Abstract

A Study on Determining Copyright Infringement in AI-Generated Outputs

Park, Yusun*

The emergence of generative AI necessitates a reexamination of copyright infringement doctrines, particularly regarding access and substantial similarity. Access, traditionally assessed based on the user's awareness and intent to use a preexisting work, now faces challenges in AI contexts.

In generative AI environments, prompts serve as points where users' creative contributions intervene, directing output generation. While user prompts may facilitate access to training data, AI models can also access copyrighted works independently of user awareness. Since access encompasses both subjective elements—user recognition and intent—and objective components, applying traditional doctrines to AI contexts becomes problematic. This ultimately raises questions about liability attribution for infringing outputs generated without user awareness.

Regarding substantial similarity, AI-driven style imitation presents novel issues within comprehensive non-literal similarity. Under the idea-expression dichotomy, style belongs to the unprotected idea realm, premised on human creative processes that generate new expressions from ideas through independent creativity. However, unlike human creation, generative AI statistically analyzes and mechanically reproduces expressive patterns embedded in training data. AI internalizes recurring expressive elements—compositional methods, visual characteristics, color schemes, technical features—

* Professor, Kangwon National University, Data·IP Convergence Study, Ph.D., U.S. Attorney.

from specific artists' works, essentially performing mathematical combinations and statistical variations of expressions.

This study examines, first, the legal significance of prompts and access criteria when AI models independently access copyrighted works, and second, the need to reconceptualize style in substantial similarity assessments given AI's fundamentally different expression reproduction mechanisms.

Keywords

Generative AI, Copyright Infringement, Access, Substantial Similarity, Prompt, Style Imitation, Idea-expression dichotomy, AI memorization, Regurgitation.