Legal Protection of Big Data in Japan

Kei Teshirogi
teshirogi@ohebashi.com

A. Introduction

Today, the importance of information as an asset is increasing due to the evolution of digital network technology and its widespread use in business. The increase in the number of work-from-home situations triggered by the pandemic has also made the value of information for individuals even greater. In particular, among the different types of information, the group of data known as "big data" is becoming a source of value for companies given the remarkable development of IoT, the means of collecting data, and AI, which analyzes and utilizes such data.

The purpose of this article is to give an overview of the legal protection of big data in Japan in light of the increasing value thereof for international use. This article will first discuss the definition and key characteristics of big data, followed by a description of the two types of legal protection thereof in Japan under the Copyright Act[1] and the Unfair Competition Prevention Act.[2] As for the protection under the Unfair Competition Prevention Act, a draft revision of its related guidelines was published on March 23, 2022,[3] and the latest discussions thereon will be presented in this article.

B. Definition and Key Characteristics of Big Data

While the term "big data" has come to be used in everyday conversation, there is no clear, universal definition thereof. Big data, stated abstractly, refers to a huge group of data composed of various types and formats of data, including unstructured data. Notably, Gartner, Inc., a major U.S. IT research firm, defines big data as having the following "Three Vs" as its key characteristics:[4]

> Big data is high-volume, high-velocity and/or high-variety information assets that demand cost-effective, innovative forms of information processing that enable enhanced insight, decision making, and process automation.

In its "White Paper on Information and Communications (2017 edition)," the Ministry of Internal Affairs and Communications classifies big data into the following

[1] Chosakukenho [Copyright Act], Act No. 48 of May 6, 1970.

[2] Fuseikyosoboshiho [Unfair Competition Prevention Act], Act No. 47 of May 19, 1993.

[3] Ministry of Economy, Trade and Industry ("**METI**"), "gentei teikyo data ni kansuru shishin (kaitei-an) [Guidelines on Shared Data with Limited Access (revised draft)] (2022) (available at https://www.meti.go.jp/shingikai/sankoshin/chiteki_zaisan/fusei_kyoso/pdf/016_04_00.pdf (in Japanese)). Although this draft has not yet been finalized, it is useful in understanding the direction of the latest discussions.

[4] "Big Data," *Gartner Glossary*, Gartner, Inc. (2022) (available at https://www.gartner.com/en/information-technology/glossary/big-data).

four categories, focusing on the entities that generate the data:[5]

1. "Open data" provided by national and local governments;
2. Digitalizing and structuralizing knowhow;
3. M2M (Machine to Machine) streaming data; and
4. "Personal data" involving attributes.

With reference to the above categories, in this article, big data will be defined as the data generated by governments, companies or individuals that are difficult to manage with ordinary software because of their high-volume, high-velocity or high-variety.

C. Legal Protection of Big Data in Japan

1. Two types of protection

There are two ways to legally protect big data in Japan. First is the method of protecting data as if it were a physical object by giving it the effects of ownership. This is called the rights-granting type of protection. In contrast, the second way to protect data is by regulating the conduct related to such data rather than protecting the data itself by creating rights to it. This is called the rule-of-conduct type of protection.

In Japan, the Copyright Act protects data by granting copyrights to databases as a form of the rights-granting type of protection. In addition, the Unfair Competition Prevention Act regulates unfair acts related to certain data called "shared data with limited access" as a rule-of-conduct type of protection.

2. Data protection under the Copyright Act

The Copyright Act defines databases as follows and recognizes the copyrightability of certain databases:

> "Database" means an aggregate of data such as articles, numerical values, or diagrams, which is systematically constructed so that such data can be searched with a computer.[6]
>
> A database that, by reason of the selection or systematic construction of information contained therein, constitutes a creation, is protected as a work.[7]

Therefore, for data to be protected as a copyrightable work in Japan, (a) the data must be organized in such a way that it can be retrieved by a computer, and (b) the selection or systematic organization of the information must be creative.

In one case, the court recognized the copyrightability of a database that classified the telephone number information of businesses nationwide based on occupation.[8] In this case, the court affirmed the copyrightability of the database on the grounds that its occupational classification system was structured to cover all occupations from the viewpoint of search convenience, and that it was devised uniquely by the plaintiff.[9]

In another case, the court recognized the copyrightability of a database containing information such as tourist facility data, accommodation data, and similar data, which was used by a travel agency to

[5] At 12-13 (available at https://www.soumu.go.jp/johotsusintokei/whitepaper/eng/WP2017/chapter-2.pdf).

[6] Copyright Act, art. 2, para. 1, item 10-3.

[7] *Id.*, art. 12-2, para. 1.

[8] *See* Tokyo District Court, March 17, 2000, Hei 8 (wa) No. 9325, Saibansho Web, https://www.courts.go.jp/app/files/hanrei_jp/286/013286_hanrei.pdf (in Japanese).

[9] *See* Id., at 6.

[10] *See* IP High Court, January 19, 2016, Hei 26 (ne) No. 10038, Saibansho Web, https://www.courts.go.jp/app/files/hanrei_jp/639/085639_hanrei.pdf (in Japanese).

prepare a process chart.[10] In the decision, the court interpreted the creativity requirement in item (b) above as follows:

> If there is a range of choices in the selection or systematic composition of information, and if the selection or systematic composition of information in a particular database shows some individuality of the creator, then it can be understood that the database may be recognized as having creativity through the selection or systematic composition of information, as if the creator's thoughts or feelings were transferred during the production process and his/her thoughts or feelings were expressed in a creative manner.[11]

On the other hand, the court denied the copyrightability of a database that collected information on actual automobiles in Japan on the grounds that the automobile data items were only arranged in order from the oldest to newest automobiles, without any further classification, and that such classification had been adopted by other companies.[12]

In light of these court cases, it can be said that a simple accumulation of raw data will not be protected by the Copyright Act, and that for big data to be protected as a copyrighted work, it must be systematized in some creative or innovative way.

3. Data protection under the Unfair Competition Prevention Act

The Unfair Competition Prevention Act defines certain data as "shared data with limited access," as further described below, and regulates certain acts pertaining to such data as acts of unfair competition. The said Act does not apply to acts pertaining to information that is identical to that made available to the public free of charge.[13]

> "Shared data with limited access" as used in this Act means technical or business information that is accumulated to a significant extent and is managed by electronic or magnetic means (meaning an electronic form, magnetic form, or any other form that is impossible to perceive through the human senses alone; ...) as information to be provided to specific persons on a regular basis (excluding information that is kept secret).[14]

According to the above provision, the following six requirements must be met for certain data to qualify as shared data with limited access:

a. The data is provided to a specific person on a regular basis;
b. The data is accumulated to a significant extent by electronic or magnetic means;
c. The data is managed by electronic or magnetic means;
d. It is technical or business data;
e. The data is not kept secret; and
f. The data is not the same as any information that has been made available to the public without compensation.

[11] *Id*., at 37.
[12] *See* Tokyo District Court, May 25, 2001, Hei 8 (wa) No. 10047, Saibansho Web, https://www.courts.go.jp/app/files/hanrei_jp/333/034333_hanrei.pdf (in Japanese).
[13] Unfair Competition Prevention Act, art. 19, para. 1, item 8-Ro.
[14] *Id*., art. 2, para. 7.
[15] METI, "gentei teikyo data ni kansuru shishin" [Guidelines on Shared Data with Limited Access] (2019) (available at https://www.meti.go.jp/english/policy/economy/chizai/chiteki/pdf/guidelines_on_shared_data_with_limited_access.pdf).

METI established the Guidelines on Shared Data with Limited Access on January 23, 2019[15] to interpret the above requirements. It also published its draft revisions on March 23, 2022.[16] In the proposed revisions, the rationale for the requirements was further clarified. To introduce some important points in the draft guidelines, first, the term “provided” in the first requirement includes not only the case of an actual provision of data but also covers cases where an intention to provide the data is acknowledged. This requirement is satisfied, for example, when the data holder allows customers to access data in its cloud. As to the second requirement, satisfaction of the “significant extent” requirement depends on the nature of the subject data. As an example, in the case of a business operator that accumulates cell phone location information based on a nationwide area and then extracts and sells such information in units of specific areas, it has been pointed out that the data for such specific areas is highly likely to satisfy this requirement. In the third requirement, the fact that the data is “managed” is required to ensure that third parties can foresee that such data may constitute shared data with limited access. Thus, measures such as access restrictions must be in place so that third parties can become aware of the data provider’s intention to control the data. The fourth requirement ensures that information that is illegal or offensive to public order and morals, such as child pornography and information about prohibited drugs, will be excluded from protection. The fifth requirement is designed to avoid duplication of protection with “trade secrets” that are already protected by the Unfair Competition Prevention Act.

When data constitutes shared data with limited access, the following acts with respect to such data are regulated as acts of unfair competition:[17]

a. Obtaining, using or disclosing the data through theft, fraud, threats or other wrongful means;
b. Use or disclosure of the data by a person who has received the data from the data holder for the purpose of obtaining unjust profits or causing damage to the holder (however, with respect to the act of use, it is limited to one in violation of a duty relating to the management of the data);
c. Obtaining, using or disclosing the data with knowledge that the data has been unlawfully obtained or unlawfully disclosed; and
d. Disclosure of the data by a person who had bona fide intentions at the time of acquisition of the data, but who learned thereafter that an act of unlawful acquisition or disclosure had intervened (excluding, however, acts of disclosure within the scope of the title acquired to the data).

A person whose business interests are infringed by any of the above-mentioned acts may file a claim for injunction or damages against the offender. The draft revisions to the guidelines also refer to the relationship between a claimant and a platform provider. For example, if a platform provider played a role in providing the environment that mediated and facilitated the provision of data by a data holder to an acquirer, then the platform provider can also be a claimant because it is the entity that stored and managed the data electromagnetically.

D. Conclusion

As mentioned above, big data in Japan is mainly protected by two laws with each of them seeking to provide a more appropriate form of protection as they are supplemented by the accumulation of court precedents and the formulation of guidelines. However, striking a balance between the protection of

[16] *See* https://www.meti.go.jp/shingikai/sankoshin/chiteki_zaisan/fusei_kyoso/pdf/016_04_00.pdf (in Japanese).
[17] Unfair Competition Prevention Act, art. 2, para. 1, items 11-16.

investments in data collection and the sharing of information can still be difficult depending on the circumstances of each case. At any rate, since Japan was the first country to introduce the rule-of-conduct type of protection of big data under the Unfair Competition Prevention Act, the accumulation of cases in Japan is expected to provide a valuable source of examples of rule-of-conduct type of protection that will benefit the international community.

DISCLAIMER

The contents of this Newsletter are intended to provide general information only, based on data available as of the date of writing. They are not offered as advice on any particular matter, whether legal or otherwise, and should not be taken as such. The authors and Oh-Ebashi LPC & Partners expressly disclaim all liability to any person in respect of the consequences of anything done or omitted to be done wholly or partly in reliance upon the whole or any part of the contents of this Newsletter. No reader should act or refrain from acting on the basis of any matter contained in this Newsletter without seeking specific professional advice.