

UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK

MIKE HUCKABEE, RELEVATE GROUP,
DAVID KINNAMAN, TSH OXENREIDER,
LYSA TERKEURST, and JOHN BLASE
on behalf of themselves and all others
similarly situated,

Plaintiffs,

- against -

META PLATFORMS, INC., BLOOMBERG
L.P., BLOOMBERG FINANCE, L.P.,
MICROSOFT CORPORATION, and
THE ELEUTHERAI INSTITUTE,

Defendants.

Case No. 23-cv-09152-MMG

**ORAL ARGUMENT
REQUESTED**

**MEMORANDUM OF LAW
IN SUPPORT OF BLOOMBERG L.P. AND BLOOMBERG FINANCE L.P.'S
MOTION TO DISMISS FIRST AMENDED CLASS ACTION COMPLAINT**

TABLE OF CONTENTS

- I. INTRODUCTION..... 1
- II. RELEVANT FACTUAL BACKGROUND 2
- III. LEGAL STANDARD 3
- IV. ARGUMENT 4
 - A. Plaintiffs’ Copyright Claim Fails to State a Claim Upon Which Relief May Be Granted. 4
 - B. Plaintiffs’ Copyright Claim Is Barred by the Doctrine of Fair Use. 9
 - C. Plaintiffs’ Amended Complaint Should Be Dismissed with Prejudice. 22
- V. CONCLUSION 23

TABLE OF AUTHORITIES

	Page(s)
Cases	
<i>Abdin v. CBS Broad. Inc.</i> , 971 F.3d 57 (2d Cir. 2020).....	4
<i>Apple Inc. v. Corellium, Inc.</i> , No. 21-12835, 2023 WL 3295671 (11th Cir. May 8, 2023).....	14, 18
<i>Ashcroft v. Iqbal</i> , 556 U.S. 662 (2009).....	3
<i>Authors Guild, Inc. v. HathiTrust</i> , 755 F.3d 87 (2d Cir. 2014).....	<i>passim</i>
<i>Authors Guild v. Google, Inc.</i> , 804 F.3d 202 (2d Cir. 2015).....	<i>passim</i>
<i>Bell Atl. Corp. v. Twombly</i> , 550 U.S. 544 (2007).....	3, 4, 5, 8
<i>Bellikoff v. Eaton Vance Corp.</i> , 481 F.3d 110 (2d Cir. 2007).....	22
<i>Bespaq Corp. v. Haoshen Trading Co.</i> , No. 04-cv-3698, 2005 WL 14841 (N.D. Cal. Jan. 3, 2005).....	8
<i>Bill Graham Archives v. Dorling Kindersley Ltd.</i> , 448 F.3d 605 (2d Cir. 2006).....	19, 20
<i>Blanch v. Koons</i> , 467 F.3d 244 (2d Cir. 2006).....	19
<i>Brody v. Fox Broad Co., LLC</i> , No. 22-cv-6249, 2023 WL 2758730 (S.D.N.Y. Apr. 3, 2023).....	9
<i>Brown v. Netflix, Inc.</i> , 462 F. Supp. 3d 453 (S.D.N.Y. 2020).....	9
<i>Cambridge Univ. Press v. Patton</i> , 769 F.3d 1232 (11th Cir. 2014)	19
<i>Campbell v. Acuff-Rose Music, Inc.</i> , 510 U.S. 569 (1994).....	9, 17, 18

Cariou v. Prince,
714 F.3d 694 (2d Cir. 2013).....20

Clark v. Transp. Alts., Inc.,
No. 18-cv-9985, 2019 WL 1448448 (S.D.N.Y. Mar. 18, 2019).....9

Cole v. John Wiley & Sons, Inc.,
No. 11-cv-2090, 2012 WL 3133520 (S.D.N.Y. Aug. 1, 2012).....8

Concord Music Group, Inc. et al. v. Anthropic PBC,
No. 23-cv-1092, ECF No. 1 (M.D. Tenn. Oct. 18, 2023).....21

Feist Publications, Inc. v. Rural Tel. Serv. Co.,
499 U.S. 340 (1991).....14

Google LLC v. Oracle Am., Inc.,
141 S. Ct. 1183 (2021).....20, 21

Hartmann v. Popcornflix.com LLC,
No. 20-cv-4923, 2023 WL 5715222 (S.D.N.Y. Sept. 5, 2023)4

Hughes v. Benjamin,
437 F. Supp. 3d 382 (S.D.N.Y. 2020).....9

Kadrey et al. v. Meta Platforms, Inc.,
No. 23-cv-03417, ECF No. 69 (N.D. Cal. Dec. 22, 2023).....8

Kaye v. Cartoon Network Inc.,
297 F. Supp. 3d 362 (S.D.N.Y. 2017).....23

Kelly v. Arriba Soft Corp.,
336 F.3d 811 (9th Cir. 2003)14, 15, 17

Kelly v. L.L. Cool J.,
145 F.R.D. 32 (S.D.N.Y. 1992)4

Lamda Sols. Corp. v. HSBC Bank USA, N.A.,
574 F. Supp. 3d 2057 (S.D.N.Y. 2021).....4

United States ex rel. Levine v. Vascular Access Ctrs., L.P.,
No. 12-cv-5103, 2020 WL 5534670 (S.D.N.Y. Sept. 15, 2020)22

Lindsay v. Wrecked & Abandoned Vessel R.M.S. Titanic,
No. 97-cv-9248, 1999 WL 816163 (S.D.N.Y. Oct. 13, 1999).....7

Lopez v. Bonanza.com, Inc.,
No. 17-cv-8493, 2019 WL 5199431 (S.D.N.Y. Sept. 30, 2019)7

Microsoft Co. et al., The New York Times,
 No. 23-cv-11195, ECF No. 1 (S.D.N.Y. Dec. 27, 2023).....21

Nat’l Basketball Ass’n v. Motorola, Inc.,
 105 F.3d 841 (2d Cir. 1997).....14

Perfect 10, Inc. v. Amazon.com, Inc.,
 508 F.3d 1146 (9th Cir. 2007)14

Shady Recs., Inc., v. Source Enters. Inc.,
 371 F. Supp. 2d 394 (S.D.N.Y. 2005).....6, 13

Sony Corp. of Am. v. Universal City Studios, Inc.,
 464 U.S. 417 (1984).....9

Swatch Grp. Mgmt. Servs. Ltd. v. Bloomberg L.P.,
 756 F.3d 73 (2d Cir. 2014).....11

A.V. ex rel. Vanderhye v. iParadigms, LLC,
 562 F.3d 630 (4th Cir. 2009) *passim*

Warner Bros. Ent. Inc. v. RDR Books,
 575 F. Supp. 2d 513 (S.D.N.Y. 2008).....10

Wheeler v. Topps Co.,
 652 F. Supp. 3d 426 (S.D.N.Y. 2023).....22

White v. W. Pub. Corp.,
 29 F. Supp. 3d 396 (S.D.N.Y. 2014).....13, 20

Wilder v. Hoiland,
 No. 22-cv-1254, 2024 WL 382141 (S.D.N.Y. Feb. 1, 2024)11, 15, 18, 19

Wright v. Warner Books, Inc.,
 953 F.2d 731 (2d Cir. 1991).....11

Yang v. Mic Network, Inc.,
 405 F. Supp. 3d 537 (S.D.N.Y. 2019).....9

Statutes

17 U.S.C. § 102(b).....13

17 U.S.C. § 107.....10, 11

Copyright Act.....15, 18, 21

Fed. R. Civ. P. 12(b)(6).....9

Fed. R. Civ. P. 15(a)22
U.S. Const. art. I, § 8.....9
U.S. Const. art. I, § 8, cl. 8.....1

I. INTRODUCTION

The core purpose of copyright is “[t]o promote the Progress of Science and useful Arts” U.S. Const. art. I, § 8, cl. 8. Copyright law thus draws a delicate balance between protecting creators’ rights to their expressive works and encouraging society to build freely upon the work of others.

Bloomberg, a leading creator and distributor of news and analysis that has long understood this balance, conducted a research project analyzing the potential uses for generative artificial intelligence (“AI”). After doing so, Bloomberg released a research paper detailing the development, performance and potential of BloombergGPT, a generative AI model developed to research potential uses of generative AI in the financial industry.

Seizing on nothing more than Bloomberg’s press release and news articles regarding that paper, Plaintiffs allege that Bloomberg violated their copyrights. But Plaintiffs’ vague and conclusory statements are insufficient bases on which to state a viable claim. And what Plaintiffs fail to plead speaks volumes.

Indeed, after having a chance to review Bloomberg’s original pre-motion letter and arguments in support of dismissal (Dkt. No. 54), and after having an opportunity to amend their original complaint, Plaintiffs still fail to allege that Bloomberg released or published BloombergGPT to the public or to any third-party, much less for profit. Nor do they allege that anyone outside Bloomberg has ever tested or used BloombergGPT. Plaintiffs do not even allege which of their copyrighted works were purportedly used by Bloomberg to train BloombergGPT.

Even viewing the Amended Complaint’s allegations in the light most favorable to Plaintiffs, it is clear that Plaintiffs’ claim is fundamentally flawed and must be dismissed as a matter of law. Bloomberg’s alleged “use” of Plaintiffs’ copyrighted works as part of a research project into the capabilities of generative AI falls squarely within the ambit of the fair use doctrine,

a statutory exception to infringement. Indeed, there is no plausible allegation that BloombergGPT has in any way displaced or replaced Plaintiffs' books in the literature market. Accordingly, Bloomberg's training of BloombergGPT could not have constituted copyright infringement, even if Plaintiffs had pleaded their complaint with sufficient detail.

Because the Amended Complaint does not state a viable claim for copyright infringement, because Plaintiffs have already had an opportunity to cure their deficiencies (with the benefit of Bloomberg's arguments outlined in its pre-motion letter), and because it is clear that any future amendment would be futile, the Amended Complaint should be dismissed in its entirety, with prejudice.

II. RELEVANT FACTUAL BACKGROUND

Plaintiffs filed an initial complaint in this matter on October 17, 2023. At that time, Plaintiffs asserted claims for copyright infringement (direct and indirect); for violations of the Digital Millennium Copyright Act ("DMCA"); and for conversion, negligence, and unjust enrichment under state law. (Dkt. No. 1.) On December 15, 2023, Bloomberg filed a pre-motion letter informing the Court of its intent to move to dismiss all causes of action. (Dkt. No. 54, "First Dismissal Letter.") The Court ordered Plaintiffs to file an amended complaint. (Dkt. No. 73.) Plaintiffs filed the Amended Complaint on January 24, 2024. (Dkt. No. 74.) Plaintiffs have now withdrawn their indirect copyright infringement, DMCA and state-law claims, but maintain their claim for direct copyright infringement.

Like their original complaint, Plaintiffs' Amended Complaint concerns BloombergGPT. (Am. Compl. ¶ 44.) BloombergGPT is a "large language model" ("LLM")—a term that refers broadly to a class of generative artificial intelligence ("AI") models "designed to understand and generate human language." (*Id.* ¶¶ 1, 44.) To "train" a LLM, the model must be exposed to a vast amount of material. In doing so, the goal is not for the model to memorize text. Instead, "training

data is used to teach the model grammar, vocabulary, context, and various language patterns.” (*Id.* ¶ 27.a.) LLMs thus learn much in the same way a student learns through repeated exposure to material. Plaintiffs vaguely allege that copies of books they wrote were included in a publicly available “data set” that was used to “train” BloombergGPT.¹ (*Id.* ¶ 38, 50.)

On March 30, 2023, Bloomberg released a research paper detailing the development and training of BloombergGPT. (*Id.* ¶¶ 45–47.) Despite the detailed information Bloomberg made publicly available through this paper, Plaintiffs’ allegations specifically regarding BloombergGPT are limited to twelve paragraphs of the Amended Complaint, which appear to be sourced entirely from news articles and press releases describing the paper, rather than the paper itself. (*Id.* ¶¶ 43–54.) Plaintiffs do not allege that BloombergGPT has been publicly released, that it has been used in any commercial capacity, or that it has even been used by anyone outside of Bloomberg at all. (*Id.* ¶¶ 43–54.)

III. LEGAL STANDARD

To avoid dismissal, a complaint must plead “enough facts to state a claim to relief that is plausible on its face.” *Bell Atl. Corp. v. Twombly*, 550 U.S. 544, 570 (2007). “Where a complaint pleads facts that are merely consistent with a defendant’s liability, it stops short of the line between possibility and plausibility of entitlement to relief.” *Ashcroft v. Iqbal*, 556 U.S. 662, 678 (2009) (internal quotation marks and citation omitted). A complaint “requires more than labels and conclusions, and a formulaic recitation of the elements of a cause of action will not do” *Twombly*, 550 U.S. at 555. Instead, the complaint must show that the assertions contained within

¹ Although Plaintiffs allege ownership interests in certain copyrighted works (*id.* ¶¶ 16–21), they do not allege which, if any, of these works was used to train BloombergGPT. Instead, Plaintiffs vaguely allege simply that “[r]esearchers and journalists have dug into the Books3 data set and verified the presence of hundreds of thousands of books in that data set, including books authored by Plaintiffs and Class Members.” (*Id.* at ¶ 38.) Notably, none of the named plaintiffs is identified in the article cited in the Amended Complaint. (*Id.* at ¶ 38 n.11.)

are “plausible,” not just “conceivable.” *Id.* at 570. Finally, “[a]lthough factual allegations of a complaint are normally accepted as true on a motion to dismiss, that principle does not apply to general allegations that are contradicted by more specific allegations in the [c]omplaint.” *Lamda Sols. Corp. v. HSBC Bank USA, N.A.*, 574 F. Supp. 3d 205, 216–17 (S.D.N.Y. 2021) (quoting *DPWN Holdings (USA), Inc. v. United Air Lines, Inc.*, 747 F.3d 145, 151–52 (2d Cir. 2014)).

IV. ARGUMENT

Plaintiffs summarily allege, with no specificity, that Bloomberg directly infringed their copyrights in various unspecified works (which they allege are included in the Books3 dataset) when Bloomberg “created, copied, maintained and/or utilized the Books3 dataset” to train BloombergGPT. (Am. Compl. ¶ 73.) Plaintiffs’ bald and conclusory assertions are not sufficient to satisfy the *Twombly-Iqbal* pleading standard. However, even accepting Plaintiffs’ insufficient allegations, it is clear that Bloomberg has not and could not have engaged in copyright infringement. Rather, the Amended Complaint demonstrates that Bloomberg’s alleged use of Plaintiffs’ works in connection with the BloombergGPT research project was a fair use, and thus is not copyright infringement.

A. Plaintiffs’ Copyright Claim Fails to State a Claim Upon Which Relief May Be Granted.

To establish a claim for copyright infringement, Plaintiffs must prove “(1) ownership of a valid copyright, and (2) copying of constituent elements of the work that are original.” *Abdin v. CBS Broad. Inc.*, 971 F.3d 57, 66 (2d Cir. 2020) (citation omitted). In addition, a plaintiff must allege “by what acts during what time the defendant infringed the copyright.” *Kelly v. L.L. Cool J.*, 145 F.R.D. 32, 36 (S.D.N.Y. 1992); *see also Hartmann v. Popcornflix.com LLC*, No. 20-cv-4923, 2023 WL 5715222, at *3 (S.D.N.Y. Sept. 5, 2023).

Plaintiffs' Amended Complaint fails to plausibly allege that Bloomberg directly infringed Plaintiffs' copyrights. Even after availing themselves of the opportunity to amend their complaint, Plaintiffs fail to provide any factual allegations, beyond mere "labels and conclusions" and a "formulaic recitation of the elements," that could establish a claim for copyright infringement. *Twombly*, 550 U.S. at 555. That is insufficient.

First, Plaintiffs' Amended Complaint consists largely of general allegations, not specific to BloombergGPT, regarding generative AI and the functioning of LLMs, which have nothing to do with alleged copyright infringement. (*See* Am. Compl. ¶¶ 24–42.) For example, Plaintiffs allege generally that "LLMs represent a significant advancement in the field of AI" and that they are "designed to understand and generate human language." (*Id.* ¶¶ 24–25.) Plaintiffs further offer their interpretations of how "LLMs work," including how "training data is used to teach the model grammar, vocabulary, context, and various language patterns," to accomplish the ultimate goal of "content creation" (*id.* ¶¶ 27–28), as well as the general advantages LLMs may provide in a corporate context (*id.* ¶ 29). Plaintiffs also discuss the publicly available "Books3" dataset, without reference to which, if any, of Plaintiffs' copyrighted works are included in the compilation. (*Id.* ¶¶ 31–42.) Absent from the Amended Complaint, however, is any allegation of how any of this applies to Bloomberg. Indeed, many of Plaintiffs' general allegations appear to relate to non-parties EleutherAI and OpenAI, which are not alleged to have any affiliation with Bloomberg at all.² (*Id.* at ¶¶ 34, 37, 41, 42, 55.)

² Much of this material made more sense in the context of the original complaint, but even then it did not relate to Bloomberg. (Dkt. No. 1.) Meta, Microsoft and EleutherAI were originally named as defendants in this action. Plaintiffs' claims against Meta and Microsoft have been severed and transferred to the Northern District of California. (Dkt. No. 69.) And Plaintiffs' claims against EleutherAI have been voluntarily dismissed. (Dkt. No. 68.)

Ultimately, Plaintiffs’ allegations specifically about BloombergGPT are limited to *twelve scant paragraphs* of their Amended Complaint.³ (*Id.* ¶¶ 43–54.) None alleges when or how BloombergGPT was developed, let alone whether or how such development involved the use of Plaintiffs’ copyrighted works. A plain reading of the Amended Complaint proves the point. Plaintiffs allege that on March 30, 2023, Bloomberg “announced” that it “began work on its own LLM,” which Bloomberg “built . . . for finance.” (*Id.* ¶¶ 43–44.) Based entirely on reporting about a research paper about BloombergGPT, Plaintiffs further allege that the model was developed “to support a diverse set of natural language processing (NLP) tasks within the financial industry” and that Bloomberg “touted its performance.”⁴ (*Id.* ¶¶ 45–47.) While Plaintiffs allege that Bloomberg “[u]s[ed] data from Books3 . . . to create its LLM” (*id.* ¶¶ 48–49), they also allege that Bloomberg has stated that it will “not include the Books3 dataset among the data sources used to train future versions of BloombergGPT” (*id.* ¶ 50).⁵ Plaintiffs then conclude, without

³ Plaintiffs did not amend their allegations with respect to “Defendants” between their original complaint—in which “Defendants” was defined to refer to Meta, Microsoft and EleutherAI in addition to Bloomberg—and the Amended Complaint. It is thus unsurprising that such allegations are not specific to BloombergGPT.

⁴ Plaintiffs’ allegation that “Bloomberg boasted that it had been ‘a trailblazer in its application of AI, Machine Learning, and NLP in finance’” (*id.* ¶ 46) is unrelated to BloombergGPT. The full sentence from the press release reads: “For more than a decade, Bloomberg has been a trailblazer in its application of AI, Machine Learning, and NLP in finance.” *Introducing BloombergGPT, Bloomberg’s 50-billion parameter large language model, purpose-built from scratch for finance*, BLOOMBERG PROFESSIONAL SERVICES (Mar. 30, 2023), <https://www.bloomberg.com/company/press/bloomberggpt-50-billion-parameter-llm-tuned-finance/> (cited by Am. Compl. ¶ 46 n.13).

⁵ Plaintiffs’ allegation that their “copyright-protected works have been baked into Bloomberg’s LLM, and all *subsequent versions*” (Am. Compl. ¶¶ 51–52 (emphasis added)) is both hypothetical and implausible. Indeed, Plaintiffs’ allegation appears to assume without basis that future research efforts by Bloomberg in the field of generative AI would involve nothing more than further training of BloombergGPT, rather than, e.g., completely new training of a completely new model. Such attempts to predict the future are inherently implausible and should be disregarded. *Shady Recs., Inc., v. Source Enters. Inc.*, 371 F. Supp. 2d 394, 397 (S.D.N.Y. 2005) (“It would, indeed, be highly inappropriate for this Court, whether now or after a jury verdict in

explanation, that “Bloomberg’s LLM is tainted with illicitly obtained, copyright-protected material” (*id.* ¶ 54), and that “Bloomberg did not compensate Plaintiffs for their copyrighted material while creating a lucrative product” (*id.* ¶ 53).

Plaintiffs’ entire copyright infringement claim, therefore, rests on the vague, conclusory allegation that Bloomberg, at some unspecified time and in some unspecified manner, used the Books3 database. (*Id.* ¶¶ 49, 73.) Nowhere do Plaintiffs allege any facts explaining “by what acts during what time” Bloomberg allegedly infringed their copyrights. And while Plaintiffs allege that Bloomberg has somehow “create[ed] a lucrative product” (*id.* ¶ 53), Plaintiffs do not actually allege that Bloomberg has in any way profited (directly or indirectly) from BloombergGPT. Indeed, Plaintiffs do not allege that BloombergGPT has been released, made available for use by others, or been used for any commercial purpose within Bloomberg. Plaintiffs’ allegations, accordingly, are insufficient to plausibly allege direct copyright infringement. *Lopez v. Bonanza.com, Inc.*, No. 17-cv-8493, 2019 WL 5199431, at *23 (S.D.N.Y. Sept. 30, 2019) (dismissing copyright infringement claim because bare allegation that defendant “copied, reproduced, exploited, advertised, marketed, sold and offered for sale clothing items” using plaintiff’s copyrighted design was mere legal conclusion); *Lindsay v. Wrecked & Abandoned Vessel R.M.S. Titanic*, No. 97-cv-9248, 1999 WL 816163, at *4 (S.D.N.Y. Oct. 13, 1999) (dismissing copyright claim comprised of “vague and conclusory allegations”).

Second, Plaintiffs, fail to allege **which** of their copyrighted works were actually included in Books3, nor do they allege that **all** works identified in their Amended Complaint were included in the dataset. Indeed, Plaintiffs allege only that Books3 contains “hundreds of thousands of books

defendants’ favor, to issue an advisory opinion about any particular hypothetical use of the material in the future”).

in that data set, *including books authored by Plaintiffs and Class Members.*”⁶ (*Id.* ¶ 38 (emphasis added); *see also id.* ¶ 10 (alleging that the datasets used to train Bloomberg’s “LLM were assembled from copyrighted works, *including copyrighted works of the Plaintiffs and Members of the Class.*”) (emphasis added).) But Plaintiffs’ failure to allege which of their copyrights were allegedly infringed is fatal to their claim, as Plaintiffs have failed to allege which of their works were actually copied. *See Cole v. John Wiley & Sons, Inc.*, No. 11-cv-2090, 2012 WL 3133520, at *13 (S.D.N.Y. Aug. 1, 2012) (copyright infringement claim concerning photographs insufficiently pleaded where plaintiffs offered only “vague and expansive allegations regarding which works are the subjects of Plaintiff’s claims”); *Bespaq Corp. v. Haoshen Trading Co.*, No. 04-cv-3698, 2005 WL 14841, at *2 (N.D. Cal. Jan. 3, 2005) (dismissing copyright infringement claims where plaintiff “[did] not identify which preexisting works in the registered catalog have been infringed by the defendants. Therefore, on the face of the complaint, the court cannot determine whether the registered catalog depicts any individual piece of miniature furniture at issue in this suit.”).

Simply put, even after availing themselves of the opportunity to amend their complaint, Plaintiffs fail to provide any factual allegations beyond mere “labels and conclusions” and the “formulaic recitation of the elements” that could establish a claim for copyright infringement. *Twombly*, 550 U.S. at 555. This is insufficient to withstand dismissal.

⁶ Plaintiffs’ pleading deficiency is not the result of a lack of publicly-available information. Plaintiffs readily could have investigated—and pleaded—which of their copyrighted works, if any, are included within the Books3 dataset. They did not do so.

Plaintiffs’ failure here is readily contrasted with the affirmative pleadings in other, recently-filed complaints challenging the training of generative AI models. For example, in *Kadrey et al. v. Meta Platforms, Inc.*, the plaintiffs attached as an exhibit to their complaint “[a] list of Plaintiffs’ books currently known to exist in the Books3 dataset” and specifically alleged Meta infringed the corresponding copyrights. No. 23-cv-03417, Am. Compl. (ECF No. 69) ¶ 39 (N.D. Cal. Dec. 22, 2023). Plaintiffs here make no similar identification.

B. Plaintiffs' Copyright Claim Is Barred by the Doctrine of Fair Use.

Where fair use is readily apparent, claims of copyright infringement are properly dismissed at the pleading stage pursuant to Rule 12(b)(6). *See, e.g., Brown v. Netflix, Inc.*, 462 F. Supp. 3d 453, 464 (S.D.N.Y. 2020), *aff'd*, 855 F. App'x 61 (2d Cir. 2021). It is well-settled that a court may evaluate fair use at the motion to dismiss stage “where the facts necessary to establish the defense are evident on the face of the complaint.” *Yang v. Mic Network, Inc.*, 405 F. Supp. 3d 537, 542 (S.D.N.Y. 2019) (granting motion to dismiss on fair use grounds); *Brody v. Fox Broad Co., LLC*, No. 22-cv-6249, 2023 WL 2758730, at *4 (S.D.N.Y. Apr. 3, 2023) (same); *Clark v. Transp. Alts., Inc.*, No. 18-cv-9985, 2019 WL 1448448, at *5 (S.D.N.Y. Mar. 18, 2019) (same); *Hughes v. Benjamin*, 437 F. Supp. 3d 382, 394 (S.D.N.Y. 2020) (same). Here, even Plaintiffs' sparse allegations demonstrate that the fair use doctrine applies and closes the book on Plaintiffs' claim.

i. The Fair Use Analysis

Congress's authority to provide for and enforce copyright protection stems from the Constitution's mandate that it shall “*promote the Progress of Science and useful Arts*, by securing for limited times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries.” (U.S. Const. art. I, § 8, cl. 8.) Accordingly, from the very “infancy of copyright protection, some opportunity for fair use of copyrighted materials has been thought necessary to fulfill copyright's very purpose” *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 575 (1994). Such principles are necessary to “balance between the interests of authors and inventors in the control and exploitation of their writings and discoveries on the one hand, and society's competing interest in the free flow of ideas, information, and commerce on the other hand” *Sony Corp. of Am. v. Universal City Studios, Inc.*, 464 U.S. 417, 429 (1984). Indeed, “[t]he fair use doctrine seeks to protect a secondary work if it ‘adds value to the original—if [copyrightable expression in the original work] is used as raw material, transformed in the creation of new

information, new aesthetics, new insights and understandings,’ because such a work contributes to the enrichment of society.” *Warner Bros. Ent. Inc. v. RDR Books*, 575 F. Supp. 2d 513, 541 (S.D.N.Y. 2008) (quoting *Castle Rock Entm’t., Inc. v. Carol Pub. Grp., Inc.*, 150 F.3d 132, 141 (2d Cir. 1998)). Following the Framers’ intent, Section 107 of the Copyright Act provides that “the fair use of a copyrighted work . . . for purposes such as criticism, comment, **news reporting, teaching . . . , scholarship, or research, is not an infringement of copyright.**” (emphasis added).

Section 107 sets forth four primary factors courts use to determine whether a particular use falls within the fair use framework: (1) the purpose and character of the use; (2) the nature of the copyrighted work; (3) the amount of the work used; and (4) the effect of the use on the primary market of the copyrighted work. 17 U.S.C. § 107; *Authors Guild v. Google, Inc.*, 804 F.3d 202, 212–13 (2d Cir. 2015) (“*Google Books*”). All four factors weigh in favor of a finding of fair use here.

ii. A Fair Use Analysis Confirms that Dismissal Is Warranted

Plaintiffs’ Amended Complaint alleges that Bloomberg, a provider of financial **news** and analytics, used Plaintiffs’ works to **teach** a generative AI model as part of a non-commercial **research** initiative. It is hard to imagine a clearer example of fair use.

a. Purpose and Character of Use

As an initial matter, Plaintiffs allege that BloombergGPT is the product of **internal research** conducted by Bloomberg researchers to test the viability of an LLM on tasks in a finance setting. (Am. Compl. ¶¶ 44–45.) Plaintiffs further allege that Bloomberg’s research involved limited use of the Books3 dataset, not to reproduce their works or to create competing works, but to “**teach** the model grammar, vocabulary, context, and various language patterns.” (*Id.* ¶ 27.a.) And Plaintiffs allege that Bloomberg is a “software, data and media company which provides a raft of media services and related products, including TV, internet and radio **news programming**,”

and a “global business and financial information and *news leader*” (*id.* ¶¶ 22–23).⁷ These alleged uses are *presumptively* of a “purpose and character” within the meaning of fair use. 17 U.S.C. § 107; *Wright v. Warner Books, Inc.*, 953 F.2d 731, 736 (2d Cir. 1991) (“[T]here is a strong presumption that factor one favors the defendant if the allegedly infringing work fits the description of uses described in § 107.”); *see also A.V. ex rel. Vanderhye v. iParadigms, LLC*, 562 F.3d 630, 638 (4th Cir. 2009) (“In assessing the ‘character’ of the use, we should consider the specific examples set forth in section 107’s preamble, ‘looking to whether the use is for criticism, or comment, or news reporting, and the like.’” (citation omitted)).

Diving more deeply, the presumption of fair use holds. Of particular significance, the Court must consider “whether [the alleged] use is of a commercial nature or is for nonprofit educational purposes.” 17 U.S.C. § 107; *Google Books*, 804 F.3d at 219. “Academic, non-commercial uses are more likely to be found to be fair uses than commercial uses.” *Wilder v. Hoiland*, No. 22-cv-1254, 2024 WL 382141, at *40, *42 (S.D.N.Y. Feb. 1, 2024) (“non-commercial, educational use that benefitted the public interest” constituted fair use).

Here, Plaintiffs do not allege that Bloomberg has made *any commercial use* of their copyrighted works. The word “commercial” does not appear even once in the Amended Complaint.⁸ Nor does the Amended Complaint include any allegation of any revenues Bloomberg

⁷ The Second Circuit has recognized that Bloomberg’s status as a reporter of financial information is significant to a fair use analysis concerning Bloomberg’s activities. *Swatch Grp. Mgmt. Servs. Ltd. v. Bloomberg L.P.*, 756 F.3d 73, 82 (2d Cir. 2014). The court in *Swatch Group* recognized that Bloomberg’s business of the “public dissemination of financial information” serves “the public purpose in the nature of news reporting,” which “lies at the core of the First Amendment [and] would be crippled if the news media and similar organizations were limited to sources of information that authorize disclosure.” *Id.* at 82–84.

⁸ In their most recent pre-motion letter to Judge Schofield, Plaintiffs repeatedly assert that BloombergGPT is somehow a “commercial product” (Dkt. No. 77 at 2 & n.1). But the Amended Complaint does not include any such allegation.

has realized (directly or indirectly) as a result of BloombergGPT.⁹ Nor does the Amended Complaint include any allegation that BloombergGPT has been made available to anyone else who has commercialized it—or even that it has been released outside Bloomberg at all.

What the Amended Complaint does allege, however, is that “***Bloomberg released a research paper.***” (Am. Compl. ¶ 45.) Not that Bloomberg released BloombergGPT, but that Bloomberg released a research paper. And although the Amended Complaint references Bloomberg’s research paper, Plaintiffs have studiously avoided citing to the paper itself. For good reason.

A review of Bloomberg’s research paper readily confirms Bloomberg’s sensitivity to “the ethics, uses, and limitations of these models,” and explains that ***Bloomberg has “not releas[ed] our model.”*** Shijie Wu, et al., “BloombergGPT: A Large Language Model for Finance,” at 37–38, available at <https://arxiv.org/pdf/2303.17564.pdf>. As Bloomberg explains: “[O]ur insights and experiences in training and evaluating BloombergGPT contribute to the developing understanding of these models.” *Id.* at 38. And the paper concludes with a brief summary of “***new research directions we hope to pursue*** with BloombergGPT.” *Id.* at 39. Accordingly, Plaintiffs are well aware that Bloomberg has not released BloombergGPT, that the model was developed for research purposes, and that the model has not been put to commercial use.¹⁰

⁹ Plaintiffs’ most recent pre-motion letter to Judge Schofield also refers to “significant revenue” and boldly asserts that Bloomberg has “ultimately reap[ed] millions of dollars” from BloombergGPT (*Id.*) But again, the Amended Complaint does not include any such allegations. The Amended Complaint’s only allegations regarding revenues relate to OpenAI, (*id.* ¶ 55), not Bloomberg.

¹⁰ The Court need not treat the research paper as incorporated by reference to reach the conclusion that Plaintiffs failed to plead—much less plausibly plead—commercial use. It is, however, significant that Plaintiffs had additional information available to them that they chose not to rely upon. Indeed, a link to the research paper appears at the bottom of the press release to which Plaintiffs cite. *Introducing BloombergGPT, Bloomberg’s 50-billion parameter large*

Regardless, even if Plaintiffs did allege that BloombergGPT was developed with a longer-term commercial motivation (and they do not),¹¹ it is well-settled that such motivation cannot “outweigh a convincing transformative purpose and absence of a significant substitutive competition with the original.” *Google Books*, 804 F.3d at 219; *iParadigms*, 562 F.3d at 639 (affirming that “the commercial aspect [of the defendant’s use] was not significant in light of the transformative nature.”); *White v. W. Pub. Corp.*, 29 F. Supp. 3d 396, 399 (S.D.N.Y. 2014) (holding that defendants’ copying, storage, and distribution of legal briefs to create “an interactive legal research tool” was fair use, despite commercial purpose).

“A use is transformative if it does something more than repackage or republish the original copyrighted work,” *Authors Guild, Inc. v. HathiTrust*, 755 F.3d 87, 96 (2d Cir. 2014), and, importantly, “[t]he more transformative the new work, the less will be the significance of other factors . . . that may weigh against a finding of fair use” *id.* (quoting *Campbell*, 510 U.S. at 579).

Here, Plaintiffs allege that Bloomberg’s use of the Books3 dataset was to “*teach* the model grammar, vocabulary, context, and various language patterns.” (Am. Compl. ¶ 27.a.) This allegation, taken as true, confirms that Bloomberg’s use of Plaintiffs’ works was not intended to copy their expression, but rather to gain the benefit of their *unprotectable* facts and ideas, such as grammar and vocabulary. It is well-settled that it is not an act of copyright infringement to extract and use information, ideas, or concepts from otherwise copyrighted works. *See* 17 U.S.C. §

language model, purpose-built from scratch for finance, BLOOMBERG PROFESSIONAL SERVICES (Mar. 30, 2023), <https://www.bloomberg.com/company/press/bloomberggpt-50-billion-parameter-llm-tuned-finance/> (cited by Am. Compl. ¶ 46 n.13). Plaintiffs’ pleading failure is of their own making.

¹¹ Any such allegation, if made, would necessarily be speculative in nature, rather than a plausible allegation of fact. And any request that a court act on such speculation would be a request for an advisory opinion. Courts are not in the business of issuing advisory opinions. *Shady Records, Inc.*, 371 F. Supp. 2d at 397.

102(b); *Feist Publications, Inc. v. Rural Tel. Serv. Co.*, 499 U.S. 340, 362 (1991) (noting that “basic information” such as name, town and telephone number are not copyrightable); *Nat’l Basketball Ass’n v. Motorola, Inc.*, 105 F.3d 841, 847 (2d Cir. 1997) (finding that facts regarding basketball game were not copyrightable).

The result of this research—BloombergGPT—is an artificial intelligence model; more specifically, it is an LLM that “support[s] a diverse set of natural language processing (NLP) tasks within the financial industry.” (Am. Compl. ¶ 45.) Plaintiffs’ copyrighted works, on the other hand, are various books intended to inform, persuade and/or entertain readers. (*Id.* ¶¶ 16–21.) And, unlike allegations presented in other lawsuits regarding generative AI technologies,¹² Plaintiffs do not allege that BloombergGPT makes their works (or any portions thereof) available to the public.¹³ The purpose and character of BloombergGPT is therefore vastly different from that of Plaintiffs’ copyrighted works.

Indeed, the purpose and character of Bloomberg’s alleged use here more readily falls within the scope of fair use than other uses that have been deemed fair. For example, in *A.V. v. iParadigms, LLC*, the Fourth Circuit determined that the copying and archiving “as digital code” of students’ essays—in their entirety—for inclusion in a database used to evaluate the originality

¹² See *infra* pp. 20–21.

¹³ Even had Plaintiffs so-alleged, Bloomberg’s use would still be transformative. Courts have repeatedly upheld findings of fair use, based on findings that uses were “transformative,” even where (unlike here) the relevant protected materials are reproduced and made available to the public. See, e.g., *Kelly v. Arriba Soft Corp.*, 336 F.3d 811, 818 (9th Cir. 2003) (finding transformative use where “search engine functions as a tool to help index and improve access to images on the internet and their related web sites.”); *Perfect 10, Inc. v. Amazon.com, Inc.*, 508 F.3d 1146, 1165 (9th Cir. 2007) (“Google’s use of thumbnails is highly transformative” where “a search engine transforms the image into a pointer directing a user to a source of information.”); *Apple Inc. v. Corellium, Inc.*, No. 21-12835, 2023 WL 3295671, at *6 (11th Cir. May 8, 2023) (holding that despite copying, storing, and displaying Apple’s iOS software, defendants’ use was transformative because its end product had a “different character” and “help[ed] security researchers do their work in a way that physical iPhones just can’t.”).

of other students' works was a "highly transformative" use, because the database "use[d] the papers for an entirely different purpose, namely, to prevent plagiarism" 562 F.3d at 638 (finding iParadigms' use of students' essays for "Turnitin" system fair use). Here, similarly, BloombergGPT's alleged use of Plaintiffs' works during the training process was for an entirely different purpose than that served by the works themselves. And unlike in *iParadigms*, Plaintiffs do not allege that Bloomberg's use even extended as far as archiving their works.

More recently, in *Google Books*, the Second Circuit held that Google's "making of a digital copy of Plaintiffs' books for the purpose of enabling a search for identification of books containing a term of interest to the searcher involves a highly transformative purpose," even where such search function "allows searchers to read snippets from the book searched" 804 F.3d at 216–17. And in *HathiTrust*, the Second Circuit determined that the creation of digital copies of entire books to create a "full-text searchable database is a quintessentially transformative use," as there is "little or no resemblance between the original text and the results of the HDL full-text search." 755 F.3d at 97. Here, again, Plaintiffs do not allege that Bloomberg's use even extended as far as creating a searchable database of their works, let alone making snippets of those works available to others.

Simply put, BloombergGPT does not "supersede[] the objects [or purposes] of [Plaintiffs'] original creation," nor does it "repackage" or "republish" Plaintiffs' works. *Id.* (citing *Campbell*, 510 U.S. at 579); *see also Kelly v. Arriba Soft Corp.*, 336 F.3d 811, 820 (9th Cir. 2003) (defendant's use "promotes the goals of the Copyright Act and the fair use exception" because it "do[es] not supplant the need for the originals" and "benefit[s] the public by enhancing information-gathering techniques on the internet."). The first fair use factor thus unequivocally weighs in favor of a finding of fair use. *Wilder*, 2024 WL 382141, at *17 (granting summary judgment where an allegedly infringing presentation "was transformative and . . . the educational,

non-commercial purpose for which [the presentation] used the [original] work weigh[ed] heavily in favor of fair use.”).

b. Nature of the Copyrighted Work

The “nature of the copyrighted work” factor “considers whether the copyrighted work ‘is of the creative or instructive type that the copyright laws value and seek to foster.’” *HathiTrust*, 755 F.3d at 96 (citation omitted). Courts evaluating fair use arguments based on diverse collections of copyrighted works have acknowledged that the nature of such works is not—and indeed rarely is ever—significant. *Google Books*, 804 F.3d at 220; *HathiTrust*, 755 F.3d at 98. To the extent the nature of Plaintiffs’ copyrighted works has any influence on the fair use analysis, it favors a finding of fair use, given that Plaintiffs explicitly allege that BloombergGPT “has an objective that differs from [Plaintiffs’] original” works, which are books, not artificial intelligence models. *Google Books*, 804 F.3d at 220; Am. Compl. ¶¶ 28, 45. Here, like in *Google Books*, the “nature” of Plaintiffs’ works necessarily combines with the “purpose and character” of BloombergGPT to support a finding that Bloomberg’s “use” is transformative and does not “replicat[e] protected expression in a manner that provides a meaningful substitute for the original.” 804 F.3d at 220.

c. Amount of Work Used

With respect to the third factor, the focus is not on “the amount and substantiality of the portion used’ *in making a copy*, but rather the amount and substantiality of *what is thereby made accessible* to a public for which it may serve as a compelling substitute.” *Google Books*, 804 F.3d at 222. Here, unlike in *Google* and *HathiTrust*, Plaintiffs do not allege that Bloomberg made any copies of their works (*or even any portions thereof*) available to the public. Rather, Plaintiffs merely allege that Bloomberg used the Books3 dataset “to assist its LLM in learning how to recognize, parse, and respond in natural language” (Am. Compl. ¶ 48), and that “[t]he training data

is used to teach the model grammar, vocabulary, context, and various language patterns” (*id.* ¶ 27.a). But grammar and natural language themselves are not eligible for copyright protection, nor are they a substitute for Plaintiffs’ works. Accordingly, the concerns underlying the substantiality factor are eliminated altogether. That should end the inquiry.

And even if Plaintiffs had pleaded public disclosure (they did not), the focus would be on “whether the copying used more of the copyrighted work than necessary. . . .” *HathiTrust*, 755 F.3d at 98. Even when the entirety of a protected work is copied, the “substantiality” factor does not necessarily weigh against a finding of fair use. *Id.* Rather, the Supreme Court has held that “[t]he extent of permissible copying varies with the purpose and character of the use.” *Campbell*, 510 U.S. at 586–87 (substantiality analysis depends on the “degree to which the [copying work] may serve as a market substitute for the original or potentially licensed derivatives.”); *see also Kelly*, 336 F.3d at 821 (“[A]lthough Arriba did copy each of Kelly’s images as a whole, it was reasonable to do so in light of Arriba’s use of the images.”).

Here, Plaintiffs themselves allege that “vast” datasets are “*necessary* to effectively train LLMs.”¹⁴ (Am. Compl. ¶ 32.) Accordingly, this factor does not weigh against a finding of fair use. *See HathiTrust*, 755 F.3d at 99; *Google Books*, 804 F.3d at 221–23; *see also iParadigms*, 562 F.3d at 642 (finding fair use where “although iParadigms uses substantially the whole of plaintiffs’ works, iParadigms’ ‘use of the original works is limited in purpose and scope’ as a digitized record for electronic ‘comparison purposes only’”).

¹⁴ And, as explained above (*see supra* pp. 14–15), Plaintiffs allege that the “data” that is actually used to train LLMs is unprotectable facts and ideas such as “grammar, vocabulary, context, and [] language patterns,” rather than Plaintiffs’ protectible expression. (Am. Compl. ¶ 27.)

d. *Effect of Use on the Primary Market for the Work*

“Because copyright is a commercial doctrine whose objective is to stimulate creativity among potential authors by enabling them to earn money from their creations, the [market] factor is of great importance in making a fair use assessment.” *Google Books*, 804 F.3d at 223 (citing *Harper & Row Publishers, Inc. v. Nation Enterprises*, 471 U.S. 539, 566 (1985)). Bloomberg’s experimental AI model has no effect on the market for Plaintiffs’ books, as confirmed by both what is and what is not alleged in the Amended Complaint.

The “market” factor “focuses on whether the copy brings to the marketplace a competing substitute for the original, or its derivative, so as to deprive the rights holder of significant revenues because of the likelihood that potential purchasers may opt to acquire the copy in preference to the original.” *Google Books*, 804 F.3d at 223. The market factor, accordingly, is concerned only with “harm that results because the secondary use serves as a substitute for the original work,” rather than economic harm generally. *HathiTrust*, 755 F.3d at 99. It follows that there is a close relationship between the first and fourth fair use factors, as the more “transformative” a copy is, the less likely it is to usurp the market for the original. *See Campbell*, 510 U.S. at 591; *see also Apple*, 2023 WL 3295671, at *12 (“[T]he Copyright Act doesn’t afford creators a monopoly over transformative markets.”).

Further, there must be a concrete, meaningful effect on the potential market for the copyrighted work to weigh against a finding of fair use. “The possibility, or even the probability or certainty, of some loss of sales does not suffice to make the copy an effectively competing substitute that would tilt the weighty fourth factor in favor of the rights holder in the original.” *Google Books*, 804 F.3d at 224; *see also Wilder*, 2024 WL 382141, at *19–21. Moreover, “a copyright holder cannot prevent others from entering fair use markets merely ‘by developing or licensing a market for parody, news reporting, educational or other transformative uses of its own

creative work.” *Bill Graham Archives v. Dorling Kindersley Ltd.*, 448 F.3d 605, 614–15 (2d Cir. 2006) (citation omitted); *see also Cambridge Univ. Press v. Patton*, 769 F.3d 1232, 1278 (11th Cir. 2014) (“Plaintiffs may not head off a defense of fair use by complaining that every potential licensing opportunity represents a potential market for purposes of the fourth fair use factor”).

Here, Plaintiffs have not plausibly alleged any economic harm at all, let alone any harm resulting from market competition with a substitute for their works. As an initial matter, Plaintiffs still fail to allege that Bloomberg has released BloombergGPT (or even a single output from BloombergGPT) to the public—BloombergGPT cannot displace a market for published books without being publicly available. (Am. Compl. ¶¶ 43–54.) This alone dictates that the market factor favors a finding of fair use. In *Wilder*, Judge Castel determined that a presentation delivered in a “non-commercial, educational” setting was a fair use of a copyrighted text the presentation incorporated, where “[t]here [was] no record evidence that the Presentation was actually downloaded or viewed, nor that it was disseminated outside the conference attendees,” and where “[t]he text was shown only briefly, and the total number of people who attended the Presentation itself was no greater than twenty.” 2024 WL 382141, at *15, *21; *see also Blanch v. Koons*, 467 F.3d 244, 258 (2d Cir. 2006) (finding artwork’s use of photograph fair where photographer acknowledged that defendant’s use “did not cause any harm to her career or upset any plans she had for [her photograph],” and “the value of [her photograph] did not decrease as the result of [defendant’s] alleged infringement.”). Bloomberg’s use here was even more circumscribed.

Instead, Plaintiffs allege that “Bloomberg did not compensate Plaintiffs for their copyrighted material while creating a lucrative project.” (Am. Comp. ¶ 53.) But the allegation that BloombergGPT is “lucrative” is not only entirely speculative, it cannot plausibly support a conclusion that BloombergGPT harmed the market for Plaintiffs’ works. Plaintiffs also assert that

“Defendants have illicitly gained an enormous amount of value from their unauthorized use of Plaintiffs’ copyright-protected works” and “profited enormously.” (*Id.* ¶¶ 57, 60.) Again, neither allegation has anything to do with harm to the market for Plaintiffs’ copyrighted works. And to the extent Plaintiffs argue that Bloomberg has harmed the market for their works simply by not paying them licensing fees, that is insufficient to tilt the market factor in Plaintiffs’ favor, as it is well-settled that Plaintiffs must demonstrate “impairment to a traditional, *as opposed to a transformative market.*” *Bill Graham Archives*, 448 F.3d at 614 (emphasis added). In other words, Plaintiffs may not avoid a fair use finding merely by arguing that Bloomberg should have compensated Plaintiffs for its transformative use of Plaintiffs’ works.

And even if BloombergGPT had been released, it would not displace Plaintiffs’ works in the market: Plaintiffs allege that Bloomberg is a “financial, software, data and media company which provides . . . financial services” (Am. Compl. ¶ 22), and that BloombergGPT was built for “tasks within the financial industry” (*id.* ¶ 45), but Plaintiffs allege they are authors of books (*id.* ¶¶ 16–21), not architects of artificial intelligence systems. Since BloombergGPT and Plaintiffs operate in totally different spaces, BloombergGPT has done nothing to impact the primary market for Plaintiffs’ works. *See Cariou v. Prince*, 714 F.3d 694, 709 (2d Cir. 2013) (an accused infringer “usurp[s]” an existing market only “where the infringer’s target audience and the nature of the infringing content is the same as the original.”); *Google LLC v. Oracle Am., Inc.*, 141 S. Ct. 1183, 1206–07 (2021) (finding “market effects” weighed in favor of finding of fair use where “devices using Google’s Android platform were different in kind from those that licensed Sun’s technology.”); *White*, 29 F.Supp.3d at 400 (legal research database comprised of copied briefs “is in no way economically a substitute for the use of briefs in their original market”). Notably, Plaintiffs’ allegations in this case are quite different from other plaintiffs’ allegations in other cases

addressing generative AI technologies. For example, in *The New York Times Co. v. Microsoft Co. et al.*, The New York Times alleges that “[the accused] GenAI tools can generate output that recites Times content verbatim . . . as demonstrated by scores of examples.” No. 23-cv-11195, Compl. (ECF No. 1) ¶ 4 (S.D.N.Y. Dec. 27, 2023). And in *Concord Music Group, Inc. et al. v. Anthropic PBC*, the plaintiffs allege that “Anthropic’s AI models generate identical or nearly identical copies of [plaintiffs’ copyrighted] lyrics.” No. 23-cv-1092, Compl. (ECF No. 1) ¶ 8 (M.D. Tenn. Oct. 18, 2023). By contrast, here, there is no allegation that BloombergGPT in any way makes available to the public the contents of Plaintiffs’ copyrighted works. Simply put, there is not one single person who has not bought the Plaintiffs’ books because of Bloomberg’s LLM research project; and Plaintiffs have not pleaded otherwise.

Finally, in evaluating its effect on the relevant market, the Court should also consider the “public benefits” associated with the use of the protected work. *Google LLC*, 141 S.Ct. at 1206. Bloomberg developed the “world’s first LLM ‘built from scratch for finance’” (Am. Compl. ¶ 44) and published a research paper explaining the benefits that can be achieved by training an AI model on both general and industry-specific materials (*id.* ¶ 45), thereby sharing its research with the public and allowing it to be considered and expanded upon. The public benefit of Bloomberg’s research efforts undeniably weighs in favor of a finding of fair use.

* * *

Ultimately, Plaintiffs fail to present allegations that would be sufficient to support a finding of anything other than fair use of their copyright works. Such use does not constitute copyright infringement within the meaning of the Copyright Act or established precedent. Dismissal is therefore warranted.

C. Plaintiffs' Amended Complaint Should Be Dismissed with Prejudice.

Plaintiffs filed their original complaint on October 17, 2023, advancing six causes of action against Bloomberg, including for direct copyright infringement, vicarious copyright infringement, violation of the Digital Millennium Copyright Act, conversion, negligence, and unjust enrichment. (Dkt. No. 1.) On December 15, 2023, Bloomberg filed a letter identifying the deficiencies in Plaintiffs' original complaint and expressing Bloomberg's intent to seek dismissal of all causes of action asserted against it. (Dkt. No. 54, "First Dismissal Letter.") Plaintiffs responded by offering to amend their complaint. (Dkt. No. 66.) At the Court's invitation, and in response to Bloomberg's First Dismissal Letter, Plaintiffs amended their complaint, voluntarily dropping five causes of action and leaving just one claim, for direct copyright infringement. (Dkt. Nos. 74, 75.)

As set forth above, Plaintiffs' Amended Complaint, however, does not cure either the pleading deficiencies or the fair-use issue Plaintiffs face, and no further amendment will do so. On that basis, the Court should dismiss this action with prejudice.

Under Federal Rule of Civil Procedure 15(a), "leave to amend is properly denied if to do so would be futile." *Wheeler v. Topps Co.*, 652 F. Supp. 3d 426, 436 (S.D.N.Y. 2023). Where a party is given a chance to amend a complaint but does not do so, leave to amend may be denied and dismissal with prejudice may be granted. *United States ex rel. Levine v. Vascular Access Ctrs., L.P.*, No. 12-cv-5103, 2020 WL 5534670, at *8 (S.D.N.Y. Sept. 15, 2020); *see also Bellikoff v. Eaton Vance Corp.*, 481 F.3d 110, 118 (2d Cir. 2007). Indeed, courts in this district have granted a defendant's motion to dismiss with prejudice after the plaintiff failed to correct deficiencies despite having previous opportunity to amend. *Wheeler*, 652 F. Supp. 3d at 436.

Plaintiffs, on notice of the deficiencies in their original pleading, amended their complaint, yet failed to incorporate any new allegations that could overcome Bloomberg's fair use defense. In fact, shockingly, Plaintiffs did not add *any* factual allegations to their Amended Complaint,

electing instead only to remove factual allegations and to make a handful of minor editorial changes. (*See* Dkt. No. 75-1.) This suggests Plaintiffs have nothing further to say, and even if they did, no amount of artful repleading could remedy the “fundamental infirmity” and fair-use buzz saw that would cut off Plaintiffs’ copyright infringement claim, particularly given the transformative nature of Bloomberg’s use of Plaintiffs’ works and the lack of any potential harm to the market for Plaintiffs’ works. *Kaye v. Cartoon Network Inc.*, 297 F. Supp. 3d 362, 371 (S.D.N.Y. 2017).

In light of the fact that Plaintiffs have already had a chance to cure, and in light of the fact that any further amendment would be futile as the fair-use reality remains, dismissal should be with prejudice.

V. CONCLUSION

Simply put, a limited and private use of copyrighted works by a news reporting enterprise to teach a not-for-commercial-use AI model as part of an internal research project into the capabilities of generative AI, is not copyright infringement. For all of the foregoing reasons, Plaintiffs’ First Amended Complaint should be dismissed in its entirety.

DATED: March 22, 2024

Respectfully submitted,

FRIED, FRANK, HARRIS, SHRIVER
& JACOBSON LLP

By: /s/ Nicole M. Jantzi

Nicole M. Jantzi
(nicole.jantzi@friedfrank.com)
Paul M. Schoenhard (*Pro Hac Vice*)
(paul.schoenhard@friedfrank.com)
FRIED, FRANK, HARRIS, SHRIVER
& JACOBSON LLP
801 17th Street NW
Washington, DC 20006
Telephone: (202) 639-7265

Amir R. Ghavi
(amir.ghavi@friedfrank.com)
FRIED, FRANK, HARRIS, SHRIVER
& JACOBSON LLP
One New York Plaza
New York, New York 10004
Telephone: (212) 859-8000

*Attorneys for Defendants Bloomberg L.P.
and Bloomberg Finance L.P.*