Algorithmic Obligations: Proactive Liability for AI-Generated Content and Copyright Compliance

Authors : Aleksandra Czubek

Abstract : As AI systems increasingly shape content creation, existing copyright frameworks face significant challenges in determining liability for AI-generated outputs. Current legal discussions largely focus on who bears responsibility for infringing works, be it developers, users, or entities benefiting from AI outputs. This paper introduces a novel concept of algorithmic obligations, proposing that AI developers be subject to proactive duties that ensure their models prevent copyright infringement before it occurs. Building on principles of obligations law traditionally applied to human actors, the paper suggests a shift from reactive enforcement to proactive legal requirements. AI developers would be legally mandated to incorporate copyright-aware mechanisms within their systems, turning optional safeguards into enforceable standards. These obligations could vary in implementation across international, EU, UK, and U.S. legal frameworks, creating a multi-jurisdictional approach to copyright compliance. This paper explores how the EU's existing copyright framework, exemplified by the Copyright Directive (2019/790), could evolve to impose a duty of foresight on AI developers, compelling them to embed mechanisms that prevent infringing outputs. By drawing parallels to GDPR's "data protection by design," a similar principle could be applied to copyright law, where AI models are designed to minimize copyright risks. In the UK, post-Brexit text and data mining exemptions are seen as pro-innovation but pose risks to copyright protections. This paper proposes a balanced approach, introducing algorithmic obligations to complement these exemptions. AI systems benefiting from text and data mining provisions should integrate safeguards that flag potential copyright violations in real time, ensuring both innovation and protection. In the U.S., where copyright law focuses on human-centric works, this paper suggests an evolution toward algorithmic due diligence. AI developers would have a duty similar to product liability, ensuring that their systems do not produce infringing outputs, even if the outputs themselves cannot be copyrighted. This framework introduces a shift from post-infringement remedies to preventive legal structures, where developers actively mitigate risks. The paper also breaks new ground by addressing obligations surrounding the training data of large language models (LLMs). Currently, training data is often treated under exceptions such as the EU's text and data mining provisions or U.S. fair use. However, this paper proposes a proactive framework where developers are obligated to verify and document the legal status of their training data, ensuring it is licensed or otherwise cleared for use. In conclusion, this paper advocates for an obligations-centered model that shifts AI-related copyright law from reactive litigation to proactive design. By holding AI developers to a heightened standard of care, this approach aims to prevent infringement at its source, addressing both the outputs of AI systems and the training processes that underlie them.

Keywords : ip, technology, copyright, data, infringement, comparative analysis

Conference Title : ICIPRL 2025 : International Conference on Intellectual Property Rights and Law
Conference Location : Tokyo, Japan
Conference Dates : March 17-18, 2025