

United States District Court
Northern District of California

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

UNITED STATES DISTRICT COURT
NORTHERN DISTRICT OF CALIFORNIA

ABDI NAZEMIAN, et al.,
Plaintiffs,
v.
NVIDIA CORPORATION,
Defendant.

Case No. 24-cv-01454-JST

**ORDER GRANTING IN PART AND
DENYING IN PART DEFENDANT'S
MOTION TO DISMISS**

Re: ECF No. 240

Before the Court is Defendant NVIDIA Corporation’s (“NVIDIA”) motion to dismiss. ECF No. 240. The Court will grant the motion in part and deny it in part.

I. BACKGROUND

Authors Abdi Nazemian, Brian Keene, and Stewart O’Nan filed this proposed class action against NVIDIA alleging that it has trained Large Language Models (LLM) on unauthorized copies of Plaintiffs copyrighted books. ECF No. 235 (“FCAC”) ¶¶ 1–5. Plaintiffs allege that NVIDIA has trained many of its Artificial Intelligence LLM models using unlawfully copied materials available on illegal pirating websites called shadow libraries. *Id.* ¶¶ 23, 27. Plaintiffs assert that NVIDIA used their copyrighted works that were contained in shadow libraries such as Anna’s Archive and other datasets. *Id.* ¶¶ 5, 33, 42, 46–47. Plaintiffs allege that the copyrighted works were used to develop multiple LLMs in the Megatron family, including Megatron 345M (also known as Megatron GPT2 345m), Nemo GPT-3 10B, InstructRetro-48B, Retro-48B, and Nemotron-4 15B. *Id.* ¶¶ 9, 38–39, 41, 59. Plaintiffs also allege that NVIDIA provided various scripts and tools that enabled its customers to download and preprocess “the Pile,” a dataset containing their copyrighted books, through the NeMo Megatron Framework and BigNLP platforms. *Id.* ¶¶ 70–72.

1 Plaintiffs filed this lawsuit on March 9, 2024. ECF No. 1. Plaintiffs filed a motion to
2 amend the scheduling order and for leave to amend the complaint on October 17, 2025. The Court
3 granted the motion on January 15, 2026. ECF No. 232. Plaintiffs filed the First Amended
4 Consolidated Complaint on January 16, 2025. ECF No. 235. NVIDIA filed a motion to dismiss
5 on January 29, 2026. ECF No. 240. Plaintiffs filed their opposition on February 12, 2026. ECF
6 No. 254. NVIDIA filed its reply on February 18, 2026. ECF No. 255. On March 6, 2026,
7 NVIDIA filed a notice narrowing the issues to be decided, informing the Court that it was no
8 longer seeking the dismissal of claims based on the Nemotron-4 15B and Nemotron-4 340 B
9 models or the Anna’s Archive, Z-Library, LibGen, SciHub, and Slimpajama datasets. ECF No.
10 271 at 2. On March 25, 2026, both parties filed a statement of recent decision regarding *Cox*
11 *Communications, Inc. v. Sony Music Entertainment*, No. 24-171, 2026 WL 815823 (U.S. March
12 25, 2026). ECF Nos. 285, 286. The Court ordered supplemental briefing on the impact of that
13 case on Plaintiffs’ contributory infringement claim, which parties filed on April 3, 2026. ECF
14 Nos. 287, 288, 289. The Court held a hearing on the motion on April 16, 2026.

15 **II. JURISDICTION**

16 This Court has subject matter jurisdiction pursuant to 28 U.S.C. §§ 1331.

17 **III. LEGAL STANDARD**

18 In considering a motion to dismiss, the Court accepts the material facts alleged in the
19 complaint, together with reasonable inferences to be drawn from those facts, as true. *Navarro v.*
20 *Block*, 250 F.3d 729, 732 (9th Cir. 2001). However, “the tenet that a court must accept a
21 complaint’s allegations as true is inapplicable to threadbare recitals of a cause of action’s
22 elements, supported by mere conclusory statements.” *Ashcroft v. Iqbal*, 556 U.S. 662, 678 (2009).

23 In addition, to survive a motion to dismiss, a plaintiff must plead “enough facts to state a
24 claim to relief that is plausible on its face.” *Bell Atlantic Corp. v. Twombly*, 550 U.S. 544, 570
25 (2007). Plausibility does not mean probability, but it requires “more than a sheer possibility that a
26 defendant has acted unlawfully.” *Iqbal*, 556 U.S. at 678. “A claim has facial plausibility when
27 the plaintiff pleads factual content that allows the court to draw the reasonable inference that the
28 defendant is liable for the misconduct alleged.” *Id.*

1 Under Federal Rule of Civil Procedure 12(f), the court may strike from the pleadings “an
 2 insufficient defense or any redundant, immaterial, impertinent, or scandalous matter.” “Rule 12(f)
 3 motions to strike are generally disfavored because the motions may be used as delay tactics and
 4 because of the strong policy favoring resolution on the merits.” *In re Amazon Serv. Fee Litig.*, 705
 5 F. Supp. 3d 1255, 1263 (W.D. Wash. 2023), *aff’d*, No. 24-5176, 2025 WL 2268252 (9th Cir. Aug.
 6 8, 2025). If a claim is stricken, the court should freely grant leave to amend when doing so would
 7 not cause prejudice to the opposing party. *Wyshak*, 607 F.2d at 826.

8 **IV. DISCUSSION**

9 NVIDIA moves the Court to: (1) dismiss or strike allegations about the newly added
 10 Megatron 345M model and any unidentified models; (2) dismiss allegations about the shadow
 11 libraries Pirate Library Mirror and Bibliotik, as well as those concerning any “unspecified” library;
 12 (3) dismiss allegations that NVIDIA infringed using the BitTorrentProtocol; and (4) dismiss
 13 claims that NVIDIA engaged in contributory and vicarious infringement. ECF Nos. 240, 271.
 14 The Court addresses each issue in turn.

15 **A. Models**

16 NVIDIA moves to dismiss allegations concerning its Megatron 345M model and dismiss
 17 or strike references to other “unidentified models.” ECF No. 240 at 11–14.¹

18 **1. Megatron 345M**

19 Plaintiffs allege that NVIDIA used a dataset containing their copyrighted works, known as
 20 “The Pile” “as training data for an LLM known as Megatron 345M.” FCAC ¶ 41. The Pile
 21 dataset included another dataset called Books3. *Id.* ¶ 42. Books3 is derived from another dataset
 22 called Bibliotek. *Id.* ¶ 33. Books3 contains approximately 196,640 books, *id.* ¶ 35, and comprises
 23 12% of The Pile dataset, *id.* ¶ 33. NVIDIA argues that Plaintiffs’ Megatron 345M allegations
 24 should be dismissed because Megatron 345M was trained on portions of The Pile other than

25
 26 ¹ NVIDIA also argues that the FCAC fails to connect the newly named models specifically to
 27 infringement of Plaintiffs’ works. ECF No. 240 at 13. As evidence, NVIDIA references
 28 Plaintiffs’ discovery requests that ask for information related to all datasets or files containing
 books and defining NVIDIA’s LLMs as broadly including all LLMs by NVIDIA. *Id.* at 7, 13–14.
 Plaintiff’s discovery requests have no bearing on whether the complaint’s allegations are
 sufficiently pleaded. Nor does this order express a view on the appropriate scope of discovery.

1 Books3, the dataset that includes the copyrighted works belonging to Plaintiffs. ECF No. 240 at
 2 14. As evidence, NVIDIA asks the court to take judicial notice of a screenshot of a model card²
 3 on NVIDIA’s website, which states that The Pile was trained on text sourced from Wikipedia,
 4 RealNews, OpenWebtext, and CC-Stories. *Id.* (citing ECF No. 206-8 at 2). Plaintiffs object to the
 5 taking of judicial notice. ECF No. 254 at 11. In the alternative, they argue that just because
 6 NVIDIA has publicly proclaimed that Megatron 345M was trained on certain parts of The Pile, it
 7 does not mean that it was not trained on other parts of The Pile that included Books3. *Id.* at 11–
 8 12.

9 NVIDIA’s request for judicial notice is denied. “Ordinarily, the court considers only the
 10 complaint and documents attached thereto in deciding a motion to dismiss; however, the court
 11 may also take judicial notice of matters of public record.” *Chavez v. City of California*, No. 1:19-
 12 cv-00646-CAC-JLT, 2020 WL 1234503, at *3 (E.D. Cal. Mar. 13, 2020) (citing *Lee v. City of Los*
 13 *Angeles*, 250 F.3d 668, 688–89 (9th Cir. 2001)).

14 This Court has previously “reject[ed] the notion that a document is judicially noticeable
 15 simply because it appears on a publicly available website, regardless of who maintains the website
 16 or the purpose of the document.” *Rollins v. Dignity Health*, 338 F. Supp. 3d 1025, 1032 (N.D.
 17 Cal. 2018); *see also In re Cal. Bail Bond Antitrust Litig.*, 511 F. Supp. 3d 1031, 1040 (N.D. Cal.
 18 Jan. 5, 2021) (quoting *Rollins*). The Ninth Circuit has also warned that “the unscrupulous use of
 19 extrinsic documents to resolve competing theories against the complaint risks premature
 20 dismissals of plausible claims that may turn out to be valid after discovery.” *Id.* at 998.

21 Thus, the Court is left only with the allegations of the complaint. Plaintiffs allege that over
 22 12% of the dataset of The Pile is comprised of Books3, that Books3 contains Plaintiffs’
 23 copyrighted material, and that Megatron 345M was trained on The Pile. FCAC ¶¶ 32–36. At this
 24 stage of the litigation, Plaintiffs have adequately alleged that Megatron 345M was trained on a
 25 dataset that contained Plaintiffs’ work. FCAC ¶ 41. Courts have routinely declined to dismiss
 26

27 ² “A model card is a standardized documentation framework for machine learning models that
 28 describes a model's intended use, performance characteristics, training data, evaluation metrics,
 ethical considerations, and limitations.” Model Card, AI Wiki, https://aiwiki.ai/wiki/model_card
[\[https://perma.cc/G6TY-9HUP\]](https://perma.cc/G6TY-9HUP).

1 allegations where plaintiffs alleged that their copyrighted material was present in datasets that
 2 contained their work. *See In re Google Generative AI Copyright Litig.*, 809 F.Supp.3d 903, 914
 3 (N.D. Cal. 2025) (“The Court concludes that Plaintiffs plausibly allege copyright infringement as
 4 to the following models: PaLM, GLaM, LaMDA, Bard, Gemini, and Imagen. Plaintiffs have
 5 identified specific copyrighted works that were allegedly included in datasets used to train these
 6 models.”); *Andersen v. Stability AI Ltd.*, 700 F. Supp. 3d 853, 864 (N.D. Cal. 2023) (approving
 7 claims that defendant “downloaded or otherwise acquired copies of billions of copyrighted images
 8 without permission”). Therefore, the Court denies NVIDIA’s motion to dismiss Plaintiffs’
 9 allegations concerning Megatron 345M.

10 2. Unidentified NVIDIA Models

11 NVIDIA moves to dismiss or strike any allegations that reference “unidentified models”
 12 beyond ones in the Megatron family and the five models that the Court granted leave to include.
 13 ECF No. 240 at 11.

14 As became clear during the hearing on this motion, Plaintiffs are not presently accusing
 15 any unidentified model of infringement, other than the checkpoints for the models they do
 16 identify. Accordingly, because NVIDIA agrees that “the proper scope of the FAC is . . . the five
 17 listed LLMs *and any checkpoints* for those five models,” ECF No. 255 at 8 (emphasis added),
 18 NVIDIA’s motion as to any unidentified models is denied.

19 B. Datasets

20 NVIDIA seeks to dismiss Plaintiffs’ allegations of infringement from the shadow libraries
 21 known as Pirate Library Mirror and Bibliotik, as well as any references to unidentified datasets or
 22 shadow libraries. ECF No. 240 at 16.

23 The complaint’s references to Pirate Library Mirror are either historical or relate to the use
 24 of those libraries by other AI firms. FCAC ¶¶ 47 (“The successor to Z-library, Anna’s Archive
 25 began existence as ‘Pirate Library Mirror[.]’”); 51 (stating that “it is an industry-wide practice to
 26 use shadow libraries such as . . . Pirate Library Mirror”); 51 n.9 (referring to Anthropic’s use of
 27 Pirate Library Mirror). Because Plaintiffs do not presently allege that NVIDIA downloaded
 28 copyrighted material from Pirate Library Mirror, that portion of the motion to dismiss is denied.

1 The references to Bibliotik are also largely historical, but the complaint does contain one
2 allegation that NVIDIA copied copyrighted works from Bibliotik. FCAC ¶ 5. Given other
3 allegations in the complaint concerning the history and contents of Bibliotik, 33, 34, 35, the Court
4 finds this allegation plausible and denies the motion to dismiss concerning Bibliotik.

5 Because Plaintiffs do not currently allege that NVIDIA infringes by the use of unidentified
6 libraries, ECF No. 254 at 17–18, that portion of NVIDIA’s motion is denied.

7 C. BitTorrent Protocol

8 NVIDIA seeks to dismiss allegations concerning its “use of any [sic] ‘BitTorrent
9 Protocol.’” ECF No. 240 at 24. There is only one reference to BitTorrent Protocol in the
10 complaint: “Bibliotik is one of a number of notorious ‘shadow library’ websites which make,
11 store, and distribute huge quantities of pirated copyrighted works via the BitTorrent Protocol.”
12 FCAC ¶ 34.

13 “BitTorrent is a communication protocol for peer-to-peer file sharing (P2P), which enables
14 users to distribute data and electronic files over the Internet in a decentralized manner.”
15 BitTorrent, Wikipedia, <https://en.wikipedia.org/wiki/BitTorrent> [<https://perma.cc/Y23K-H2YP>].
16 In other words, BitTorrent is merely a tool, not a library or dataset. Asking to dismiss allegations
17 concerning BitTorrent is like asking to dismiss allegations concerning paintbrushes in a case about
18 a dolphin painting. *See Folkens v. Wyland Worldwide, LLC*, 882 F.3d 768, 771 (9th Cir. 2018).

19 The motion to dismiss the allegation concerning BitTorrent is denied.

20 D. Third Party Infringement

21 NVIDIA seeks to dismiss the contributory infringement and vicarious infringement claims.
22 ECF No. 240.

23 Preliminarily, both contributory liability and vicarious liability require, “an underlying act
24 of direct infringement.” *Perfect 10, Inc. v. Yandex N.V.*, 962 F. Supp. 2d 1146, 1158 (N.D. Cal.
25 2013). “One infringes contributorily by intentionally inducing or encouraging direct infringement,
26 and infringes vicariously by profiting from direct infringement while declining to exercise a right
27 to stop or limit it.” *Metro-Goldwyn-Mayer Studios Inc. v. Grokster*, 545 U.S. 913, 930 (2005)
28 (internal citations omitted). Plaintiffs allege that NVIDIA provided its customers Writer,

1 Persimmon AI Labs, Amazon, and “numerous other customers” with scripts “to automatically
2 download and preprocess The Pile dataset.” FCAC ¶¶ 70–71. “Meaning, NVIDIA provided tools
3 and resources for its customers to use the NVIDIA platform to download The Pile, thereby
4 infringing on Plaintiffs’ copyrights.” *Id.* ¶ 70. NVIDIA also provided the hardware for these
5 downloads: “Using the NeMo Framework, a customer could expect to quickly develop a language
6 model trained on The Pile in only 9.8 days using NVIDIA’s servers.” *Id.* ¶ 72.

7 NVIDIA argues that Plaintiffs’ allegations about third-party infringement based “on
8 information and belief” are insufficient and that the complaint fails to identify any specific
9 instance in which a customer downloaded or used The Pile. ECF No. 240 at 18. NVIDIA cites
10 *Concord Music Grp., Inc. v. Anthropic PBC* for the proposition that a plaintiff must allege a clear
11 act of direct third-party infringement, rather than identify what third parties could theoretically
12 have done with the product. *Id.* at 18–19 (citing *Concord*, No. 24-cv-03811-EKL, 2025 WL
13 1487988 (N.D. Cal. Mar. 26, 2025)).

14 First, allegations on information and belief are sufficient where they are “based on factual
15 information that makes the inference of culpability plausible,” particularly where the relevant facts
16 are within the defendant’s control. *Menzel v. Scholastic, Inc.*, 2018 WL 1400386, at *2 (N.D. Cal.
17 Mar. 19, 2018) (citing *Arista Record LLC v. Doe 3*, 604 F.3d 110, 120 (2nd Cir. 2010)). Such is
18 the case here. Second, Plaintiffs have identified specific completed instances of infringement by
19 specific customers.

20 Accordingly, the Court finds that Plaintiffs have plausibly alleged predicate acts of direct
21 infringement by third parties.

22 1. Contributory Infringement

23 NVIDIA moves to dismiss contributory infringement claim, arguing that Plaintiffs fail to
24 plead the remaining elements of that claim. ECF No. 240 at 19.

25 To state a claim for contributory infringement, Plaintiffs must plausibly allege that
26 NVIDIA (1) had knowledge of a third party’s infringement; and (2) induced that infringement or
27 provided a service that was tailored to that infringement. *Perfect 10, Inc. v. Visa Int’l Serv. Ass’n*,
28 494 F.3d 788, 795 (9th Cir. 2007). Furthermore, “[t]he provider of a service is contributorily

1 liable for the user's infringement only if it intended that the provided service be used for
 2 infringement. The intent required for contributory liability can be shown only if the party induced
 3 the infringement or the provided service is tailored to that infringement.” *Cox Commc’ns, Inc. v.*
 4 *Sony Music Ent.*, 146 S. Ct. 959, 967 (2026). The Court finds that Plaintiffs have adequately
 5 alleged contributory infringement.

6 **a. Knowledge of Specific Infringing Acts**

7 NVIDIA argues that Plaintiffs fail to allege the requisite knowledge because Plaintiffs rest
 8 their allegations only on “generalized awareness of product capabilities” as opposed to specific
 9 instances of infringement or direction to infringe. ECF No. 240 at 19. Plaintiffs respond that they
 10 allege concrete facts about NVIDIA’s assistance to specific customers who used NVIDIA scripts
 11 to download The Pile, supporting a plausible inference of NVIDIA’s knowledge of third-party
 12 infringement. ECF No. 254 at 22–24. The Court agrees with Plaintiffs.

13 The knowledge prong requires more than a generalized awareness of the possibility of
 14 infringement. *See Luvdarts, LLC v. AT&T Mobility, LLC*, 710 F.3d 1068, 1072 (9th Cir. 2013).
 15 Knowledge of a product’s capability to be used for infringement is insufficient absent awareness
 16 by the defendant of actual infringing activity by third parties utilizing the product. *See Grokster*,
 17 545 U.S. at 933. But Plaintiffs allege such awareness. They allege that NVIDIA itself
 18 downloaded, stored, and repeatedly used The Pile to develop and train several LLMs. ECF No.
 19 235 ¶¶ 38–42. Plaintiffs further allege that NVIDIA provided scripts to customers designed to
 20 enable the acquisition and pre-processing of The Pile to develop their own LLMs. *Id.* ¶¶ 54–58,
 21 69–72. In short, Plaintiffs have alleged that NVIDIA knew that its scripts and other assistance
 22 were directly contributing to infringement by third parties. This satisfies the knowledge element.
 23 *See Napster*, 239 F.3d at 1021–22.

24 NVIDIA cites several inapposite cases in which plaintiffs either alleged only a product’s
 25 abstract capability to infringe or made conclusory assertions untethered to factual allegations of
 26 actual infringing use. *See e.g. Luvdarts*, 710 F.3d at 1072–73 (notices failed to identify “which of
 27 [the] titles were infringed, who infringed them, or when”); *YZ Prods., Inc. v. Redbubble, Inc.*, 545
 28 F. Supp. 3d 756, 764 (N.D. Cal. 2021) (allegations that defendant had “specific knowledge” were

1 “threadbare” and unsupported by facts); *Nat’l Photo Grp., LLC v. Allvoices, Inc.*, 2014 WL
 2 280391, at *6 (N.D. Cal. Jan. 24, 2014) (“There are no alleged facts plausibly suggesting that
 3 Defendant was on any notice . . . of the infringing [content] on its website.”). By contrast, the
 4 allegations here are neither abstract nor conclusory. The Court finds that Plaintiffs have plausibly
 5 alleged that NVIDIA had knowledge of infringing activity.

6 **b. Inducement Theory**

7 NVIDIA claims that Plaintiffs have failed to establish that it induced infringement through
 8 specific acts because “[t]he FAC does not allege that NVIDIA promoted or marketed the NeMo
 9 Megatron Framework as a tool to download copyrighted content.” ECF No. 288 at 3 (citing *Cox*,
 10 146 S.Ct. at 967). NVIDIA emphasizes that the complaint lacks allegations NVIDIA promoting
 11 the NeMo Framework in its entirety as a tool for copyright infringement, instead alleging only that
 12 an optional configuration in the Framework permitted downloading and preprocessing of the Pile.
 13 *Id.* at 3–4. Plaintiffs respond that NVIDIA misrepresents the complaint by framing Plaintiffs’
 14 allegations in terms of the entire NeMo Framework as opposed to the specific scripts that NVIDIA
 15 provided to automatically download and preprocess The Pile. ECF No. 289 at 6. Plaintiffs allege
 16 that these scripts are tools “specifically designed for copyright infringement,” which induces
 17 infringement by NVIDIA’s customers. *Id.* at 7

18 To allege contributory infringement under an inducement theory, a plaintiff must allege
 19 that the defendant “actively encourages infringement through specific acts.” *Cox*, 2026 WL
 20 815823, at *6 (citing *Grokster*, 545 U.S. at 936). Plaintiffs allege that NVIDIA took the specific
 21 steps of “develop[ing] and distribut[ing] code to ‘download and extract’” copyrighted files to its
 22 customers. FCAC ¶ 70. That is sufficient to allege that NVIDIA induced infringement. That the
 23 NeMo Megatron Framework as a whole may have other, non-infringing uses does not alter this
 24 conclusion.

25 NVIDIA also contends Plaintiffs have failed to allege inducement because they did not
 26 plead that NVIDIA advertised these infringing uses or promoted the NeMo Framework as a tool
 27 for infringement. ECF No. 280 at 3. Advertising or promotion, however, is an example of an
 28 inducing act, not a pre-requisite for alleging inducement. See *Cox* at 967 (citing *Grokster*, 545

1 U.S. 914, 941) (“A provider induces infringement if it actively encourages infringement through
 2 specific acts. . . . For example, in *Grokster*, we held that a jury could find two file-sharing
 3 software companies liable for inducement . . . [when t]he companies promoted and marketed their
 4 software as a tool to infringe copyrights.”). It is sufficient that Plaintiffs allege that defendant
 5 “instruct[ed] how to engage in an infringing use” and “express[ed] an affirmative intent that the
 6 product be used to infringe,” such as by designing a tool specifically to promote the infringement.
 7 *Grokster*, 545 U.S. 914, 915. Therefore, the Court finds that Plaintiffs have alleged that NVIDIA
 8 engaged in contributory liability by inducing infringement.

9 **c. “Service Tailored to Infringement” Theory**

10 NVIDIA also argues that Plaintiffs’ contributory infringement claim fails because the
 11 NeMo Megatron Framework is capable of substantial non-infringing uses and is not a service
 12 tailored to infringement. ECF No. 240 at 21; ECF No. 288 at 6. Plaintiffs again argue that
 13 NVIDIA misrepresents the scope of the complaint as applying to the entire Framework rather than
 14 the specific tools NVIDIA utilized to infringe. Plaintiffs argue that the scripts and material
 15 assistance provided by NVIDIA to its customers is the service tailored to infringement, as it had
 16 no purpose except to infringe. ECF No. 254 at 27–28. The Court agrees with Plaintiffs.

17 To allege that a defendant is committing copyright infringement by providing a service
 18 tailored to infringement, the complaint must allege that the defendant offered products “not
 19 capable of ‘substantial’ or ‘commercially significant’ noninfringing uses.” *Cox*, 146 S.Ct. at 967
 20 (quoting *Grokster*, 545 U.S. 914, 942 (Ginsburg, J., concurring)). Further, a product that is
 21 “capable of substantial noninfringing uses” does not give rise to contributory liability based solely
 22 on its distribution. *Sony Corp. of Am. v. Universal City Studios, Inc.*, 464 U.S. 417, 442 (1984).
 23 Here, Plaintiffs allege that NVIDIA provided scripts that enabled customers to download The Pile,
 24 a dataset alleged to contain copyrighted works. FCAC ¶¶ 31–38, 70–71. The scripts are alleged
 25 to have no other purpose than to speed up the process of infringement, unlike the digital video
 26 recorder systems at issue in *Sony Corp.* or the internet service provided in *Cox*. 464 U.S. at 442;
 27 146 S.Ct. at 968. Therefore, the Court finds that Plaintiffs allege a contributory liability claim
 28 under both the “inducement” and “service tailored to infringement” theories and denies NVIDIA’s

1 motion to dismiss the contributory infringement claim.

2 **2. Vicarious Infringement**

3 Plaintiffs also allege that NVIDIA is liable for vicarious infringement of their copyrighted
4 material. “To state a claim for vicarious copyright infringement, a plaintiff must allege that the
5 defendant has (1) the right and ability to supervise the infringing conduct and (2) a direct financial
6 interest in the infringing activity.” *Visa*, 494 F.3d at 802 (citing *Ellison v. Robertson*, 357 F.3d
7 1072, 1078 (9th Cir. 2004)). Stated differently, “[o]ne . . . infringes vicariously by profiting from
8 direct infringement while declining to exercise a right to stop or limit it.” *Id.* (quoting *Grokster*,
9 545 U.S at 930).

10 **a. Right and Ability to Control Infringement**

11 NVIDIA first argues that Plaintiffs fail to plausibly allege that NVIDIA had the right and
12 ability to control the alleged infringement because NVIDIA’s ability to design or modify its own
13 tools does not confer a legal right to control independent conduct by third-party users. ECF No.
14 240 at 21–22. Plaintiffs respond that NVIDIA exercised control by designing and providing
15 scripts that enabled customers to automatically download and process The Pile. ECF No. 254 at
16 29.

17 To establish vicarious liability, Plaintiffs must allege that NVIDIA had both the legal right
18 to stop or limit the directly infringing conduct and the practical ability to do so. *Perfect 10, Inc. v.*
19 *Amazon.com, Inc.*, 508 F.3d 1146, 1173 (9th Cir. 2007). To allege that defendants have the legal
20 right to stop infringement, Plaintiffs must allege that “the defendant exercises the requisite control
21 over the direct infringer.” *Id.* The Ninth Circuit has made clear that the ability to control one’s
22 own systems or services, even if doing so could affect infringement “to some degree,” does not
23 establish the right and ability to control the infringing acts of third parties. *Visa*, 494 F.3d 788 at
24 803. Nor is it enough to allege that a defendant could take steps having the “indirect effect of
25 reducing infringing activity;” the defendant must be able to supervise and control the infringement
26 itself and “not just affect it.” *Id.* at 803–05.

27 The Court finds that Plaintiffs have not alleged that NVIDIA had the requisite ability to
28 control third-party infringement. Plaintiffs assert that “NVIDIA had the right and ability to

1 control the direct infringements of customers.” ECF No. 235 ¶ 88. The allegation is conclusory
 2 and fails to identify *how* NVIDIA could exercise that control. Plaintiffs’ allegations are unlike
 3 those in *Napster*, where the defendant maintained centralized control over the infringing activity
 4 and possessed the ability to police or block user access to the very system through which
 5 infringement occurred. *Napster, Inc.*, 239 F.3d at 1023–24 (9th Cir. 2001). Plaintiffs’ reliance on
 6 *Unicolors, Inc. v. Joy 153, Inc.* is also misplaced. 2016 WL 3462128, at *5 (C.D. Cal. June 24,
 7 2016). The defendant there manufactured and sold garments bearing the infringing design to
 8 retailers and thus “had the right and ability to not engage in these transactions . . . and . . . could
 9 have prevented further infringement.” *Id.* Plaintiffs do not identify any legal right by which
 10 NVIDIA could prevent or limit its customers’ independent decisions to access or use The Pile
 11 from third-party sources.

12 Plaintiffs argue that NVIDIA controlled its training platform, including scripts that
 13 automatically downloaded The Pile, which provided it with the legal ability to control
 14 infringement. ECF No. 254 at 28–29. However, those allegations only establish control over
 15 NVIDIA’s own tools. The complaint does not plausibly allege that NVIDIA had the legal right or
 16 practical ability to stop users from obtaining or using infringing materials. *See Visa*, 494 F.3d at
 17 803 (allegations that defendant was aware of the infringement and had the ability to stop
 18 processing payments to infringing websites were not “sufficient to establish vicarious liability
 19 because even with all reasonable inferences drawn in Perfect 10’s favor, Perfect 10’s allegations of
 20 fact cannot support a finding that Defendants have the right and ability to control the infringing
 21 activity”). In *Grokster*, the Supreme Court explained that the “control” element focuses on the
 22 defendant’s “right and ability to supervise the direct infringer.” *Grokster*, 545 U.S. 914, 930 n.9.
 23 No comparable ability to supervise the alleged infringement is plausibly alleged here. Plaintiffs’
 24 allegations are insufficient to satisfy this element of vicarious liability.

25 **b. Financial Interest**

26 NVIDIA also argues that Plaintiffs fail to plausibly allege that NVIDIA derived a direct
 27 financial benefit from any alleged infringement. ECF No. 240 at 22. It contends that the NeMo
 28 Megatron Framework and associated scripts were offered for free, that the platform has substantial

1 non-infringing uses, and that Plaintiffs failed to allege any concrete connection between the
2 alleged infringement and NVIDIA’s revenue. *Id.* Plaintiffs respond that NVIDIA financially
3 benefited by attracting customers through tools that facilitate access to datasets such as The Pile
4 and by promoting the platform’s ability to enable rapid AI model development. ECF No. 254 at
5 29.

6 To state a claim for vicarious liability, Plaintiffs must plausibly allege that NVIDIA
7 received a direct financial benefit from the infringing activity. A direct financial benefit exists
8 where the availability of infringing material “acts as a draw for customers.” *Napster*, 239 F.3d at
9 1023 (citing *Fonovisa*, 76 F.3d at 263–64). The essential inquiry is whether there is a causal
10 relationship between the alleged infringement and the financial benefit received. *Ellison*, 357 F.3d
11 at 1079. The size of the benefit is immaterial; rather, the question is whether infringement itself
12 serves as a customer draw. *See Ellison*, 357 F.3d at 1079; *see also Perfect 10, Inc. v. Giganews,*
13 *Inc.*, 847 F.3d 657, 673 (9th Cir. 2017). It is not sufficient to allege that customers were drawn “to
14 obtain access to infringing material in general”; customers must be drawn to “the *plaintiff’s*
15 copyrighted material, rather than general infringement.” *Perfect 10*, 847 F.3d at 673 (emphasis
16 added).

17 The complaint alleges that “NVIDIA has directly benefitted financially from the direct
18 infringement of its customers because NVIDIA generated revenue from customers using the
19 NeMo Megatron Framework to download The Pile (and Books3) dataset (and potentially other
20 datasets containing copyrighted books as well).” ECF No. 235 ¶ 89. But this does not plead the
21 required causal relationship between the alleged infringement and any financial benefit. *Ellison*,
22 357 F.3d at 1079 (requiring a “causal relationship between the infringing activity and any financial
23 benefit a defendant reaps.”). Similarly, Plaintiffs’ allegation that “NVIDIA also caused numerous
24 third parties to download and store Plaintiffs’ copyrighted works by encouraging, facilitating, and
25 promoting its customers to download copies of The Pile dataset,” ECF No. 235 ¶ 7, says nothing
26 about whether access to The Pile “acts as a ‘draw’ for customers.” *Fonovisa, Inc. v. Cherry*
27 *Auction, Inc.*, 76 F.3d 259, 263–64 (9th Cir. 1996). The “central question ... is whether the
28 infringing activity constitutes a draw ... not just an added benefit.” *Ellison*, 357 F.3d at 1079.

United States District Court
Northern District of California

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

Finally, Plaintiffs’ allegation that, “[i]n short, The Pile (and Books3) was key to NVIDIA attracting customers, and NVIDIA materially aided its customers to infringe Plaintiffs copyrights,” ECF No. 235 ¶ 73, without supporting factual content, also does not satisfy this standard.

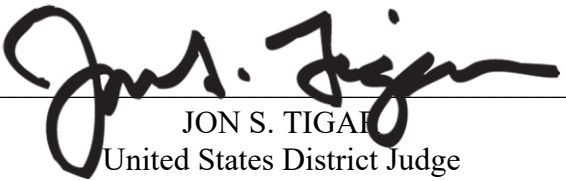
Accordingly, Plaintiffs claim for vicarious liability is dismissed with leave to amend.

CONCLUSION

For the foregoing reasons, the Court denies NVIDIA’s motion to dismiss claims about Megatron 345M, unidentified NVIDIA LLMs, Pirate Library Mirror, Bibliotik, unidentified datasets, BitTorrent, and contributory infringement. The Court grants the motion to dismiss Plaintiffs’ vicarious infringement claim. Leave is granted solely to correct the deficiencies identified in this order. Any amended complaint must be filed with 21 days of this order.

IT IS SO ORDERED.

Dated: May 5, 2026



JON S. TIGAI
United States District Judge