

FENWICK & WEST LLP

1 JEDEDIAH WAKEFIELD (CSB No. 178058)  
 jwakefield@fenwick.com  
 2 RYAN KWOCK (CSB No. 336414)  
 rkwock@fenwick.com  
 3 FENWICK & WEST LLP  
 555 California Street, 12th Floor  
 4 San Francisco, CA 94104  
 Telephone: 415.875.2300  
 5 Facsimile: 415.281.1350

6 DAVID HAYES (CSB No. 122894)  
 dhayes@fenwick.com  
 7 FENWICK & WEST LLP  
 801 California Street  
 8 Mountain View, CA 94041  
 Telephone: 650.988.8500  
 9 Facsimile: 650.938.5200

10 *Attorneys for Defendants*  
 DATABRICKS, INC. and  
 11 MOSAIC ML, LLC, formerly MOSAIC ML, INC.

12 *Additional counsel listed on signature page*

13  
 14 UNITED STATES DISTRICT COURT  
 15 NORTHERN DISTRICT OF CALIFORNIA  
 16 SAN FRANCISCO DIVISION

18 Stewart O’Nan, an individual; Abdi Nazemian,  
 an individual; and Brian Keene, an individual,  
 19 Individual and Representative  
 20 Plaintiffs,  
 21 v.  
 22 Databricks, Inc., a Delaware corporation; and  
 Mosaic ML, Inc., a Delaware corporation,  
 23 Defendants.  
 24

Case No.: 3:24-cv-01451-CRB

**DEFENDANTS’ ANSWER TO CLASS  
 ACTION COMPLAINT**

25  
 26  
 27  
 28

1 Defendants Databricks, Inc. (“Databricks”) and Mosaic ML, LLC, formerly Mosaic ML,  
 2 Inc. (“MosaicML”) (together “Defendants”) submit this Answer to the Class Action Complaint  
 3 (“Complaint”) filed by Plaintiffs Stewart O’Nan, Abdi Nazemian, and Brian Keene on behalf of  
 4 themselves and others similarly situated (collectively “Plaintiffs”). Unless specifically admitted,  
 5 Defendants deny each of the allegations in the Complaint.

## 6 OVERVIEW<sup>1</sup>

7 1. Defendants admit that “artificial intelligence” is commonly abbreviated “AI.” As  
 8 Paragraph 1 pertains generally to AI and not any particular AI technologies of Defendants,  
 9 Defendants lack knowledge and information sufficient to form a belief as to the truth of the  
 10 remaining allegations in Paragraph 1, and on that basis deny them.

11 2. Defendants admit that MosaicML created and that Databricks distributed a large  
 12 language model (“LLM”) series called MosaicML Pretrained Transformer (“MPT”). Defendants  
 13 admit that LLMs may generate text outputs in response to user prompts.

14 3. As Paragraph 3 pertains generally to LLMs and not any particular LLM of  
 15 Defendants, Defendants lack knowledge and information sufficient to form a belief as to the truth  
 16 of the allegations in Paragraph 3, and on that basis deny them. Defendants deny any suggestion  
 17 that the LLMs at issue in this litigation have the goal of imitating protected expression.

18 4. Defendants lack knowledge and information sufficient to form a belief as to the  
 19 truth of the allegations in Paragraph 4, and on that basis deny them.

20 5. Defendants lack knowledge and information sufficient to form a belief as to the  
 21 truth of the allegations in Paragraph 5, and on that basis deny them.

22 6. Defendants admit that Databricks is the corporate parent of MosaicML and also  
 23 distributed certain MPT models. Defendants deny the remaining allegations in Paragraph 6.

## 24 JURISDICTION AND VENUE

25 7. The allegations in this Paragraph state a legal conclusion to which no response is  
 26 required. To the extent a response is deemed required, Defendants lack knowledge and

27 \_\_\_\_\_  
 28 <sup>1</sup> Defendants include the headings as listed in the Complaint without any admission as to the  
 accuracy or appropriateness of the headings.

1 information sufficient to form a belief as to the truth of the allegations in Paragraph 7 and on that  
2 basis deny them.

3 8. The allegations in this Paragraph state a legal conclusion to which no response is  
4 required. To the extent a response is deemed to be required, Defendants admit that the Court has  
5 personal jurisdiction over it with respect to this litigation and that venue is proper in the Northern  
6 District of California. Defendants deny the remaining allegations in Paragraph 8.

7 9. The allegations in this Paragraph state a legal conclusion to which no response is  
8 required. To the extent a response is deemed to be required, Defendants admit that assignment of  
9 this case to the San Francisco Division is proper, and that this action purports to arise under  
10 intellectual property laws.

#### 11 **PLAINTIFFS**

12 10. Defendants lack knowledge and information sufficient to form a belief as to the  
13 truth of the allegations in Paragraph 10, and on that basis deny them.

14 11. Defendants lack knowledge and information sufficient to form a belief as to the  
15 truth of the allegations in Paragraph 11, and on that basis deny them.

16 12. Defendants lack knowledge and information sufficient to form a belief as to the  
17 truth of the allegations set forth in Paragraph 12, and on that basis deny them.

18 13. Defendants lack knowledge and information sufficient to form a belief as to the  
19 truth of the allegations in Paragraph 13 or as to the content of Exhibit A, and on that basis deny  
20 the allegations.

#### 21 **DEFENDANTS**

22 14. Defendants admit the allegations in Paragraph 14.

23 15. Defendants admit that MosaicML is a Delaware limited liability company, is  
24 located at this address, and is a subsidiary of Databricks.

#### 25 **AGENTS AND CO-CONSPIRATORS**

26 16. Defendants deny the allegations in Paragraph 16.

27 17. Defendants deny the allegations in Paragraph 17.

**FACTUAL ALLEGATIONS**

1  
2 18. Defendants admit that MosaicML was incorporated in 2020, and one of its goals  
3 was to provide tools to facilitate the efficient training of AI models.

4 19. Defendants admit that MosaicML released the MPT-7B LLM in May 2023.  
5 Defendants admit that LLMs may generate text outputs in response to user prompts. Defendants  
6 deny the remaining allegations in Paragraph 19.

7 20. Defendants admit that the corpus of material used to train an LLM may be referred  
8 to as a “training dataset.” As Paragraph 20 pertains generally to LLMs and not any particular  
9 LLM of Defendants, Defendants lack knowledge and information sufficient to form a belief as to  
10 the truth of the remaining allegations in Paragraph 20, and on that basis deny them.

11 21. Defendants admit that the MPT-7B LLM contains 6.7 billion parameters but  
12 otherwise deny the last sentence of this paragraph. As Paragraph 21 otherwise pertains generally  
13 to LLMs and not any particular LLM of Defendants, Defendants lack knowledge and information  
14 sufficient to form a belief as to the truth of the remaining allegations in Paragraph 21, and on that  
15 basis deny them. Defendants deny any suggestion that the LLMs at issue in this litigation  
16 progressively adjusted their output to more closely approximate any protected expression  
17 contained in the training dataset.

18 22. As Paragraph 22 pertains generally to LLMs and not any particular LLM of  
19 Defendants, Defendants lack knowledge and information sufficient to form a belief as to the truth  
20 of the allegations set forth in Paragraph 22, and on that basis deny them. Defendants deny any  
21 suggestion that the LLMs at issue in this litigation have the goal of imitating any protected  
22 expression ingested from the training dataset.

23 23. The allegations in this Paragraph state a legal conclusion to which no response is  
24 required. To the extent a response is deemed required, Defendants deny that MosaicML’s  
25 transformative use of works to train the MPT LLMs required consent, credit, or compensation.  
26 Defendants lack knowledge and information sufficient to form a belief as to the truth of the  
27 remaining allegations set forth in Paragraph 23, and on that basis deny them.

28

1           24. Defendants admit that MosaicML released the MPT-7B LLM in May 2023 and  
2 published a blog post titled “Introducing MPT-7B: A New Standard for Open-Source,  
3 Commercially Usable LLMs” available at <https://www.databricks.com/blog/mpt-7b>. To the  
4 extent the allegations set forth in this Paragraph purport to summarize or characterize the contents  
5 of the blog post located at <https://www.databricks.com/blog/mpt-7b>, the document speaks for  
6 itself.

7           25. To the extent the allegations set forth in this Paragraph purport to summarize or  
8 characterize the contents of the blog post located at <https://www.databricks.com/blog/mpt-7b>, the  
9 document speaks for itself. Defendants deny the remaining allegations in Paragraph 25.

10           26. To the extent the allegations set forth in this Paragraph purport to summarize or  
11 characterize the contents of the website located at  
12 <https://huggingface.co/datasets/togethercomputer/RedPajama-Data-1T>, the document speaks for  
13 itself. Defendants lack knowledge and information sufficient to form a belief as to the truth of the  
14 remaining allegations in Paragraph 26, and on that basis deny them.

15           27. To the extent the allegations set forth in this Paragraph purport to summarize or  
16 characterize the contents of the document located at <https://arxiv.org/pdf/2101.00027.pdf>, the  
17 document speaks for itself. Defendants lack knowledge and information sufficient to form a  
18 belief as to the truth of the remaining allegations in Paragraph 27, and on that basis deny them.

19           28. The allegations in this Paragraph state a legal conclusion to which no response is  
20 required. To the extent a response is deemed required, Defendants lack knowledge and  
21 information sufficient to form a belief as to the truth of the allegations in Paragraph 28, and on  
22 that basis deny them.

23           29. Defendants lack knowledge and information sufficient to form a belief as to the  
24 truth of the allegations in Paragraph 29, and on that basis deny them.

25           30. Defendants deny that they infringed Plaintiffs’ alleged copyrights. Defendants  
26 lack knowledge and information sufficient to form a belief as to the truth of the remaining  
27 allegations in Paragraph 30, and on that basis deny them.

28



1 37. Defendants lack knowledge and information sufficient to form a belief as to the  
2 truth of the allegations in Paragraph 37, and on that basis deny them.

3 38. Defendants admit that portions of the RedPajama – Books dataset were used as  
4 training data for the MPT-7B and MPT-30B LLMs. Defendants lack knowledge and information  
5 sufficient to form a belief as to the truth of the remaining allegations set forth in Paragraph 38,  
6 and on that basis deny them

7 39. Defendants admit that MosaicML released the MPT-7B-StoryWriter-65k+ LLM.  
8 To the extent the allegations set forth in this Paragraph purport to summarize or characterize the  
9 contents of the blog post located at <https://www.databricks.com/blog/mpt-7b>, the document  
10 speaks for itself. Defendants deny the remaining allegations in Paragraph 39.

11 40. Defendants deny that MosaicML needed Plaintiffs’ authorization to use any  
12 portions of the RedPajama – Books dataset to train MosaicML’s MPT models. Defendants lack  
13 knowledge and information sufficient to form a belief as to the truth of the remaining allegations  
14 in Paragraph 40, and on that basis deny them.

15 41. Defendants deny that MosaicML needed Plaintiffs’ authorization to use any  
16 portion of the RedPajama – Books dataset to train MosaicML’s MPT models. Defendants deny  
17 that MosaicML violated any exclusive rights under the Copyright Act. Defendants lack  
18 knowledge and information sufficient to form a belief as to the truth of the remaining allegations  
19 in Paragraph 41, and on that basis deny them.

20 42. Defendants deny the allegations in Paragraph 42.

## 21 **COUNT 2**

### 22 **Vicarious Copyright Infringement against Databricks**

23 43. Defendants incorporate by reference their responses to Paragraphs 1 through 42  
24 above as though fully set forth herein.

25 44. Defendants admit that Databricks acquired MosaicML in July 2023. Defendants  
26 deny the remaining allegations in Paragraph 44.

27 45. Defendants deny the allegations in Paragraph 45.

28 46. Defendants deny the allegations in Paragraph 46.

**CLASS ALLEGATIONS**

1  
2           47.     The allegations in Paragraph 47 state legal conclusions or arguments to which no  
3 response is required. To the extent a response is deemed required, Defendants deny the  
4 allegations in Paragraph 47.

5           48.     The allegations in Paragraph 48 state legal conclusions or arguments to which no  
6 response is required.

7           49.     The allegations in Paragraph 49 state legal conclusions or arguments to which no  
8 response is required.

9           50.     The allegations in Paragraph 50 state legal conclusions or arguments to which no  
10 response is required. To the extent a response is deemed required, Defendants deny that they  
11 possess information concerning the exact number of members of Plaintiffs’ putative class.  
12 Defendants lack knowledge and information sufficient to form a belief as to the truth of the  
13 remaining allegations in Paragraph 50, and on that basis deny them.

14           51.     The allegations in Paragraph 51 state legal conclusions or arguments to which no  
15 response is required. To the extent a response is deemed required, Defendants deny the  
16 allegations in Paragraph 51.

17           52.     The allegations in Paragraph 52 state legal conclusions or arguments to which no  
18 response is required. To the extent a response is deemed required, Defendants deny the  
19 allegations in Paragraph 52.

20           53.     The allegations in Paragraph 53 state legal conclusions or arguments to which no  
21 response is required. To the extent a response is deemed required, Defendants lack knowledge  
22 and information sufficient to form a belief as to the truth of the allegations in Paragraph 53, and  
23 on that basis deny them.

24           54.     The allegations in Paragraph 54 state legal conclusions or arguments to which no  
25 response is required. To the extent a response is deemed required, Defendants deny the  
26 allegations in Paragraph 54.

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

FENWICK & WEST LLP



1 **DEMAND FOR JUDGMENT**

2 Defendants deny that Plaintiffs and the proposed Class are entitled to any relief, including  
3 but not limited to the relief sought in the section of the Complaint titled “Demand for Judgment.”  
4 To the extent that this section contains any allegations, Defendants deny them.

5 **JURY TRIAL DEMANDED**

6 Defendants hereby demand a trial by jury on all claims, defenses, and issues in this action  
7 so triable.

8 **GENERAL DENIAL**

9 Unless expressly admitted above, Defendants deny each and every allegation set forth in  
10 Plaintiffs’ Complaint.

11 **DEFENSES**

12 Defendants further plead the following separate and additional defenses to the Complaint.  
13 By pleading these defenses, Defendants do not in any way agree or concede that they have the  
14 burden of proof or persuasion on any claims or defenses. Defendants reserve the right to plead  
15 any and all defenses that may be evident or revealed after investigation and discovery in this  
16 matter.

17 **FIRST DEFENSE  
(Failure to State a Claim)**

18 Plaintiffs’ and putative class members’ claims fail, in whole or in part, because the  
19 Complaint, and each purported cause of action therein, fails to state a claim upon which relief  
20 may be granted and/or to state facts sufficient to constitute a claim for relief against Defendants.

21 **SECOND DEFENSE  
(Fair Use)**

22  
23 To the extent that Defendants made any unauthorized copies of any of Plaintiffs’ or  
24 putative class members’ registered copyrighted works, such copying constituted fair use under  
25 U.S.C. § 107, given, among other factors, the purpose and transformative character of the use and  
26 the effect of the use upon the potential market for or value of the allegedly copyrighted works.  
27  
28



1 **EIGHTH DEFENSE**  
2 **(Lack of Standing)**

3 Plaintiffs' and putative class members' claims are barred to the extent they do not own a  
4 valid copyright for some or all of the works, do not hold a valid copyright registration for some or  
5 all of the works, and/or otherwise lack standing or fail to meet statutory requirements to assert  
6 their claims.

7 **NINTH DEFENSE**  
8 **(Unavailability of Injunctive Relief)**

9 Plaintiffs' and putative class members' claims fail, in whole or in part, because Plaintiffs  
10 are not entitled to injunctive relief (temporarily, preliminarily, or permanently), including because  
11 any injury to them is not immediate or irreparable, Plaintiffs would have an adequate remedy at  
12 law, the balance of hardships favors no injunction, and the public interest is best served by no  
13 injunction.

14 **TENTH DEFENSE**  
15 **(Lack of Willfulness)**

16 Plaintiffs' and putative class members' claims fail in whole or in part because Defendants'  
17 conduct was innocent, not willful.

18 **ELEVENTH DEFENSE**  
19 **(Lack of Injury)**

20 Plaintiffs' and putative class members' claims fail in whole or in part because they have  
21 not suffered injury as a result of the conduct alleged in the Complaint.

22 **TWELFTH DEFENSE**  
23 **(Failure to Mitigate Damages)**

24 Plaintiffs and putative class members are barred from recovery of damages because of and  
25 to the extent of their failure to mitigate their alleged damages (to which, in any event, they are not  
26 entitled).

**THIRTEENTH DEFENSE  
(Attorneys' Fees Improper)**

The Complaint fails to state a cause of action or allege sufficient facts to support a claim for attorneys' fees.

**RESERVATION OF ADDITIONAL DEFENSES**

Defendants do not presently know all facts concerning Plaintiffs' claims and the putative class sufficient to state all affirmative defenses at this time. Defendants will seek leave to amend this Answer should they later discover facts demonstrating the existence of additional affirmative defenses. Defendants reserve any and all additional affirmative defenses available to them.

**DEFENDANTS' DEMAND FOR JUDGMENT**

WHEREFORE, Defendants pray for judgment with respect to Plaintiffs' Complaint and Defendants' defenses as follows:

- a) A judgment in Defendants' favor denying Plaintiffs all relief requested in the Complaint and dismissing the Complaint with prejudice;
- b) For an award to Defendants of their attorneys' fees, costs, and expenses of litigation; and
- c) Such other relief as the Court shall deem just and proper.

FENWICK & WEST LLP

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

1 DATED: May 2, 2024

FENWICK & WEST LLP

3 By: /s/ Jedediah Wakefield

4 Jedediah Wakefield (CSB No. 178058)

jwakefield@fenwick.com

5 Ryan Kwock (CSB No. 336414)

555 California Street, 12th Floor

6 San Francisco, CA 94104

7 Telephone: 415.875.2300

Facsimile: 415.281.1350

8 David Hayes (CSB No. 122894)

dhayes@fenwick.com

9 801 California Street

Mountain View, CA 94041

10 Telephone: 650.988.8500

Facsimile: 650.938.5200

11 Deena Feit (admitted *pro hac vice*)

dfeit@fenwick.com

12 401 Union Street, 5th Floor

13 Seattle, WA 98101

14 Telephone: 206.389.4510

Facsimile: 206.389.4511

15 Charles Moulins (admitted *pro hac vice*)

cmoulins@fenwick.com

16 902 Broadway, Ste 14

New York, NY 10010

17 Telephone: 212.430.2600

Facsimile: 650.938.5200

18 *Attorneys for Defendants*

19 DATABRICKS, INC. and

MOSAIC ML, LLC, formerly MOSAIC ML,

20 INC.

FENWICK & WEST LLP