AI Exploitation of Intellectual Property:
Corporate Considerations for Transformative Fair Use

Amanda Cline
Senior Counsel, Quadrant Law Group, LLP

Financial institutions such as banks, hedge funds, and broker-dealers are increasingly embracing artificial intelligence (AI) to drive internal and external efficiencies ranging from IT automation and code creation to portfolio development, risk analysis, and market research. While corporations race to implement the latest and greatest technology to keep pace with competitors and customer demand, the law is only beginning to evaluate how machine-generated content may infringe on human works.

When corporations develop proprietary large language models (LLMs) or use a third-party LLM provided by the market leaders such as OpenAI, Anthropic, and Amazon, it is important to consider how data used to train the LLM may impact intellectual property rights of others to avoid becoming the next headline news infringement case.

Courts are increasingly emphasizing fair-use principles in determining if data used to train an LLM violates third-party intellectual property rights. Given the evolution of generative AI's (GenAI) ability to create novel outputs, whether the GenAI's use of copyrighted works is "transformative" is the key consideration.

Thompson Reuters v. ROSS Intelligence[1]

"There is nothing that Thompson Reuters created that Ross could not have created for itself or hired LegalEase to create for itself without infringing Thompson Reuter's copyrights." - Judge Bibas

Thompson Reuters (TR) sued Ross for direct copyright infringement alleging Ross' use of its Westlaw headnotes to train AI in Ross' legal research tool violated TR's intellectual property rights.

The court weighed the following four fair use principles:

- ***Purpose and Character of Ross's Use favors TR***: Ross's commercial use of headnotes was not transformative because it was used to develop a competing legal research tool.
- ***Nature of the Original Work favors Ross***: headnotes have some level of creativity but are not creative works.
- ***Amount and Substantiality of the Work Used favors Ross***: Ross did not make the headnotes publicly accessible.
- ***Market Impact favors TR***: Ross's use of the headnotes was meant to compete with Westlaw by developing a market substitute where legal research is the primary market and data to train legal platform AI is a secondary market. The effect on a potential market for training AI data is enough to show violation of fair use.

Corporate Takeaway: evaluate the sourcing of data used to train an LLM because direct or indirect sourcing to develop a competing product can lead to infringement where fair-use factors are weak, which may include re-purposing analyst reports or using research data in violation of license agreements to train an LLM.

Anthropic v. Bartz[2]

---

[1] *Thompson Reuters Enterprise Central GmbH and West Publishing Corp. v Ross Intelligence Inc.*; Case 1:20-cv-613-SB; February 11, 2025

[2] *Bartz v. Anthropic* PBC, No. 3:24-cv-05417-WHA (N.D. Cal. June 23, 2025).

Anthropic, the creator of Claude models, was sued for copyright infringement because it used both pirated books and purchased books converted to a digital format to train its models.

Again, the court evaluated transformative fair use for the purchased books:

- ***Purpose and Character favors Anthropic*** – Anthropic's use of the books to "iteratively map statistical relationships between text fragments between every text-fragment and every sequence of text-fragments so that a completed LLM could receive new text inputs and return new text outputs as if it were a human reading prompts and writing responses" was spectacularly transformative
    - ***Note***: plaintiffs did not allege AI outputs were infringing
- ***Nature of the Original Work slightly favors Bartz***: the books have some expressive elements
- ***Amount and Substantiality of the Work Used favors Anthropic***: copying the purchased books to convert them to digital print to use for LLM training was necessary for the transformative use, and copying the books in their entirety was reasonable because there was no surplus copying
- ***Market Impact favors Anthropic***: demand for the authors printed books is not displaced by the copies used to train Anthropic's LLMs

While Anthropic's use of purchased printed books converted to digital formats used to train its LLMs did not violate fair-use principles, pirating books used for the same purpose was not justifiable. Further, it remains to be determined by the court whether storing copies of the purchased books that were ultimately not used for training but still made accessible to hundreds of engineers was fair use.

Corporate Takeaway: remain vigilant in knowing where and how data used to train an LLM was sourced, implement auditing standards to track ongoing data use, and seek contractual protections such as lawful sourcing and IP infringement protections from third-party providers.

Financial institutions should also be cognizant that financial data may contain material non-public information, trade secrets, and/or client data which need to be assessed for SEC and other regulatory compliance, contractual confidentiality obligations, and internal compliance policies when being used to train an LLM or ongoing as inputs that could produce infringing outputs.

Kadrey v. Meta Platforms[3]

Meta, the creator of llama models, was sued for copyright infringement and concealment of copyright management information in violation of the Digital Millennium Copyright Act (DCMA).

Meta's mass downloading of various authors' books from shadow libraries, which are online platforms that provide access to copyrighted materials where permission from the copyright holder may not have been obtained, and using those books to train its LLMs was not copyright infringement because such copying may have been transformative and there was a lack of evidence of harm to the market.[4] Market harm remains the most critical factor in evaluating fair use in recent AI cases.

Corporate Takeaway: fair-use analyses are fact-specific, and corporations should not assume purchased data provides free reign for use in training LLMs nor that outputs generated are clear of infringement. Reproducing large data sets of copyrighted material via AI output could lead to market harm. Incorporating guardrails to limit the reproduction of copyrighted material may help mitigate that risk. Derivative works may also be infringing if a court determines they lead to market dilution.

---

[3] *Richard Kadrey et. al v. Meta Platforms, Inc.*; Case 3:23-cv-03417-VC; June 25, 2025.

[4] Jackson Walker News; Arthur Gollwitzer, July 11, 2025.

Hot News

"Hot News" doctrine protects time-sensitive factual information from misappropriation.[5] It originated in the 1918 case *International News Service v, Associated Press*.[6] The 2011 case *Barclays Capital v. Theflyonthewall.com* clarified its relevance to the financial industry, where the court ruled that a website's re-distribution of securities-trading recommendations to its subscribers before the institutions could release it to their clients was *not* hot news misappropriation.[7]

Technology such as AI—which enhances the speed and efficiency of large data aggregation—could enable customers to redistribute proprietary financial data quicker than the publishers. Financial institutions subscribing to news services and data feeds should evaluate whether their tools are extracting and re-distributing licensed content too close in time such that a "hot news" infringement claim could be made.

Conclusion

Just as many organizations have "Know Your Client" requirements, so too should organizations adopt "Know Your AI Training Data" protocols. These should include documenting the source of the data, copyright and use restrictions of that data, and whether the desired use case could cause market harm or displacement. Enterprises that purchase high volumes of market data and other information may consider negotiating contracts with express use rights of purchased or licensed data in AI models and avoid using data to train models that compete with the publisher's products. AI will continue to challenge established legal frameworks and will shape new principles in various legal domains.

For additional information, please contact:

Amanda Cline | Senior Counsel
Quadrant Law Group, LLP
T +1-949-396-1026
acline@quadrantlaw.com

This Bulletin is provided for informational and educational purposes only. It does not constitute legal advice or establish an attorney-client relationship. This Bulletin may be considered attorney advertising in some states.

Quadrant Law Group, LLP © 2025. All rights reserved.
203 N. La Salle St. Suite 2100, Chicago, IL 60601

---

[5] The Hot News Doctrine Analyzed; Steve Vondran; September 12, 2023.

[6] *International News Service v. Associated Press*, 248 U.S. 215 (1918).

[7] *Barclays Capital Inc. v. Theflyonthewall.com, Inc*, 650 F.3d 876 (2d Cir. 2011).