

# Who owns the oil that fuels AI machines: Japan's evolving IP protection for data

*Morgan Lewis & Bockius lawyers Akiko Araki and Christopher Wells in Tokyo analyse changes to Japan's IP regime designed to better protect AI businesses' data ownership.*

With the development of emerging technologies including AI in recent years, data is now recognised as one of the most important raw materials for many future industries. AI-related technologies such as “deep learning” require massive and continuously updated data sets from a variety of sources to create useful recursive algorithms to guide prediction and decision-making in emerging technologies from medical diagnosis to autonomous driving.

Because of these technologies, the right to use and exploit large sets of data collected from sensors, points of sale, business decision-making and other information sources has become a key area of intellectual property law developments in Japan. Unfortunately, many of the issues related to the rights in parties collecting and seeking to exploit ownership in such data are not currently well developed under Japanese law.

For example, the question of ownership rights in data, and whether data holders can control the further use of data shared with a third-party user – and to what extent – remains uncertain. There are both upsides and downsides to protecting this data. Overly strong protection of data could impede utilisation of data



**AKIKO ARAKI**



**CHRISTOPHER WELLS**

as it is impractical to get consent or permission from all rights holders of huge data sets. On the other hand, less protection may also impede utilisation, as data holders could be unwilling to share such data with third parties, since maintaining the secrecy of the data may be the only method of recovering the cost of collection. In recent years, the Japanese government has been seeking an appropriate balance.

A related and more difficult question for the AI and big data industries is whether data holders have any rights in the AI products derived from the data using AI programmes or recursive algorithms as a form of derivative work, and to what extent data holders can control derivative works. This issue has also been the

subject of intense debate in Japan in recent years.

## **Balancing the protection and utilisation of data resources**

With respect to the data utilisation/protection issue, one recent Japanese government policy thrust is directed at strengthening the protection of data through the introduction of a new protected data system, which came into effect on 1 July 2019.

Since data is an intangible asset, it is not subject to property or possessory rights under Japanese law. This means that in general, no one can obtain exclusive rights to the use of factual data unless it is covered by another recognised IP right, such as copyright or patent. However, in principle, factual data that the market might consider a valuable data set cannot be copyrighted or patented, as it generally lacks the required creativity elements. Nevertheless, data that has been sufficiently processed – such as compilations – may become eligible for copyright protection if sufficient creativity in its organisation can be demonstrated.

However, in the context of AI businesses, it is still often difficult to obtain copyright protection because



## RELATED ARTICLES

Mixed messages from Poland: the uneasy relationship between regulations on reuse and the GDPR  
[gdr.news/1203436](#)

Data breaches: beyond the ICO into the big bucks  
[gdr.news/1176506](#)

the data sets used are sometimes created in general form, meaning sufficient creativity in the organisation of the data cannot be found. In addition, some data is collected automatically by sensors or other means (not through human intervention). When that is the case, these data sets are not copyrighted because the required human contribution to the creative work cannot be found.

Protection by contractual undertakings between parties is possible, but these are relatively difficult to enforce owing to the absence of a robust disclosure regime. Recognising the difficulties facing persons wishing to provide collected data under contractual arrangements, Japan's Ministry of Economy, Trade and Industry (METI) in 2018 issued "Contract Guidelines on Utilisation of AI and Data", which includes sample legal language to help facilitate uniformity in interpretation.

An alternative method for seeking protection for rights in factual data collected by data holders would be through trade secret protection under

the Unfair Competition Prevention Act (UCPA). However, the ability to obtain conventional trade secret protection of data under that legislation has also been limited.

Under the Agreement on Trade-Related Aspects of Intellectual Property Rights, as in many other jurisdictions, seeking trade secret protection in Japan means the relevant data must not be publicly available, must be subject to reasonable efforts to maintain its secrecy and must have economic value. Although Japanese requirements for trade secret protection are similar to those in other jurisdictions, in practice, the courts tend to interpret the "reasonable efforts to maintain secrecy" requirement more strictly than in many jurisdictions (including notably the US). As a result, the willingness of courts to enforce trade secrets against misappropriation has been relatively limited in comparison.

In the 10 years between 2003 and 2012, only 76 cases of trade secret misappropriation litigation were reported

nationwide in Japanese district courts. In the context of the massive data sets used by the AI industry, attempts to obtain trade secret protection against misappropriation may have been largely unsuccessful because of the difficulty in demonstrating that there has been a “reasonable effort to maintain secrecy” when seeking damages or injunctions for the misappropriation of such data.

### **The “protected data” system**

Because of the above issues, and to create an environment where businesses are incentivised to incur the substantial costs of collecting and compiling data that can be licensed to AI industry participants, METI has recently amended the coverage of the UCPA and created a new civil remedy against misappropriation of “protected data”. These new rules came into force on 1 July 2019. The main purpose of this revision was to protect the economic value of big data shared with a limited group of permitted users engaged in AI businesses and whose data currently enjoys only limited protection under conventional trade secret principles.

This UCPA amendment introduces the concept of “protected data” as the subject matter of the protection, which is defined to mean technical or business information accumulated in a reasonable amount and managed by electronic or magnetic means, meaning an electronic form, magnetic form or any other form that is impossible to perceive through human senses alone (“electromagnetic management”), that is provided to specified persons regularly. According to the “Guidelines on Protected Data” published by METI on 23 January 2019, for the electromagnetic management requirement to be satisfied, it is necessary to ensure that third parties could recognise the intention of data holders to control the use of the relevant data as something to be provided exclusively to permitted users.

The UCPA defines several categories of acts of misappropriation,

including unauthorised acquisition and significant violation of the principle of good faith, and provides that these are subject to civil remedies including injunctions and damages. This new system of protecting data has also opened the gate to seeking protection of data holders when they share their data with permitted users – which previously could not be protected as a trade secret. Data holders will need to satisfy the requirements for protection, including electromagnetic management, to ensure that the data shared with permitted users falls within the definition of protected data.

While this new system will be helpful in protecting the rights of data

of information such as personal data and otherwise copyrighted works receive such protection. In policy discussions, these strong protections were considered as obstacles to the beneficial utilisation of such data by AI businesses. In view of these considerations, Japanese government policy has aimed to loosen traditional protections through recent amendments to the Act on the Protection of Personal Information (PIPA) and the Copyright Act.

When data provided for use in AI and deep learning contains personal data, the consent of each individual contributing data is generally required under the PIPA. Since obtaining

## **IN THE 10 YEARS BETWEEN 2003 AND 2012, ONLY 76 CASES OF TRADE SECRET MISAPPROPRIATION LITIGATION WERE REPORTED NATIONWIDE IN JAPANESE DISTRICT COURTS.**

holders sharing data with permitted users, the acts of misappropriation that will be viewed as constituting unfair competition with respect to protected data are limited to the specific acts specified in the revised UCPA. For example, disclosure of protected data by a subsequent acquirer who is in good faith at the time of acquisition largely remains outside the scope of civil remedies under these provisions.

### **Promoting AI businesses’ use of data sets**

While factual data has generally not been easily protected under Japanese IP laws, policymakers appear to have recognised that certain data sets have traditionally been afforded strong protection. For example, special categories

individual consent is costly for AI application developers, 2015 revisions to the PIPA (effective 30 May 2017) introduced a new “anonymously processed information” exemption from PIPA coverage, under which a business operator can provide such information to a third party without obtaining prior consent if the operator fulfils certain protective obligations. In addition, a new law, the Next Generation Medical Infrastructure Act, permitting the transfer of anonymised medical data for R&D contribution purposes, became effective in May 2018. The law enables the utilisation of medical data where certified anonymised data agents ensure proper anonymisation. However, even this law allows individual patients to expressly opt out.

In addition, in May 2018, the Copyright Act was amended (effective 1 January 2019) to facilitate the use of technology involving data mining. The Copyright Act now provides that to the extent it is considered necessary, it is permissible to exploit certain copyrighted works in data analysis and other ways where the exploitation is not directed at enjoying the idea or emotion of the copyrighted work. This provision permits copying and other exploitation of copyrighted works without obtaining prior permission. This provision is unique in that it is applicable to for-profit activities, as well as not-for-profit activities. Professor Ueno of Waseda University in Tokyo has labelled Japan as the “paradise of machine learning”.

Another major intellectual property issue is whether a derivative work developed in an AI business can benefit from IP rights. More specifically, one of the most controversial issues in Japan is whether AI products, which constitute the outputs of data input to recursive algorithms, can be protected by IP rights – and if so, who owns the rights. Under current Japanese law, where a human being creates the relevant AI products using AI as a tool, that product may be eligible for IP protection, but if AI creates products without the contribution of a human being they are not eligible for protection.

While this seems a relatively straightforward rule, it is unfortunately not that easy to distinguish whether such products were created with a contribution by a human being. Further, if there is evidence of a contribution by a human being, and the product is eligible for protection through traditional IP rights, it is often not necessarily clear who owns these rights: in the course of producing the final creation, several actors may have been involved, including humans participating in the process of collecting data, creating the relevant AI programme, using the AI, and so on.

### **Best practices for data holders**

Based on the default rules described above, while the strong protection of certain categories of data such as personal data and copyrighted works used in machine learning in AI businesses have been modestly reduced, protection of general data used for machine learning has now expanded with the introduction of the UCPA’s protected data system. Unfortunately, the system remains imperfect because of the limited scope of the misappropriation acts that constitute unfair competition. In addition, legal protection is limited and ownership rules remain unclear for AI products resulting from recursive algorithms.

Japanese government is continuing to discuss the developments in data protection and utilisation in this industry. Based on the status quo, however, it is important for data holders to review and take protective action before sharing their data with a third party by:

- retaining sufficient evidence of either a reasonable effort to maintain secrecy or electromagnetic management of data to ensure the data will be considered a trade secret or protected data under the UCPA if it is misappropriated;
- preparing agreements permitting the use of data that specify the terms of use, legal liability for

## **UNDER CURRENT JAPANESE LAW, IF AI CREATES PRODUCTS WITHOUT THE CONTRIBUTION OF A HUMAN BEING THEY ARE NOT ELIGIBLE FOR PROTECTION.**

Under these circumstances, data holders should prepare agreements that specify the terms of use, legal liability for unauthorised disclosure, and rights in derivative data and AI products resulting from machine learning or recursive algorithms. The “Contract Guidelines on Utilisation of AI and Data” published by METI in 2018, which summarises factors to be considered when drafting agreements regarding the utilisation of data and AI, provides considerable guidance and sample language for these agreements that appears to be becoming more widely adopted.

### **Other considerations and recommendations**

As the scope of legal protection afforded to data holders and AI businesses remains uncertain, the

unauthorised disclosure, and any rights in derivative data and AI creations resulting from the machine learning or trained models using such data, and so on; and

- closely monitoring legislative and regulatory developments in Japan concerning the protection of AI and deep learning-related data as recent changes have narrowed certain protections while expanding others. Agreements should be drafted flexibly with a view to allowing amendments to take into consideration continuing changes to Japanese IP laws.