Abstract

AI Learning Data and Copyright:
Analysis of EU and France's TDM Regulations and Their Implications

Won Lee*

This study aims to analyze the legal trends of the European Union (EU) and France regarding the regulation of text and data mining (TDM), a core process in AI learning, and to derive implications for policy development. The 2019 EU Copyright Directive allowed limitations on copyright for scientific research purposes by research institutions and cultural heritage organizations while enabling TDM unless rights holders explicitly opt out, creating a developer-friendly system. Subsequently, the 2024 EU AI Act strengthened copyright compliance obligations for General-Purpose Artificial Intelligence (GPAI) models, requiring not only the confirmation of rights holders' opt-out intentions but also the disclosure of detailed summaries of training data. This represents a significant step in enhancing transparency and oversight while building upon the legal principles of the existing copyright directive. France incorporated the EU directive into its domestic law, clarifying TDM exceptions, and further proposed legislation to regulate AI through copyright, demonstrating a strong commitment to protecting creators. In conclusion, the TDM regulations of the EU and France provide a precedent for fostering a mutually beneficial relationship between AI advancement and copyright protection. This underscores the urgent need for South Korea to align its institutional frameworks with international trends.

* Professor, Department of Culture Contents, Incheon Catholic University

Keywords

Text and Data Mining (TDM), EU Artificial Intelligence Act, EU Copyright Directive, General Purpose Artificial Intelligence (GPAI), French Copyright Law, Opt-out, AI Training Data