AIPLA QUARTERLY JOURNAL

Volume 54, Number 2 Page 155 Spring 2026

IS RETRIEVAL-AUGMENTED GENERATION FAIR USE?

*Jake Glendenning**

I. INTRODUCTION

The United States Copyright Office recently released a controversial[1] pre-publication report addressing the legality of training generative AI systems on copyrighted works.[2] The report primarily addressed whether training AI models is lawful, and training data licensing.[3] It also included a cursory mention of a nascent "important feature of many AI products" called Retrieval-Augmented Generation ("RAG").[4] The report did not analyze the new technology in detail, however.[5] While model training and licensing are essential and relevant questions for generative AI, they have been explored in great detail.[6] RAG, on the other

---

1 Pre-publication reports are unusual from the Copyright Office, and the report was published one day after the Librarian of Congress was dismissed, and one day before the Register of Copyrights was dismissed. *See* Richard M. Assmus et al., *United States Copyright Office Weighs in on Fair Use Defense for Generative AI Training*, MAYER BROWN (May 15, 2025), https://www.mayerbrown.com/en/insights/publications/2025/05/united-states-copyright-office-weighs-in-on-fair-use-defense-for-generative-ai-training [https://perma.cc/Z7BQ-A262]; *U.S. Copyright Office Issues Pre-Publication Version of 3rd Report on AI Training and Fair Use. AI Training is Transformative, but Degree Depends on How AI Functions. Supports New Market Dilution Theory of Harm Under Factor 4*, CHATGPT IS EATING THE WORLD (May 10, 2025), https://chatgptiseatingtheworld.com/2025/05/10/us-copyright-issues-pre-publication-version-of-3rd-report-on-ai-training-and-fair-use-ai-training-is-transformative-but-degree-depends-on-how-ai-functions-supports-new-market-dilution-theory-of-ha/ [https://perma.cc/U6PV-K3XM] [hereinafter *U.S. Copyright Office Issues Pre-Publication Version*].

2 U.S. COPYRIGHT OFFICE, COPYRIGHT AND ARTIFICIAL INTELLIGENCE PART 3: GENERATIVE AI TRAINING, 1 (May 2025).

3 *See generally id.*

4 *Id.* at 31.

5 *Id.* at 30–31.

6 *See generally* Mark A. Lemley & Bryan Casey, *Fair Learning*, 99 TEX. L. REV. 743, 748–49 (2021); Benjamin L. W. Sobel, *Artificial Intelligence's Fair Use Crisis*, 41 COLUM. J.L. & ARTS 45 (2017); Matthew Sag, *The New Legal Landscape for Text Mining and Machine Learning*, 66 J. COPYRIGHT SOC'Y U.S.A. 291, 292–93 (2019); Jake Glendenning, *Machine Learning: An Alternative Approach to the Copyright Infringement Question*, 51 AIPLA Q.J. 1, 4–5 (2023); Bartz v. Anthropic PBC, 787 F. Supp. 3d 1007, 1014 (N.D. Cal. 2025); Kadrey v. Meta Platforms, Inc., 788 F. Supp. 3d 1026, 1034 (N.D. Cal. 2025).

hand, is significantly rising in popularity and is drastically underexplored in comparison.[7] There are currently over ninety pending lawsuits against AI companies in the United States, yet only five focus on RAG.[8] This number will likely increase as RAG becomes increasingly common in AI services offered by well-known brands. It warrants serious consideration.

RAG is an increasingly ubiquitous innovation that substantially mitigates hallucination,[9] a common problem in which generative AIs[10] include plausible

---

7 *See* Emily Winks, *What is Retrieval-Augmented Generation (RAG)?*, ATLAN (Apr. 3, 2026), https://atlan.com/know/what-is-retrieval-augmented-generation/ [https://perma.cc/UAQ7-8HLK].

8 *See* Compl. ¶ 9, Encyclopaedia Britannica, Inc. v. Perplexity AI, Inc., No. 1:25-cv-07546 (S.D.N.Y. Sept. 10, 2025); 2d Am. Compl. ¶ 6, Dow Jones & Co. v. Perplexity AI, Inc., No. 24-cv-7984 (S.D.N.Y. Oct. 21, 2024); Compl. ¶¶ 4–5, Chicago Tribune Co., LLC v. Perplexity AI, Inc., No. 1:25-cv-10094 (N.D. Ill. Dec 4, 2025); Compl. ¶ 4, The New York Times Co. v. Perplexity AI, Inc., No. 1:25-cv-10106 (S.D.N.Y. Dec. 5, 2025); Compl. ¶ 4, Advance Local Media LLC v. Cohere Inc., No. 1:25-cv-01305 (S.D.N.Y. Feb. 13, 2025). The New York Times' suit against OpenAI and Microsoft also appears to challenge RAG, though the gravamen of its complaint challenges model training. *See also,* Compl. ¶ 108, The New York Times Co. v. Microsoft Corp., No. 23-cv-11195 (S.D.N.Y. Dec. 27, 2023). This author surveyed the operative complaint in each non-dismissed lawsuit on Professor Edward Lee's database of AI lawsuits to determine that only the aforementioned cases focus on RAG. *See also AI Litigation Tracker: Copyright Suits v. AI Companies,* CHATGPT IS EATING THE WORLD (last updated Mar. 27, 2026), https://chatgptiseatingtheworld.com/aicopyrightcasetracker/ [https://perma.cc/D3JT-AA2C].

9 The use of the word "hallucination" is debated. Some commentators argue that the term unnecessarily anthropomorphizes AI and argue instead for the term "confabulation." *See* Chetan Conkee, *Stop Saying AI Hallucinates—It's Confabulating, and We're Way Worse at it,* SUBSTACK (July 4, 2025), https://conikeec.substack.com/p/stop-saying-ai-hallucinates-its-confabulating [https://perma.cc/R5Y4-QZ3F]. This Article uses "hallucination" because it is currently the more common industry-accepted term. *Id.*

10 Specifically, RAG is commonly used in large language models like AI assisted web-search. *See infra* notes 12-13. Further, RAG also solves additional problems with traditional generative AIs, in addition to hallucination. *See infra* Part II.

sounding yet false information in their outputs.[11] In basic terms, RAG works by allowing a large language model ("LLM") to retrieve specific knowledge on an issue instead of "making up" information. This approach can reduce hallucinated responses by up to eighty-five percent.[12] Google Search now uses RAG,[13] as does Perplexity,[14] a web search startup credited with threatening Google's search dominance for the first time since the giant's founding.[15] In addition to these simple search services, newer "deep research" and "agentic search" services offered by OpenAI, Google, and xAI incorporate advanced versions of RAG as well.[16]

---

[11] *What are AI Hallucinations?*, GOOGLE CLOUD, https://cloud.google.com/discover/what-are-ai-hallucinations?hl=en [https://perma.cc/Z2LX-LDVZ].

[12] Kurt Shuster et al., *Retrieval-Augmented Generation Reduces Hallucination in Conversation*, at 2 (Apr. 15, 2021), https://arxiv.org/pdf/2104.07567 [https://perma.cc/CMF4-YUR3].

[13] Krishnan Kumar et al., *Building a Google-quality Search System with Vertex AI*, GOOGLE FOR DEVELOPERS, https://codelabs.developers.google.com/build-google-quality-rag#0 [https://perma.cc/3MV8-4VM7] . ("Google Search is a powerful tool that uses a massive index of web pages and other content to provide relevant results to user queries. This is made possible by a technique called Retrieval Augmented Generation (RAG), which is a key technique in modern AI.").

[14] *RAG 2.0: Improved Retrieval Techniques*, PERPLEXITY (Jun. 22, 2024), https://www.perplexity.ai/page/rag-2-0-improved-retrieval-tec-OA0X2PggTBCgZmL0F_FyLA [https://perma.cc/A4JA-WSRX].

[15] Marco Quiroz-Gutierrez, *AI Startups Believe Google's Chrome is Vulnerable to a New Wave of Intelligent Browsers*, FORTUNE (Jul. 13, 2025), https://fortune.com/2025/07/13/ai-web-browsers-perplexity-comet-openai-google-chrome-search/ [https://perma.cc/NS9P-YSKA].

[16] Although some companies do not refer to these services as RAG, their iterative retrieval of indexed information makes them indistinct from RAG for the purposes of this Article. *See* Timothy B. Lee, *These Experts Were Stunned by OpenAI Deep Research*, UNDERSTANDING AI (Feb. 24, 2025), https://www.understandingai.org/p/these-experts-were-stunned-by-openai [https://perma.cc/2YLH-ULDE] (comparing OpenAI and Google's Deep Research services with traditional RAG techniques); *see also Grok-3 Beta – The Age of Reasoning Agents*, XAI BLOG (Feb, 19, 2025), https://x.ai/news/grok-3 [https://perma.cc/8RBS-VQ6V] ("On the LOFT (128k) benchmark, which targets long-context RAG use cases, Grok 3 achieved state-of-the-art accuracy (averaged across 12 diverse tasks), showcasing its powerful information retrieval capabilities."); *Gemini Deep Research*, GEMINI,

The new technology is not without controversy, however. News organizations are already taking legal action against web-search services using RAG. Concerns about RAG form the basis for lawsuits by the owners of the Wall Street Journal[17] and New York Times.[18] Organizations that have not yet taken direct legal action have expressed concerns via the United States Copyright Office's comment process.[19] These news publishers are concerned that RAG services provide summaries of their articles without directing users to webpages displaying the content.[20] They argue that this lack of web traffic threatens to deprive news providers of ad revenue or subscription payments, thus creating an existential threat to their businesses.[21]

Although RAG may seem similar to the generative AI technology at issue in some high-profile, fast-moving AI lawsuits,[22] it is distinct in several significant

https://gemini.google/overview/deep-research/ [https://perma.cc/B7MD-YP8A] ("Gemini Deep Research is designed to tackle your complex research tasks by breaking them down, exploring the web to find answers, and synthesizing findings into comprehensive results."); *Introducing Deep Research*, OPENAI (Feb. 2, 2025), https://openai.com/index/introducing-deep-research/ [https://perma.cc/556X-QE9F] ("[I]t leverages reasoning to search, interpret, and analyze massive amounts of text, images, and PDFs on the internet . . . .").

17 2d Am. Compl. ¶ 15, *Dow Jones & Co.*, No. 24-cv-7984.

18 Compl. ¶ 4, *New York Times*, 1:25-cv-10106; *see also* Alexandra Bruell, *New York Times to Bezos-Backed AI Startup: Stop Using Our Stuff*, WALL ST. J. (Oct. 15, 2024), https://www.wsj.com/business/media/new-york-times-to-bezos-backed-ai-startup-stop-using-our-stuff-20faf2eb?st=J4nJaK&reflink=desktopwebshare_permalink [https://perma.cc/3JUL-A4B7].

19 *See* Letter from News/Media All. to U.S. Copyright Off., *Summary of Ex Parte Meeting on Apr. 29, 2024 Regarding the Office's AI Study* (May 3, 2024), https://www.copyright.gov/policy/artificial-intelligence/ex-parte-communications/letters/NewsMedia-Alliance-May-3-2024.pdf [https://perma.cc/TN6D-56XV].

20 *See id.*

21 The president and CEO of the News/Media Alliance referred to RAG services as "an existential crisis for journalism." Ray Schultz, *The Ultimate Threat: Content Scraping by AI Companies*, MEDIAPOST (Jan. 19, 2025), https://www.mediapost.com/publications/article/402687/the-ultimate-threat-content-scraping-by-ai-compan.html [https://perma.cc/9XFH-2DBX].

22 *See, e.g.*, *Bartz*, 787 F. Supp. 3d at 1014; *Kadrey*, 788 F. Supp. 3d at 1034.

ways. These differences will change how courts determine whether RAG is fair use. For example, RAG used for internet research may be less transformative than standard generative text AIs like ChatGPT when being used as a substitute for news articles.[23] The narrow scope of internet research may also provide a more concrete market harm than the harms that courts have analyzed in high-profile AI training lawsuits, like those against Meta and Anthropic.[24] While these arguments, among others, may support the conclusion that RAG is less likely to be a fair use than conventional LLMs, the opposite may also be true. RAG is arguably similar to conventional search engines like Google or Bing, which index websites and provide previews and links to web content.[25] It may therefore be transformative for the same reasons as Google Search and Google Books, and thus a fair use.[26]

This Article asks whether RAG and related deep research and agentic research services are fair use. It first describes how RAG services work. It then summarizes relevant United States copyright doctrine and concludes that RAG services likely infringe copyright on their face, thus prompting the question of whether they are fair use. Next, it analyzes how notable relevant cases have applied the four statutory fair use factors. Finally, it concludes that courts will likely hold that RAG is not fair use.

II. HOW RETRIEVAL-AUGMENTED GENERATION ("RAG") WORKS

RAG is a technique in which a generative AI[27] fetches specific information to inform its response to a query. In other words, it allows AI chatbots like

---

[23] *See* U.S. COPYRIGHT OFFICE, *supra* note 2, at 47.

[24] *See Bartz*, 787 F. Supp. 3d at 1031; *Kadrey*, 788 F. Supp. 3d at 1056.

[25] *See In-Depth Guide to How Google Search Works*, GOOGLE SEARCH CENTRAL https://developers.google.com/search/docs/fundamentals/how-search-works; https://support.microsoft.com/en-us/topic/how-bing-delivers-search-results-d18fc815-ac37-4723-bc67-9229ce3eb6a3 [https://perma.cc/QF8U-9VX8].

[26] *See* Kelly v. Arriba Soft Corp., 336 F.3d 811, 818 (9th Cir. 2003); Authors Guild v. Google Inc., 804 F.3d 202, 207 (2d Cir. 2015); Field v. Google Inc., 412 F. Supp. 2d 1106, 1119 (D. Nev. 2006).

[27] Generative AI services like ChatGPT are built on computational models called large language models. *How ChatGPT and Our Foundation Models Are Developed*, OPENAI HELP, https://help.openai.com/en/articles/7842364-how-chatgpt-and-our-foundation-models-are-developed [https://perma.cc/KM36-7B3M].

Large language models are artificial intelligence systems designed to simulate human language by encoding "how to structure sentences, connect words,

OpenAI's ChatGPT[28] or Google's Gemini[29] to reference an external knowledge base,[30] like the internet, for example, and incorporate information from that knowledge base into its answer.[31]

Fundamentally, RAG works by supplementing the information already internalized within a foundation model.[32] Foundation models[33] underlie well-known generative AI chatbots like ChatGPT.[34] Such models are trained on massive quantities of data[35] at a discrete point in time.[36] RAG services retrieve information

---

and follow the rules of grammar. This ability comes from their exposure to vast amounts of text, allowing them to pick up on patterns and structures." *See* Kristian Hammond, *The Hallucination Problem: A Feature, Not a Bug*, NW. UNIV. CTR. FOR APPLIED & SCI. MODELING & INQUIRY (Aug. 26, 2024), https://casmi.northwestern.edu/news/articles/2024/the-hallucination-problem-a-feature-not-a-bug.html [https://perma.cc/6HXM-Q4UP].

28 *Introducing Deep Research, supra n*ote 16.

29 Kumar et al., *supra* note 13.

30 For clarity, the knowledge base is external to the information contained within the model itself, though the knowledge base may be internal to the company offering a RAG service.

31 *What Is RAG (Retrieval-Augmented Generation)*?, AMAZON WEB SERVS., https://aws.amazon.com/what-is/retrieval-augmented-generation/ [https://perma.cc/B9FN-Z6BX].

32 *See id.*

33 The term "Foundation Models" was coined by the Stanford Institute for Human-Centered Artificial Intelligence to refer to models trained on a large corpus of data that can be used generally for different tasks with just some fine-tuning. Researchers "chose the term 'foundation' to connote the significance of architectural stability, safety, and security: poorly-constructed foundations are a recipe for disaster and well-executed foundations are a reliable bedrock for future applications." *See* Rishi Bommasani et al., *On the Opportunities and Risks of Foundation Models*, STANFORD UNIV. CTR. FOR RES. ON FOUNDATION MODELS, at 7 (2021), https://crfm.stanford.edu/assets/report.pdf [https://perma.cc/86UK-FCKU].

34 *How ChatGPT and Our Foundation Models Are Developed*, *supra* note 27.

35 One leading AI lab, Anthropic, trained a prior version of its flagship "Claude" model on roughly 7 million book files. *See Bartz*, 787 F. Supp. 3d at 1015.

36 *What Are the Limitations of Large Language Models (LLMs)?*, PROMPTDRIVE (Apr. 23, 2024), https://promptdrive.ai/llm-limitations/ [https://perma.cc/X8GB-6UTX].

from a new data source, which is usually regularly updated with up-to-date information, to supplement the knowledge base internalized within a foundation model based on a user's input.[37] The user's query and the retrieved information are then provided to the model, which uses the retrieved information to supplement its trained knowledge to create a response to the user's query.[38] The retrieved information can come from multiple data sources, including document repositories or, as Dow Jones & Co alleges, from the internet, including paywalled webpages.[39]

Specifically, RAG services generally use a technique called "embedding language models" to convert data, including documents, into numerical representations that can be stored in a specialized database useful for large language models, called a vector database.[40] This vector database is then continuously updated with relevant data that can be retrieved by the RAG service in response to user queries.[41] Then, when a user asks the RAG service a question, the query is converted to a vector representation and compared to the vector database to find relevant data in the database.[42] One exemplar use case for the technology is a company that sets up a RAG chatbot for employees to ask H.R. questions.[43] In such a case, an employee could ask the chatbot a question like "how

---

For more information on how AI models are trained and the implications for copyright law, *see generally* Lemley & Casey, *supra* note 6, at 748–49; Sobel, *supra* note 6, at 45; Sag, *supra* note 6, at 292–93; Glendenning, *supra* note 6, at 4–5; Benjamin L. W. Sobel, *Artificial Intelligence's Fair Use Crisis*, 41 COLUM. J.L. & ARTS 45 (2017); Matthew Sag, *The New Legal Landscape for Text Mining and Machine Learning*, 66 J. COPYRIGHT SOC'Y U.S.A. 291 (2019).

37 *What Is RAG (Retrieval-Augmented Generation)?*, *supra* note 31.

38 *Id.*

39 *See* 2d Am. Compl. ¶ 110, *Dow Jones & Co.*, No. 24-cv-7984.

One plausible source of the paywalled content may be the Common Crawl database, upon which many foundation LLMs are trained, and which allegedly contains many news articles that are normally only accessible to humans by paying for a subscription. *See* Alex Reisner, *The Company Quietly Funneling Paywalled Articles to AI Developers*, ATLANTIC (Nov. 4, 2025), https://www.theatlantic.com/technology/2025/11/common-crawl-ai-training-data/684567/ [https://perma.cc/BP8N-4E84].

40 *What Is RAG (Retrieval-Augmented Generation)?*, *supra* note 31.

41 *Id.*

42 *Id.*

43 *Id.*

much annual leave do I have?" The RAG chatbot would convert that question into a vector representation and compare it to a vector database of company leave policy documents and that particular employee's past leave record.[44] After comparing the vectorized query to the database, the RAG service augments the user's initial query by adding its own relevant data to create a new synthesized query.[45] This synthesized query is then sent to the foundation model, which returns an answer in the form of a text response.[46]

RAG addresses three related problems that occur with the large language models ("LLMs") that power generative AIs. First, LLMs are trained at a specific point in time, and therefore have no knowledge of factual developments that occur after that point.[47] In other words, if a user asks an AI trained in 2021 if there is a constitutional right to abortion, it will likely say yes, citing *Roe v. Wade*.[48] Of course, that answer is now incorrect because *Roe* was overruled in 2022 by *Dobbs v. Jackson Women's Health Organization*.[49] RAG allows AIs to supplement their knowledge with more up-to-date information before providing an answer to the user.[50]

Second, LLMs have a tendency to make up information, or "hallucinate."[51] In many contexts, these hallucinations are a feature of LLMs and not a bug. They are a necessary byproduct of LLMs' abilities to simulate creativity.[52] They are very useful, for example, for creating first drafts and plot outlines for fiction writing.[53] In other applications, however, this feature can be dangerous. Many attorneys

---

44 *Id.*

45 *Id.*

46 *What Is RAG (Retrieval-Augmented Generation)?*, *supra* note 31.

47 *Tutorial: ChatGPT Over Your Data*, LANGCHAIN BLOG (Feb. 5, 2023), https://blog.langchain.com/tutorial-chatgpt-over-your-data/ [https://perma.cc/P8KA-5FAW].

48 Roe v. Wade, 410 U.S. 113, 154 (1973).

49 Dobbs v. Jackson Women's Health Org., 597 U.S. 215, 302 (2022).

50 *What Is RAG (Retrieval-Augmented Generation)?*, *supra* note 31.

51 *See* Shuster et al., *supra* note 12, at 15.

52 *See* Hammond, *supra* note 27.

53 Uri Baruchin, *Turn AI Chatbots into Better Writing Partners: A Step-by-Step Guide for the Less Famous*, MEDIUM (Sept. 19, 2024), https://uriba.medium.com/turn-ai-chatbots-into-better-writing-partners-a-step-by-step-guide-for-the-less-famous-c0e29722b00d [https://perma.cc/V877-XB2F].

have already been caught submitting false citations generated by AI.[54] Defamation lawsuits[55] have been filed because of allegedly hallucinated information. Further, one study evaluating AI-prepared medical research proposals found that ChatGPT hallucinated a significant amount of references.[56] Of 178 references generated by ChatGPT, sixty-nine could not be verified with Digital Object Identifiers ("DOIs"), and twenty-eight references did not appear in Google searches or have an existing DOI at all.[57] These examples highlight a particular danger of relying on AI for fact-intensive tasks. AI often produces convincing, plausible seeming, yet completely false answers.

Third, RAG services allow traceability by providing sources or hyperlinks to the webpages on which their answers are based.[58] These links allow users to navigate to the webpage that AI search tools rely on when providing their textual response to a user's question.[59] The ability to navigate to source webpages allows users to double-check the AI's interpretation of the source, thereby further decreasing the propensity for factual error. In sum, without RAG, LLMs are like "over-enthusiastic new employee[s] who refuse[] to stay informed with current events but will always answer every question with absolute confidence."[60] RAG is designed to mitigate the pitfalls of such an approach.

---

54 *See* Damien Charlotin, *AI Hallucination Cases*, DAMIENCHARLOTIN.COM, https://www.damiencharlotin.com/hallucinations/ [https://perma.cc/XY2E-DPKB].

55 *See* Yaron Dori et al., *Georgia Court Dismisses Defamation Suit Against AI Developer OpenAI*, COVINGTON & BURLING, LLP (June 16, 2025), https://www.insideglobaltech.com/2025/06/16/georgia-court-dismisses-defamation-suit-against-ai-developer-openai/ [https://perma.cc/HKL2-J6HX].

56 *See* Sai Anirudh Athaluri et al., *Exploring the Boundaries of Reality: Investigating the Phenomenon of Artificial Intelligence Hallucination in Scientific Writing Through ChatGPT References*, CUREUS (Apr. 11, 2023), https://www.cureus.com/articles/148687-exploring-the-boundaries-of-reality-investigating-the-phenomenon-of-artificial-intelligence-hallucination-in-scientific-writing-through-chatgpt-references [https://perma.cc/UT9B-K3HC]; *How Does Perplexity Work?*, PERPLEXITY HELP CTR., https://www.perplexity.ai/help-center/en/articles/10352895-how-does-perplexity-work [https://perma.cc/QCD3-KXB8].

57 *See* Athaluri et al., *supra* note 56; *How Does Perplexity Work?*, *supra* note 56.

58 *See AI Mode - Search Like Never Before*, GOOGLE SEARCH, https://search.google/ways-to-search/ai-mode [https://perma.cc/H2Y4-F5FV].

59 *See id.*

60 *What Is RAG (Retrieval-Augmented Generation)?*, *supra* note 31.

The process outlined above describes the original version of RAG, first explored in a seminal 2020 paper that coined the term.[61] More advanced implementations of the technology have developed since 2020. "Deep Research" products from leading AI labs describe similar processes in which models "search, interpret, and analyze massive amounts of text, images, and PDFs on the internet"[62] and "embed . . . graphs and images from websites in . . . responses, and cite specific sentences or passages from . . . sources."[63] Notably, when automated tools "read" or "analyze" the internet, it necessarily involves copying content from pages on the internet.[64] Therefore, when RAG-based AI and similar "deep research" or "agentic research" services operate, they necessarily copy those webpages, which may be copyright infringement under the Copyright Act, depending on the circumstances.[65]

Notably, however, the RAG services described above are similar to the common search engine technology that has been powering internet search services like Google and Bing for decades. Such search engines, in their basic form, work by first "crawling" websites on the internet and downloading copies of those webpages.[66] The engines then display links to those webpages in response to a user's search query, sometimes also including a snippet of the text available on the website, or, in the case of images, a low-resolution version of the image.[67] Users can then click the link and navigate to the website where they can view the content.

---

61 Patrick Lewis et al., *Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks*, NEURIPS (Apr. 14, 2021), https://arxiv.org/pdf/2005.11401 [https://perma.cc/A8FW-PN6Y].

62 *Deep Research System Card*, OPENAI (Feb. 25, 2025), https://cdn.openai.com/deep-research-system-card.pdf [https://perma.cc/DC6M-GJ23].

63 *Introducing Deep Research*, *supra* note 16.

64 *In-depth Guide to How Google Search Works*, GOOGLE FOR DEVELOPERS, https://developers.google.com/search/docs/fundamentals/how-search-works [https://perma.cc/SH3A-Y38C] ("Google downloads text, images, and videos from pages it found on the internet with automated programs called crawlers" and "analyzes the text, images, and video files on the page, and stores the information in the Google index, which is a large database.").

65 *See* 17 U.S.C. § 501.

66 *See Field*, 412 F. Supp. 2d at 1110; *see also* Bobby Allyn & John Ruwitch, *Online News Publishers Face 'Extinction-Level Event' From Google's AI-Powered Search*, NPR (Jul. 31, 2025) https://www.npr.org/2025/07/31/nx-s1-5484118/google-ai-overview-online-publishers [https://perma.cc/4K7M-3JJM].

67 *See Field*, 412 F. Supp. 2d at 1111.

Courts have deemed these well-established versions of web-search technology fair use.[68]

There are key differences between these established search engine technologies and the new RAG services, however, that may compel different outcomes. Principally, instead of providing a small snippet of the website with a link to navigate to it, RAG services provide substantial information and context, allowing a user to find all the information they need without navigating to the website hosting the information.[69] Further, the links to websites hosting the information are displayed smaller and less obviously in the user interface of many RAG and deep research services than in traditional web search, potentially leading to fewer click-throughs.[70]

Although RAG services resemble traditional web search tools, however, they also resemble LLM-powered chatbots like ChatGPT, which have a mixed record in front of courts so far.[71] Therefore, it is not clear, based on existing analogous technologies, whether RAG services are lawful. It is therefore necessary to examine the relevant copyright law and survey key authorities on fair use.

III. RELEVANT LAW

Next, this Article summarizes relevant United States copyright doctrine. It concludes that RAG services are likely prima facie infringing, and then outlines

---

[68] *See Kelly*, 336 F.3d at 815; *Field*, 412 F. Supp. 2d at 1123.

[69] *See* 2d Am. Compl. ¶ 1, *Dow Jones & Co.*, No. 24-cv-7984.

[70] *See generally* Allyn & Ruwitch, *supra* note 66; Hema Budaraju, *New Ways To Connect To The Web With AI Overviews*, GOOGLE (Aug. 15, 2024) https://blog.google/products-and-platforms/products/search/new-ways-to-connect-to-the-web-with-ai-overviews/ [https://perma.cc/3757-EYU2]; *Using AI Tools for Literature Discovery*, STORYBOOK UNIVERSITY, https://tcsedsystem.libguides.com/ai_tools_for_lit_discovery/perplexity [https://perma.cc/DZ4A-AN8V].

[71] *See* Thomson Reuters Enters. v. Ross Intelligence, 765 F. Supp. 3d 382, 401 (D. Del. 2025); *Kadrey*, 788 F. Supp. 3d at 1039–40, 1059–60.

the fair use defense and surveys relevant authority to determine whether the defense applies to RAG services.

A. General Copyright Law

To be eligible for a copyright, a work must be (1) an original work of authorship, and (2) fixed in a tangible medium of expression.[72] Works covered by copyright law include literary works, among others.[73] Once the above eligibility requirements are satisfied, authors receive several exclusive rights.[74] These rights include, in relevant part, the rights to reproduce the work in copies and prepare derivative works based on the copyrighted work.[75]

The owner of such exclusive rights can sue for any infringement of that particular right committed while they are the owner of it.[76] To prove infringement, the owner of a copyright must show that (1) there was actual copying and (2) there was a misappropriation that violates a statutory exclusive right.[77] The first element requires the plaintiff to prove that copying occurred and that the elements copied were expressions rather than ideas.[78] To show the requisite misappropriation, the plaintiff must show that protected elements of the copyrighted works were appropriated, and if so, that the allegedly infringing work was substantially similar to the copyrighted work.[79] Indirect evidence may be used to show copying

---

72 Feist Publ'ns, Inc. v. Rural Tel. Serv. Co., 499 U.S. 340, 355 (1991) (discussing the intention and impact of the Copyright Act of 1976).

73 Copyright Act of 1976, 17 U.S.C. § 102(a).

74 *Id.* § 106 (listing the exclusive rights bestowed upon the owner of copyrighted material).

75 *Id.*

76 *Id.* § 501(b).

77 *See* Arnstein v. Porter, 154 F.2d 464, 468 (2d Cir. 1946); Castle Rock Enters. v. Carol Publ'g Grp., 150 F.3d 132, 137 (2d Cir. 1998) (quoting Laureyssens v. Idea Grp., Inc., 964 F.2d 131, 139–40 (2d Cir. 1992)).

78 *See* 17 U.S.C. § 102(a). For example, a form used to notate a bookkeeping system is an expression that may be copyrightable, but the bookkeeping system, itself, is an uncopyrightable idea. *See* Baker v. Selden, 101 U.S. 99, 104 (1879).

79 *See* Steinberg v. Columbia Pictures Indus., 663 F. Supp. 706, 711 (S.D.N.Y. 1987) (stating the necessary elements of a copyright infringement claim).

and misappropriation by demonstrating that (a) the defendant had access to the copyright owner's work and (b) the two works are substantially similar.[80]

To meet the originality requirement to attain copyright, a work must only contain "at least some minimal degree of creativity."[81] The "requisite level of creativity is extremely low; even a slight amount will suffice. The vast majority of works make the grade quite easily as they possess some creative spark, no matter how crude, humble, or obvious it might be."[82] For example, phrases sufficiently creative for copyright protection include: "I may not be totally perfect, but parts of me are excellent," and "I have abandoned my search for truth and am now looking for a good fantasy."[83] While originality is required, novelty is not. [84] A new work may be copyrightable despite similarities to a prior work so long as the similarities are not due to copying."[85] "Words and short phrases such as names, titles, and slogans," however, do not meet this minimum threshold and are not subject to copyright.[86]

In addition to the right to reproduce their original works, authors also receive an exclusive right to create derivative works of their originals, which include "abridgement[s] and "condensation[s]."[87] Further, the "reporting of facts is not protectable under the Copyright Act since facts are 'never original to an author.'"[88] Compilations of facts, however, may be protected because the arrangement or presentation of facts "can display originality."[89] Descriptions of

---

80 *See* Price v. Fox Ent. Grp., Inc., 473 F. Supp. 2d 446, 454 (S.D.N.Y. 2007); Bright Tunes Music Corp. v. Harrisongs Music, Ltd., 420 F. Supp. 177, 180–81 (S.D.N.Y. 1976).

81 *Feist*, 499 U.S. at 345.

82 *Id.*

83 Brilliant v. W.B. Prods., Inc., 1979 U.S. Dist. LEXIS 9092, at *1 (C.D. Cal. Oct. 17, 1979).

84 *See id.*

85 *See id.*

86 37 C.F.R. § 202.1(a) (1992); *see also* Magic Mktg., Inc. v. Mailing Servs. of Pittsburgh, Inc., 634 F. Supp. 769, 772 (W.D. Pa. 1986).

87 17 U.S.C. §§ 101, 106(2).

88 Assoc. Press v. Meltwater U.S. Holdings, Inc., 931 F. Supp. 2d 537, 549 (S.D.N.Y. 2013) (quoting Nihon Keizai Shimbun, Inc. v. Comline Bus. Data, Inc., 166 F.3d 65, 70 (2d Cir.1999)).

89 Nihon Keizai Shimbun, Inc. v. Comline Bus. Data, Inc., 166 F.3d 65, 70 (2d Cir. 1999).

facts are also copyrightable because they can include originality.[90] "Thus, news articles may be entitled to protection under the Copyright Act to the extent they contain original expression."[91]

Even if a work is infringing per the law described above, however, such infringement may be deemed a fair use, which is a defense to copyright infringement.[92] Under this doctrine, otherwise infringing activity may be deemed non-infringing so long as it satisfies a four-factor test.[93] The test evaluates: "(1) the purpose and character of the use; (2) the nature of the copyrighted work; (3) the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and (4) the effect of the use upon the potential market for or value of the copyrighted work."[94]

Based on the foregoing, RAG services are likely prima facie infringing in two ways.[95] First, the process of building a database on which the RAG service can access information requires verbatim copying of webpages or documents available on the internet, which are usually copyrighted.[96] Second, RAG services generate outputs that likely contain sufficient copyrightable expression to raise a prima facie case of infringement. Such outputs can produce substantially similar or verbatim segments of expressive text, or derivative works in the form of abridgments and summarizations.[97] It is important, then, to determine whether RAG services are fair uses of the copyrighted works that inform their responses.

---

90 *See id.*

91 *Meltwater U.S. Holdings*, 931 F. Supp. 2d at 550.

92 *See* 17 U.S.C. § 107.

93 *See id.* (listing the four factors used to assess whether certain behavior is or is not protected as fair use of copyrighted work).

94 *Id.*

95 The process of training foundation models on which RAG services rely may also be infringing, but many courts and commentators have already rigorously analyzed whether training LLMs infringes copyright. *See generally* Lemley & Casey, *supra* note 6, at 748–49; Sobel, *supra* note 6, at 45; Sag, *supra* note 6, at 292–93; Glendenning, *supra* note 6, at 4–5.

96 *See supra* Part II, describing the compilation of vector databases.

97 *See* 2d Am. Compl. ¶¶ 5, 7–8, *Dow Jones & Co.*, No. 24-cv-7984 (alleging that Perplexity's outputs "summarize and paraphrase original, human-generated content, even at times reproducing that content *verbatim*.") (emphasis in original); *see also* Compl. ¶¶ 99–100, 107, 113, *New York Times v. Microsoft*, No.

B. FAIR USE LAW

> "[T]he fair use of a copyrighted work . . . for purposes such as criticism, comment, news reporting, teaching (including multiple copies for classroom use), scholarship, or research, is not an infringement of copyright. In determining whether the use made of a work in any particular case is a fair use the factors to be considered shall include —
>
> > (1) the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes;
> >
> > (2) the nature of the copyrighted work;
> >
> > (3) the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and
> >
> > (4) the effect of the use upon the potential market for or value of the copyrighted work.[98]

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."[99] The fair use test is a "flexible concept"[100] and each factor "may prove more important in some contexts than in others."[101] With that said, the fourth factor tends to be afforded the most weight.[102] Further, Fair use is meant to be a flexible doctrine that takes account of "significant changes in technology."[103]

Three distinct groups of cases are helpful to understand how courts may decide whether RAG is fair use: (1) recent AI cases, (2) search engine cases, and (3) news media cases.

23-cv-11195 for examples of LLM outputs allegedly containing verbatim text snippets of training data.

98 17 U.S.C. § 107.

99 Google LLC v. Oracle America, Inc., 593 U.S. 1, 18 (2021) (quoting Stewart v. Abend, 495 U.S. 207, 236 (1990)).

100 Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith, 598 U.S. 508, 527 (2023).

101 *Google*, 593 U.S. at 19.

102 *See* Harper & Row Publishers, Inc. v. Nation Enters., 471 U.S. 539, 566 (1985).

103 *Google*, 593 U.S. at 19.

1. *Recent AI Cases Differ on Whether Training AI Models is Fair Use*

Three recent orders on motions for summary judgment reached different conclusions when deciding whether training AI models is fair use.

First, in *Bartz et al. v. Anthropic PBC*, a class[104] of authors alleged that Anthropic trained its foundation LLM, Claude, on their copyrighted books.[105] The defendant, AI lab Anthropic, initially used a dataset of roughly seven million "pirated" books obtained from file-sharing websites before subsequently paying for scans of legitimately purchased physical books.[106] Judge William Alsup, a jurist with a wealth of experience in cutting-edge technology cases,[107] issued a multi-faceted ruling. First, he held that training Claude on the purchased book and "pirated" digital copy databases was fair use.[108] He then ruled that the act of building and retaining "forever" a "general purpose" library of the illegitimately obtained copies was not fair use.[109] To arrive at this cross-cutting decision, the court emphasized the recent Supreme Court fair use decision, *Warhol v. Goldsmith*, which instructs courts to "decide whether a 'copyrighted work has been used in multiple ways,' then evaluate each."[110]

Regarding the pirated copies, the court held that the first factor favored the authors because using the works for an indefinite general purpose library was not transformative.[111] The court noted that Anthropic did not pay for the copies, did not transform them into a significantly altered form, and lacked internal controls for their access and use.[112] Regarding the second factor, the nature of the copyrighted work, the court found against fair use because the books were

---

104 When Judge Alsup decided the motion for summary judgment, he had not yet ruled on plaintiffs' motion for class certification. He then certified a class thereafter. *Bartz*, 787 F. Supp. 3d at 1018.

105 *See id.*

106 *Id.* at 1016.

107 *See, e.g.*, Oracle America, Inc. v. Google Inc. 872 F. Supp. 2d 974 (N.D. Cal. 2012); Waymo LLC v. Uber Technologies, Inc. 256 F. Supp. 3d 1059 (N. D. Cal. 2017).

108 *Bartz*, 787 F. Supp. 3d at 1033.

109 *Id.*

110 *Id.* at 1020 (quoting *Andy Warhol Found.*, 598 U.S. at 533).

111 *See id.* at 1029.

112 *See id.* at 1025.

expressive works selected for their expressive qualities.[113] On the third factor, the amount and substantiality of the portion used, the court held that the amount of works was not "reasonable in relation to the purpose of the copying" because Anthropic "lacked any entitlement to hold copies of the books at all," concluding that almost any copying in this context was excessive.[114] Fourth, regarding market harm, the court held that an internal library comprised of illegitimately obtained copies could "destroy the [book] publishing market."[115]

Regarding the court's finding that training Anthropic's foundation model was fair use, the court noted that Anthropic implemented "limits," also known as guardrails, to mitigate outputs that include verbatim memorized text snippets.[116] The Court then proceeded through the factors, first noting that training Claude was "quintessentially transformative" and, that Claude was "trained upon works not to race ahead and replicate or supplant them—but to turn a hard corner and create something different."[117] The second factor analysis did not differ from that of the illegitimate copies.[118] Third, Judge Alsup held that, although Anthropic used the books at issue because they provided strong examples of writing, and therefore improved the quality of the model, using millions of works was reasonably necessary to train a high-quality LLM, and using "any one work for actually training LLMs was about as reasonable as the next."[119] Fourth, the Court held that no market existed for licensing data for AI-training purposes, and that there was no legal protection affording such a right, nor a strong likelihood that one would develop.[120] Finally, the court addressed plaintiffs' "market dilution" concern, that Claude can produce an influx of competing works that will decrease demand for future works by the plaintiff authors. It held that "[t]his is not the kind of competitive or creative displacement that concerns the Copyright Act" which "seeks to advance original works of authorship, not to protect authors against competition."[121]

---

113 *See id.* at 1029.

114 *Bartz*, 787 F. Supp. 3d at 1029–30.

115 *Id.* at 1033.

116 *See id.* at 1021.

117 *Id.* at 1022.

118 *See id.* at 1029.

119 *Id.* at 1030.

120 *See Bartz*, 787 F. Supp. 3d at 1032.

121 *Id.* (citing *Sega*, 977 F.2d at 1523–24.).

Next, in a similar case, *Kadrey v. Meta Platforms*,[122] a class of authors sued Meta, the parent company of Facebook and Instagram, claiming that it used the millions of the authors' books, also obtained through nefarious means, to train Meta's LLM, Llama.[123] Judge Chhabria held that Meta's training of Llama was fair use, but dedicated a substantial part of the opinion to an issue that was not robustly briefed: market dilution.[124] Judge Chhabria noted that market dilution will be sufficient to render model training *not* fair use "in most cases" and that the "upshot is that in many circumstances it will be illegal to copy copyright-protected works to train generative AI models without permission.[125] Which means that the companies, to avoid liability for copyright infringement, will generally need to pay copyright holders for the right to use their materials."[126] This fourth factor analysis differs significantly from *Bartz* on the market dilution theory.[127] *Kadrey* also differs from *Bartz* because Judge Chhabria found no separate infringing use in *Kadrey*; only the fair use of training an LLM.[128] The last significant difference in the order deals with whether the new use provides a "public benefit," which is sometimes analyzed as a subfactor[129] of the fourth factor.[130] Judge Chhabria held that Llama provides a slight public benefit by "making it better at helping users generate creative text or by improving its 'memory' and thereby making it more useful to the researchers who use it to develop software."[131] Aside from these three points,

122 Note: As of the time of publication, the author of this Article serves as counsel for plaintiffs in this case. This Article was prepared independently of that representation based solely on public information, and all discussions of *Kadrey* in the Article were substantially completed prior to the representation.

123 *Kadrey*, 788 F. Supp. 3d at 1034.

124 *See id.* at 1054–56.

125 *Id.* at 1034,1036.

126 *Id* at 1036.

127 *Bartz*, 787 F. Supp. 3d at 1032.

128 Limited only to the record before Judge Chhabria. *See Kadrey*, 788 F. Supp. 3d at 1058–59.

129 For a comprehensive background on the "public benefit" subfactor, see Amanda Levendowski, *Fairer Public Benefit in Copyright Law*, 47 CARDOZO L. REV. 119, 119 (2025).

130 *See Kadrey*, 788 F. Supp. 3d at 1058.

131 *Id.* at 1058.

Judge Chhabria's order generally came out similarly to Judge Alsup's "purchased library copy" analysis in *Bartz*.[132]

Finally, in a somewhat less similar case dealing with non-generative AI, Circuit Judge Stephanos Bibas, sitting by designation in the District of Delaware, held that training an AI model was not fair use.[133] In *Thomson Reuters Enterprise Centre GmbH v. Ross Intelligence Inc.*, Thomson Reuters, owner of the Westlaw research service, alleged that ROSS Intelligence, a competitor developing an AI-based legal research tool, infringed its copyrights in headnotes, short summaries of legal opinion sections.[134] ROSS obtained summaries of legal issues and relevant cases known as "Bulk Memos" from a third-party vendor that copied Westlaw's headnotes to create them.[135] ROSS then used these memos to train its AI model.[136] Unlike Claude and Llama, above, this model is not a generative AI, but rather an AI-based research tool.[137] The training process is nonetheless quite similar and involves providing large quantities of high-quality text.[138]

The court first established that many of Westlaw's headnotes met the "minimal threshold" of creativity for copyright protection and found that 2,243 headnotes had been directly infringed.[139] The court then turned to the first fair use factor, analyzing the purpose and character of the use, and held that ROSS's use was not transformative.[140] Although ROSS converted headnotes into data for its AI, its ultimate objective was to "make it easier to develop a competing legal research tool," which served a very similar purpose to Westlaw.[141] The court also held that ROSS's commercial use heavily weighed against a finding of fair use.[142] Second, the court held that the nature of the copyrighted work slightly favored ROSS Westlaw's headnotes were minimally creative and factually based. This

---

132 *See id.* at 1059.; *Bartz*, 787 F. Supp. 3d at 1031–32.

124 *See Thomson Reuters*, 765 F. Supp. at 390.

134 *See id.*

135 *Id.* at 391.

136 *See id.*

137 *See id.* at 399.

138 *See id.* at 391.

139 *Thomson Reuters*, 765 F. Supp. at 395.

140 *See id.* at 398–99.

141 *Id.* at 399.

142 *See id.*

factor was not afforded significant weight.[143] Next, the court held that the third factor favored ROSS and deemed it irrelevant that substantial parts of the headnotes were copied internally, finding the third factor focuses on the amount and substantiality of what is "made accessible to the public" as a competing substitute, not what is copied internally.[144] In other words, because ROSS provided legal opinions as its outputs and not headnotes, the amount and substantiality was minimal.[145] Although this factor favored ROSS, it was outweighed by factors one and four.[146] Fourth, Judge Bibas held that the effect of the use upon the potential market for or value of the copyrighted work weighed toward Thomson Reuters.[147] He reasoned that ROSS's use of the headnotes was not fair. ROSS "meant to compete with Westlaw by developing a market substitute" and had deprived Thomson Reuters of licensing revenue based on evidence that ROSS had initially approached Thomson Reuters to license its data.[148] As part of his fourth factor analysis, Judge Bibas also held that the prospect of a public benefit does not save ROSS, holding that the public has no right to Thomson Reuters's summaries.[149]

2. *Search Engine Cases May Favor AI Companies, Though Recent Trends May Favor Plaintiffs*

Next, several search engine cases provide insight into how courts may analyze RAG services. These cases generally support the inference that RAG is fair use, though more recent cases indicate a trend away from pro-fair-use arguments.

In the first case, *Kelly v. Arriba Soft Corp*., a professional photographer sued an internet image search engine called Arriba Soft Corp.[150] Arriba's system "crawled" the web, automatically downloading full-sized images, processing them to generate smaller, lower-resolution "thumbnails," and then storing these thumbnails in its database.[151] The Ninth Circuit found that the creation and display

143 *See id.*

144 *Id.* at 400.

145 *See Thomson Reuters*, 765 F. Supp. at 400.

146 *See id.* at 401.

147 *See id.* at 400.

148 *Id.*

149 *See id.*

150 *See generally* Kelly v. Arriba Soft Corp., 336 F.3d 811 (9th Cir. 2003).

151 *Id.* at 815.

of thumbnail images on the internet was fair use.[152] On the first factor, the court characterized Arriba's use of the plaintiff's photos as "incidental and less exploitative" than traditional commercial uses, and noted that Arriba did not directly promote its own site with the images or sell them.[153] The court emphasized that the images were a small piece of a large database.[154] Commerciality weighed slightly against a finding of fair use.[155] Further, the court found Arriba's internet search technology transformative.[156] While the thumbnails were exact replications of the original images, they were significantly smaller, lower-resolution, and served a fundamentally different function.[157] While Kelly's original images were artistic works intended for aesthetic experience and informative purposes, Arriba's use was for indexing and improving internet access to images.[158] In short, Arriba's use transformed the images from expressive content to functional information-gathering tools.[159] The court also noted that Arriba's search tool provided a public benefit, ultimately finding that factor one weighed toward Arriba.[160]

On factor two, the court held that the photos were creative works and thus subject to the "core of copyright protection," though this factor was mitigated by the presence of the photos on the internet, separate from Arriba's indexing of them.[161] The court held that the third factor favored neither party because copying the entire image was reasonable and necessary for the search engine to work.[162] Fourth, the court held that Arriba's use of thumbnails did not harm the market for, or value of, Kelly's images because the thumbnails would guide interested customers to Kelly's website.[163]

---

152 *See id.* at 822.

153 *Id.* at 818.

154 *See id.*

155 *See id.*

156 *See Kelly*, 336 F.3d at 819.

157 *See id.* at 818.

158 *See id.* at 820.

159 *See id.* at 818.

160 *See id.* at 820.

161 *Id.*

162 *See Kelly*, 336 F.3d at 821.

163 *See id.* at 821.

Following *Kelly*, the Ninth Circuit relied on and supplemented much of the reasoning above in a similar case, *Perfect 10, Inc. v. Amazon.com, Inc*.[164] In this case, a publisher of nude model photos, Perfect 10, sold full-resolution images on its website along with reduced-size images for download on cell phones.[165] It sued Google, which operated an image search engine, and Amazon.com, which used the search functionality.[166] Google's image search functioned similarly to Arriba's, and had inadvertently included unauthorized republications of Perfect 10 images without authorization in its search results.[167] Perfect 10 argued that internet users could use Google's search service as a substitute for its own subscriptions and cell-phone image download services.[168]

The Ninth Circuit disagreed with Perfect 10, finding the search service was a fair use.[169] The court held that Google's use of thumbnails was "highly transformative" for broadly the same reasons as in *Kelly*; web search was a functional search service and Perfect 10 provided aesthetic services.[170] The court noted that although Google's service was commercial, there was no evidence of market harm.[171] The court's reasoning on the second and third factors broadly mirrored its rationale in *Kelly*.[172] On the fourth factor, the court reaffirmed its reasoning in *Kelly* regarding full-sized images and held that any market harm to reduced-size cell phone images was "hypothetical."[173]

Next, in *Authors Guild v. Google, Inc*., the Authors Guild sued Google concerning its Google Books and Library Project, alleging Google had digitized tens of millions of books without obtaining permission from rights holders.[174] The service enabled users to search the full text of books and view "snippets"—small portions of the books' texts.[175] Google also provided digitized copies to

---

164 *See* Perfect 10, Inc. v. Amazon.com, Inc., 508 F.3d 1146, 1154 (9th Cir. 2007).

165 *See id.* at 1157.

166 *See id.* at 1155.

167 *See id.* at 1169.

168 *See id.* at 1165–66.

169 *See id.* at 1168.

170 *Perfect 10*, 508 F.3d at 1165.

171 *See id.* at 1168.

172 *See id.* at 1167–68; *Kelly*, 336 F.3d at 820–821.

173 *Perfect 10*, 508 F.3d at 1168.

174 *Authors Guild*, 804 F.3d at 207.

175 *Id.*

participating libraries.[176] The Second Circuit held that Google's service was fair use.[177] First, the court found that Google's unauthorized digitizing, search functionality, and snippet display were "highly transformative," because the search tool that enabled new forms of research and discovery was distinct from reading books.[178] The court held that the transformativeness of the services outweighed the commercial nature of the product, emphasizing the goal of copyright law to, and the benefit of, expanding "public knowledge."[179] The second and third factors weighed toward the Authors Guild because Google copied the entire copyrighted works, though these were outweighed by the first and fourth factors.[180] Regarding factor four, the court held that Google Books was not a market substitute because it limited public display of the books' text to small snippets.[181] It held that any effect of the use upon the potential market for or value of the copyrighted work was minimal because the snippets did not provide a "significant market substitute for the protected aspects of the originals."[182] Further, the court held that Google's provision of digitized copies to libraries did not constitute infringement because Google understood that the libraries would use them in accordance with the law.[183] Notably, the Second Circuit stated, and has since emphasized that whether Google Books was fair use was a close call, noting that it "tested the boundaries of fair use."[184]

Recently, in a somewhat-comparable digital book library lawsuit, the Second Circuit held that a similar book digitization scheme was not fair use.[185] In *Hachette Book Group, Inc. v. Internet Archive*, the non-profit Internet Archive created a "controlled digital lending" program in which it digitized lawfully-acquired physical books and lent them to users on a one-to-one basis.[186] First, the Second Circuit held that the service was not transformative because digital book lending

---

176 *See id.*

177 *See id.*

178 *Id.* at 216–17.

179 *Id.* at 207.

180 *See Authors Guild*, 804 F.3d at 220–22.

181 *See id.* at 223–25.

182 *Id.* at 229.

183 *See id.* at 226.

184 Fox News Network, LLC v. TVEyes, Inc., 883 F.3d 169, 174 (2d Cir. 2018).

185 Hachette Book Grp, Inc. v. Internet Archive, 115 F. 4th 163, 174 (2d Cir. 2024).

186 *See id* at 176.

served the same fundamental purpose, reading, of the physical versions of the books.[187] Notably, the court acknowledged that the Internet archive did not stand to profit off the digital lending program, but nonetheless held that the first factor favored the plaintiff.[188] The court held that the second factor favored the plaintiff as well because the books at issue were valuable for their expressive content.[189] Similarly, the court held that the third factor favored the plaintiffs because the Internet Archive copied and revealed to the public the entirety of the books.[190] Finally, the court held that the digital library caused market harm by supplanting the existing licensing market for e-books, in which authors and publishers license out their copyrights for digital distribution.[191] It accepted the publishers' assertion of market harm from lost licensing fees based on a "common-sense inference" that the publishers would lose licensing fees from the Internet Archive's practice.[192] When addressing whether it is acceptable to analyze potential, rather than actual, market harm under the fourth factor, the court specified that "only an impact on potential licensing revenues for traditional, reasonable, or likely to be developed markets should be legally cognizable."[193]

Finally, in *Field v. Google Inc.*, a poet sued Google alleging that it provided a "cached," or saved, copy of the poet's website to web search users, of his website without his permission in connection with its now-well-known Google Search service.[194] Judge Jones of the District of Nevada held that Google's caching activities constituted fair use.[195] The first factor weighed in Google's favor because the plaintiff poet's purpose was to "enrich and entertain others," while Google's was to facilitate search engine operation, serve archival purposes, track website changes, and explain search results.[196] Google's commercial status was of "less importance" due to the transformative nature of its use, and because Google did

187 *See id.* at 181.

188 *See id.* at 185.

189 *See id.* at 187.

190 *See id.* at 188–89.

191 *See Hachette Book Grp.*, 115 F.4th at 189–196.

192 *Id.* at 192. Notably, this may set up a circuit split between the Second Circuit and the Ninth Circuit, which required specific evidence of lost low-resolution thumbnail sales in *Perfect 10*. *Id.*

193 *Id.* at 192 (quoting *Fox News Network*, 883 F.3d at 180).

194 *Field*, 412 F. Supp. 2d at 1110–11.

195 *See id.* at 1123.

196 *Id.* at 1118–19.

not place ads on the cached version of Field's webpage.[197] Second, the nature of the copyrighted work also weighed in Google's favor because Field made the works available for free and, notably, added a robots.txt file to his website without noting in the file that caching that website was prohibited. The robots.txt file facilitated search engines' abilities to make his site more discoverable.[198] Third, the court found the amount and substantiality prong favored neither party because the plaintiff made his website available to the world and because Google's purpose was transformative.[199] Finally, factor four weighed in favor of Google because the plaintiff provided no evidence of a market for his specific works or of a licensing market for cached links, and admitted he had not profited from the works at issue.[200] Therefore, there was no evidence that caching harmed any such market for the plaintiff's work. The court continued by analyzing whether Google acted in good faith,[201] noting that the four fair use factors above are not exclusive, and that some courts in the Ninth Circuit "evaluate whether an alleged copyright infringer has acted in good faith as part of a fair use inquiry."[202] The court found that Google acted in good faith by abiding by industry standards.[203] Those standards included (1) not scraping or caching websites that provided anti-scraping instructions via a robots.txt file, and (2) adding instructions on its own website explaining how to prevent one's website from being cached.[204]

3. News Media Cases Disfavor RAG

Finally, cases involving traditional media provide arguments that generally disfavor RAG services.

---

197 *Id.* at 1119–20.

198 *See id.* at 1120.

199 *See id.* at 1120–21.

200 *See Field*, 412 F. Supp. 2d at 1121–22.

201 It is currently unclear whether the fair use test allows a "good faith" inquiry. Although the Supreme Court has held that "[f]air use presupposes good faith and fair dealing," *Harper & Row*, 471 U.S. at 562, it more recently expressed "skepticism about whether bad faith has any role." *Google,* 593 U.S. at 32–33. Even more recently, the Supreme Court has held that it is the "subjective intent" of a copyist is not relevant, but the "objective . . . use" of the copy is. *Andy Warhol Found.*, 598 U.S. at 544–45.

202 *Field*, 412 F. Supp. 2d at 1122.

203 *Id.*

204 *See id.* at 1122–23.

In *Associated Press v. Meltwater U.S. Holdings, Inc.*, the Associated Press sued Meltwater, a news monitoring service, alleging that Meltwater utilized automated web crawlers to electronically clip AP news articles, including concise summaries of key information, and distributed these clips to paid subscribers.[205] Meltwater raised the fair use defense and cited the search engine cases summarized above, among others.[206]

Judge Cote of the Southern District of New York held that Meltwater's service was not fair use.[207] The court first found that Meltwater's use of the plaintiff's articles was not transformative because, while Meltwater claimed to provide search functionality, the court held that it "uses its computer programs to automatically capture and republish designated segments of text from news articles, without adding any commentary or insight in its News Reports."[208] The court also examined the frequency with which users clicked through links on Meltwater's site to view the content on the Associated Press's site, which was 0.08% of the time, and noted that Meltwater provided no data comparing its clickthrough rates with those of traditional search engines.[209] The court concluded that Meltwater's product served as a substitute for accessing AP's articles rather than for search purposes.[210] The commercial nature of Meltwater's service, operating as a closed, paid subscription unlike public search engines, also weighed against fair use.[211]

Second, the nature of the copyrighted work weighed slightly toward fair use, because the Associated Press's news articles were published news articles, which are "more vulnerable to application of the fair use defense than works of fiction" due to their factual nature.[212] Third, the amount and substantiality of the portion used weighed against fair use because Meltwater copied significant portions (between 4.5% and 61%) of each article, including the highly journalistic "lede," which is meant to convey the "heart of the story."[213] Fourth, Judge Cote held that the effect on the market for or value of the copyrighted work was

---

205 *See Meltwater U.S. Holdings*, 931 F. Supp. 2d at 544–46.

206 *See id.* at 555–56, 564.

207 *See id.* at 561.

208 *Id.* at 552, 556–58.

209 *See id.* at 554.

210 *See id.*

211 *See Meltwater U.S. Holdings*, 931 F. Supp. 2d at 551–53.

212 *Id.* at 557.

213 *Id.* at 556–58.

substantial and weighed "strongly" against fair use.[214] The plaintiff generated significant revenue from licensing its articles and Meltwater's service by bypassing these established licensing fees, "cheapen[ed] the value of the AP's work by competing with companies that *do* pay a licensing fee in the way that Meltwater does."[215]

Next, in *Fox News Network, LLC v. TVEyes, Inc.*, TVEyes continuously recorded the audiovisual content of over 1,400 television and radio channels, including the plaintiff Fox News.[216] It imported the content into a database that users could pay $500 a month to access.[217] Users could then use a keyword search to view or download 10-minute clips of content.[218] The Second Circuit held that TVEyes's service was not fair use.[219] Regarding the first factor, the court found TVEyes's service had "a modest transformative character" because it enabled users to isolate specific material and access it efficiently, but provided no new "expression, meaning, or message."[220] The commercial nature of TVEyes' service further weighed against fair use, leading this factor to only slightly favor TVEyes overall.[221] The court held that the second factor was neutral and had "rarely played a significant role," while rejecting the argument that "the factual nature of [] content militates in favor of a finding of fair use."[222] Third, the court held that factor three "clearly favors Fox because TVEyes makes available virtually the entirety of the Fox programming that TVEyes users want to see and hear."[223] The 10-minute clips were substantial enough to convey the "entirety of the message" of news segments, unlike the small snippets in Google Books.[224] Finally, the court held that the effect of the use upon the potential market of the work was "the single most important element of fair use."[225] The court held that TVEyes's "success is evidence

---

214 *Id.* at 560–61.

215 *Id.* at 561 (emphasis in original).

216 *See Fox News Network*, 883 F.3d at 175.

217 *See id.*

218 *See id.* at 173–74.

219 *See id.* at 182.

220 *Id.* at 176.

221 *See id.* at 178.

222 *Fox News Network*, 883 F.3d at 178.

223 *Id.* at 179.

224 *Id.*

225 *Id.* (quoting *Harper & Row*, 471 U.S. at 566).

that consumers are willing to pay for a service that allows them to search for and view selected television clips, which is worth millions of dollars."[226]

IV. APPLYING THE LAW TO RAG

Given the authority above, a court would likely determine that RAG services, as they generally exist in their current state, are not fair use. Each factor would disfavor fair use or be neutral, at best. This Article analyzes each factor in turn, below.[227]

C. FACTOR ONE – THE PURPOSE AND CHARACTER OF THE USE, INCLUDING WHETHER SUCH USE IS OF A COMMERCIAL NATURE OR IS FOR NONPROFIT EDUCATIONAL PURPOSES

First, a court deciding whether RAG is fair use would first determine if the RAG service achieves the same purpose as the sources of its information. Courts would likely hold that RAG is mildly to moderately transformative.

RAG services are transformative because they allow users to locate and organize information from all across the internet efficiently, and in a manner that was not possible before RAG services.[228] For example, it is not currently possible to ask the New York Times website a question and receive a multi-paragraph written response synthesizing multiple articles and providing links to the source material. A user could, however, achieve this task with several RAG services currently available at no cost. This innovation is akin to the services analyzed in the search engine cases and news media cases above because RAG provides a useful information-organizing service.[229] Of course, as with *Thomson Reuters*, *Meltwater*, and *TVEyes*, providing a transformative organizational service is not always sufficient to establish fair use.[230]

---

226 *Id.* at 180.

227 Importantly, fair use is a fact specific inquiry that will depend on the facts of a given case and potentially the circuit in which the case is filed. *See* 4 MELVILLE B. NIMMER & DAVID NIMMER, NIMMER ON COPYRIGHT § 13F.03 (2026). Therefore, this Article does not predict the outcome of any particular case or set of facts.

228 *See supra* Part II.

229 *See What Is RAG (Retrieval-Augmented Generation)?*, *supra* note 31.

230 The courts in *Meltwater* and *TVEyes*, specifically, noted that those services did not offer any new "commentary or insight," *Meltwater U.S. Holdings*, 931 F. Supp. 2d at 552, or "expression, meaning, or message." *See TVEyes*, 883 F.3d at 176. Therefore, more advanced "deep research" services like those offered

RAG is likely only mildly to moderately transformative, rather than "spectacularly" transformative, as Judge Alsup referred to foundation models.[231] RAG services provide only a modest innovation in addition to the foundation models on which they are built; automatically citing and linking to sources. RAG services also have a narrower purpose than foundation models and are more similar to the works they copy than foundation LLMs are to their training data.[232] RAG services generate outputs that summarize specific information in their database, while foundation models proficiently generate any text in response to a query.[233] A court may even find that RAG services and the works that they copy share the same purpose because they both inform and educate users.[234]

RAG services are capable of additional purposes as well, like creating new expression.[235] It is not clear, however, whether users actually employ them for those potential purposes and how responsible the RAG service is for that purpose, as opposed to the underlying foundation model. While RAG services can generate new expression, commentary, and insight, largely due to their underlying foundation LLMs,[236] users may primarily use them only as a method to access the information, expression, commentary, and insight from their source documents. They might, therefore, be more similar to Meltwater's service in the eponymous case, in which its news clipping service was a substitute for Associated Press's

---

by OpenAI and Google may be more transformative than comparatively basic RAG services like Perplexity and Google's AI Search service because the former can perform more expression, commentary, and insight than the latter. *Compare* the example search results on *Gemini Deep Research, supra* note 16; *Introducing Deep Research, supra* note 16 *with* the example search results at *Using AI Tools for Literature Discovery*, *supra* note 70.

231 *Bartz*, 787 F. Supp. 3d at 1021.

232 *See RAG 2.0: Improved Retrieval Techniques,* PERPLEXITY (Jun. 22, 2024), https://www.perplexity.ai/page/rag-2-0-improved-retrieval-tec-OA0X2PggTBCgZmL0F_FyLA [https://perma.cc/A4JA-WSRX]; *Using AI Tools for Literature Discovery*, *supra* note 70.

233 *See supra* Part II.

234 *Id.*

235 For clarity, RAG services create new expression because they generate paragraphs of text in response to the user query, similar to chatbots built on foundation models. *See What Is RAG (Retrieval-Augmented Generation)?*, *supra* note 31.

236 *See, e.g., Using AI Tools for Literature Discovery*, *supra* note 70.

articles, than the Claude foundation LLM at issue in *Bartz*.[237] Therefore, courts may analyze evidence like the click-through rate analysis conducted by the *Meltwater* court to determine whether RAG services are actually used for a new purpose, or primarily as a substitute for source documents.[238] Evidence on Google AI overviews, which employ RAG,[239] strongly suggests that the service reduces clickthrough rates.[240] Consequently, RAG services may be deemed not transformative at all, or moderately transformative at most.

As the next part of the first factor analysis, courts also analyze whether a service is commercial. Most public RAG services are commercial services, operating on either a subscription model[241] or an ad model.[242] As with *Authors Guild v. Google*, *Kadrey*, and *Bartz*, a commercial use is not necessarily sufficient to weigh against fair use, though it would certainly not help the RAG service.[243]

In response to the above, RAG service providers might counter that transformativeness heavily influences other factors, including factors three and

---

237 *Compare Meltwater U.S. Holdings*, 931 F. Supp. 2d at 544–46, *with Bartz*, 787 F. Supp. 3d at 1021

238 *Meltwater U.S. Holdings*,, 931 F. Supp. 2d, at 554.

239 Oleg Zendel, *Should You Trust Google's AI Overview?*, RMIT UNIVERSITY (June 19, 2025), https://www.rmit.edu.au/news/media-releases-and-expert-comments/2025/jun/google-ai-overview [https://perma.cc/4PDN-SNCK]; Krishnan Kumar et al., *Building a Google-quality Search System with Vertex AI*, GOOGLE FOR DEVELOPERS, https://codelabs.developers.google.com/build-google-quality-rag#0 [https://perma.cc/3MV8-4VM7] . ("Google Search is a powerful tool that uses a massive index of web pages and other content to provide relevant results to user queries. This is made possible by a technique called Retrieval Augmented Generation (RAG), which is a key technique in modern AI.").

240 Ryan Law & Xibeijia Guan, *Update: AI Overviews Reduce Clicks by 58%*, AHREFS BLOG (Feb. 4, 2026) https://ahrefs.com/blog/ai-overviews-reduce-clicks-update/ [https://perma.cc/Z67B-4J2P].

241 *See Which Perplexity Subscription Plan is Right for You*, PERPLEXITY (last updated Feb. 19, 2026), https://www.perplexity.ai/help-center/en/articles/11187416-which-perplexity-subscription-plan-is-right-for-you [https://perma.cc/ZJ56-CJBE]; *ChatGPT Pricing*, OpenAI, https://openai.com/chatgpt/pricing/ [https://perma.cc/7CJV-25E9].

242 *See About Ads and AI Overviews*, GOOGLE ADS HELP, https://support.google.com/google-ads/answer/16297775?hl=en [https://perma.cc/CMB5-GKNV].

243 *See supra* Section III.B.1-2.

four. This trend was evident in cases like *Kelly*, *Perfect 10,* and *Google Books*, but has decreased in recent years.[244] One Second Circuit judge has criticized this trend in a concurrence,[245] and courts are now more likely to evaluate each factor independently from one another.[246] Overall, because RAG services are likely only mildly to moderately transformative, and because they are commercial in nature, the first factor would likely only slightly favor a RAG service provider, if at all.

Furthermore, even if maintaining a database for RAG services is transformative, courts may follow Judge Alsup's interpretation of the Supreme Court's *Warhol* decision and also scrutinize any other uses of the RAG service's database.[247] If RAG service providers do not have strong controls deciding who can access a database and for what purpose it may be accessed, a court could find no fair use for such a "general-purpose library," even if it finds fair use for the RAG service itself.[248]

D. FACTOR TWO – THE NATURE OF THE COPYRIGHTED WORK

The outcome of this factor would likely depend on the facts of the case and the circuit in which the court sits for two reasons. First, the availability of the works on the internet may be relevant, as they were in *Field*.[249] For example, if the RAG service provider honors robots.txt files and therefore does not crawl/scrape content from websites that opt out, this factor may favor the RAG service provider.

---

244 *See* Jane C. Ginsburg. *Fair Use Factor Four Revisited: Valuing the Value of the Copyrighted Work*. 67 J. COPYRIGHT SOC'Y U.S.A. 19, 19 (2020).

245 *Fox News Network*, 883 F.3d at 183 (Kaplan, J. Concurring) ("As noted commentators have observed, courts "appear to label a use `not transformative' as a shorthand for `not fair,' and correlatively `transformative' for `fair.' Such a strategy empties the term of meaning." Indeed, as will appear, some of our own decisions on the issue are at least in tension with one another.").

246 This is particularly the case after recent Supreme Court precedent doing the same. *See Google*, 593 U.S. at 10; *Andy Warhol Found.*, 598 U.S. at 1262.

247 *Bartz*, 787 F. Supp. 3d at 1020 (quoting *Andy Warhol Found.*, 598 U.S. at 533).

248 *See Bartz*, 787 F. Supp. 3d at 1033. Further, Judge Alsup's interpretation of Warhol might imply that other decisions have been abrogated. For example, in *Authors Guild v. Google*, the court noted that Google keeps a centralized database of book scans. If Google maintained that database without strong controls over how internal employees could use it, could such activity constitute an infringing general-purpose database like the one in *Bartz*. *Id.*

249 *See Field*, 412 F. Supp. 2d at 1120.

Conceivably, a court may also analyze whether the content is kept behind a paywall, as is common practice among news websites, for example. Second, if the content that the RAG service provider copies is news content, the court may find that "news stories [are] therefore more vulnerable to application of the fair use defense than works of fiction" as the court did in *Meltwater*.[250] Notably, however, the Second Circuit, in *TVEyes*, rejected this argument,[251] so the outcome of this factor is less clear than the first. In any case, courts do not generally afford this factor much weight.[252]

E. FACTOR THREE – THE AMOUNT AND SUSTAINABILITY OF THE PORTION USED IN RELATION TO THE COPYRIGHTED WORK AS A WHOLE

Factor three would also likely depend on the facts of the case. It may disfavor the RAG service provider in some cases because RAG services necessarily copy entire works in the form of webpages when creating their database.[253] Further, RAG outputs likely often contain the "heart" of the work, which the *Meltwater* court held disfavored fair use, citing the Supreme Court's *Harper & Row* decision.[254]

RAG service providers may implement safeguards, however, to prevent verbatim snippets longer than a certain length from being included in outputs, like the defendant in *Kadrey*.[255] Even if RAG services limit verbatim outputs to fifty words, however, like the *Kadrey* defendant, such a limitation may not be enough if the verbatim copy is the heart of a relatively short news story, or includes the lede.[256] Conversely, like *Kelly*, a court may find that full copying is proportional and necessary for RAG to work at all, and thus fulfill its transformative purpose.[257] It is not clear whether courts still allow the first factor to influence the third factor

---

250 *Meltwater U.S. Holdings*, 931 F. Supp. 2d at 557.

251 *Fox News Network*, 883 F.3d at 178.

252 *Thomson Reuters*, 765 F. Supp. at 401 ("Factor two matters less than the others . . . .").

253 *See supra* Part II.

254 *Meltwater U.S. Holdings*, 931 F. Supp. 2d at 558 (citing *Harper & Row*, 471 U.S. at 564–65).

255 *Kadrey*, 788 F. Supp. 3d at 1041–42.

256 *Meltwater U.S. Holdings*, 931 F. Supp. 2d at 556–58.

257 *See Kelly*, 336 F.3d at 820.

in this way, however.[258] Like the second factor, though, courts generally do not afford this factor significant weight, especially when the use at issue is transformative, and it is unlikely to be dispositive on its own.[259]

F. FACTOR FOUR – THE EFFECT OF THE USE UPON THE POTENTIAL MARKET FOR OR VALUE OF THE COPYRIGHTED WORK

This factor is often cited as the "most important factor" in the fair use test.[260] It would likely disfavor a RAG service provider for three reasons.

First, RAG services would likely harm the licensing markets of news outlets or websites from which they retrieve data. Many media organizations quickly began licensing their data to AI companies for model training purposes in the last few years, establishing a strong existing licensing market.[261] Some media companies have even begun licensing their data specifically for RAG purposes.[262]

---

258 *See Kelly*, 336 F.3d at 183 (Kaplan, J. Concurring).

259 Whether or not a use is transformative, which is usually evaluated under the first factor, carries significant weight. NIMMER & NIMMER, *supra* note 227, §§ 13F.05, 13F.10 ("Transformative use [] emerged as the dominant framework in fair use analysis."). The Fourth factor also carries significant weight, and the Supreme Court has twice called it the most important element of fair use. Stewart v. Abend, 495 U.S. 207, 238 (1990); *Harper & Row,* 471 U.S. at 566. In comparison, the second and third factors are often afforded less weight. *See, e.g., See* Warren Publ'g Co. v. Spurlock, 645 F. Supp. 2d 402, 423 (E.D. Pa. 2009) ("This Court finds that this [second] factor weighs slightly in favor of Plaintiffs, but it is of limited relevance because of the prior finding that [the] work is transformative."); Belmore v. City Pages, Inc., 880 F. Supp. 673, 678–79 (D. Minn. 1995) ("although wholesale copying militates against a finding of fair use, this [third] factor does not have significant weight when applied to the unique facts presented in this case").

260 *See Thomson Reuters*, 765 F. Supp. at 400 (quoting *Harper & Row*, 471 U.S. at 566); *Fox News Network*, 883 F.3d at 179 (quoting *Harper & Row*, 471 U.S. at 566).

261 *See* Bill Rosenblatt, *The Media Industry's Race To License Content For AI*, FORBES (July 18, 2024), https://www.forbes.com/sites/billrosenblatt/2024/07/18/the-media-industrys-race-to-license-content-for-ai/ [https://perma.cc/35EE-PXDV].

262 *See License your Content to Expand Your Reach,* DOW JONES, https://www.dowjones.com/business-intelligence/content-licensing/ [https://perma.cc/G65D-NWQG] ("Segmented and protected within secured database retrieval (RAG) models.").

Therefore, although some courts have been reluctant to find market harm for licensing markets that do not yet exist,[263] this is no longer an issue because licensing markets now exist.[264] Even if a court were to limit the licensing market specifically to RAG licensing, such a market would likely meet *Hachette's* standard, which states "only an impact on potential licensing revenues for traditional, reasonable, or likely to be developed markets should be legally cognizable."[265]

Second, RAG services would likely directly harm web traffic or subscription costs. Although causal evidence may be hard to demonstrate in some cases, plaintiffs could use evidence similar to that employed in *Meltwater* to demonstrate low click-through rates from RAG services to show that users use the RAG services in lieu of the sources.[266] Plaintiffs could potentially uncover data on how often their articles are cited and the corresponding click-through rates. They could supplement this information with data on decreasing subscriptions and/or page views to provide a strong inference of harm.

Finally, RAG services threaten to dilute the market for web news or similar informational websites for the same reason that foundation LLMs threaten to dilute the market for books. They significantly decrease the time and effort required to create competitive works.[267] Because news providers dedicate significant time and resources to gathering news, market dilution may be particularly harmful to such services.[268] Therefore, courts may agree with *Kadrey's* market dilution analysis and, assuming plaintiffs litigate the issue thoroughly, find no fair use.

Further, regarding the evidentiary burden to plaintiffs attempting to show market harm, courts seem to be trending away from the burdensome and exacting evidentiary burdens of *Perfect 10*.[269] Even the *Kadrey* court, which is located within

---

263 *See Bartz*, 787 F. Supp. 3d at 1032; *Kadrey*, 788 F. Supp. 3d at 1056–57.

264 *See* Sara Guaglione, *A Timeline of the Major Deals Between Publishers and AI Tech Companies in 2025*, DIGIDAY (Jan. 1, 2026), https://digiday.com/media/a-timeline-of-the-major-deals-between-publishers-and-ai-tech-companies-in-2025/ [https://perma.cc/Q26R-F6A9]; Rosenblatt, *supra* note 261 ("Segmented and protected within secured database retrieval (RAG) models.").

265 *Hachette Book Grp.*, 115 F.4th at 192 (quoting *Fox News Network*, 883 F.3d at 180).

266 *Meltwater U.S. Holdings*, 931 F. Supp. 2d at 132.

267 *See Kadrey*, 788 F. Supp. 3d at 1054.

268 *See* 2d Am. Compl. ¶¶ 1–5, *Dow Jones & Co.*, No. 24-cv-7984.

269 *See Fox News Network*, 883 F.3d at 179–80; *Hachette Book Grp.*, 115 F.4th at 192–93.

the Ninth Circuit, followed the Second Circuit in *Hachette*, rather than *Perfect 10*, holding that "rightsholder[s] need not prove or present evidence of market harm, they 'may bear some initial burden of *identifying* relevant markets.'"[270] Notably, however, *Kadrey* stopped short of applying *Hachette's* inference of market harm because courts may only infer market harm without any evidence when the copying at issue is "mere duplication for commercial purposes," as was arguably the case in *Hachette*.[271]

G. ADDITIONAL FACTORS AND SUBFACTORS

Two additional analyses might influence courts deciding whether RAG is a fair use: (1) public benefits and (2) good faith.[272] While both are worth mentioning, it is unlikely that either would significantly bear on the ultimate determination of fair use.

First, some courts may analyze the public benefits provided by RAG. *Kadrey* analyzed public benefits under the fourth factor and determined that it provided some public benefits to creative writers and software developers.[273] Ultimately, however, the court held that market dilution effects would outweigh public benefits in most circumstances.[274] In the case of RAG, the benefits would likely be even more marginal and would be limited to helping researchers find sources of information and confirm that information is not hallucinated. Therefore, most courts would likely hold that market harm outweighs any marginal public benefits.

Second, some courts inquire into whether the putatively infringing copying was made in good faith. While *Perfect* 10 and *Field* analyzed this pseudo-factor, it appears to be falling out of favor.[275] Both *Bartz* and *Kadrey* noted that the Supreme Court expressed skepticism of any good faith factor in *Google v. Oracle*,

---

[270] *Kadrey*, 788 F. Supp. 3d at 1038 (quoting *Hachette Book Grp.*, 115 F.4th at 194).

[271] *Id*. at 1057.

[272] *See Kelly*, 336 F.3d at 820; *Kadrey*, 788 F. Supp. 3d at 1060; *Perfect 10*, 508 F.3d at 1164; *Field*, 412 F. Supp. 2d at 1122 applying these analyses.

[273] *Kadrey*, 788 F. Supp. 3d at 1058; *Compare Kelly* which analyzed public benefits under its first factor transformativeness analysis and held that image searches provide significant public benefits. It is not clear, however, to what extent public benefits drove the ultimate determination on the first factor. In any case, courts would likely not assign this factor much weight for the reasons stated above. *Kelly*, 336 F.3d at 820.

[274] *Kadrey*, 788 F. Supp. 3d at 1060.

[275] *Perfect 10*, 508 F.3d at 1164; *Field*, 412 F. Supp. 2d at 1122.

and declined to analyze the issue in depth.[276] Therefore, this factor is unlikely to carry significant weight in any RAG cases.

V. CONCLUSION

In recent years, lawsuits related to generative AI have exploded, though so far, few of those lawsuits address RAG. Despite its increasing prevalence, RAG may infringe copyrights for similar reasons to the foundation LLMs at the heart of many well-known lawsuits, but to an even greater degree. RAG is different from the technology at issue in those lawsuits in meaningful ways and will warrant a distinct fair use analysis. Due to those meaningful differences, courts are less likely to find that RAG services are fair uses than their foundation LLM counterparts, which already have a mixed record in recent groundbreaking cases. RAG services are less likely to have a transformative purpose and character, generate outputs that might copy the "heart" of the copied work, and present greater threats to the potential use and market of the original works than foundation LLMs. Therefore, RAG services are likely not fair use.

[276] *Kadrey*, 788 F. Supp. 3d at 1046; *Bartz*, 787 F. Supp. 3d at 1026 n.5.