Managing the Risk of Using AI-Generated Content in a World of Copyright Uncertainty: Are AI Content Generators and AI Generated Expressions "Derivatives" of Copyrighted Works?

August 28, 2023

The use of artificial intelligence (AI) to create new written, visual and audio expression is rapidly expanding. AI-enabled content generation is being added to many products and services every day. Each new use raises a different question about the application of copyright law to AI-generated content.

In the United States, copyright law protects original works of authorship, including literary, dramatic, musical, and artistic works. If an original work is used to train a "large language model" ("LLM"), like ChatGPT or Google's Bard, or to create training sets or algorithms for an AI image generator, or if a copyrighted work is used as an input prompt, does the resulting data set, algorithm, or expression infringe the copyright on the original works? This article focuses on the linchpin of that question, whether the AI generator or AI-generated content is a "derivative work" of a copyrighted work and so an infringement of the original work.

The United States Copyright Act defines a "derivative work" as "a work based upon one or more preexisting works, such as a translation, musical arrangement, dramatization, fictionalization, motion picture version, sound recording, art reproduction, abridgment, condensation, or any other form in which a work may be recast, transformed, or adapted." (17 U.S.C. § 101). Copyright holders have the right to create, control and license "derivative works" based on their copyrighted work, and the "derivative works" are also copyrightable, to the extent that they add new, copyrightable expression. (17 U.S.C. § 106).

The definition of a "derivative work" is broad and encompasses a wide range of works. Some examples of derivative works include:

- A movie adaptation of a book

- A song based on a poem

- A painting based on a photograph

- A remix of a song

- A translation of a book

- A parody of a work of literature

- A sequel to a movie

- A comic book adaptation of a novel

- A video game based on a movie

Historically, derivative works were created directly by a person writing or drawing a new work, perhaps with the use of a tool or technique that created some predictable effect or change. The question of whether AI data sets, generators, and AI-generated content are derivative works of copyrighted works is at the core of the class-action lawsuits filed by writers (such as the group led by Sarah Silverman) and visual artists (such as the group led by Sarah Anderson) against companies that provide LLMs and AI art generators like OpenAI (creator of ChatGPT), Meta and *Stability AI* (US and UK entities), *Midjourney* and *DeviantArt*.

Legally, to be considered a derivative work, the new work must "copy", in some way, from an original work. However, a work can be copied from one medium to another, and be adapted for the new medium in the process, and still be a "copy." *ABS Entm't, Inc. v. CBS Corp.*, 908 F.3d 405, 416, 418 (9th Cir. 2018) (citing the treatise Nimmer on Copyright). The amount of incorporation required to create a derivative work is not always clear. Existing court decisions have generally held that the new work must incorporate a substantial amount of the original work in order to be considered a derivative work. See, e.g. *Caffey v. Cook*, 409 F. Supp. 2d 484, 496 (S.D.N.Y. 2006) (citing Nimmer on Copyright).

But the law has never applied the legal principles defining a derivative work to an unpredictable technology like LLMs where a huge number of works are input to create one output work that necessarily is based only on the input works. For example, the data sets of images compiled by Stability AI include hundreds of millions of images scraped from the web with their accompanying text captions.

The question of whether AI-generated content is a derivative work will depend largely on the technical details of its creation, as well as how the courts interpret the relevant statutory language and decisional precedent. At a very high level, LLMs and AI image generators take apart the works they are trained on, transforming them into component parts of a neural network that are then weighted using mathematical principles. These AI-powered engines can then create new expression by breaking an input prompt into weighted tokens that are run though the engine.

Certainly, it will take careful analysis and expert testimony for a court to understand how LLMs and AI image generators work, before the court can determine whether a particular application of this new technology is creating an infringing derivative work, or a copyrightable new work. Notwithstanding the term "AI," courts and the Copyright Office have taken the position that only a human author, and not an AI generator, can be an "author" of a copyrightable work. How much human input into the creation or revision of an AI generated work would be sufficient to create a copyrightable work is an open question.

Even if an LLM or other AI-generated content is found to be a derivative work, the courts will then have to consider whether fair use principles apply that excuse the infringement. For example, if the AI-generated work is considered "transformative" or it is used for educational purposes, then fair use may apply. Conversely, if the AI-generated work is directly competitive with the original works that were used to train the AI or used as an input, or the market for derivative works authorized by the copyright holder, then fair use would likely not apply. *See Andy Warhol Foundation for the Visual Arts v. Goldsmith*, 143 S. Ct. 1258(2023). The recent author and artist class-actions argue that AI generators compete with and divert the market for commissions and licenses by appropriating the style and content of copyright holders.

Any company that is planning to incorporate AI-generated content in their products or services should be aware of the potential copyright implications and risk. They should take steps to avoid liability for copyright infringement, such as:

- Obtaining permission from the copyright holder before using a set of copyrighted material to train an AI model.

- Using only public domain material to train an AI model.

- Investigating the processes used by AI vendors, especially before asserting a copyright claim to material developed with AI inputs.

- Requiring contractual terms that shift the risk of copyright infringement to the provider of the AI functionality being added to the product or service.

Owners of copyrighted works should take steps now to control access to their copyrighted works. This can include:

- Reminding licensed users that the licensed content cannot be used to train an LLM or be input into an LLM as a prompt.

- Setting up policing systems that scan new content released on the internet for evidence that it was based on the owner's copyrighted content by looking for unique information or visual clues in the new AI-generated content. This can include reference to specific facts (names of people and places) as well as watermark information that is replicated in the AI-generated content.

- Employing cloaking technology that can disrupt the ability of AI generators to use content, such as Glaze (a University of Chicago project) and Mist for images, and monitoring the arms race of cloaking technologies against AI generators.

- Filing applications to register copyrighted works, to provide greater leverage against potential infringement.

Businesses and IP owners should not throw up their hands – they can take practical steps to minimize risk and maximize recovery now. The potential value of AI generated works is hard to overstate. But so are the legal and business risks accompanying use of AI-generated content and integration of AI technology into a company's existing products and services.

* * *

related professionals

John M. Griem, Jr. / Partner
D 212-238-8659
griem@clm.com

Judith Wallace / Partner
D 212-238-8743
wallace@clm.com