

Journals take up arms against AI-written text

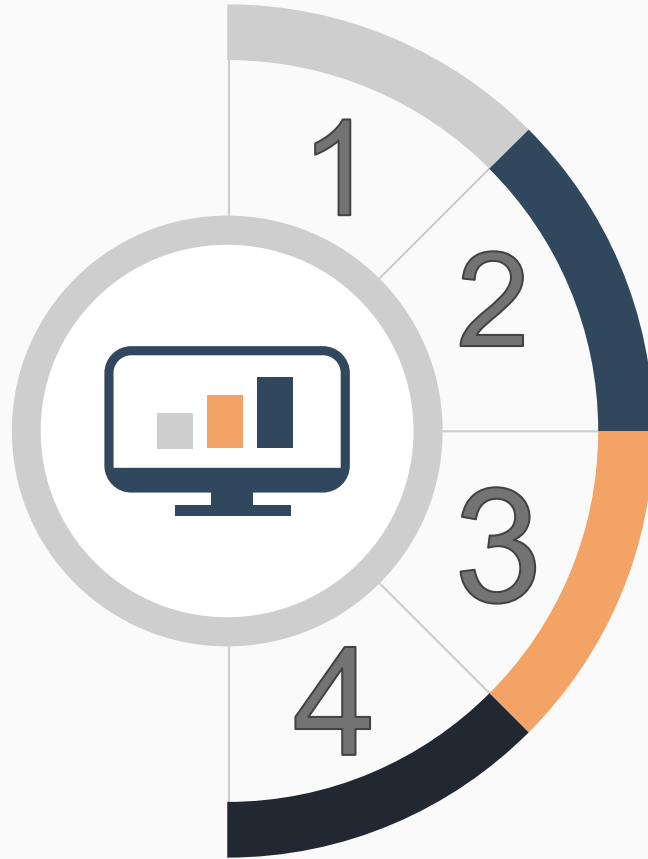
Author: Jeffrey Brainard
Presented by: Kaiden Liu
CS846 - Software Engineering for Big
Data and AI



Outline

- Benefits of Generative AI
- Controversies in Generative AI
- Influences of ChatGPT in Scientific Research
- Concerns of ChatGPT in Scientific Research
- Approaches to AI-Generated Content in Journals
- Approaches to using AI in Research

Benefits of Generative AI



01

Content Generation

02

Human-like

03

Knowledge Retrieval

04

Accessible to Public

Controversies in Generative AI

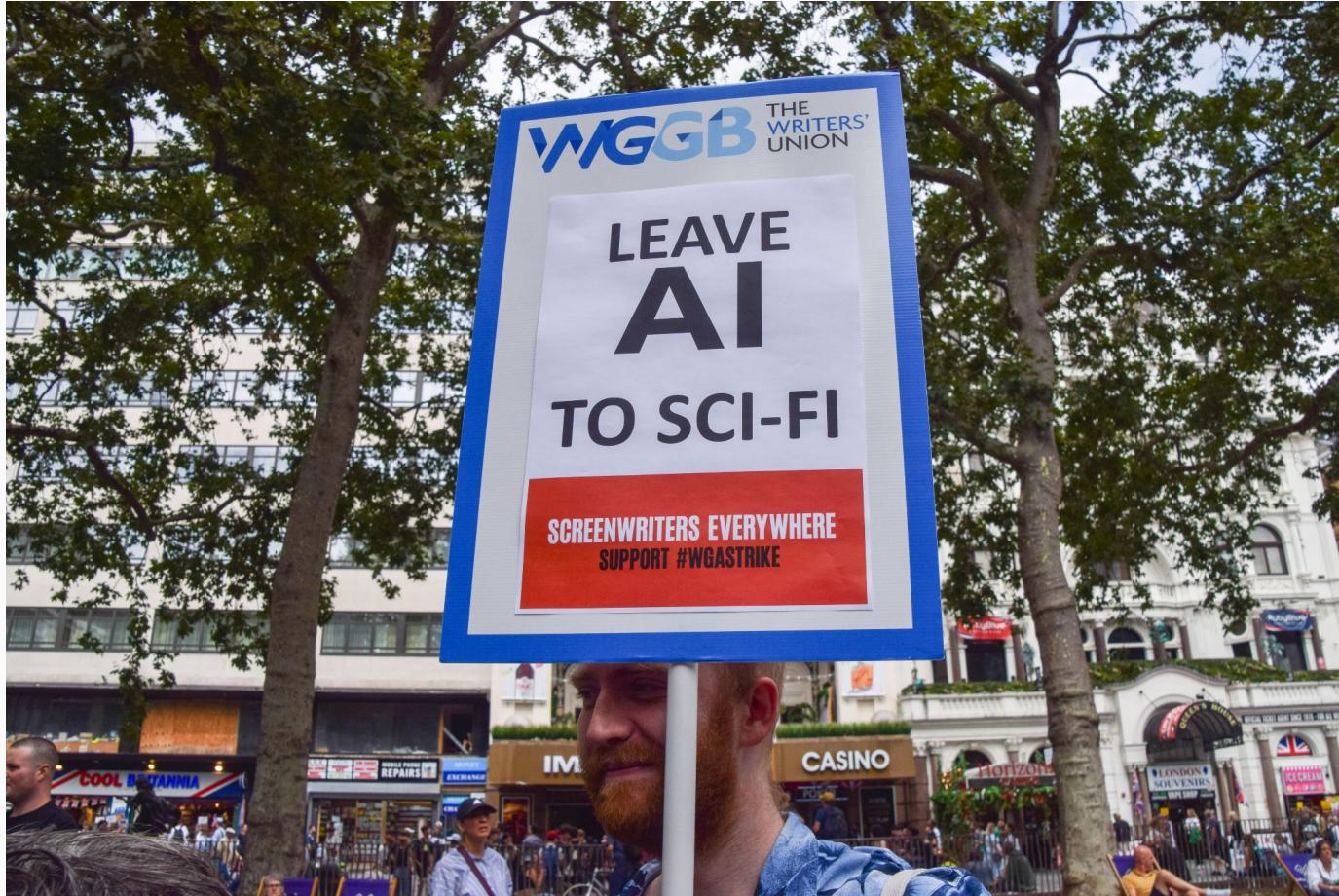


Image source: <https://newrepublic.com/article/174650/hollywoods-fight-ai-will-affect-us>

Controversies in Generative AI (cont'd)

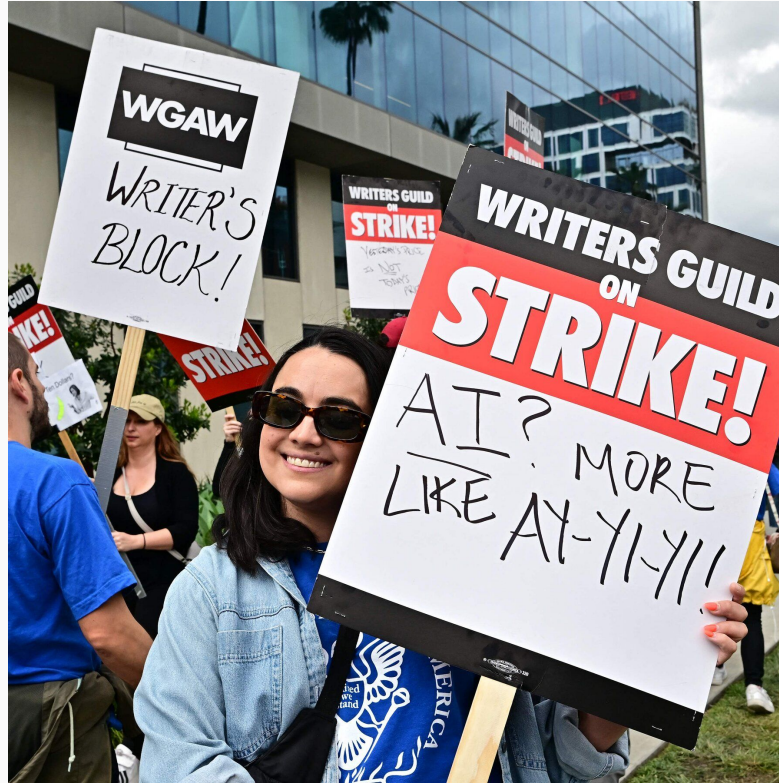


Image source: <https://www.nytimes.com/2023/05/10/arts/television/writers-strike-artificial-intelligence.html>

Controversies in Generative AI (cont'd)



NEWS

[Home](#) | [War in Ukraine](#) | [Climate](#) | [Video](#) | [World](#) | [US & Canada](#) | [UK](#) | [Business](#) | [Tech](#) | [Science](#)

[Tech](#)

Sarah Silverman sues OpenAI and Meta

🕒 12 July

Controversies in Generative AI (cont'd)



PRESS RELEASES

The Authors Guild, John Grisham, Jodi Picoult, David Baldacci, George R.R. Martin, and 13 Other Authors File Class-Action Suit Against OpenAI

Artificial Intelligence

September 20, 2023

Source: <https://authorsguild.org/news/ag-and-authors-file-class-action-suit-against-openai/>

Controversies in Generative AI (cont'd)

Allegation to OpenAI

-14-

Case 1:23-cv-08292 Document 1 Filed 09/19/23 Page 15 of 47

109. In short, OpenAI admits it needs¹⁷ and uses¹⁸ “large, publicly available datasets that include copyrighted works”¹⁹—and specifically, “high-quality”²⁰ copyrighted books—to “train” its LLMs; pirated sources of such “training” data are readily available; and one or more of these sources contain Plaintiffs’ works.

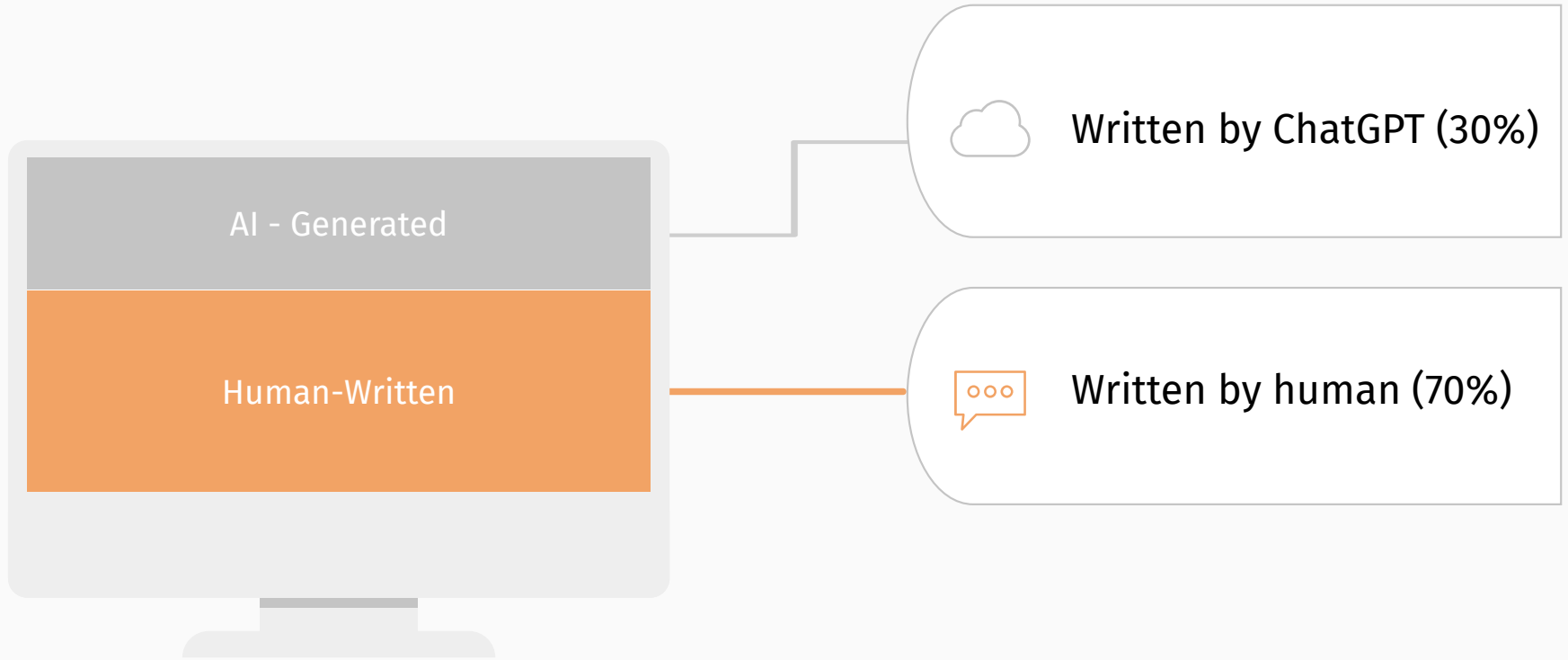
110. Defendants knew that their “training” data included texts protected by copyright but willfully proceeded without obtaining authorization.

Controversies in Generative AI (cont'd)

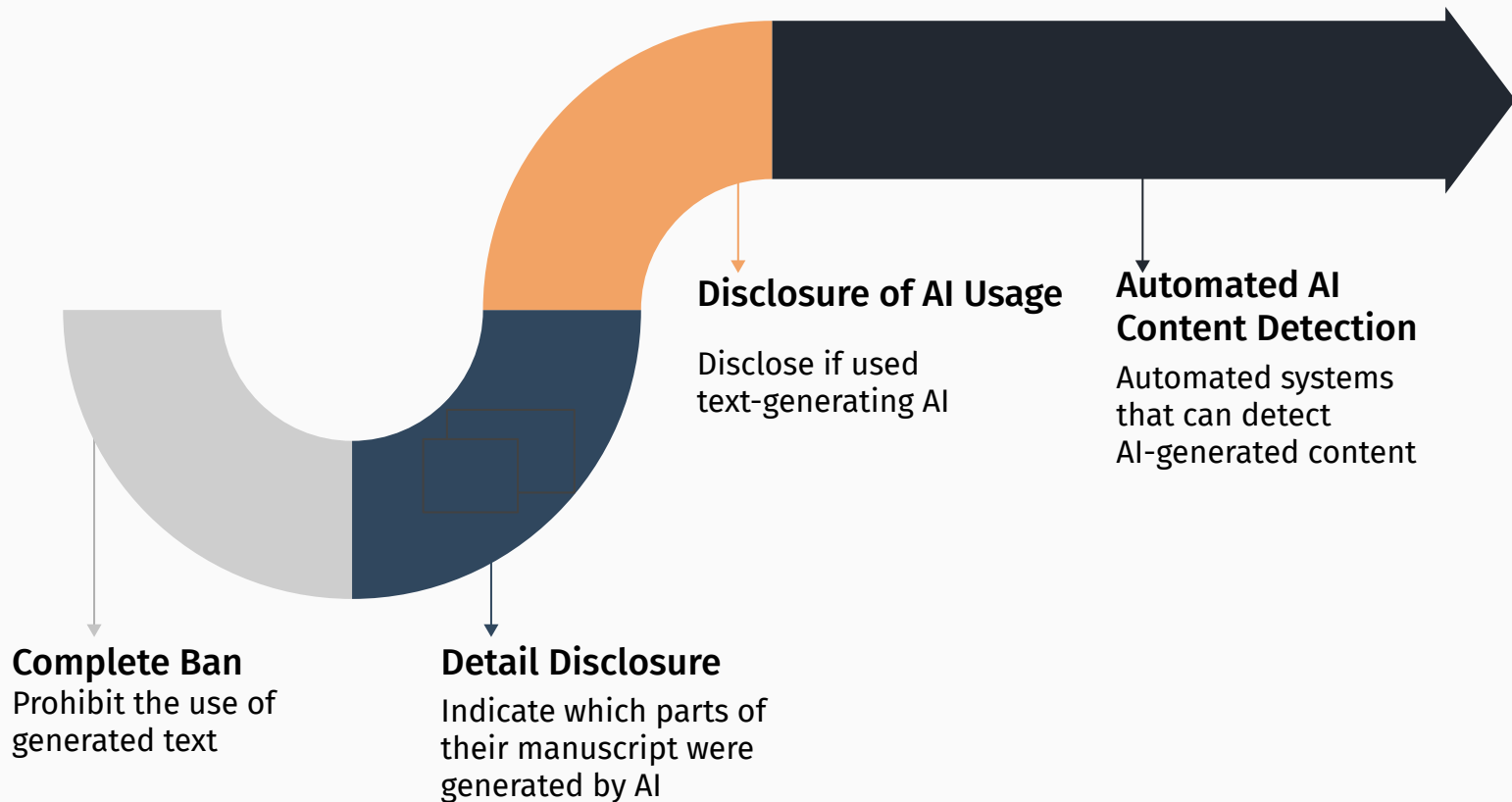
OpenAI's Defense

- 1) LLM is transformative, that it changes and repurposes the data it's trained on in a unique way
- 2) Intention of copyright laws: promoting scientific and technological progress

Influence of ChatGPT in Scientific Research



Influence of ChatGPT in Scientific Research (cont'd)



Concerns of ChatGPT in Scientific Research

- Hallucination: Produce outputs that are not real
- Bias: Toward established scientific ideas and hypotheses contained in the train set
- AI-giarism: Authors presenting the output of the AI as if they wrote them

Concerns of ChatGPT in Scientific Research

Examples

- Meta took down Galactica after days after its launch
- Sparrow by Google Deepmind, found up to 20% of its responses contained errors [4]

Concerns of ChatGPT in Scientific Research (cont'd)



>wiki article on bears living in space

Bears living in space

Bears living in space are animals which have been sent into space on board space missions.

History

Soviet Union

The Soviet Union was the first country to put a bear into space. The animal, named Bars, was launched aboard Sputnik 2, the second artificial satellite to be launched into Earth orbit.

Bars was a Karelian bear (a small, sleek brown bear), which was selected by S. P. Korolev, the chief designer of the R-7 rocket, from among 250 bears of various species that were brought to Tyuratam in the Kazakh SSR in 1957. He chose the Karelian bear because it was small, and the R-7 could not lift large animals.

Bars was a three-year-old female bear who weighed 40 kg (88 lb). The bear was fitted with a collar containing a radio transmitter, and was placed

Approaches to using AI in Research



OpenAI's Classifier

Disabled due to
bad performance



Watermark

Detectable
secret code



Stanford

DetectGPT by
ranking



TurnItIn

Detector trained on
academic writing

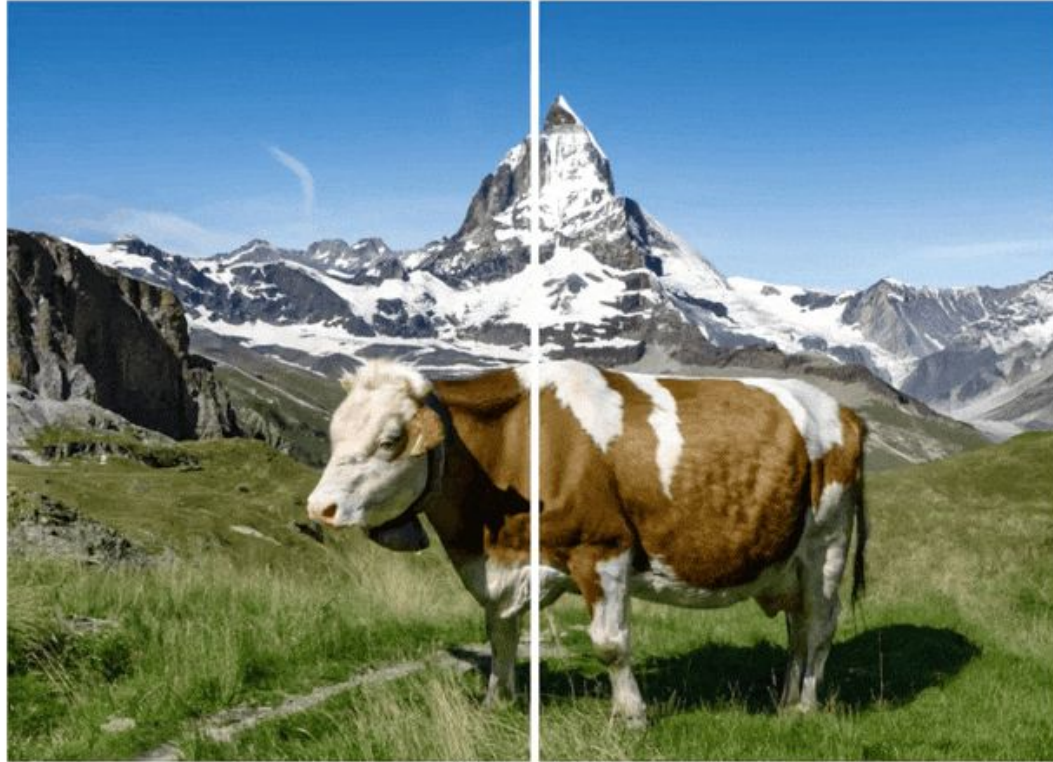


scite.ai(Domenic
Rosati)

Factual accuracy



Approaches to using AI in Research (Watermark Example)



Watermarked

Non-watermarked

Conclusion

- 1) Ethical concerns arise when AI-generated content is used
- 2) Scientific journals are not clear about using AI generated contents in scholarly submissions
- 3) Current tools to detect AI-generated content are not fully reliable

“I can see scenarios where there will be causes for concern, but then I also see a tremendous amount of opportunity with these types of tools”

Magdalena Skipper, editor-in-chief of Nature

References

1. "The Authors Guild, John Grisham, Jodi Picoult, David Baldacci, George R.R. Martin, and 13 Other Authors File Class-Action Suit Against OpenAI," *The Authors Guild*, Sep. 20, 2023. [Online]. Available: <https://authorsguild.org/news/ag-and-authors-file-class-action-suit-against-openai/>.
2. B. Masse, "OpenAI seeks to dismiss majority of Sarah Silverman's and authors' claims in ChatGPT lawsuits," *VentureBeat*, Aug. 29, 2023. [Online]. Available: <https://venturebeat.com/ai/openai-seeks-to-dismiss-majority-of-sarah-silvermans-and-authors-claims-in-chatgpt-lawsuits/>.
3. C. Vallance, "Sarah Silverman sues OpenAI and Meta," *BBC*, Jul. 12, 2023. [Online]. Available: <https://www.bbc.com/news/technology-66164228>.
4. J. BRAINARD "As scientists explore AI-written text, journals hammer out policies," *Science*, Feb. 22, 2023. [Online]. Available: <https://www.science.org/content/article/scientists-explore-ai-written-text-journals-hammer-policies>
5. Case 3:23-cv-03416, Document 1, at 6 ll. 30-31 (filed July 7, 2023).
6. Case 1:23-cv-08292 Document 1 Filed 09/19/23

Thank You!

Any questions?

Interesting points found in the legal complaint document

87. ChatGPT itself admits as much. In response to a query submitted to it in January 2023, the chatbot responded,

It is possible that some of the books used to train me were under copyright. However, my training data was sourced from various publicly available sources on the internet, and it is likely that some of the books included in my training dataset were not authorized to be used. ... If any copyrighted material was included in my training data, it would have been used without the knowledge or consent of the copyright holder.

1 contains long stretches of contiguous text, which allows the generative model to learn to condition on
2 long-range information.” Hundreds of large language models have been trained on BookCorpus,
3 including those made by OpenAI, Google, Amazon, and others.

4 30. BookCorpus, however, is a controversial dataset. It was assembled in 2015 by a team of
5 AI researchers for the purpose of training language models. They copied the books from a website
6 called Smashwords that hosts self-published novels, that are available to readers at no cost. Those
7 novels, however, are largely under copyright. They were copied into the BookCorpus dataset without
8 consent, credit, or compensation to the authors.

9 31. OpenAI also copied many books while training GPT-3. In the July 2020 paper
10 introducing GPT-3 (called “Language Models are Few-Shot Learners”), OpenAI disclosed that 15% of
11 the enormous GPT-3 training dataset came from “two internet-based books corpora” that OpenAI
12 simply called “Books1” and “Books2”.