

**UNITED STATES DISTRICT COURT
NORTHERN DISTRICT OF CALIFORNIA
SAN FRANCISCO DIVISION**

ANDREA BARTZ, ANDREA BARTZ, INC.,
CHARLES GRAEBER, KIRK WALLACE
JOHNSON, and MJ + KJ, INC., individually and on
behalf of others similarly situated,

Plaintiffs,

v.

ANTHROPIC PBC,

Defendant.

Case No. 3:24-CV-05417-WHA

Action Filed: August 19, 2024

**DECLARATION OF TOM TURVEY IN
SUPPORT OF DEFENDANT
ANTHROPIC PBC'S
ADMINISTRATIVE MOTION FOR
SUMMARY JUDGMENT**

I, Tom Turvey, hereby declare as follows:

A. Personal Background

1. I am the Vice President of Product Partnerships at Anthropic. I have worked in this position since February 2024. I submit this declaration in support of Anthropic's Motion for Summary Judgment. Unless stated otherwise, all facts stated herein are within my personal knowledge. If called upon, I would be willing to testify as set forth in this declaration.

2. I have a Bachelor of Arts in political science and government from Oklahoma State University.

3. For the first 11 years of my career, I worked in the book publishing industry.

4. From December 1993 to January 1997, I was the Director of Online Sales & Marketing at HarperCollins Publishers. In this role, I pioneered the online bookselling channel as early as 1995 by establishing trading terms and relationships with Barnes & Noble.com, Amazon.com, and other online book retailers.

5. From January 1997 to January 2001, I was the Director of Merchandising & Publisher Relations at Barnes & Noble ("B&N"). In this role, I built the first online presence for B&N, including building key publisher relationships for the largest brick and mortar bookseller in the world. At B&N, my teams and I negotiated both deals for physical book sales and early digital book efforts that preceded dedicated digital book readers.

6. From January 2001 to January 2004, I was the Vice President of Content Development at ebrary, a book publisher-funded content start-up. In this role, I created the company's digital book products by licensing tens of thousands of digital books from hundreds of global book publishers of all kinds to libraries and institutions around the world.

7. After working in the book publishing industry, I spent 20 years working at Google in various capacities. At Google, in 2004, I helped create Google Books, which is Google's massive book digitization project in partnership with publishers and libraries around that world that has scanned over 40 million books over the years for many purposes, including providing full viewing of public domain books, partial view & snippet view of in-copyright books, searching capabilities, and more. Over 20 years, my team concluded thousands of licensing deals between Google and

1 traditional publishing partners to support products like Google Books, Play Books, Google Scholar,
2 Google News, Search, and other products.

3 8. Over time, I took on an increasingly diverse set of responsibilities at Google related
4 to commercial partnerships with newspapers, journals, magazines, and other types of traditional
5 publishing partnerships.

6 9. As such, before I came to Anthropic, I had extensive experience negotiating deals
7 related to books from sales agreements for physical books to licenses of digital books in different
8 contexts.

9 10. I have also served on the book industry publishing standards board, Book Industry
10 Study Group.

11 11. As a result of the sum of the work described above, I have an extensive network of
12 contacts inside the book publishing industry. If there is business to be conducted between book
13 publishers and tech companies, I am extremely likely to be contacted and believe that I am highly
14 qualified to negotiate such deals efficiently.

15 **B. My Role at Anthropic and Data Acquisition for LLM Training**

16 12. As Vice President of Product Partnerships, my job is to build and maintain
17 commercial partnerships related to development of Anthropic's LLM technology. This includes
18 partnerships related to the acquisition of data for Anthropic's LLM training, as well as a diverse
19 array of other product partnerships, such as [REDACTED], as one
20 example.

21 13. When I started at Anthropic in February 2024, [REDACTED]
22 [REDACTED]
23 [REDACTED] Other than that, to the best of my knowledge, at that
24 point, Anthropic had acquired all or substantially all of its LLM training data from publicly-
25 available sources on the Internet and not through any third party deals.

26 14. Over time, I have communicated with various third parties to explore potential
27 avenues to acquire LLM training data. My team maintains a document tracking deals and
28 negotiations, including deals and potential deals for data acquisition, a version of which I understand

has been produced as ANT_BARTZ_000437329. To the best of my knowledge, among many other deals, this tracker identifies all or substantially all of the potential deals that Anthropic has evaluated concerning the acquisition of book datasets to use as LLM training data as of the date of that document.

15. During the first few months of my employment, I had many communications with book publishers about potentially acquiring electronic book datasets as LLM training data. After several months of discussing potential deals such as this, I concluded that, while there are some circumstances where it makes sense for Anthropic to enter third party deals to acquire data, there is not a viable market to license book datasets at any scale remotely approaching Anthropic's complete needs for the use of books as LLM training data. At this time, we have not entered into any deal to acquire electronic book datasets

16. With respect to book datasets specifically, I found that, relative to the size of the global book market, which includes tens of millions of books, very few books are available to be licensed in datasets for use as LLM training data.

17. Further, in the cases that books were purportedly available, potential licensors were frequently demanding pricing that was not economically justifiable for Anthropic to pay based on the incremental value those books would provide with respect to its LLM training needs.

(ANT_BARTZ_000279410).

Exhibit 45

18. Additionally, putting aside the issue of cost, if licenses are not available from book publishers at scale, then there is no possible way for Anthropic to negotiate licenses for tens of millions of books by attempting to seek licenses from one book author at a time—let alone for all

the other different types of training data that Anthropic uses (e.g., trillions of tokens of data pulled from source code repositories, the government, and innumerable websites). Anthropic does not have the resources, manpower, or time that would be required to negotiate millions of individual licenses for books, or the many more millions of licenses it would take to negotiate licenses to all training data of other types.

19.

C.

20.

21.

22.

23.

24.

25.

[REDACTED]

[REDACTED]

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28

26.

[REDACTED]

27.

[REDACTED]

1 [REDACTED]

2 [REDACTED]

3 **D. Other Discussions Related to Acquisition of Books for LLM Training**

4 28. During the time I was actively exploring potential data acquisition deals for books, I

5 had discussions with many publishers concerning possible deals related to books that did not

6 advance beyond preliminary discussions.

7 29. As noted above, given my long career working in and around publishing, I have

8 primary or secondary contacts at most of the major book publishing houses and I reached out to a

9 number of those contacts for this purpose, as described in more detail below.

10 **E. US Trade Publishing Houses**

11 30. The US consumer book publishing market is dominated by five major publishing

12 companies, often referred to as the “Big Five,” (HarperCollins, Penguin Random House, Simon &

13 Schuster, MacMillan, and Hachette Books Group). This is sometimes called the “trade” publishing

14 side of the industry. Plaintiffs’ Books are published by three of these publishers: Penguin Random

15 House, Simon & Schuster and Hachette Books Group. In my experience, around half of the output

16 of these trade publishing houses is fiction, which Anthropic does not value as highly, so I spent less

17 time in general engaging with trade publishing houses.

18 31. To the best of my knowledge, none of the five major trade publishing houses in the

19 US make anything close to the entirety of their catalogues available for LLM licensing. In fact (as

20 discussed below), HarperCollins may be the only US trade publishing house that even licenses a

21 subset of its books for this purpose.

22 32. [REDACTED]

23 [REDACTED]

24 [REDACTED]

25 [REDACTED]

26 [REDACTED]

27 [REDACTED]

28 [REDACTED]

33.

34.

35.

36. I am aware of only one reported license agreement for LLM training data by those US trade publishing houses. In November 2024 [REDACTED] HarperCollins confirmed reports that it had entered a deal offering licenses for certain backlist¹ nonfiction books as LLM training data at a rate of \$5,000 per book on a three year term.² The deal is reported to be an opt-in deal, meaning that each individual book author must give permission for his or her work to be included in the license. I am also aware that the Author's Guild has taken the position that this deal does not provide a fair split of revenue between authors and publishers.³ Based on these terms, I would be shocked if any LLM company had licensed a significant amount of books at prices like that, and I am dubious about whether this license has been duplicated since or, instead, is intended

¹ In publishing, a backlist book is a book that is still in the publisher's active catalog but not a new release. Publishers typically do not invest money in actively marketing backlist books and, therefore, they typically do not charge the same price premiums for backlist books as new releases.

² Jay Peters, *HarperCollins Is Asking Authors to License their Books for AI Training*, The Verge, <https://www.theverge.com/2024/11/18/24299882/harpercollins-authors-license-books-ai-training> (Nov. 18, 2024) (Exhibit 46)

³ The Authors Guild, *HarperCollins AI Licensing Deal*, <https://authorsguild.org/news/harpercollins-ai-licensing-deal/> (Nov. 19, 2024) ("We believe that a 50-50 split for a mere AI training license gives far too much to the publisher.") (Exhibit 47).

[REDACTED]

1 to generate publicity.

2 [REDACTED]

3 [REDACTED]

4 [REDACTED]

5 **F. Academic and Professional Publishers**

6 37. In contrast to the trade publishers, in my experience, academic and professional
7 publishers often acquire more expansive licensing rights from authors. However, while some
8 academic publishers purport to be offering licenses for LLM training data, this is far from uniformly
9 true, and even those publishers that do purport to have such rights are often not attempting to license
10 their full catalogues.

11 38. [REDACTED]

12 [REDACTED]

13 [REDACTED]

14 [REDACTED]

15 [REDACTED]

16 [REDACTED]

17 [REDACTED]

18 [REDACTED]

19 [REDACTED]

20 [REDACTED]

21 [REDACTED]

22 [REDACTED]

23 39. [REDACTED]

24 [REDACTED]

25 [REDACTED]

26 [REDACTED]

27 [REDACTED]

28 ⁴ [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

43.

[REDACTED]

44.

[REDACTED]

45.

[REDACTED]

G. Self-Publishers

46. I also had some discussions with self-publishing platforms (which allow authors to publish without working with traditional publishers) about the acquisition of books datasets, but those conversations were even less productive.

47.

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

48. [REDACTED]

[REDACTED]

[REDACTED]

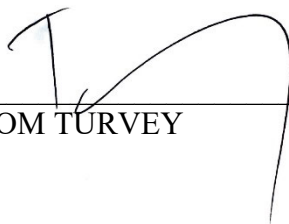
49. [REDACTED]

[REDACTED]

H. Purported Third Party Rights Aggregators

50. There are also times that Anthropic has been approached by companies that allegedly aggregate rights from third parties including books, but based on my industry experience, I am not sufficiently confident that these companies are legitimate rightsholders for books to engage in negotiations with them.

I declare under penalty of perjury under the laws of the United States of America that the foregoing is true and correct. Executed on March 27, 2025, in New York City, New York.


TOM TURVEY



1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28

CERTIFICATE OF SERVICE

I, Douglas A. Winthrop, am the ECF user whose identification and password are being used to file the foregoing **DECLARATION OF TOM TURVEY IN SUPPORT OF DEFENDANT ANTHROPIC PBC’s MOTION FOR SUMMARY JUDGMENT.**

Dated: March 27, 2025

/s/ Douglas A. Winthrop