# Authors sue Meta over alleged use of pirated materials for AI training



In a significant legal controversy, notable authors including Sarah Silverman and Ta-Nehisi Coates have initiated a lawsuit against Meta, alleging that the tech giant utilised pirated materials to develop its AI models. The focal point of the accusations revolves around claims that Meta's Llama AI models were trained using content from LibGen, a website known for hosting pirated eBook copies.

The legal proceedings uncover a broader issue surrounding copyright infringement and the ethical implications of using stolen material in AI training. Meta has been under scrutiny, as LibGen frequently faces litigation and financial penalties due to its distribution of copyrighted works without permission.

During his deposition, Meta's CEO, Mark Zuckerberg, acknowledged that the company utilised LibGen for training at least one iteration of its Llama models, despite the potential legal ramifications that such a practice entails. When asked about his familiarity with LibGen, Zuckerberg stated, “I haven’t heard of that specific thing.” He mentioned that while it is crucial for Meta to navigate the complexities of using copyrighted materials in AI training cautiously, he remains in support of employing extensive datasets during the training process.

This lawsuit not only highlights the ongoing debates surrounding AI automation and the ethics of data usage but also raises questions about the future trends in AI technology and how businesses navigate legal challenges in an increasingly automated landscape. As industry forecasts suggest a rising reliance on AI for various business practices, the outcomes of this case may have significant implications for how companies approach data sourcing and the potential risks associated with automation in the future.

Source: [Noah Wire Services](https://www.noahwire.com)

## Bibliography

* <https://ai.plainenglish.io/metas-ai-accused-of-theft-lawsuits-loom-over-copyright-violations-4049377cd50b?gi=0765d6725649> - Corroborates the allegations that Meta trained its AI models on pirated materials, including books from LibGen, and the subsequent lawsuits and ethical debates.
* <https://www.bitdegree.org/crypto/news/meta-faces-heat-authors-claim-ai-trained-on-pirated-books> - Supports the claim that authors, including Sarah Silverman, have accused Meta of using pirated books from LibGen to train its AI models, and that Mark Zuckerberg approved the use of this dataset.
* <https://www.bitdegree.org/crypto/news/meta-faces-heat-authors-claim-ai-trained-on-pirated-books> - Details the lawsuit filed by authors alleging Meta used their books without permission and the concealment of critical information until a discovery deadline.
* <https://www.siliconrepublic.com/business/meta-llama-lawsuit-copyright-authors> - Confirms that Meta allowed its Llama AI to commit copyright infringement on pirated data and upload it for commercial gain, and the removal of copyright management information (CMI) from the works.
* <https://www.siliconrepublic.com/business/meta-llama-lawsuit-copyright-authors> - Provides evidence that Meta downloaded content from LibGen for Llama’s training and acted as a 'leech' on pirated data, and a Meta representative admitted to uploading pirated files.
* <https://ai.plainenglish.io/metas-ai-accused-of-theft-lawsuits-loom-over-copyright-violations-4049377cd50b?gi=0765d6725649> - Highlights the broader issue of copyright infringement and the ethical implications of using stolen material in AI training, as well as potential financial and reputational damages for Meta.
* <https://www.bitdegree.org/crypto/news/meta-faces-heat-authors-claim-ai-trained-on-pirated-books> - Mentions internal messages showing hesitation among engineers about downloading pirated content and the continued distribution of pirated content through torrenting networks.
* <https://www.aiaaic.org/aiaaic-repository/ai-algorithmic-and-automation-incidents/library-genesis-sued-for-staggering-copyright-infringement> - Details the litigation and financial penalties faced by LibGen due to its distribution of copyrighted works without permission.
* <https://ai.plainenglish.io/metas-ai-accused-of-theft-lawsuits-loom-over-copyright-violations-4049377cd50b?gi=0765d6725649> - Discusses the allegations that Mark Zuckerberg personally approved the decision to use copyrighted materials for training and the potential implications for AI ethics and regulation.
* <https://www.siliconrepublic.com/business/meta-llama-lawsuit-copyright-authors> - Notes the judge's decision to deny Meta's request to redact portions of the new filing, indicating that Meta's request was designed to avoid negative publicity rather than protect sensitive business information.
* <https://www.bitdegree.org/crypto/news/meta-faces-heat-authors-claim-ai-trained-on-pirated-books> - Supports the claim that the lawsuit raises questions about the future trends in AI technology and how businesses navigate legal challenges in an increasingly automated landscape.
* <https://www.analyticsinsight.net/news/zuckerbergs-youtube-defense-sparks-debate-in-metas-ai-copyright-battle> - Please view link - unable to able to access data