

Do Readers Accept AI as an Author?

A Preliminary Analysis of Reddit Discussions on AI-Generated Book Content

Author(s) anonymized for peer review

Increasing AI-generated content in books has sparked concern and pushback, as many believe these books harm the interest of human authors and erode trust in the literary marketplace (Goodfellow, 2024; Sato, 2023). Such concerns have prompted real-world action in the publishing industry, such as certifications for “human-author-only” books (Authors Guild, 2025; Loffhagen, 2025). On platforms like Goodreads, some readers have called for boycotts of books with AI-generated covers—viewed as unethical derivatives of human labor—while others argue that such backlash unfairly harms authors, particularly when production decisions are made by publishers (Goodreads, 2023). These tensions and debates call for critical investigation into the expansion of generative book content, and the perceived sociotechnical challenges it has introduced.

To this end, this research explores how books with AI-generated content have been discussed and contested online, through a case study based on data collected from Reddit, a social platform for community-based and topic-driven discussion with abundant textual data for our research questions (Proferes et al., 2021). We combined computational text analysis with qualitative investigation to analyze the data.

Figure 1 shows our workflow, from data collection to mixed-method analysis. We collected 775,640 comments in English via Reddit’s public API using 16 book-related search terms. Figure 2 presents the search terms used, and their corresponding numbers of posts collected. With both LDA topic modeling (McCallum, 2002) and dictionary-based keyword matching, we identify 386 posts about books with AI-generated content. These selected posts were tagged as “AI book posts” while the rest of the posts were tagged as “general book posts”. Next, we conducted VADER sentiment analysis (Hutto & Gilbert, 2014) to compare the sentiment of AI book posts and general book posts, and conducted qualitative analysis of content (Wildemuth, 2016) on the AI book posts.

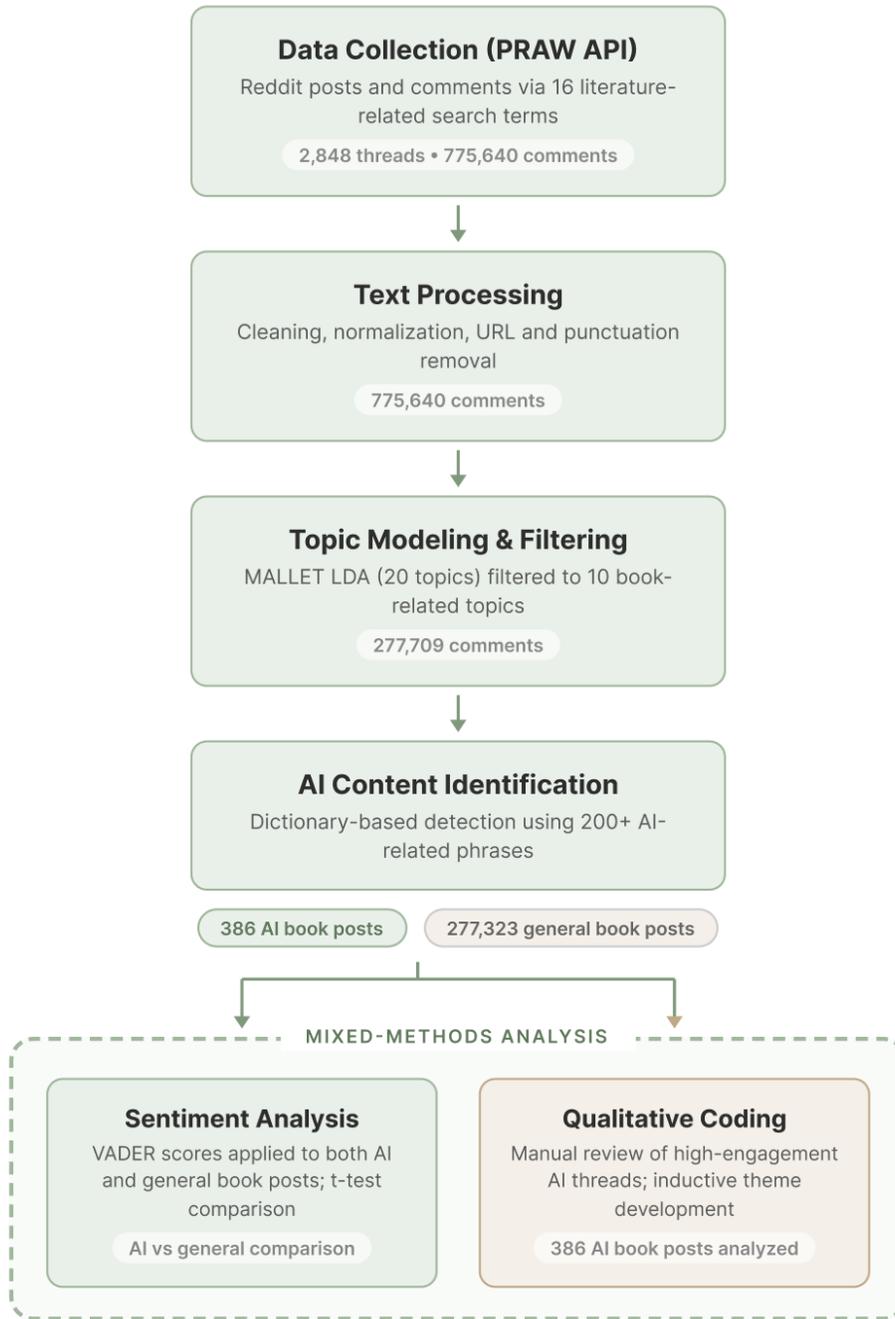


Figure 1. Overall workflow for data collection and processing

SEARCH TERM	THREADS	COMMENTS
storygraph	195	86,302
fable	182	79,325
audiobooks	165	79,143
librarything	185	78,383
booktok	177	78,257
lovelybooks	151	75,458
audible	156	77,143
kindle	200	60,218
goodreads	178	47,089
book club	182	40,353
book review	123	38,635
publishers weekly	200	11,902
book podcast	200	11,590
bookstagram	200	7,116
kirkus reviews	200	3,331
book riot	202	1,395

TOTAL THREADS **2,848** TOTAL COMMENTS **775,640**

Figure 2. Distribution of collected Reddit text data across 16 search terms

Quantitative analysis revealed that AI book posts (n=386) showed more negative sentiment than general book posts (n=277,323), with mean VADER scores of 0.115 and 0.180, respectively ($t = -2.57$, $p = 0.01$). While this difference is statistically significant, the effect size is small (Cohen's $d = -0.131$). To examine whether this result was influenced by the substantial imbalance in dataset sizes, we computed sentiment scores for 1,000 samples of 386 randomly selected general book posts. In these tests, AI book posts still showed more negative sentiment, but only 28.7% of tests reached statistical significance. This suggests that the original result may be partially influenced by the imbalanced sizes of data. For meaningful comparisons, we plan to enlarge the existing AI book posts datasets by collecting relevant posts from other platforms. We also plan to conduct sentiment analysis with word embeddings as VADER's dictionary-based approach may miss nuanced affective expressions such as sarcasm.

We complemented our quantitative analysis with qualitative analysis and identified four major clusters of topics: (1) **human value**, where readers raise ethical questions about what counts as human-made art and creative works; (2) **labor concerns**, reflecting concerns about AI replacing creative professionals such as book authors, editors, and illustrators; (3) **authenticity**, with AI writing viewed as imitation rather than genuine creativity; and (4) **calls for disclosure**, with many insisting AI-generated books must be clearly labeled to enable informed and ethical consumption. Figure 3 presents two exemplary quotes for each cluster. Across these clusters, most of the posts challenge the legitimacy of AI authorship and express ethical concerns surrounding AI-generated content.

It is important to note that even negative critiques of AI-generated book content tend to be more evaluative and discussive, rather than completely negative or hostile. Additionally, a few readers expressed openness or neutrality to AI books. For example, one user posted that “Using AI to assist in writing is very much a ‘you get out of it what you put’ sort of relationship”, and another user stated, “I am not against an author having an AI editor”.

HUMAN VALUE

"A lot of books have weight precisely because there is a human behind it."

"Why are we using AI to do the cool, artistic stuff that makes us human and life enjoyable?"

LABOR CONCERNS

"As a full time author I'm worried about the longevity of my career."

"I am also worried that writers are endangered in a way that their work and style is easier to copy or falsificate."

AUTHENTICITY

"If you couldn't be bothered to write it, why would I bother to read it?"

"I'm not sure AI will ever be able to innovate in art, but it will definitely be able to produce convincing impersonations of other people's art."

"From a reader's perspective, there's not going to be any depth to an AI novel — it could entertain, but nothing more."

CALLS FOR DISCLOSURE

"I think if it's AI written, it should be clearly said on the cover. I wouldn't read it at all."

"I will continue to disclose my use and the amount because I want to show partnership can happen with tools and they can be used responsibly."

Figure 3. Representative quotes across four major clusters of topics in AI book posts

Our mixed-method analysis of Reddit data reveals sustained concerns in discussions about AI-generated book content, with a few users viewing AI as a potentially useful tool. To deepen and enrich this preliminary investigation, we plan to collect posts on generative book content across platforms and user bases. In particular, recognizing the cultural and linguistic dependencies of book-related discourse (Hu et al., 2025), we aim to include book discourse in languages other than English and generated by geo-culturally diverse readerships.

References

- Authors Guild. (2025). Authors Guild launches "Human Authored" certification. <https://authorsguild.org/human-authored/>.
- Goodfellow, P. (2024). The Distributed Authorship of Art in the Age of AI. *Arts*, 13(5), 149. <https://doi.org/10.3390/arts13050149>.
- Goodreads. (2023). Fractal Noise (Book page on Goodreads) [Social Reviewing Platform].
- Goodreads. <https://www.goodreads.com/book/show/62711641-fractal-noise>.
- Hu, Y., Underwood, T., Layne-Worthey, G., & Downie, J. S. (2025). Comparative analysis of classics book review data created by users across Douban and Goodreads. *Digital Scholarship in the Humanities*, fqaf084.
- Hutto, C. J., & Gilbert, E. (2014). VADER: A parsimonious rule-based model for sentiment analysis. *Proceedings of the AAAI Conference on Web and Social Media*, 8(1), 216-225.
- Jiang, S. (2024). Generative-AI and the disappearance of human creative labour. *Media International Australia*, 191(1), 200-217.
- Loffhagen, E. (2025). *Certified organic and AI-Free: New stamp for human-written books launches*. The Guardian. <https://www.theguardian.com/books/2025/oct/15/books-by-people-for-people-publishers-launch-certification-human-written-ai>
- McCallum, Andrew Kachites. MALLET: A Machine Learning for Language Toolkit. <http://mallet.cs.umass.edu>, 2002.
- Proferes, N., et al. (2021). Studying Reddit: A systematic overview. *Social Media and Society*, 7(2), 1-15.
- Sato, M. (2023). AI-generated fiction is flooding literary magazines—but not fooling anyone. *The Verge*, 25.
- Wildemuth, B. M. (Ed.). (2016). Applications of social research methods to questions in information and library science. Bloomsbury Publishing USA.