UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK

MIKE HUCKABEE, *et al.*, on behalf of themselves
and all others similarly situated,

Plaintiffs,

-against-

BLOOMBERG L.P. and BLOOMBERG FINANCE,
L.P.,

Defendants.

1:23-CV-09152 (MMG)

OPINION & ORDER

MARGARET M. GARNETT, United States District Judge:

INTRODUCTION

This is one of many cases percolating through the federal courts concerning the

intersection of artificial intelligence ("AI") and copyright law. Plaintiffs are a proposed class

that alleges that Defendants Bloomberg, L.P. and Bloomberg Finance, L.P. (together,

"Bloomberg") are developing an AI program using Plaintiffs' copyrighted works, without their

permission or compensation. Currently before the Court is Bloomberg's motion to dismiss. For

the reasons discussed below, the motion is DENIED in full.

BACKGROUND

I.   RELEVANT FACTS[1]

Plaintiffs are a proposed class of individuals or corporations from across the United

States with at least two things in common. First, they hold or were licensed a copyright in at

---

[1] The following facts are taken from the first amended complaint (referred to as the "Amended
Complaint" in text and "FAC" in citations). *See* Dkt. No. 74. The Court shall refer to the parties'
memoranda of law in support of and opposition to the motion to dismiss as follows: Dkt. No. 87 ("Mot.");
Dkt. No. 90 ("Opp."); and Dkt. No. 91 ("Reply").

1

least one written work. FAC ¶¶ 16–21. And second, their copyrighted works were uploaded

without authorization onto a dataset called "Books3," itself part of a larger dataset called "The

Pile." *Id.* ¶¶ 5–7, 38. Defendant Bloomberg is a well-known media company specializing in

financial, regulatory, and political coverage. *Id.* ¶¶ 22–23. Bloomberg is developing an AI

program called BloombergGPT. BloombergGPT is a large language model ("LLM") AI

program designed specifically "for finance." *Id.* ¶ 44. To develop a sophisticated LLM AI

program, a developer must provide the program a massive trove of text data. *See id.* ¶ 27(a)–(h).

In Bloomberg's case, per a research paper that Bloomberg published, one of the data sources it

used to train BloombergGPT was Books3. *Id.* ¶¶ 47–50, ¶ 48 n.17. Bloomberg's use of the

copyrighted works within Books3 (the "Works"), without authorization, forms the basis for this

lawsuit.

II.   PROCEDURAL HISTORY

Plaintiffs filed this lawsuit on October 17, 2023, naming multiple defendants besides

Bloomberg and bringing claims for copyright infringement and violation of the Digital

Millennium Copyright Act, as well as state-law tort claims. *See* Dkt. No. 1. After dismissing

one defendant, severing claims against two defendants, and dropping certain claims, Plaintiffs

filed the Amended Complaint on January 24, 2024. *See* Dkt. Nos. 58–60, 68–69, & 74. It names

as Defendants Bloomberg, L.P. and Bloomberg Finance, L.P. and alleges a single count of direct

copyright infringement. FAC ¶¶ 22–23, 70–78.

The Amended Complaint identifies several copyrighted works of the named Plaintiffs, *id.*

¶¶ 16–21, alleges that Plaintiffs' copyrighted works were in Books3, *id.* ¶ 5–7, 38, and states that

Bloomberg used Books3 to train BloombergGPT, *id.* ¶¶ 47–50. As support, the Amended

Complaint references a research paper Bloomberg released that states Bloomberg used Books3 to

train BloombergGPT. *Id.* ¶ 48 n.17. The Amended Complaint alleges that developing an LLM

AI program takes massive amounts of text. *See id.* ¶ 27(a)–(h). "Using data from Books3," the

Amended Complaint alleges, "enabled Bloomberg to create its LLM faster and more efficiently,

and save the company a substantial amount of time, resources, and money because it would not

have to pay for source material." *Id.* ¶ 49. The Amended Complaint also alleges that companies,

including Bloomberg, will derive enormous profits from AI programs, but that authors like

Plaintiffs remain uncompensated for the use of their creative work to train LLM AI programs.

See id. ¶¶ 56–57, 59.

Bloomberg moved to dismiss the action in its entirety, arguing Plaintiffs failed to state a

claim and that the defense of fair use applied. *See* Dkt. Nos. 86 & 87.

DISCUSSION

Bloomberg's motion to dismiss rests primarily on two arguments. First, that the

Amended Complaint fails to state a claim because it does not identify precisely which Works

were within Books3. Second, that the fair use defense applies. Neither argument is meritorious

at the motion to dismiss stage, and accordingly the motion is DENIED.

I.   LEGAL STANDARD

To survive a motion to dismiss under Rule 12(b)(6), 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). [2] A claim is facially plausible "when the plaintiff pleads factual content that

allows the court to draw the reasonable inference that the defendant is liable for the misconduct

alleged." *Ashcroft v. Iqbal*, 556 U.S. 662, 678 (2009). A complaint is properly dismissed where,

---

[2] Unless otherwise indicated, case quotations omit all internal citations, quotation marks, footnotes and omissions, and adopt alterations.

as a matter of law, "the allegations in a complaint, however true, could not raise a claim of

entitlement to relief." *Twombly*, 550 U.S. at 558. The Court must assume all well-pled facts to

be true, "drawing all reasonable inferences in favor of the plaintiff." *Koch v. Christie's Int'l

PLC*, 699 F.3d 141, 145 (2d Cir. 2012); *see also A.I. Trade Fin., Inc. v. Petra Bank*, 989 F.2d 76,

79–80 (2d Cir. 1993) ("[A]ll allegations are construed in the light most favorable to the plaintiff

and doubts are resolved in the plaintiff's favor, notwithstanding a controverting presentation by

the moving party."). However, the Court need not accept conclusory assertions. *Whiteside v.

Hover-Davis, Inc.*, 995 F.3d 315, 321 (2d Cir. 2021).

Finally, on a Rule 12 (b)(6) motion, a court "may review only a narrow universe of

materials." *Goel v. Bunge, Ltd.*, 820 F.3d 554, 559 (2d Cir. 2016). That universe includes "the

facts alleged in the complaint, documents attached to the complaint as exhibits, and documents

incorporated by reference in the complaint." *United States ex rel. Foreman v. AECOM*, 19 F.4th

85, 106 (2d Cir. 2021).

II.   **THE FIRST AMENDED COMPLAINT PLAUSIBLY STATES
        A CLAIM OF COPYRIGHT INFRINGEMENT**

Bloomberg first argues that Plaintiffs fail to allege "by what acts during what time"

Bloomberg committed the alleged infringement and, accordingly, that the Amended Complaint

fails to state a claim. Mot. at 7. The Court has little trouble rejecting this argument. Federal

Rule of Civil Procedure 8 requires "a short and plain statement of the claim showing that the

pleader is entitled to relief." The rule is intended to ensure that a complaint gives fair notice of a

claim, *see Swierkiewicz v. Sorema N.A.*, 534 U.S. 506, 512–13 (2002); it does not require every

factual detail of every allegation. Nor does it require a court to ignore common sense. The

Amended Complaint (1) alleges that Books3 contains Plaintiffs' copyrighted works; (2) cites a

paper, released by Bloomberg, that stated Bloomberg used the data in Books3 to train

BloombergGPT; and (3) states that this use occurred without Plaintiffs' authorization or permission. FAC ¶¶ 5–8, 47–50. This provides fair notice and a plausible allegation of the conduct Plaintiffs contend was copyright infringement.

Bloomberg also argues that Plaintiffs never specifically identify which of their copyrighted works were in Books3, and that this purported omission justifies dismissal. Mot. at 5. The Court rejects this argument for a similar reason. Plaintiffs allege they hold copyrights in several specifically named works. FAC ¶¶ 16–21. They allege that the Books3 data set includes their copyrighted works. *Id.* ¶¶ 38–39, 41. And they allege that Bloomberg used the materials in Books3 to train BloombergGPT without permission from or compensation to Plaintiffs. At this stage of the case, "all allegations are construed in the light most favorable to the plaintiff and doubts are resolved in the plaintiff's favor, notwithstanding a controverting presentation by the moving party." *A.I. Trade Fin., Inc.*, 989 F.2d at 79–80. The clear and logical inference from Plaintiffs' allegations, viewed as a whole, is that one or more of the specific copyrighted works attributed to each named Plaintiff in the Amended Complaint were included in Books3. That is sufficient to survive a motion to dismiss.

III.   BLOOMBERG HAS NOT ESTABLISHED IT IS ENTITLED TO A FAIR USE DEFENSE ON THE FACE OF THE AMENDED COMPLAINT

Bloomberg's second and final argument is that its use of the Works constituted "fair use" under copyright law and thus Plaintiffs' claim of direct infringement must fail. Mot. at 9–21.

Fair use is an affirmative defense to a claim of copyright infringement. *Andy Warhol Foundation for the Visual Arts, Inc. v. Goldsmith* ("*Warhol*"), 598 U.S. 508, 537 n.21 (2023). By statute, a court must consider at least four factors in determining whether the defense applies: (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 whole"; and

(4) "the effect of the use upon the potential market for or value of the copyrighted work." 17

U.S.C. § 107(1)–(4). The Second Circuit has acknowledged the "possibility" that a Court may

decide a fair use defense on the face of a complaint. *TCA Television Corp. v. McCollum*, 839

F.3d 168, 178 (2d Cir. 2016). But courts "most frequently" address the defense at summary

judgment, and a defendant's entitlement to a fair use defense must be "clearly established on the

face" of a complaint "to support dismissal." *Id.* Bloomberg's claimed entitlement to a fair use

defense does not meet this high standard, in large part because the Court needs a robust factual

record to conduct the fair use analysis in this case.

Consider, as just one example, the fourth fair use factor—"the effect of the use upon the

potential market for or value of the copyrighted work," and how it applies to this case. 17 U.S.C.

§ 107(4). The Supreme Court has described the fourth factor as the "single most important

element of fair use." *Harper & Row Publishers, Inc. v. Nation Enters.*, 471 U.S. 539, 566

(1985). To apply this factor, a court must examine "not only the extent of market harm caused

by the particular actions of the alleged infringer, but also whether unrestricted and widespread

conduct of the sort engaged in by the defendant would result in a substantially adverse impact on

the potential market for the original." *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 592

(1994). In this case, analyzing this factor would conceivably require the Court to address at least

the following questions: Will widespread copying of literary works to train LLM AI programs

result in a deluge of AI-produced literary works, adversely affecting the market for the Works?

Because LLM AI programs have the potential to yield massive profits but require significant

amounts of text to train, is there a derivative market for literary works used to train AI?[3] And, if

---

[3] Second Circuit precedent has offered conflicting accounts of whether district courts may
consider damages to a potential licensing market. *Compare TCA Television Corp.*, 839 F.3d at 186 ("A

so, does unlicensed copying of literary works detrimentally harm an author's ability to participate in and profit from that derivative market? Answering these questions, among many others, could be vital to the Court's fair use determination. Just these few examples illustrate the futility of conducting a fair use analysis on the face of a complaint in this complex and evolving area of both law and technology, without a factual record. The other fair use factors present similar problems.[4] Accordingly, Bloomberg's motion to dismiss must be denied.
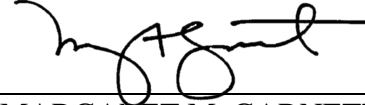
CONCLUSION

For the foregoing reasons, Bloomberg's motion to dismiss is DENIED. A separate Order will issue setting a conference date to discuss a plan for discovery.

The Clerk of Court is respectfully directed to terminate Dkt. No. 86.

Dated: November 24, 2025
   New York, New York

SO ORDERED.

_____
   MARGARET M. GARNETT
   United States District Judge

---

court considering fair use properly identifies and weighs relevant harm to the derivative market for a copyrighted work, which market includes uses that creators of original works might license others to develop.'"), *with Authors Guild, Inc. v. HathiTrust*, 755 F.3d 87, 100 (2d Cir. 2014) ("Lost licensing revenue counts under Factor Four only when the use serves as a substitute for the original . . . ."). Nevertheless, even if the Court could not consider a potential derivative market for literary works to train LLM AI programs, the Court would still be unable to determine if the fair use defense applied to this case absent a more robust evidentiary record.

[4] For example, in examining the first factor, the Court must assess whether the allegedly infringing use was commercial in nature. *See Warhol*, 598 U.S. at 536–37. The Court lacks a sufficient factual record to make that determination. And, drawing all inferences in Plaintiffs' favor, the Amended Complaint suggests that Bloomberg's development of BloombergGPT was for a commercial purpose. Bloomberg is a finance news company. It allegedly built BloombergGPT as the world's first LLM AI program "for finance." FAC ¶ 44. The plausible inference is that BloombergGPT will assist Bloomberg in conducting financial reporting, the core of Bloomberg's line of business, or developing new information products to sell to its financial industry customers.