Generative AI Systems Tee Up Fair Use Fight

Article By:

Ariba A. Ahmad

Andrew M. Gross

The surge of generative artificial intelligence ("AI") systems entering the market faces a barrage of intellectual property challenges in the courts. In one particular flavor, copyright holders allege that the generative AI systems infringe the owners' copyrighted works. This infringement, copyright owners contend, happens both on the "input side" and the "output side" of the systems. On the input side, copyright owners contend that the process of "training" the systems' algorithms or models by ingesting large swaths of publicly-available works infringes the owners' copyrights. On the output side, copyright owners contend that the secondary works created by generative AI systems infringe the owners' copyrights. The defendants contend (or are anticipated to contend) that the accused activities are permissible fair use, teeing up whether fair use shields AI companies from copyright infringement claims. The following are some of currently-pending cases that interested parties may wish to monitor for developments on this issue:

- ***Tremblay v. OpenAI, Inc.*, No. 3:23-cv-03223 (N.D. Cal.):**[1] A collection of authors sued OpenAI in the U.S. District Court for the Northern District of California, alleging OpenAI infringed plaintiffs' copyrighted books by training OpenAI's ChatGPT and other AI products with those works. Defendants filed a motion to dismiss all causes of action except the direct copyright infringement claim. The court dismissed certain challenged claims but granted leave to amend plaintiff's complaint.
- ***Andersen v. Stability AI Ltd.*, No. 3:23-cv-00201 (N.D. Cal.):** Three visual artists sued Stability AI Ltd., Stability AI, Inc., Deviant Art, Inc., and Midjourney, Inc. on behalf of a putative class in the U.S. District Court for the Northern District of California, alleging defendants infringed plaintiffs' copyrighted images by training their respective generative AI systems with those works. Each defendant moved to dismiss, and the court dismissed all claims against all three defendants with leave to amend except the claim of direct infringement against Stability AI. Plaintiffs filed an amended complaint, adding Runway AI, Inc. to the complaint. As of the date of this article, each defendant has moved to dismiss plaintiffs' amended complaints.
- ***Authors Guild v. OpenAI, Inc.*, No. 1:23-cv-08292 (S.D.N.Y.):** Authors of registered

copyrights sued OpenAI in the U.S. District Court for the Southern District of New York, alleging OpenAI infringed the authors' copyrighted works by training ChatGPT with those works.As of the date of this article, defendant has filed its answer and asserted numerous defenses including fair use.

- ***Getty Images (US), Inc. v. Stability AI, Inc.*, No. 1:23-cv-00135 (D. Del.):** Getty Images sued Stability AI in the U.S. District Court for the District of Delaware, alleging Stability AI infringed Getty's copyrighted works by training Stability AI's accused AI with more than 12 million of Getty's copyrighted images. Defendant moved to dismiss on multiple grounds and moved to transfer. As of the date of this article, the court has not ruled on those motions.
- ***The New York Times Co. v. Microsoft Corp.*, No. 1:23-cv-11195 (S.D.N.Y.):** The New York Times sued OpenAI and Microsoft (and related corporate entities) in the U.S. District Court for the Southern District of New York, alleging Microsoft and OpenAI infringed the Times' copyrighted newspaper articles by training the accused chatbots with the Times' articles. Defendants moved to dismiss, and Microsoft moved to intervene and dismiss, stay, or transfer. As of the date of this article, the court has not ruled on those motions.

These cases are in the early stages, but it is likely that each of the defendants will assert a fair use defense, which will make these cases among the first cases that tee up the applicability of the fair use defense in the context of AI. While some believe that training AI models on copyrighted works is clearly fair use,[2] the unique nature of AI and the inherent lack of transparency into how the systems work could make that issue less clear and potentially present greater challenges to accused infringers.

One particular challenge could stem from the fact that accused infringers do not know exactly how their AI systems work. For example, deep learning models—one of the most prevalent forms of modern AI—learn much the same way that humans learn.[3] In such models, companies train their algorithms with correct examples of something they want the algorithm to recognize and eventually the algorithms develop a "neural network" capable of categorizing things to which the algorithms have not been exposed.[4] Indeed, website users have likely unwittingly participated in such training by completing a "reCAPTCHA" to access a website (e.g. click on all of the street signs in a given image). Because of the way these deep learning models work, AI companies do not know precisely how their systems make decisions or come to conclusions—a phenomenon known as the "black box" problem.[5]

Fair use is an affirmative defense to copyright infringement, and the accused infringer bears the burden of proving the defense. In deciding whether a use is fair use, courts consider the following factors: (1) the purpose and character of the use; (2) the nature of the copyrighted work; (3) the amount and substantiality of the portion used; and (4) the effect of the use upon the potential market for the copyrighted work.[6] The black box problem may well be something accused infringers need to navigate in articulating how these factors weigh in favor of fair use, potentially presenting challenges to demonstrating fair use. That issue could affect all four factors but could be particularly challenging for factor (3). For example, if AI defendants do not know how their systems learn, it could be challenging to explain to a court the "amount and substantiality of the portion used." Interested parties will want to monitor the above-listed cases to see whether the black box problem becomes an issue, as well as to gain insight into how courts deal with fair use in the AI context.

1. Two other cases involving authors were filed in the U.S. District Court for the Northern District of California against OpenAI and then consolidated into *Tremblay* in October 2023. *See Chabon v. OpenAI, Inc.*, No. 3:23-cv-04625 (N.D. Cal.) and *Silverman v. OpenAI, Inc.*, No. 3:23-cv-03416 (N.D. Cal.). ??
2. *See* Katherine Klosek, *Training Generative AI Models on Copyrighted Works is Fair Use*, Association of Research Libraries, last updated Jan. 23, 2024 (accessed 2/28/2024), https://www.arl.org/blog/training-generative-ai-models-on-copyrighted-works-is-fair-use/. ??
3. Lou Blouin, *AI's Mysterious 'Black Box' Problem, Explained*, UM-Dearborn News, Mar. 6, 2023 (accessed 2/28/2024), https://umdearborn.edu/news/ais-mysterious-black-box-problem-explained. ??
4. *Id.* ??
5. *Id.* ??
6. *See Andy Warhol Foundation for the Visual Arts, Inc. v. Goldsmith*, 143 S. Ct. 1258 (2023).

Source URL:https://natlawreview.com/article/generative-ai-systems-tee-fair-use-fight