



The Copyright Office on Fair Use and AI

By Brandon Butler

Re:Create Executive Director

The Copyright Office's third report on copyright and artificial intelligence has been hotly anticipated, and the political storm unfolding around the President's attempted dismissal of the Librarian of Congress and the Register of Copyrights has only heightened the drama. The [draft version](#) of the report published last week does not live up to this hype. Its influence on courts is likely to be limited, its conclusions are not as stark or unequivocal as some press coverage suggests, and when the report does take sides against fair use for AI development, its main arguments fall flat.

Limited influence on courts

Several factors limit the report's likely influence on the courts. For one, the pre-publication version is not yet official. To the (albeit limited) extent that the imprimatur of the Copyright Office adds weight to the report, this draft version does not yet have that imprimatur. It remains to be seen whether an official version will be published, or what that version, if any, will say. However, a note on the report's title page reads "A final version will be published in the near future, without any substantive changes expected in the analysis or conclusions."

More importantly, the Copyright Office has no special authority to proclaim the metes and bounds of fair use, and its opinions on the subject are no more legally binding than those of any interested party. Only courts have the power to shape fair use by applying it to new facts in specific cases. The draft report will only be influential to the extent that courts find its arguments persuasive. In that sense, the report is essentially an amicus brief, and courts are free to disagree with its conclusions or to ignore them altogether.

For example, Judge Leval of the Second Circuit was completely unpersuaded by the Copyright Office's arguments that sound recordings protected (at that time) by state copyright laws were not subject to the federal safe harbors for internet service providers in § 512 of the Copyright Act. In [Capitol Records, LLC v. VIMEO, LLC](#), Leval wrote, "[The Copyright Office's] reading of § 512(c)... is based in major part on a misreading of the statute." Leval also took the Office to task for misapplying multiple canons of statutory interpretation.

Another attempt by the Copyright Office to predict the application of fair use in the midst of ongoing litigation had no apparent impact on the courts or the development of the law. In 2011, the Copyright Office issued a ["preliminary analysis"](#) of mass digitization in parallel to the then-ongoing litigation over the [Google Books project](#) and affiliated [HathiTrust Digital Library](#). The Office's analysis focused almost entirely on licensing, spending a scant four pages

discussing fair use, casting doubt on its applicability, and emphasizing the tension between fair use and the markets for licensing. In the end, the courts found both [Google](#) and [HathiTrust](#) had engaged in fair use. The Office's report was not cited in either decision.

No need for legislation or government intervention in markets

If the Copyright Office's report has an official target audience, it's not the courts; it's Congress. And if the report is meant to answer a specific question, it's not "is AI training a fair use"—it's "Do we need to pass a new copyright law to address AI?" The report's widely overlooked response is clear:

While the use of copyrighted works to power current generative AI systems may be unprecedented in scope and scale, **the existing legal framework can address it** as in prior technological revolutions. The fair use doctrine in particular has served to flexibly accommodate such change. We believe it can do so here as well.

Thus, there is no support in the report for federal bills like the [TRAIN Act](#) or [COPIED Act](#), which create broad new rights that preempt fair use and upset the balance of the copyright system. Nor is there any support in the report for state bills like [California's AB 412](#), which assumes that all training is infringement - a conclusion the report rejects.

The report also explores several potential legislative changes to encourage licensing, but rejects them and instead "recommends allowing the licensing market to continue to develop without government intervention." Since both the TRAIN and COPIED Acts represent such interventions, they are inconsistent with the Office's conclusion.

Recognition of Transformativeness for Foundation Models

The draft report is hardly a slam dunk for copyright maximalists, who have taken the position that all copying for AI training is unlawful, period. One of the most striking conclusions in the report is that, "In the Office's view, training a generative AI foundation model on a large and diverse dataset will often be transformative." The paragraph that follows is an endorsement of the core fair use argument for every major foundation model, from ChatGPT to Gemini to Claude. As the Office explains:

The purpose of creating works of authorship is to disseminate them for human enjoyment and education. Many AI models, however, are meant to perform a

variety of functions, some of which may be distinct from the purpose of the copyrighted works they are trained on.

This is the cornerstone of the fair use case for AI training. If courts agree with the Office that the use of copyrighted works in AI training is transformative, the AI developers have all but won the fair use argument. As the Supreme Court has said, transformative uses are “[at the heart](#)” of fair use, and they are [almost always favored](#) by every other element of the fair use analysis. Unfortunately, the Office fails to follow this argument through to its logical conclusion, turning to an entirely unprecedented theory of market harm rather than face the consequences of the caselaw.

“Market dilution” turns copyright on its head

The report’s major misstep is its endorsement of “market dilution,” a theory of market harm that turns copyright law on its head. Under this novel theory, which the report itself characterizes as “uncharted territory,” a use would be considered less fair if it results in the creation of new creative works, because such new works may compete with previous works. While the report is correct that this is a “market effect” in the literal sense, it is not a market effect that any appellate court has ever recognized as relevant to fair use. Unlike market substitution (offering a work’s protected expression in a copy or a derivative as a substitute for that work, as the Supreme Court found the Andy Warhol Foundation had done in the [Warhol v. Goldsmith](#) case), “market dilution” is caused by completely new, non-infringing works that share no protected expression with any previous work.

It is hard to overstate how bizarre this theory is from the point of view of established copyright doctrine. Market dilution isn’t uncharted territory. The courts have encountered it many times, and have said that copyright encourages market dilution, also known as creativity and competition.

Copyright’s purpose, stated in Article I § 8 of the Constitution, is to “promote the Progress of Science,” i.e., the growth of knowledge and culture. As the Supreme Court wrote in [Fogerty v. Fantasy](#), “[C]opyright law ultimately serves the purpose of enriching the general public through access to creative works.” Or, as Justice Hughes wrote nearly a century ago in [Fox Film Corp. v. Doyal](#), “The sole interest of the United States and the primary object in conferring the [copyright] monopoly lie in the general benefits derived by the public from the labors of authors.” Copyright has never protected authors from competition from new works. It encourages authors to bring new works to market.

Justice Sandra Day O’Connor expressed this idea eloquently in [Feist v. Rural Telephone](#). An upstart telephone directory publisher had copied the name and telephone number data from an established publisher, sparking a copyright lawsuit and a fierce debate in the copyright community. Some argued that even though facts are not protected by copyright, permitting them to be freely copied would dilute the market for information-based publications like phone

directories and databases. Writing for the majority, Justice O'Connor explained that in fact this kind of competition is fair, and indeed it is exactly what copyright intends:

“It may seem unfair that much of the fruit of the compiler’s labor may be used by others without compensation. As Justice Brennan has correctly observed, however, this is not “some unforeseen byproduct of a statutory scheme.” It is, rather, “the essence of copyright,” and a constitutional requirement... To this end, copyright assures authors the right to their original expression, but encourages others to build freely upon the ideas and information conveyed by a work. This principle, known as the idea/expression or fact/expression dichotomy, applies to all works of authorship.... This result is neither unfair nor unfortunate. It is the means by which copyright advances the progress of science and art.” (Internal citations omitted.)

The idea that “market dilution” would count against fair use is particularly bizarre because fair use “permits courts to avoid rigid application of the copyright statute when, on occasion, it would stifle the very creativity which that law is designed to foster.” [Stewart v. Abend](#). Justice Breyer expounded on this idea in [Google v. Oracle](#), explaining that fair use:

“can focus on the legitimate need to provide incentives to produce copyrighted material while examining the extent to which yet further protection creates unrelated or illegitimate harms in other markets or to the development of other products.”

As the Copyright Office report acknowledges, AI models are exactly the kind of transformative ‘other products’ favored by fair use.

To the extent that one use of an AI model might be to facilitate the creation of new creative works, the situation is almost perfectly analogous to [Sega v. Accolade](#), in which a competing video game publisher copied Sega’s protected video game software as part of the process of developing new competing video games. The Ninth Circuit explained that Accolade’s copying:

has led to an increase in the number of independently designed video game programs.... It is precisely this growth in creative expression, based on the dissemination of other creative works and the unprotected ideas contained in those works, that the Copyright Act was intended to promote.... [A]n attempt to monopolize the market by making it impossible for others to compete runs counter to the statutory purpose of promoting creative expression and cannot constitute a strong equitable basis for resisting the invocation of the fair use doctrine.

This mistake—treating creativity and competition as if they were inconsistent with copyright—is at the root of the rest of the report’s legal errors. Notably, it underlies the report’s mistaken claim that training an AI model may be more or less transformative depending on whether the model can be used to facilitate the creation of new works of the same kind as in its training data.

Because the report treats these creations as unfair competition, it views the purpose of facilitating creativity as non-transformative. Once the “market dilution” theory is abandoned, however, we see that the ability of a model to facilitate the creation of new works is fully transformative and consistent with the creativity-promoting purpose of copyright. A guitar is not a song, a typewriter is not a book, and an AI model is not a Reddit post or a newspaper article or any other of the billions of things it’s trained on. It’s a tool for creativity (and often many other purposes), that adds something genuinely new to the world relative to its training data. That’s a quintessential fair use.

Conclusion

The Copyright Office’s draft report on fair use and AI training is a mish-mash. On one hand, it recognizes that foundational AI models, which are the ones at issue in most of the ongoing copyright litigation around AI, are textbook examples of transformative use. On the other hand, it credits a bizarre and unprecedented theory of market harm that short-circuits what would otherwise be a straightforward path from years of fair use precedent to a finding that AI training is generally fair use. The substance of its legal analysis has no binding effect on courts, AI developers, or copyright holders. Its only power over the courts comes from its persuasiveness, and history suggests that the draft report is unlikely to move courts one way or the other. Despite this equivocation on the substance, the draft report reaches a fairly clear and reasonable conclusion on the question it was tasked with answering for Congress, which is whether any new law is needed to accommodate AI in the copyright system. The answer is an unequivocal “no,” because fair use is up to the task of protecting transformative uses, and the licensing market appears to be developing naturally to accommodate non-transformative uses. While the Office’s reasoning may be murky, it reaches the correct conclusion on this ultimate question.