



December
2017

Legal issues regarding the intellectual property of Artificial Intelligence (A.I.)

Shinnosuke Fukuoka

This newsletter outlines the relevant legal issues in developing and using A.I. Although A.I. is being widely discussed in the media and by the general public, it seems that the various legal issues relating to A.I. are barely being acknowledged or analyzed at the present. In this newsletter, we discuss the type of legal issues that are relevant to the intellectual property of A.I., how to protect the Intellectual Property of A.I., and what you should pay attention to when developing and using A.I. I would like to provide a general explanation on the legal issues that arise and need to be considered. The detailed explanation of legal issues of A.I. will be provided in the future newsletter.

1. Process and development in using A.I.

Typically, the process of developing and using A.I. is as follows.

- ① User provides data to the developer
↓
- ② Developers develop A.I. (sometimes it is already developed)
↓
- ③ The developer prepares the learning dataset so that it conforms to the A.I. program developed by the developer about the received data
↓
- ④ The developer learns the A.I. program using the learning dataset
↓
- ⑤ Learned models are completed, the developer delivers the product or provides services to users
↓
- ⑥ The user or developer adds the new learning dataset from the additional data to the learned model and provides the additional learning

This newsletter is the product of its authors and does not reflect the views or opinion of Nishimura & Asahi. In addition, this newsletter is not intended to create an attorney-client relationship or to be legal advice and should not be considered to be a substitute for legal advice. Individual legal and factual circumstances should be taken into consideration in consultation with professional counsel prior to taking any action related to the subject matter of this newsletter.

Such a series of flows can be divided into (1) a learning dataset preparation stage, (2) a learned model creation stage, and (3) a learning model use stage.

2. Legal Issues regarding the Intellectual Property of A.I.

(1) Preparing the dataset for learning

In order for A.I. to learn, a learning dataset is necessary, which is created from raw data. The types of raw data used include, for example, the data of a website or SNS, data of a user such as a customer's purchase records or factory operation data. Therefore, when using another person's work such as a photos or text data uploaded to the website for creating a learning dataset, care must be taken to ensure the copyright is not infringed.

In addition, the learning dataset is created by collecting raw data, selecting it, and processing it. For machine learning, it is necessary to attach teaching data to raw data, but this requires a considerable amount of labor. In addition, having the requisite know-how for collecting, selecting and processing raw data to a learning dataset is also important, and this know-how comes from a deep understanding of A.I. and much trial and error and experience. Therefore, protection of the intellectual property of the learning dataset is an issue that needs to be considered.

(2) Learned model creation stage

In machine learning, a program imports a learning dataset as the input data and the parameters, such as the weighting parameters, are adjusted so as to minimize the error of the output data of the input data. A learned model is generated by iteratively calculating such adjustment. Providing learning to A.I. requires a huge amount of labor, cost, machine power and know-how. For that reason, the question of who has the right to the learned model and how to protect that right is important.

The learned model is comprised of a program and numerical data, such as tensors, of learned parameters. If the A.I. program is generic or OSS, since there is no great difference in the program part, the learned parameters can be the source of competitiveness. However, since the learned parameters consist of data that is automatically generated by A.I., the issue of whether there is copyright in the learned parameters needs to be considered.

(3) Use stage of learning model

When the learning model is made, it will be used for practical purposes. However, this is not the end. It is supposed to continue learning by being provided with new data, to adjust the learned parameters so that its performance is improved, and to evolve the learned model. Therefore, the issue of what rights the right holder of the original learned model and the data provider have to the learned model (derivative model) after the evolution need to be considered. Additionally, those who do not have a right of the learned model alter the learned parameters to create the derivative model.

3. How to deal with legal issues concerning the intellectual property of A.I.

Although copyright law, patent law, and unfair competition prevention law can be considered as the main methods of protecting

A.I., this newsletter focuses on copyright. Under Japanese copyright law, a work is protected when it (1) expresses thoughts or feelings (2) possesses creativity (3) is an expression and (4) belongs to the realm of literature, academic, art or music (Copyright Law 2.1 item 1).

(1) Preparing a dataset for learning

In principle, using the work of another person without their permission can constitute copyright infringement; therefore, creating a learning dataset using another persons' work may infringe copyright and alteration rights. However, according to Article 47-7 of the Copyright Law, as an exception, it is permissible to record on the recording medium and alter it to the extent it is deemed necessary for the purpose of computer information analysis.

Please note that under Article 47-7 of the copyright law such permission is limited to recording to the recording medium and alteration, and it is not permissible to assign it or upload it to the web, which is an infringement of the copyright or the public transmission right. The learning dataset is protected by copyright if it is a database work. In order to be recognized as a database work, it is necessary that (1) the information is systematically organized so that it can be retrieved, and (2) it possesses creativity due to the selection of information and its systematic composition (Copyright Law 12, Article 2, paragraph 1). Among these requirements, whether the requirement of "creativity" is satisfied is often argued.

(2) Learned model creation stage

It can be said that the learned model is a combination of a program and a learned parameter. With regard to programs, the copyright law generally protects a program work; therefore, a large scale and complex program is likely to have a copyright. On the other hand, since learned parameters are those that A.I. automatically creates, it is a question of whether it is an "expression of thought or emotion" or "creativity", which is a requirement for copyright.

In this regard, no criteria have yet to be established, since there are no precedents at the moment and it is not clearly stipulated by the copyright law. Since learned parameters are created automatically by A.I., it is questionable whether they would be recognized as copyrighted works because they are not expressions of "thought or emotion" or are not "creative". However, there is room to believe that expressions of "thought or emotion" and "creativity" exist because the learned parameters are created by neural networks and a learning method that reflects a human's original idea.

At the present, under Japanese copyright law, it is uncertain whether copyright would be granted for the learned parameters of a learned model. Therefore, to protect A.I. legally, it is important to protect these rights by contract or the unfair competition prevention law.

(3) Use stage of a learning model

If the derivative model is a modified version of the learned parameters' part of the original learned model, the copyright holder of the program of the original learned model will have copyright in the program of the derivative model. On the other hand, with respect to the learned parameters, if there is no copyright in that part, anyone can use and modify it. Those who create the original learned model do not have the copyright in the learned parameters of the derivative model. On the other hand, if there is copyright in the learned parameters, in principle, it cannot be used or modified without the consent of the holder of the copyright. Those who

create the original learned model may have the copyright in the learned parameters of the derivative model. However, even if the learned parameters are protected by copyright, the owner of a copy of the learned parameters can duplicate and alter it to the extent it is deemed necessary for him/her to use it on his/her own computer (Article 47-3, paragraph 1 of the Copyright Law).

4. Conclusion

Finally, as mentioned above, the legal issues concerning developing and using A.I. are diverse. In order to promote the development and use of A.I., it is important that many people understand how to deal with these legal issues of A.I. and that the legal system and the laws where the rights are appropriately protected and people can use A.I. at ease are established.



Shinnosuke Fukuoka

Partner

E-mail: s_fukuoka@jurists.co.jp

Shin Fukuoka is a partner of N&A. In the area of Robotics/Artificial Intelligence, he mainly handles A.I., Big Data and IoT. He has also contributed to many publications, including “Law and Strategy of the Internet of Things and Artificial Intelligence”.