Abstract

Data and Fair use

Kim, Yun-Myung*

Data collection and use are the beginning and end of machine learning. Looking at ChatGPT, data is making machines comparable to human capabilities. Commercial purposes are not naturally rejected in the judgment of fair use of the process of producing or securing data for system learning. The UK, Germany, and the EU are also introducing copyright restrictions for data mining for non-profit purposes such as research studies, and Japan is more active. Japan's active legislation is the reason why there are no comprehensive fair use regulations like Korea and the United States, but it shows its willingness to lead the artificial intelligence industry. In 2020, a revision to the Copyright Act was proposed in Korea to introduce restrictions for information analysis. It will be able to increase the predictability for operators. However, the legislation of the amendment is expected to be opposed by the right holder and may take time. Therefore, it was examined whether machine learning such as data crawling and TDM corresponds to fair use through fair use under the current copyright law. In conclusion, it was considered that it may correspond to fair use, citing that it is different from human use behavior. However, it is questionable whether it is reasonable to attribute all exclusive negligence to the business operator by using the works of others according to fair use. The reason why the compensation system for profits earned by operators through the use of machine works generated by TDM or machine learning cannot be excluded from the possibility of

* Kiyul Patent Office Research Fellow, Ph.D. in Law.

serious consequences for a fair competitive environment.

Keywords

fair use, learning data, culture enjoyment, innovation, machine learning, crawling, ChatGPT