (Article 39).

Article 39(1) of KPIL recognizes the special jurisdiction of Korean courts over claims for infringement of intellectual property rights in the following cases: (1) the defendant committed the infringing act in Korea; (2) the defendant committed the infringing act outside Korea but the result of the infringement occurred in Korea; and (3) the defendant committed the infringing act from outside Korea and directing to Korea.

Here, the proviso to Article 39(1) of KPIL only recognizes jurisdiction over the results of intellectual property infringement occurring in Korea (quantitative limitation), and Article 39(2) of KPIL does not recognize the related case jurisdiction (Article 6(1)) of Korean courts over claims involving the results of intellectual property infringement occurring in a foreign country (i.e., if infringement occurs in Korea and Japan for the same copyright, Korean courts cannot merge Korean and Japanese copyright infringement claims). This adopts the "mosaic" approach of the CJEU's Shevill decision. However, as an exception to the 'mosaic' approach, if the main infringement of an intellectual property right occurs in Korea, a claim for all results of the infringement, including those arising in a foreign country, may be brought in a Korean court (Article 39(3) of KPIL).

The jurisdiction of the place of perpetration (No. 1) and the jurisdiction of the place of results (No. 2) rules have been influenced by the EU's interpretation of the Brussel I Regulation. The meaning of "place of perpetration" and "place of result" is discussed later in the tort section of this study.

A typical example of copyright infringement that occurs during the AI learning phase is the unauthorized copying of works by data scraping. If a Korean work on a Korean website is scraped without permission by a U.S. AI development company, Article 39(1) of KPIL recognizes the Korean court's international jurisdiction over the copyright infringement lawsuit. Here, the act of data scraping occurred in the United States, so it is difficult for the Korean court to recognize jurisdiction over the place of perpetration (No. 1). However, since the place of legal interest at the time of copyright infringement can be considered to be Korea, it falls under the jurisdiction of the place of results (No. 2), and international jurisdiction can be recognized by the Korean court.

From the perspective of emphasizing the principle of territoriality of copyright, it is difficult to pinpoint the situation under Article 39(1)(2) of KPIL(No.2). For example, if a U.S. resident uses

a U.S. server to transmit a Korean work, it is clear that the copyright (transmission right) infringement occurs in the U.S., but there are different opinions as to whether the result of the infringement occurs in Korea. According to some opinions, the mere possibility of accessing a U.S. server from Korea is not sufficient to constitute the "result of the infringement," and the result of downloading and copying an illegal work by a Korean resident must occur in order for 'results' jurisdiction (No. 2) to be recognized.[8]

'Directional' jurisdiction (No. 3) is understood to adopt the "purposefully directed" doctrine of U.S. courts.

1.3. Governing Law

KPIL provides that "the protection of intellectual property rights shall be governed by the law of the place of infringement" (Article 24). Although copyright infringement is a tort in nature, KPIL distinguishes it from torts (Article 44). The above provisions do not mention the relationship with international treaties, so the relationship with Article 5(2) of the Berne Convention on Copyright is ambiguous, and there are different opinions on whether the principle of the place of infringement applies to issues other than infringement of intellectual property rights. For example, when it comes to the initial attribution and transfer of copyright, there is a division between the view that it should be governed by the law of the place of infringement and the view that it should be governed by a single law in that the first copyright holder of a work should be determined on a worldwide basis.

In this context, "the place of the infringement" has the same meaning as " the country where protection is claimed ", i.e., the country that exploits the intellectual property at issue in any form within its territory or seeks to defend it against third parties, or the country whose intellectual property protection is being claimed on its territory.

[8] Young Gi Kim, "Interpretation and Application of Private International Law in International Intellectual Property Disputes.", Juris Vol 63, Judicial Development Foundation, 2023, p. 499

If a copyright infringement claim is filed in a Korean court for scraping data for AI model development under Article 24 of KPIL, the Korean Copyright Act shall be the governing law.

2. AI output phase (= OUTPUT phase)

2.1 substantive law issues

Upon inputting specific comments or function names into the GitHub Copilot model, a portion of the source code used in its training data is reproduced almost verbatim in the output. However, even the plaintiffs' complaint in the GitHub Copilot case references research suggesting that merely 1% of Copilot 's output consists of open-source code from the training data.

There is research revealing the Diffusion model 's ability to memorize training data to the extent that it can extract the original image from the AI-generated output.

Generative AI models such as GAN, Stable Diffusion, Copilot, ChatGPT, etc. mimic the probability distribution of training data or generate content based on the probability of occurrence, so it is possible that content that is identical or similar to the training data is generated by the user's prompt, although the frequency of occurrence is minimal (this is most likely to occur when the user enters a prompt for 'in the style of' content).

There is some debate about whether AI 's output can be deemed a derivative work of the training data. (e.g., Stable Diffusion)

With current technology, determining which input value, node, or parameter produced a specific AI output remains challenging, which makes it difficult to prove 'access' requirement in the copyright infringement.

2.2. International Jurisdiction

AI systems such as ChatGPT, Stable Diffusion, and Midjourney provide AI services to members. In the course of AI services, there may be a result that the works of third parties that existed in the training data are exposed to members who live in Korea almost as they are, although the frequency cannot be said to be high.

If the defendant's center of operations is located in Korea, the "general jurisdiction" rule applies and the Korean court has jurisdiction (the same as described in the AI learning phase).

If the AI service is provided to members in Korea and allows them to pay with credit cards issued in Korea, such as the ChatGPT plus and DeepL Pro services, such acts are deemed to provide services directed to Korea, and therefore jurisdiction over the place of destination (Article 39(1)(3) of KPIL) is recognized. If the output of the AI service is generated on a Korean site, the copyright holder's Korean copyright is infringed from the perspective of territorial principle, and the jurisdiction of the place of result (Article 39(1)(2) of KPIL) is recognized. If the AI service is provided on a server outside of Korea, such as a server in the United States, the copyright infringement itself is committed in the US, so jurisdiction over the place of perpetration (Article 39(1)(1) of KPIL) is not established.

2.3. Governing Law

This is the same as what we discussed in the AI training phase (INPUT phase).

3. Service phase of the AI model

Some AI systems, such as Midjourney, allow Service Members to upload images created by Service Members and make them visible to other Members. These AI services are Online Service Providers (OSPs) and are subject to the Digital Millennium Copyright Act (DMCA) if the AI service company is located in the United States.

The DMCA provides that a subscriber must consent to the jurisdiction of a federal district court within the United States when sending a Counter Notice to challenge an OSP's takedown action. (Section 512(g)(3)(D)). In this case, the jurisdiction is as follows.

(1) if the subscriber's address is within the United States, the judicial district having jurisdiction over the subscriber's address; or

(2) if the subscriber's address is outside the United States, a court of competent jurisdiction within the United States where the OSP is located.

Therefore, if a subscriber who is a resident of South Korea wishes to challenge a takedown by a U.S. OSP (e.g., Midjourney) to restore an image post, he or she should choose a U.S. court as the forum for any future disputes with the copyright holder.

III . International Jurisdiction and Choice of Law Issues in Privacy Breaches

1. Determining the characteristics of Privacy Violations and Types of AI Litigation Cases

The exercise of the rights of the data subject can take many forms, but the main ones are the exercise of the data subject's right of access, the right to request the details of third-party provision, the right to restrict the processing of personal information, and the right to claim damages in case of personal information infringement. It is generally accepted that the 'special jurisdiction of claims in tort' (Article 44 of KPIL) applies to the international jurisdiction of all non-contractual claims for personal information protection, such as data subjects' requests for access and claims for damages against data controllers.[9]

[9] Jong-Hyuk Lee, "A theoretical study on the international private law issues of privacy-related litigation - focusing on the types of privacy-related litigation and the debate on international jurisdiction", International Trade Law Review Vol.28 No.2, 2019, p.151; Maja Brkan, "Data protection

Currently, there is a lawsuit against AI systems for privacy infringement (more precisely, privacy or communication secrecy protection). Privacy infringement lawsuits against AI products such as OpenAI's ChatGPT and Google's Bard are underway. In the AI learning phase (INPUT phase), the problem is that personal information is collected and used without the consent of the data subject during the data scraping process for machine learning. In the AI output phase, similar to the copyright case, the problem is when data identical to a third party's personal information in the training data is generated (exposed) at the prompting of the user, although less frequently. In the AI service phase, data breaches that illegally leak service members' information are a problem, as shown in the recent ChatGPT case in Italy.

2. AI training phase (= INPUT phase)

2.1. substantive law issues

In the process of collecting and using personal information for machine learning purposes, if the AI development company fails to secure the legal basis for processing personal information (such as the consent of the data subject), personal information infringement occurs.

2.2. International Jurisdiction

If the defendant's center of operations is located in Korea, the "general jurisdiction" rule(Article 3) applies and the Korean court has jurisdiction (the same as explained in the copyright section).

---

and European private international law: observing a bull in a China shop", International Data Privacy Law Vol.5 No.4, Oxford University Press, 2015, p.270

The unauthorized collection of personal information constitutes a tort, and under the Special Jurisdiction Rules for Tort Claims, it can be brought before a Korean court "if the tort is committed in or directed to Korea, or if the result of the tort occurs in Korea". However, Korean courts do not have jurisdiction if it was not foreseeable that the results of the tort would occur in Korea (Article 44 of KPIL).

The special jurisdiction rules for tort claims are similar to the special jurisdiction rules for intellectual property infringement actions (Article 39) in that they recognize jurisdiction over the place of perpetration, place of results and place of destination, but there are some differences. First, the jurisdiction of the place of results is not quantitatively limited. KPIL has not adopted the EU's "mosaic" approach to the jurisdiction of tort claims. Second, it requires foreseeability to recognize the jurisdiction of the forum state. This is the same as the provisions of Article 3(3)(8) of the Japanese Code of Civil Procedure.

Assuming that a U.S.-based AI developer has unauthorizedly scraped the personal information of a Korean resident from the Internet, the Korean resident may be able to sue in a Korean court, at least because Korea is the seat of the legal interest(the place of results).

2.3. Governing Law

KPIL provides the following governing law for torts (Article 52)

Article 52 Paragraph (1) provides that the law of the place where perpetrated and the place where the result occurs shall govern the tort. The place of perpetration is the place where the constituent acts are performed and is usually the place where the actor is present at the time of the tort. The place of results is the place where the protected legal interest is directly violated by the tort, i.e., the location of the legal interest at the time of the violation.[10] Unlike in the case of international jurisdiction , the law of the place of destination cannot be the governing law. Article 17 of Japan's General Rules for Application of Law stipulates that the law of the place of the result is the governing law and requires that the occurrence of the

[10] Kwang-Hyun Seok, "Commentary on International Private Law", Park Young-Sa, 2013, p. 394

result is ordinarily foreseeable in the place of the result, but KPIL does not require foreseeability.

Article 52 Paragraph (2) adopts a common personal law (lex personalis) based on habitual residence. Paragraph (3) provides for subordinate or secondary connections. Paragraph (4) provides for the limitation of damages where foreign law is the governing law.

The illegal act in question in the AI learning phase (INPUT phase) is the act of an AI development company collecting personal information from the Internet without the consent of the data subject and using it for machine learning. Here, the place of perpetration is the location of the AI development company, and the place of result is the habitual residence of the data subject, which is the location of legal interest. If the AI development company is located in the United States and the data subject's habitual place of residence is in Korea, the place of perpetration and the place of result are different, and the question arises as to which law should be the governing law. The theories include the view that the protection of legal interests takes precedence over the illegality of the act, so the place of the result should be prioritized, the view that the victim should choose between the place of perpetration and the place of the result and the court should find the place more closely related to the matter under Article 52 of KPIL. [11] Korean court precedent is understood to recognize the victim's right to choose.[12] Based on these discussions, it is likely that Korean law will be the governing law in cases where the personal information of data subjects residing in Korea is breached. Japan's General Rules for Application of Law also stipulates the law of the place of the result as the governing law for torts, so the law of the data subject's habitual place of residence will be designated as the governing law, but if the result is not foreseeable in the place of the result, the law of the place of perpetration will be the governing law.

3. AI output phase (= OUTPUT phase)

[11] Kwang-Hyun Seok, "Commentary on International Private Law", Park Young Sa, 2013, p. 394

[12] Seoul High Court, judgment 2002na32662, decided on January 26, 2006.

In the AI OUTPUT stage, the case of privacy violation is that the same or similar data learned in the AI learning phase is generated by the user's prompt. However, as mentioned above, the frequency is considered to be minimal due to the constraints of the algorithm and the filtering measures of the AI developer. The characteristic of tortious acts occurring in the AI output phase is almost the same as in the case of the AI input phase, so the discussion of international jurisdiction and governing law is the same as in the case of the AI input phase.

4. AI service phase

4.1. Where the problem is

The types of privacy violation cases that occur in the AI service phase are different from those in the AI learning and AI output phases. In this study, we focus on cases where personal information of members using the AI service was leaked, such as the ChatGPT case in Italy. It is also possible to discuss cases where members using the service exercise the rights of data subjects such as right to access based on the Personal Information Protection Act.

During the AI service phase, the data subject, who is a member, will enter into a service use agreement with the AI service company. In general, service agreements or terms of use contain (exclusive) forum selection clauses and governing law provisions. For example, OpenAI's Terms of Service provides for exclusive jurisdiction in the courts of San Francisco and governed by the laws of the State of California, except for certain arbitration matters, and Midjourney's Arbitration Agreement provides for the resolution of all disputes by arbitration and specifies the laws of the State of California as the governing law. As the service use agreement between the AI service company and the data subject may constitute a consumer contract under private international law, the scope of the validity of the forum selection clause and the choice of law clause is at issue from the perspective of a consumer contract.

The service use agreement or terms of use concluded between the service member and the AI service company stipulates the content and scope of the services that the AI service company must provide to the member. It is also possible for a member to file a claim against

an AI service company seeking fulfillment of the service obligations set forth in the contract. In such cases, the provisions of Article 41 Special Jurisdiction over Lawsuit Regarding Contract of KPIL should be reviewed. If the conduct at issue also constitutes a tort (e.g., breach of privacy), then the special jurisdiction and governing law of the tort will be an issue, but this is redundant to the foregoing and will not be discussed here.

4.2. Special Rule of international jurisdiction over consumer contracts

KPIL makes special provisions regarding jurisdiction and choice of law to protect passive consumers (Articles 42, 47). For reference, the consumer contract section of KPIL is influenced by Articles 15 to 17 of the Brussel I Regulation. The Jurisdiction Provision on Intellectual Property Infringement and Tort also recognizes jurisdiction over "acts directed to Korea," but this provision is based on the practice of U.S. courts and has a different legislative history than the "targeted activity criterion" or "directed to criterion" stipulated in the article 42.

As current AI services such as ChatGPT and Midjourney are all based on the Internet, the interpretation of Article 42(1)(1) of KPIL is important. Article 42(1)(1) of KPIL provides that business activities, including the solicitation of transactions by means of advertisements, are carried out outside the country of the consumer's habitual residence towards the country of the consumer's habitual residence, and the contract falls within the scope of the business activities of the business operator. This part is intended for consumer contracts concluded over the Internet.

Since KPIL on consumer contracts has been influenced by the EU's Brussel I Regulation, Korean theories refer to the CJEU's jurisprudence on the interpretation of this regulation, such as the Pammer and Hotel Alpenhof judgment.[13] In other words, the mere accessibility of a business's website from the consumer's address is not enough to recognize the existence of the business's directed activity, but the business must be able to do business with consumers that the business envisages doing business with. This restrictive interpretation makes sense because

[13] cjeu c-585/08, c-144/09

if the mere fact that a consumer can access a business's website is enough to recognize international jurisdiction, the business would be at risk of being sued around the world.

In the past, KPIL was insufficient to protect active(mobile) consumers by requiring that "the consumer perform the acts necessary to conclude the contract in the country of his habitual residence" (former Article 27(1) of KPIL). The new law removes this requirement and provides for the inclusion of active consumers under certain conditions. These amendments also reflect the attitude of the Brussel I Regulation.[14] However, the above amendment does not include all active consumers, which is different from the Japanese Civil Procedure Law (Article 3.4).

Looking at the current state of AI services, most services such as ChatGPT, Copilot, Midjourney, and Stable Diffusion support the Korean language and accept payments with credit cards issued in Korea, so these AI service companies are considered to be operating 'toward Korea'. According to this interpretation, even if an AI service company stipulates a foreign court as the exclusive forum in its terms of use, a user residing in Korea can file a lawsuit in a Korean court under Article 42(1) of KPIL.

4.3. Applicable Law for Consumer Contracts

KPIL makes special provisions on the governing law for consumer contracts (Article 47), where "consumer contract" means "consumer contract" in Article 42 , which provides jurisdictional special provisions for consumer contracts.

Consumers enjoy special protection under the substantive law of each country, and the legislative purpose of substantive law would be defeated if parties could circumvent the restrictions of substantive law by designating a foreign law as the governing law by agreement. In order to protect consumers, private international law has limited the principle of party autonomy (paragraph 1) and modified the general principles for the determination of the governing law and the method of contract so that the law of the consumer's habitual residence should govern (paragraph 2, 3).

[14] Kwang-Hyun Seok, 'Commentary on the Revisedl Private International Law of 2022', Park Young sa, 2022, p. 227.

The mandatory provision referred to in paragraph (1) of this Article refers to domestic mandatory legislation, i.e. legislation that cannot be excluded by the parties' contract, and generally includes laws enacted for the protection of consumers. Such mandatory provisions are distinct from "internationally mandatory rules" under Article 20 of KPIL, which must be applied regardless of the governing law in light of the legislative purpose.

As mentioned above, AI service use agreements are likely to be consumer contracts, so Korean consumers may be protected by Korean mandatory law even if the terms of use specify a foreign country's law as the governing law.

4.4. Supreme Court of Korea 2017da219232, decided April 13, 2023 (Google ruling)[15]

4.4.1. Facts

Plaintiffs are six natural persons, two of whom use Gmail personally, two of whom are affiliated with Amnesty International Korea and have created Google personal email accounts with the organization's name, "aikorea," in their email account and use Gmail primarily in connection with their work, and two of whom use Google's corporate email service.

Google's Terms of Service provide that any action arising out of or relating to these Terms or the Services shall be brought exclusively in the federal or state courts of Santa Clara County, California, U.S.A., and that the laws of the State of California, U.S.A., shall govern any dispute arising out of or relating to these Terms or the Services.

The plaintiffs filed a lawsuit against Google LLC (U.S.) and Google Korea in a South Korean court, seeking to compel Google to disclose the details of the plaintiffs' personal information and service usage to third parties under the Personal Information Protection Act.[16]

[15] First instance Seoul Central District Court, October 16, 2015, Decision 2014ga38116, and on appeal Seoul High Court, February 16, 2017, Decision 2015na2065729.

[16] At the time of the lawsuit, the Information and Communications Network Act applied to internet service providers such as Google with respect to personal information protection, so the judgment listed the Information and Communications Network Act as the relevant law. However, since all the privacy provisions of the Information

4.4.2. Effect of Exclusive International jurisdiction Agreement

The Korean Supreme Court has previously held that "in order for an exclusive international jurisdiction agreement to exclude the jurisdiction of the Korean courts and make a foreign court the court of jurisdiction, it is required that the case does not fall within the exclusive jurisdiction of the Korean courts and that the designated foreign court has jurisdiction over the case under the foreign law, and that the case has a reasonable connection to the foreign court('rational relevance' requirement), and the exclusive jurisdiction agreement is valid unless such exclusive jurisdiction agreement is grossly unreasonable and unfair and constitutes a legal act contrary to public policy and morals".[17] However, the 'rational relevance' requirement of the above legal doctrine is expected to be deleted in the future under the revised Article 8 of KPIL.

In this case, the Supreme Court affirmed the trial court's decision to validate the plaintiffs' exclusive international jurisdiction agreement with Google LLC under existing law (no exclusive jurisdiction in the Korean courts, jurisdiction in the federal or state courts of Santa Clara County, California, USA, reasonable connection between the Santa Clara County courts and the dispute at issue, and no violation of public policy in light of the fact that the services are provided free of charge).

4.4.3. Whether it is a Consumer Contract

The Court of Appeals held that Google LLC operates a separate domain address for domestic users, provides services displayed in Korean, and engages in business activities such as soliciting transactions for service use contracts through advertisements on the Internet from areas outside Korea towards Korea, as well as obtaining revenue by receiving advertisements

and Communications Network Act have since been transferred to the Personal Information Protection Act, the Personal Information Protection Act will be listed as the applicable law in this study for convenience.

[17] Supreme Court, Aug. 26, 2010, Decision 2010da28185

from companies or individuals operating in Korea, and it's conduct falls under the definition of 'consumer contract'("professional or business activities, including the solicitation of transactions by advertisements, from a country outside the country to the country of the consumer" under Article 27(1)(1) of the former KPIL).[18]

In addition, domestic users sign up for Google services using a computer terminal connected to the Internet network in Korea and create a Google account, which constitutes 'a consumer' under Article 27(1)(1) of the former KPIL, which states that "the consumer has performed an act necessary to conclude a contract in the country." [19]

However, the court held that the two plaintiffs who created personal emails by including the name of the organization to which they belonged in their Gmail account and used the email accounts mainly in connection with their work, and the two plaintiffs who used Google's corporate email service, were not within the scope of the consumer contract because they were using Google's services for professional activities, and only the two plaintiffs who used Gmail for personal use were within the scope of the consumer contract. The Supreme Court affirmed the appellate decision.

4.4.4 Effect of Exclusive International Jurisdiction Agreement for Consumer Contracts

The Court of Appeals held that even if the parties to a consumer contract had agreed to exclude the jurisdiction of the Korean courts before the dispute was issued, such an agreement was invalid in violation of Article 27(6) of the former KPIL (now Article 42 of the Act), and that the two plaintiffs, who used Gmail for personal use, could sue in the Korean courts despite the exclusive international jurisdiction provision. The Supreme Court affirmed the appellate decision.

[18] Article 42(1) of the current Private International Law

[19] The requirement in Article 27(1)(1) of the old Private International Law Code that the "consumer has performed an act necessary for the formation of a contract in the country" has been deleted from the current Private International Law Code.

4.4.5. Whether the Privacy Act falls under the internationally mandatory provisions of Article 20 of KPIL

The Court of Appeals held that, in view of the unclear legislative intentions regarding the international mandatory nature of the Personal Information Protection Act and the fact that Article 27 of the former KPIL (now Article 47) separately provides for mandatory rules for consumer contracts with respect to the designation of governing law for the protection of consumers, the Personal Information Protection Act cannot be considered as an international mandatory law under Article 7 of the former KPIL (now Article 20), and that the Personal Information Protection Act cannot be applied as an international mandatory law to contracts that agree on foreign law as the governing law.

The Supreme Court did not decide this ground of appeal because it was raised by the plaintiffs who denied the jurisdiction of the Korean court. The appellate court's reference to the mandatory regulation of consumer contracts as a basis for denying the international mandatory regulation of privacy laws has been criticized by scholars.

4.4.6. Effect of Choice of Law Agreements on Consumer Contracts

The Court of Appeals considered Article 35(1) and (3) of the Personal Information Protection Act to be domestic mandatory rules under Article 27(1)(1) of the former KPIL (now Article 47(1)) and held that the two plaintiffs were protected by the provisions of the Personal Information Protection Act, which are mandatory rules of their country of habitual residence, Korea. The Supreme Court also determined that the provisions of the Personal Information Protection Act constitute (domestic) mandatory regulations, considering the purpose and objectives of the Personal Information Protection Act, the functions and roles of the relevant provisions for the protection of personal information, and the sanctions imposed on business operators in case of violation, and approved the judgment of the trial court. Accordingly, it

was determined that the two plaintiffs could exercise their right to access personal information against Google LLC (USA) and Google Korea.

4.4.7. Whether the right to access personal information is limited by U.S. law

As shown above, the Court of Appeals and the Supreme Court did not differ significantly in the content of the judgment, but they differed in their conclusions on the scope of the right to access personal information that the two plaintiffs (who used Gmail personally) could exercise.

The Court of Appeals held that Google LLC is required to comply with U.S. laws in addition to Korean laws, including 18 U.S.C. §2709(c)(1) and 50 U.S.C. §1861(d), 50 U.S.C. §1861(d), which imposes an obligation on Google LLC not to disclose that a U.S. government agency has accessed a user's personal information, subject to certain requirements, and therefore Google LLC is obligated to disclose the details of the third-party provision except for those that it is obligated not to disclose under these laws.

On the other hand, the Supreme Court held that a business that must comply with foreign laws in addition to Korean laws is not justified in principle in refusing to provide access to personal information solely because the foreign law restricts the disclosure of such information, but that the contents of such foreign laws may be considered to determine whether there are legitimate reasons for the refusal.

In conclusion, the Supreme Court held that whether Google LLC has taken all the necessary measures under the Personal Information Protection Act requires a comprehensive examination and review of whether the non-disclosure obligation under the foreign law is consistent with the content and purpose of the Korean Constitution, laws, etc., whether the need to respect the foreign law is significantly superior to the need to protect personal information, and whether the non-disclosure requirements under the foreign law have been met with respect to the information requested by the user, and whether the service provider

has actually borne the non-disclosure obligation. For this specific examination, the Supreme Court remanded the appellate judgment.

In the case of Korean domestic mandatory of consumer contracts, the question arises as to which country's law should be prioritized for foreign businesses that must comply with both Korean and foreign laws. The Korean Supreme Court's ruling above is a good example of the court's consideration of both Korean and foreign laws in this case, while specifying the issues to be considered.

IV. Conclusion

In order to discuss private international law issues related to generative AI models and its services, it is first necessary to understand the operating principles of generative AI and the types of disputes (or characterizations). Since generative AI is designed to generate outputs similar to the probability distribution of training data, there is a risk of copyright and privacy infringement at the AI output phase. Therefore, it is necessary to consider the AI output phase separately from the perspective of the type of dispute. However, AI development companies are also aware of this problem, and as a solution, they are removing duplicate data from the training data (if there is a lot of duplicate input data in the machine learning process, the AI system is more likely to generate output similar to duplicate data) and taking filtering measures at the AI output phase to prevent the creation of other people's works or personal information. The AI output phase overlaps with the AI service stage, but to avoid redundant review, the legal issues in the AI service stage can be reviewed centering on the relationship between the service member and the AI service company.

The possibility of copyright or privacy infringement in the AI learning phase is a difficult problem for machine learning algorithms to avoid. From the perspective of AI development, it is difficult to fully secure the consent of copyright holders and data subjects in the process of collecting and training a large amount of data from the Internet, and even if there is a commercial market for training data, it is difficult to obtain a license for the data and develop an AI model due to the burden of huge license fees. As a result, it is necessary to collect and

use training data by utilizing fair use defenses under copyright laws, TDM exceptions, and lawful basis under privacy laws, and until best practices are established, it is expected that international litigation will continue to occur over the AI training phase.

If we look at the stages of AI model development and its service sequentially as described above, the phase where legal issues are raised can be divided into (1) AI learning phase (INPUT phase), (2) AI output phase (OUTPUT phase), and (3) AI service phase (SERVICE phase), and the type (or characterization) of the dispute and the issues of private international law can be discussed for each of the above phases. Furthermore, since generative AI models and systems are entirely dependent on the Internet environment, existing theories and precedents developed in the context of the Internet are largely applicable to the analysis of the above issues.

(end of document)