

1 JOSEPH C. GRATZ (CA SBN 240676)
JGratz@mofo.com
2 TIFFANY CHEUNG (CA SBN 211497)
TCheung@mofo.com
3 MORRISON & FOERSTER ^{LLP}
425 Market Street,
4 San Francisco, California 94105-2482
Telephone: 415.268.7000
5 Facsimile: 415.268.7522

6 *Attorneys for Defendant*
7 SALESFORCE, INC.

8 UNITED STATES DISTRICT COURT
9 NORTHERN DISTRICT OF CALIFORNIA
10 SAN FRANCISCO DIVISION

11
12 *In re Salesforce LLM Copyright Infringement*
13 *Litigation*

Master File No. 3:25-cv-08862-CRB

14 **SALESFORCE'S ANSWER TO**
15 **PLAINTIFFS' FIRST AMENDED**
16 **CONSOLIDATED COMPLAINT**

**SALESFORCE, INC.'S ANSWER TO
PLAINTIFFS' FIRST AMENDED CONSOLIDATED COMPLAINT**

Defendant Salesforce, Inc. ("Salesforce") hereby answers Plaintiffs' First Amended Consolidated Complaint ("Consolidated Complaint"). Any and all allegations not specifically admitted herein are denied. To the extent the Consolidated Complaint asserts conclusions of law, such conclusions of law require no response. To the extent any response is required to headings or other unnumbered paragraphs in the Consolidated Complaint, Salesforce denies the factual allegations, if any, contained in such headings or unnumbered paragraphs.

RESPONSE TO OVERVIEW

1. Salesforce admits that it developed its XGen series of large language models ("LLMs"). To the extent that the allegations of this paragraph purport to quote or reference Salesforce's public statements, the full text of those statements speaks for itself. To the extent the allegations of this paragraph purport to summarize or characterize or are inconsistent with the full text available at that link, Salesforce denies those allegations. Salesforce denies any remaining allegations in this paragraph.

2. Denied.

3. Salesforce lacks knowledge or information sufficient to admit or deny the allegations of this paragraph about Plaintiffs and third parties, and on that basis denies them. Salesforce denies any remaining allegations in this paragraph.

4. Denied.

5. Denied.

6. Denied.

RESPONSE TO JURISDICTION AND VENUE

7. The allegations of this paragraph state a legal conclusion to which no response is required. To the extent a response is deemed required, Salesforce admits that the federal district courts have subject matter jurisdiction over this action pursuant to 28 U.S.C. § 1331. Salesforce denies any remaining allegations in this paragraph.

8. The allegations of this paragraph state legal conclusions to which no response is required. To the extent a response is deemed required, Salesforce admits that it is headquartered in this District. Salesforce admits that it developed its XGen series of LLMs. For the purpose of this matter only, Salesforce admits that personal jurisdiction and venue are proper. Salesforce denies any remaining allegations in this paragraph.

9. The allegations of this paragraph state legal conclusions to which no response is required. To the extent a response is deemed required, Salesforce admits that venue is proper. Salesforce denies any remaining allegations in this paragraph.

RESPONSE TO PARTIES

A. RESPONSE TO PLAINTIFFS

10. Salesforce lacks knowledge or information sufficient to admit or deny allegations of this paragraph about Plaintiff, and on that basis denies them. Salesforce denies any remaining allegations in this paragraph.

11. Salesforce lacks knowledge or information sufficient to admit or deny allegations of this paragraph about Plaintiff, and on that basis denies them. Salesforce denies any remaining allegations in this paragraph.

12. Salesforce lacks knowledge or information sufficient to admit or deny allegations of this paragraph about Plaintiff, and on that basis denies them. Salesforce denies any remaining allegations in this paragraph.

13. Salesforce lacks knowledge or information sufficient to admit or deny allegations of this paragraph about Plaintiff, and on that basis denies them. Salesforce denies any remaining allegations in this paragraph.

14. Salesforce lacks knowledge or information sufficient to admit or deny allegations of this paragraph about Plaintiff, and on that basis denies them. Salesforce denies any remaining allegations in this paragraph.

15. Salesforce lacks knowledge or information sufficient to admit or deny allegations of this paragraph about Plaintiffs, and on that basis denies them. Salesforce denies any remaining allegations in this paragraph.

B. RESPONSE TO DEFENDANT

16. Salesforce admits that Salesforce, Inc. is a Delaware corporation. Salesforce admits that its principal place of business is at 415 Mission Street, 3rd Floor, San Francisco, CA 94105.

C. RESPONSE TO AGENTS AND CO-CONSPIRATORS

17. Denied.

18. Denied.

RESPONSE TO FACTUAL ALLEGATIONS**A. Response to “The Pile” and “RedPajama” Contain Plaintiffs’ and Class Members’ Works.**

19. Salesforce lacks knowledge or information sufficient to admit or deny the allegations of this paragraph about Plaintiffs and third parties, and on that basis denies them. Salesforce denies any remaining allegations in this paragraph.

20. Salesforce admits that the paper *The Pile: An 800GB Dataset of Diverse Text for Language Modeling*, available at <https://arxiv.org/abs/2101.00027> was submitted to arXiv in December 2020. To the extent the allegations of this paragraph purport to quote from or reference the paper, the full text of that paper speaks for itself. To the extent the allegations of this paragraph purport to summarize or characterize or are inconsistent with the paper, Salesforce denies those allegations. Salesforce denies any remaining allegations in this paragraph.

21. Salesforce lacks knowledge or information sufficient to admit or deny the allegations of this paragraph about third parties, and on that basis denies them. Salesforce denies any remaining allegations in this paragraph.

22. To the extent the allegations of this paragraph purport to reference public statements of a third party, those statements speak for themselves. To the extent the allegations of this paragraph purport to summarize or characterize or are inconsistent with public statements, Salesforce denies those allegations. Salesforce lacks knowledge or information sufficient to admit or deny the allegations of this paragraph about a third party, and on that basis denies them. Salesforce denies any remaining allegations in this paragraph.

1 23. Salesforce lacks knowledge or information sufficient to admit or deny the
2 allegations of this paragraph about third parties, and on that basis denies them. Salesforce denies
3 any remaining allegations in this paragraph.

4 24. To the extent the allegations of this paragraph purport to quote from or reference the
5 Hugging Face website, the full text of that website speaks for itself. To the extent the allegations
6 of this paragraph purport to summarize or characterize or are inconsistent with the full text available
7 at that link, Salesforce denies those allegations. Salesforce lacks knowledge or information
8 sufficient to admit or deny the allegations of this paragraph about a third party, and on that basis
9 denies them. Salesforce denies any remaining allegations in this paragraph.

10 25. To the extent the allegations of this paragraph purport to quote from or reference the
11 Hugging Face website, the full text of that website speaks for itself. To the extent the allegations
12 of this paragraph purport to summarize or characterize or are inconsistent with the full text available
13 at that link, Salesforce denies those allegations. Salesforce lacks knowledge or information
14 sufficient to admit or deny the allegations of this paragraph about third parties, and on that basis
15 denies them. Salesforce denies any remaining allegations in this paragraph.

16 26. Salesforce lacks knowledge or information sufficient to admit or deny the
17 allegations of this paragraph about third parties, and on that basis denies them. Salesforce denies
18 any remaining allegations in this paragraph.

19 27. Salesforce lacks knowledge or information sufficient to admit or deny the
20 allegations in this paragraph about Plaintiffs, and on that basis denies them.

21 **B. Response to AI Models Require Datasets.**

22 28. Salesforce admits that LLMs are a type of AI software that generate new content in
23 response to user inputs. Salesforce denies any remaining allegations in this paragraph.

24 29. Salesforce admits that LLMs are software programs that are part of the category of
25 generative AI. Salesforce admits that LLMs are trained on a massive amount of text data, typically
26 referred to as a training dataset. Salesforce denies any remaining allegations in this paragraph.

27 30. Salesforce admits that generally, training an LLM involves a multi-stage process
28 that may include data acquisition and curation. Salesforce admits that the training process

1 sometimes involves ablation studies. Salesforce admits that pre-training is a process that uses
2 sophisticated statistical techniques to adjust the LLM's parameters (also known as weights) by
3 analyzing a massive corpus of diverse data. Salesforce admits that a pre-trained model may be
4 fine-tuned and that fine-tuning a model may involve training a model on additional data. Salesforce
5 denies any remaining allegations in this paragraph.

6 31. Salesforce admits that acquiring and curating data is part of training an LLM.
7 Salesforce denies any remaining allegations in this paragraph.

8 32. Salesforce admits that training an LLM requires a massive corpus of diverse data.
9 Salesforce admits that selection and curation of training data is part of training an LLM. Salesforce
10 lacks knowledge or information sufficient to admit or deny the allegations of this paragraph about
11 third parties, and on that basis denies them. Salesforce denies any remaining allegations in this
12 paragraph.

13 33. Salesforce admits that pretraining is part of training an LLM. Salesforce admits that
14 pre-training requires many training steps, which include using sophisticated statistical techniques
15 to gradually adjust a model's weights (or parameters) by analyzing a massive corpus of diverse
16 data. Salesforce denies the remaining allegations in this paragraph.

17 34. Salesforce admits that engineers may conduct ablation experiments as part of
18 training an LLM. Salesforce denies any remaining allegations in this paragraph.

19 35. Salesforce admits that LLMs generate new content in response to user inputs.
20 Salesforce admits that the weights of the LLMs are used to process a prompt and generate an output.
21 Salesforce denies any remaining allegations in this paragraph.

22 36. Salesforce lacks knowledge or information sufficient to admit or deny the
23 allegations of this paragraph about third parties, and on that basis denies them. Salesforce denies
24 any remaining allegations in this paragraph.

C. Response to Salesforce Creates a Library of Pirated Books to Develop its Large Language Models

37. Salesforce admits that it was founded in 1999. Salesforce admits that it provides a cloud-based customer relationship management technology to its clients. Salesforce denies any remaining allegations in this paragraph.

38. Salesforce admits that it released CodeGen 1.0 in March 2022. Salesforce admits that it submitted to arXiv a paper titled *CodeGen: An Open Large Language Model for Code with Multi-Turn Program Synthesis*, available at <https://arxiv.org/pdf/2203.13474.pdf>. To the extent the allegations of this paragraph purport to quote from or reference the paper, the full text of that paper speaks for itself. To the extent the allegations of this paragraph purport to summarize or characterize or are inconsistent with the paper, Salesforce denies those allegations. Salesforce lacks knowledge or information sufficient to admit or deny the allegations of this paragraph about third parties, and on that basis denies them. Salesforce denies any remaining allegations in this paragraph.

39. Denied.

40. Salesforce admits that it released its XGen series in June 2023. Salesforce admits that the XGen series consists of the XGen-7B-4K-base, XGen-7B-8K-base, and XGen-7B-{4K, 8K}-inst models. To the extent the allegations of this paragraph purport to quote from or reference a blog post, the full text of that post speaks for itself. To the extent the allegations of this paragraph purport to summarize or characterize or are inconsistent with the blog post, Salesforce denies those allegations. Salesforce denies any remaining allegations in this paragraph.

41. Salesforce lacks knowledge or information sufficient to admit or deny the allegations of this paragraph about Plaintiffs and third parties, and on that basis denies them. Salesforce denies any remaining allegations in this paragraph.

42. Salesforce admits that it uploaded its XGen models to GitHub as open-source models. Salesforce lacks knowledge or information sufficient to admit or deny the allegations of this paragraph about third parties, and on that basis denies them. To the extent the allegations of this paragraph purport to quote from or reference a publicly available webpage, the full text of that

1 page speaks for itself. To the extent the allegations of this paragraph purport to summarize or
2 characterize or are inconsistent with the full text available at that webpage, Salesforce denies those
3 allegations. Salesforce denies any remaining allegations in this paragraph.

4 43. Salesforce admits that Tian Xie is a former employee. To the extent the allegations
5 of this paragraph purport to quote from or reference a publicly available webpage the full text of
6 that webpage speaks for itself. To the extent the allegations of this paragraph purport to summarize
7 or characterize or are inconsistent with the full text available at that webpage, Salesforce denies
8 those allegations. Salesforce denies any remaining allegations in this paragraph.

9 44. Denied.

10 45. To the extent the allegations of this paragraph purport to quote from or reference a
11 publicly available webpage the full text of that webpage speaks for itself. To the extent the
12 allegations of this paragraph purport to summarize or characterize or are inconsistent with the full
13 text available at that webpage, Salesforce denies those allegations. Salesforce denies any remaining
14 allegations in this paragraph.

15 46. Salesforce admits that on September 7, 2023, Salesforce submitted to arXiv a paper
16 titled *XGen-7B Technical Report*, available at <https://arxiv.org/pdf/2309.03450.pdf>. To the extent
17 the allegations of this paragraph purport to quote from or reference the paper, the full text of that
18 paper speaks for itself. To the extent the allegations of this paragraph purport to summarize or
19 characterize or are inconsistent with the paper, Salesforce denies those allegations. Salesforce
20 denies any remaining allegations in this paragraph.

21 47. Salesforce admits that on December 12, 2023, Salesforce published *Developing the*
22 *New XGen: Salesforce's Foundational Large Language Models*, available at
23 [https://engineering.salesforce.com/developing-the-new-xgen-salesforces-foundational-large-](https://engineering.salesforce.com/developing-the-new-xgen-salesforces-foundational-large-language-models/)
24 [language-models/](https://engineering.salesforce.com/developing-the-new-xgen-salesforces-foundational-large-language-models/). To the extent the allegations of this paragraph purport to quote from or reference
25 a publicly available webpage the full text of that webpage speaks for itself. To the extent the
26 allegations of this paragraph purport to summarize or characterize or are inconsistent with the full
27 text available at that webpage, Salesforce denies those allegations. Salesforce denies any remaining
28 allegations in this paragraph.

1 48. Salesforce admits that it published a blog post *xGen-Sales: Next-Gen AI for Sales*
2 *Cloud*, available at <https://www.salesforce.com/blog/xgen-sales/>. To the extent the allegations of
3 this paragraph purport to quote from or reference the blog post, the full text of that post speaks for
4 itself. To the extent the allegations of this paragraph purport to summarize or characterize or are
5 inconsistent with the blog post, Salesforce denies those allegations. Salesforce denies any
6 remaining allegations in this paragraph.

7 49. Salesforce admits that on May 2, 2025, Salesforce published a blog post *xGen-small:*
8 *Enterprise-Ready Small Language Models*, available at [https://www.salesforce.com/blog/xgen-](https://www.salesforce.com/blog/xgen-small-enterprise-ready-small-language-models/)
9 [small-enterprise-ready-small-language-models/](https://www.salesforce.com/blog/xgen-small-enterprise-ready-small-language-models/). To the extent the allegations of this paragraph
10 purport to quote from or reference the blog post, the full text of that post speaks for itself. To the
11 extent the allegations of this paragraph purport to summarize or characterize or are inconsistent
12 with the blog post, Salesforce denies those allegations. Salesforce denies any remaining allegations
13 in this paragraph.

14 50. Denied.

15 51. Denied.

16 52. Denied.

17 **D. Response to Salesforce's Piracy And Use of Copyrighted Works Is Not Fair Use.**

18 53. Denied.

19 54. Denied.

20 55. Denied.

21 56. Denied.

22 57. To the extent the allegations of this paragraph purport to reference a publicly
23 available study, the full text of that study speaks for itself. To the extent the allegations of this
24 paragraph purport to summarize or characterize or are inconsistent with the full text of that study,
25 Salesforce denies those allegations. Salesforce lacks knowledge or information sufficient to admit
26 or deny the allegations of this paragraph about third parties, and on that basis denies them.
27 Salesforce denies any remaining allegations in this paragraph.
28

1 58. Salesforce admits that the statutory fair use factors are: (1) the purpose and character
2 of the use, including whether such use is of a commercial nature or is for nonprofit educational
3 purposes; (2) the nature of the copyrighted work; (3) the amount and substantiality of the portion
4 used in relation to the copyrighted work as a whole; and (4) the effect of the use upon the potential
5 market for or value of the copyrighted work. Salesforce denies any remaining allegations in this
6 paragraph.

7 59. Denied.

8 60. Denied.

9 61. Denied.

10 62. Denied.

11 63. Denied.

12 64. Denied.

13 65. Denied.

14 66. Salesforce admits that Marc Benioff is the CEO of Salesforce. Salesforce admits
15 that Mr. Benioff participated in an interview with Bloomberg. To the extent that the allegations of
16 this paragraph purport to quote or reference Mr. Benioff's interview, the full text of the interview
17 speaks for itself. To the extent that the allegations of this paragraph purport to summarize or
18 characterize or are inconsistent with the full content of that interview, Salesforce denies those
19 allegations. Salesforce denies any remaining allegations of this paragraph.

20 67. Denied.

21 **RESPONSE TO CLASS ALLEGATIONS**

22 68. This paragraph recites a definition, to which no response is required. To the extent
23 a response is deemed required, denied.

24 69. This paragraph recites a definition, to which no response is required. To the extent
25 a response is deemed required, denied.

26 70. This paragraph recites a definition, to which no response is required. To the extent
27 a response is deemed required, denied.

28 71. Denied.

1 72. Denied.

2 73. Denied.

3 74. Denied.

4 75. Denied.

5 **RESPONSE TO CAUSE OF ACTION**

6 **RESPONSE TO COUNT 1**

7 **Direct Copyright Infringement (17 U.S.C. § 501) against Defendant**

8 76. This paragraph incorporates by reference the preceding allegations, and therefore no
9 response is required. To the extent a response is deemed required, Salesforce incorporates by
10 reference its responses to paragraphs 1–75 as if fully set forth herein.

11 77. Denied.

12 78. Denied.

13 79. Denied.

14 80. Denied.

15 81. Denied.

16 82. Salesforce lacks knowledge or information sufficient to admit or deny the
17 allegations of this paragraph about Plaintiffs and third parties, and on that basis denies them.
18 Salesforce denies any remaining allegations in this paragraph.

19 83. Denied.

20 84. Denied.

21 85. Denied.

22 86. Denied.

23 87. Denied.

24 88. Denied.

25 89. Denied.

26 90. Denied.

RESPONSE TO DEMAND FOR JUDGMENT

In response to the demand for judgment, Salesforce denies that Plaintiffs are entitled to the requested relief, or to any relief whatsoever.

RESPONSE TO JURY TRIAL DEMANDED

With respect to the jury demand contained in the Consolidated Complaint, Salesforce states that no response is required but likewise demands a trial by jury.

AFFIRMATIVE DEFENSES

In further answer to the allegations made by Plaintiffs in the Consolidated Complaint, Salesforce asserts the following affirmative defenses. Salesforce does not concede that it has the burden of proof on the defenses listed below.

FIRST AFFIRMATIVE DEFENSE

Plaintiffs' claim is barred, in whole or in part, because Plaintiffs do not own or hold exclusive rights under 17 U.S.C. § 106 or any copyright law over each work that was allegedly infringed by Salesforce.

SECOND AFFIRMATIVE DEFENSE

Plaintiffs' claim is barred, in whole or in part, because the copyright registrations purporting to cover some or all the works in dispute are invalid or otherwise do not satisfy the requirements of 17 U.S.C. §§ 411–412.

THIRD AFFIRMATIVE DEFENSE

To the extent there is copying of copyrightable expression, that copying constitutes fair use.

FOURTH AFFIRMATIVE DEFENSE

Plaintiffs' claim fails in whole or in part because the alleged use was validly licensed by express or implied license.

FIFTH AFFIRMATIVE DEFENSE

Plaintiffs' claim is barred, in whole or in part, by the doctrine of estoppel.

SIXTH AFFIRMATIVE DEFENSE

Plaintiffs' claim is barred, in whole or in part, to the extent they claim rights to elements of works or to works which are not protectable under copyright law, or that are in the public domain,

1 lack the requisite originality, are unregistered, or are works to which copyright protection has been
2 abandoned.

3 **SEVENTH AFFIRMATIVE DEFENSE**

4 Plaintiffs' claim of copyright infringement is barred or limited by the doctrine of merger.

5 **EIGHTH AFFIRMATIVE DEFENSE**

6 Plaintiffs' claim of copyright infringement is barred or limited by the idea/expression
7 dichotomy.

8 **NINTH AFFIRMATIVE DEFENSE**

9 Plaintiffs' claim of copyright infringement is barred or limited because the material in which
10 Plaintiffs claim copyright constitutes "*scènes à faire*."

11 **TENTH AFFIRMATIVE DEFENSE**

12 To the extent there is actionable copying of copyrightable expression, that copying is *de*
13 *minimis*.

14 **ELEVENTH AFFIRMATIVE DEFENSE**

15 Plaintiffs' claim of copyright infringement is barred or limited because the material in which
16 Plaintiffs claim copyright constitutes unprotectable facts.

17 **TWELFTH AFFIRMATIVE DEFENSE**

18 Plaintiffs' claim for damages is barred or limited by the Due Process Clause.

19 **THIRTEENTH AFFIRMATIVE DEFENSE**

20 Salesforce reserves all affirmative defenses under Rule 8(c) of the Federal Rules of Civil
21 Procedure, and any other defenses at law or in equity, that may now exist or in the future become
22 available based on discovery and further factual investigation.
23
24
25
26
27
28

1 Dated: February 19, 2026

Respectfully submitted,

2 MORRISON & FOERSTER LLP

3
4 By: /s/ Joseph C. Gratz

5 Joseph C. Gratz
Tiffany Cheung

6 *Attorneys for Defendant*
7 SALESFORCE, INC.