GRAND PARK LAW GROUP
A PROFESSIONAL CORPORATION

WESTSIDE TOWERS
11835 W. Olympic Boulevard #870E
Los Angeles, California 90064

O:+1 888.629.4754
F: +1 310.900.0805

INSIGHT · ENTERTAINMENT & TALENT

Q1 2023 Legal Review: Generative AI on Trial — Training Data, Fair Use, and Creator Rights

Grand Park Law Group · March 31, 2023 · 10 min read



Generative AI training on copyrighted materials prompted nationwide IP litigation, pitting the Fair Use doctrine against creators' exclusive reproduction and derivative rights.

In the opening quarter of 2023, the meteoric rise of large language models and diffusion image generators sparked the first major wave of intellectual property litigation against artificial intelligence companies. Class-action lawsuits filed in the Northern District of California (including *Andersen v. Stability AI* and *Tremblay v. OpenAI*) initiated a legal clash over whether scraping copyrighted works for model training constitutes copyright infringement or transformative fair use.

I. The Theory of Infringement: Scraping & Reproduction

Plaintiffs—comprising visual artists, authors, and photographers—alleged that AI developers violated 17 U.S.C. § 106(1) by scraping billions of copyrighted works from the open web to train foundational models without authorization or compensation. The complaints asserted three primary theories of liability:

1. **Direct Input Infringement:** The unauthorized intermediate copying of expressive works into training datasets (such as LAION-5B).
2. **Derivative Work Creation:** The claim that the trained neural network models and their downstream outputs represent unauthorized derivative works under 17 U.S.C. § 106(2).
3. **Digital Millennium Copyright Act (DMCA) Violations:** The intentional removal or alteration of Copyright Management Information (CMI) under 17 U.S.C. § 1202(b).

II. The Fair Use Battleground Under 17 U.S.C. § 107

AI developers mounted a vigorous defense anchored in the doctrine of transformative fair use, citing *Authors Guild v. Google, Inc.* (2d Cir. 2015) 804 F.3d 801 (the Google Books case) and *Kelly v. Arriba Soft Corp.* (9th Cir. 2003) 336 F.3d 811. Developers argued that training neural networks extracts uncopyrightable statistical patterns, linguistic rules, and conceptual associations rather than copying artistic expression.

However, plaintiffs drew sharp distinctions between search engine indexing—which directs users back to the original source—and generative AI tools that generate synthetic substitutes capable of directly cannibalizing the commercial market for human creative output under the fourth statutory fair use factor (17 U.S.C. § 107(4)).

III. Strategic Guidance for Content Creators and AI Developers

The emerging AI litigation landscape requires proactive IP portfolio management:

- **Copyright Registration Timing:** Creators must obtain timely copyright registrations from the U.S. Copyright Office under 17 U.S.C. § 412 before public dissemination to preserve statutory damages (up to $150,000 per willful infringement) and attorney fee shifting.
- **Dataset Provenance & Licensing:** Technology companies training proprietary models must audit training pipelines, document data lineage, and transition toward opt-in licensed repositories to mitigate catastrophic class-action exposure.
- **Contractual AI Warranties:** Enterprise software contracts now routinely require representations, warranties, and IP indemnification covenants regarding whether software features incorporate generative AI trained on unvetted third-party materials.

Practice	Entertainment & Talent
Topics	Artificial Intelligence · Copyright · Tech Litigation
Publication	Legal News & Opinion · March 31, 2023

Available online at grandparklawgroup.com/insights/articles/quarterly-2023-q1-generative-ai-copyright-battle. This publication is provided for general information and is not legal advice.
© 2026 Grand Park Law Group, A.P.C. Attorney Advertising.