

# Sarah Silverman sues OpenAI and Meta claiming AI training infringed copyright

Publication Date: 2023-07-10

Author: Dan Milmo

Section: Technology

Tags: Artificial intelligence (AI), Sarah Silverman, ChatGPT, Meta, OpenAI, Computing, Intellectual property, news

Article URL: <https://www.theguardian.com/technology/2023/jul/10/sarah-silverman-sues-openai-meta-copyright-infringement>



The US comedian and author Sarah Silverman is suing the ChatGPT developer OpenAI and Mark Zuckerberg's Meta for copyright infringement over claims that their artificial intelligence models were trained on her work without permission. Silverman has filed the suits along with two authors, Christopher Golden and Richard Kadrey, in which they claim the AI models developed by OpenAI and Meta used their work as part of their training data. Tools like ChatGPT, a highly popular chatbot, are based on large language models that are fed vast amounts of data taken from the internet in order to train them to give convincing responses to text prompts from users. The lawsuit against OpenAI claims the three authors "did not consent to the use of their copyrighted books as training material for ChatGPT. Nonetheless, their copyrighted materials were ingested and used to train ChatGPT." The lawsuit concerning Meta claims that "many" of the authors' copyrighted books appear in the dataset that the Facebook and Instagram owner used to train LLaMA, a group of Meta-owned AI models. The suits claim the authors' works were obtained from "shadow library" sites that have "long been of interest to the AI-training community". The OpenAI suit includes exhibits claiming that, when prompted, it summarised three books: Silverman's *The Bedwetter*, Ararat by Golden, and Kadrey's *Sandman Slim*. The Meta suit cites multiple works by Kadrey and Golden, alongside *The Bedwetter*, and flags a Meta paper that indicates LLaMA's training datasets included material taken from shadow libraries the suit describes as "flagrantly illegal". The lawyers representing the three authors, Joseph Saveri and Matthew Butterick, have written that since the release of ChatGPT they have been hearing from writers, authors and publishers expressing concern about the tool's "uncanny" ability to generate text similar to copyrighted material. Saveri and Butterick are also representing two more US authors, Mona Awad and Paul Tremblay, who have filed a separate class action lawsuit against OpenAI claiming ChatGPT was trained on their work without the writers' consent. Getty Images, the stock photo company, is suing the company behind AI image generator Stable Diffusion over alleged breach of copyright. Saveri and Butterick are representing three artists – Sarah Andersen, Kelly McKernan and Karla Ortiz – in a lawsuit against image generators Stability AI, Midjourney, and DeviantArt. The lawsuits over AI models also extend to the false answers, or "hallucinations", they can be prone to issuing. A radio host in the US state of Georgia is suing OpenAI for defamation after it falsely stated he had been accused of fraud. OpenAI and Meta have been approached for comment.