

Harvesting Library Labour in an AI World

The Grim Reaping of Library Work(ers) and What We Can Do About It

Tim Ribaric - Brock University

Cecile Farnum - Toronto Metropolitan University



Agenda

- How Tim and Cecile entered this conversation
- The Library Labour Context
- The ‘Death of the Author’ Transition to ‘The Blob Theory’
- Research Life Cycle implications
- What do we do about this (if anything?)

How Did Tim and Cecile
Enter This Conversation?

—

Scholars Portal



OCUL Artificial Intelligence/Machine Learning Report and Strategy

OCUL Task Force on Machine Learning/AI

Mark Asberg, Queen's University
Talia Chung, University of Ottawa
Scott Gillies, Wilfrid Laurier University
Vivian Lewis, McMaster University
Mark Robertson, Toronto Metropolitan University
Catherine Steeves, Western University (Chair)
Amy Greenberg, OCUL
Kate Davis, Scholars Portal
Michael Ridley, University of Guelph

Final Report – for Public Distribution

May 2024

Relevant excerpts that freaked (some) people out

LibraryH3lp, OCUL's Ask a Librarian chat reference platform, is in the early stages of developing a separate VR chat product

OCUL has been asked to provide Ask a Librarian logs for training and to work with LibraryH3lp as beta testers of their VR product.

Objective: A chatbot, owned by LibraryH3lp, that has experience with university website data. The objective here is not to begin using a chatbot, but to participate in its development and then perform an analysis of its ability. This tool must be **human mediated**, give consistently high-quality responses, and have some type of information literacy objectives built into its process.

Origin Story Timeline

March 11 2024 - post by Jane Schmidt in the CAUTLib listserv, responding to the OCUL Task Force on Machine Learning/AI [draft report](#) (final report [here](#)):

“...I am curious as to **why we are so eager to outsource one of our core functions to bots, on our own dime. **Invest in people, please. Not bots.**”**

Tim's response: “We can **collectively (and rightly) say no to having our labour subsumed...**but what is the alternative? Perhaps a meaningful response will be to do what the Actors Guild did and explicitly bargain language around AI. Strong collective agreements (in my mind) are what are keeping the boat afloat...**protections against AI might be the next big win labour has to make.**”



Jane Schmidt Mar 12, 1:36 PM

I was intrigued by Tim's suggestion to follow the Actor's Guild and build language to protect us from being replaced by AI!

Mar 12, 1:37 PM

Yes, I agree, that's a great idea...it would be interesting to get a hold of that collective agreement language, and somehow suggest what might work in our context...

SEPTEMBER 2024

Commentary / AI protections for librarian work: put it in the collective agreement?



**By Tim Ribaric and
Cecile Farnum**

There is a chorus in the media focused on the impact artificial intelligence will have on people's working lives. Pundits are

attempting to make calls about the future of work, motivated by the rapid and seemingly sudden explosion of generative AI tools that kicked off in the fall of 2022, when ChatGPT captured the imagination of the world.

Human in the Loop Defense

Human-in-the-loop refers to learning models that require human interaction, allowing humans to modify the output of the system. This approach involves human input in simulations, enabling the identification of model shortcomings that may not be evident before real-world testing.

AI generated definition based on the book: [Artificial Intelligence and Deep Learning in Pathology, 2021](#)

“Human-in-the-loop machine learning is a set of strategies for combining human and machine intelligence in applications that use AI”.

Monarch, R. (2021). Human-in-the-loop machine learning: active learning and annotation for human-centered AI. Manning Publications.

Why is librarian work vulnerable?

- Our work is often done ‘out in the open’. E.g. instructional content + open repositories / archives - https://learn.library.torontomu.ca/LAW535/cases_notingup
- The reference interview, the question-answer engagement between librarian and user, can be mimicked in a chatbot interaction (we now live in a *post-reference* world)
- Libraries are perpetually cash-strapped, and there is appeal in purchasing a product or tool vs. investing in the labour of people
- Our piece imagines the not-to-distant future, where we may need to bargain protections to our labour, similar to the language won by the Writers and Screen Actors Guilds in the U.S.

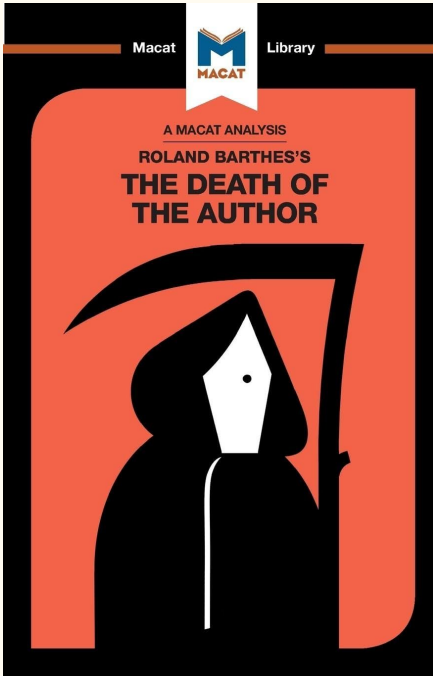
The Labour Behind Scholarly Publishing

- Often invisible
- Scholarly communication support at libraries can be uneven, and precarious
- Often understaffed or emerging areas of support in libraries
- Yoonhee Lee's 'Invisible Labor of Digitizing' paper
- Let's Talk About the Volunteers in Scholarly Publishing - many aspects of research life cycle involve volunteerism - editorial review, peer reviews, presenting research at conferences, membership to societies, etc.

AI and Authorship

—

The Death of the Author



- Seminal work from Roland Barthes
- Shout out to Johnny Soraker, AI Ethics Lead @ Google
- Barthes - the concept of authorship is 'dead' - Interpretation of the text is more important than the intentions of the author
- Use of grim reaper image, mentioned in our abstract

The Death of the Author in the Age of AI

- John Potts *Near Death of the Author: Creativity in the Internet Age* extends Barthes' argument to the AI age
- AI as author? Will it supplant human authorship? What about copyright?
- NaNoMoWri AI controversy
- *The Atlantic* article about LibGen, where Meta used a database of pirated books/articles to train its AI

Search LibGen, the Pirated-Books Database That Meta Used to Train AI

Millions of books and scientific papers are captured in the collection's current iteration.

By Alex Reisner



Machinations already underway



The screenshot shows the homepage of The Bookseller website. The header includes the site's name, a tagline, a search bar, a login link, and a navigation menu with categories like News, Rights, Comment, Bestsellers, Books, Spotlight, Events, Jobs, and Directory. A sidebar on the right promotes a subscription. The main content area features a large article titled 'Academic authors 'shocked' after Taylor & Francis sells access to their research to Microsoft AI', dated July 19, 2024, by Matilda Battersby. Below the title is a large portrait of a woman with brown hair and bangs, smiling, with a yellow background and a green leaf graphic. To the right of the main article is a 'Latest News' section with five smaller articles, each accompanied by a small portrait of the author or a relevant image.

← → ↺ 🔒 https://www.thebookseller.com/news/academic-authors-shocked-after-taylor-francis-sells-ac 80% ☆ ⬇ 📄 📧

🔍 👤 LOGIN

THE BOOKSELLER
At the heart of the book trade since 1858

Subscribe from less than £3.50 a week
SUBSCRIBE

☰ TOPICS

NEWS RIGHTS COMMENT BESTSELLERS BOOKS SPOTLIGHT EVENTS JOBS DIRECTORY

Home > News

Academic authors 'shocked' after Taylor & Francis sells access to their research to Microsoft AI

NEWS JUL 19, 2024 BY MATILDA BATTERSBY

in 🔍 f

EXCLUSIVE

Latest News

Books in the Media: Robert Macfarlane's 'timely' book a reminder of humans' interconnectedness with nature

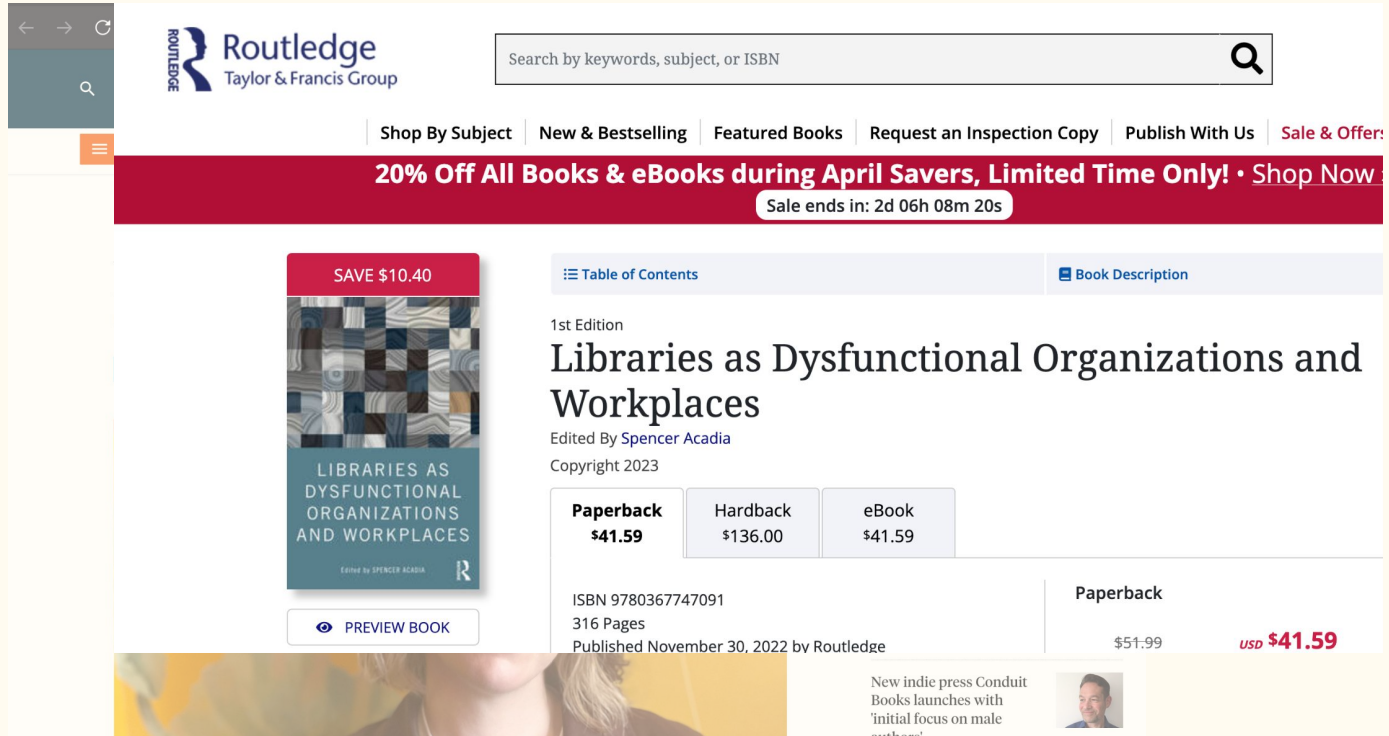
Hachette parent company Lagardère reports strong results after 'solid momentum' in the UK

Bethany Carter appointed head of PR at Hachette Children's Group

French indie bookshop openings remain buoyant but post-Covid-19 boom slows, CNL figures reveal

New indie press Conduit Books launches with 'initial focus on male authors'

Machinations already underway



The screenshot shows the Routledge website interface. At the top, there's a navigation bar with the Routledge logo (Taylor & Francis Group) and a search bar. Below the navigation bar, a prominent red banner announces a 20% discount on all books and eBooks during the April Savers sale, which is limited time only and ends in 2 days, 6 hours, 8 minutes, and 20 seconds. The main content area features a book listing for 'Libraries as Dysfunctional Organizations and Workplaces' by Spencer Acadia. The book cover is displayed on the left, showing a grid pattern. To the right of the cover, the book's title, edition (1st), and editor are listed. Below this, the pricing for different formats (Paperback, Hardback, eBook) is shown. The paperback price is \$41.59, which is highlighted as a savings of \$10.40. The ISBN, page count (316), and publication date (November 30, 2022) are also provided. A 'Preview Book' button is visible below the cover. At the bottom of the page, there's a section for a new indie press, Conduit Books, which is launching with an initial focus on male authors. A small photo of a man is included in this section.

Routledge
Taylor & Francis Group

Search by keywords, subject, or ISBN

Shop By Subject | New & Bestselling | Featured Books | Request an Inspection Copy | Publish With Us | Sale & Offers

20% Off All Books & eBooks during April Savers, Limited Time Only! • Shop Now

Sale ends in: 2d 06h 08m 20s

SAVE \$10.40

LIBRARIES AS DYSFUNCTIONAL ORGANIZATIONS AND WORKPLACES

Edited by SPENCER ACADIA

PREVIEW BOOK

Table of Contents | Book Description

1st Edition

Libraries as Dysfunctional Organizations and Workplaces

Edited By [Spencer Acadia](#)

Copyright 2023

Paperback
\$41.59


Hardback
\$136.00

eBook
\$41.59

ISBN 9780367747091
316 Pages
Published November 30, 2022 by Routledge

Paperback
\$51.99 **USD \$41.59**

New indie press Conduit Books launches with 'initial focus on male authors'



The Labour Context: AI First?

← → ↺ 🔒 https://www.forbes.com/sites/douglaslaney/2025/04/09/selling-ai-strategy-to-employees-shopify-ceos-manifesto

Subscribe To Newsletters

Forbes


LEADERSHIP > CIO NETWORK

Viral Shopify CEO Manifesto Says AI Now Mandatory For All Employees

By [Douglas B. Laney](#), Contributor. [Data, Analytics and AI Strategy Advisor an...](#) [Follow Author](#)

Apr 09, 2025, 12:43pm EDT

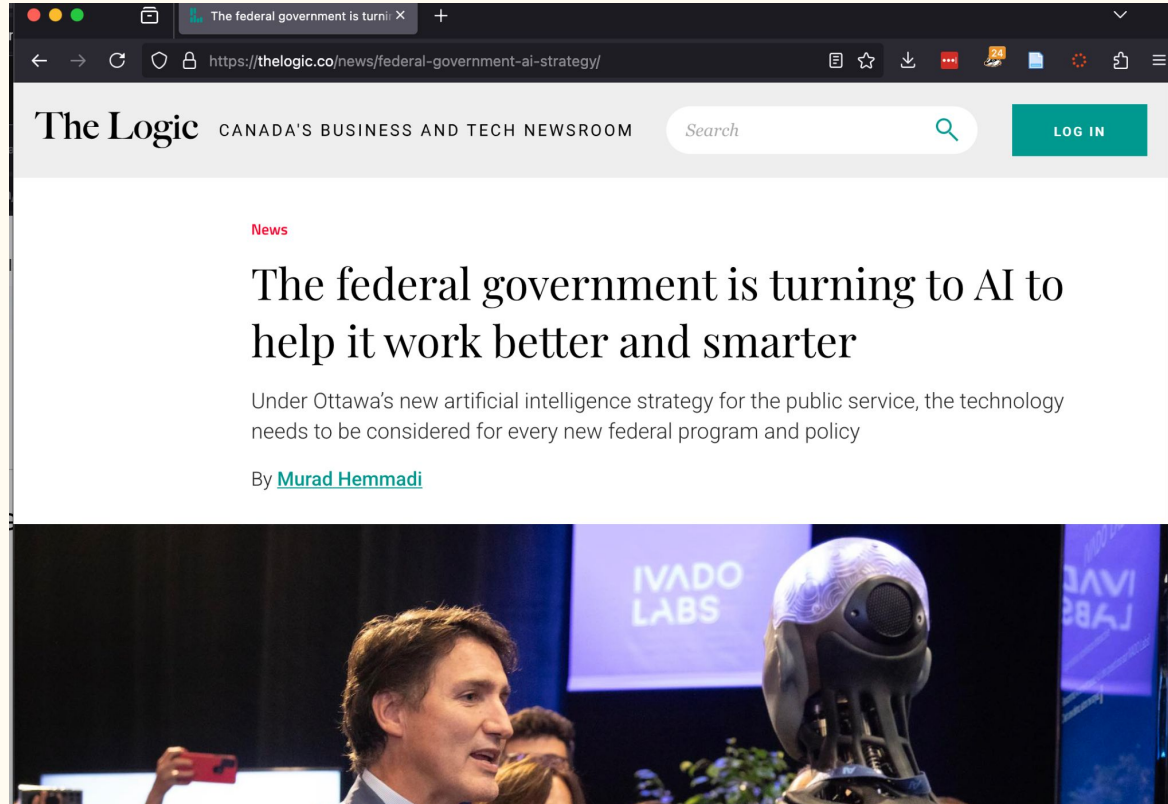
[Share](#) [Save](#)



Tobias Lütke, CEO, Shopify (Photo By David Fitzgerald/Sportsfile via Getty Images)
SPORTSFILE VIA GETTY IMAGES

A leaked [internal memo from Shopify CEO Tobias Lütke](#) has quietly gone viral in executive

Canadian Federal Government's push to AI



Is AI a Grim Reaper...or is it The Blob?



- 1950s sci-fi film about an amorphous blob that crashes onto the earth - it envelopes people, getting bigger and bigger, and cannot be killed
- Also a children's game, where players try to avoid being tagged by a "blob" of players who have already been tagged. When a player is tagged, they join the blob by holding hands, and the blob grows as more players are tagged.
- Only way to stop The Blob is to freeze it, and send it to Antarctica - 'as long as it stays cold'

AI Implications

- Can A.I. be stopped? Probably not...
- Do we want it stopped? Probably not?
- Will it continue to envelope us all? (We both relied on A.I. at points to develop this presentation...)
- If it envelopes us all, and can't be stopped...can we somehow affect it from within?
- Can WE change The A.I. Blob?

ScholComm & AI

LPC Community

You are Scholarly Communication librarians, open advocates, repository managers, journal publishers, journal editors, metadata creators, RDM mavens...

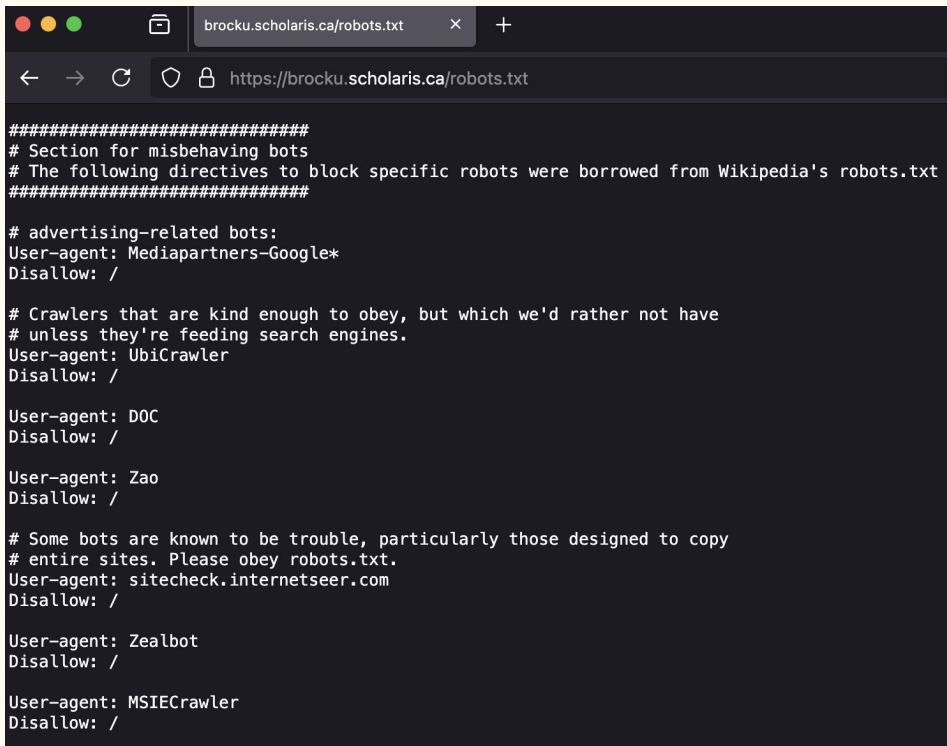
Do you use AI in your daily work? If yes, how? If not, are you planning to?

Do you have policies on AI use?

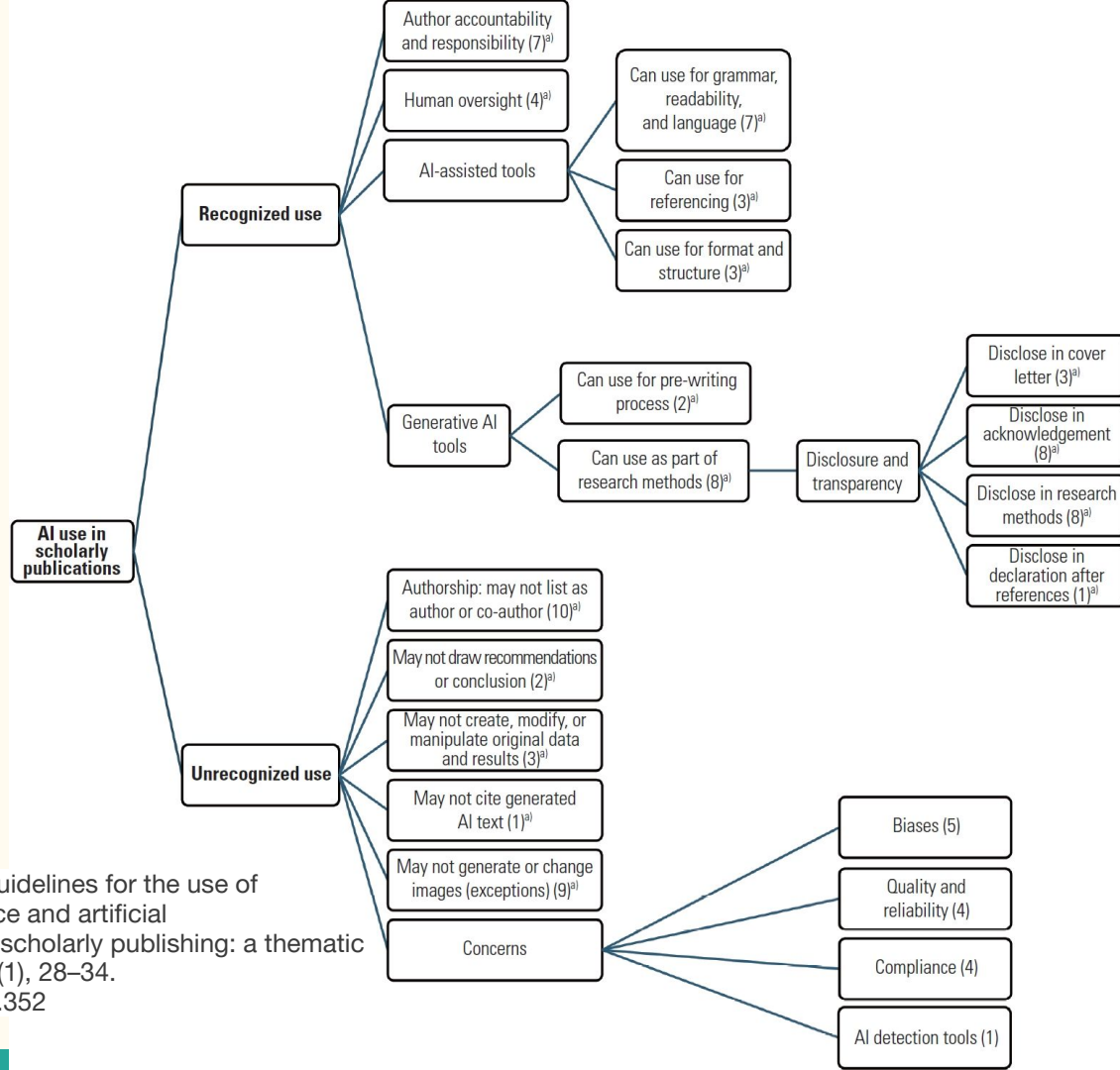
In the Library Sphere

AI Companies are already harvesting...

- Repositories
- Digital Archives
- Journals
- LibGuides

A screenshot of a web browser window showing the robots.txt file for the domain brocku.scholaris.ca. The browser's address bar displays the URL https://brocku.scholaris.ca/robots.txt. The page content is a text file with directives for search engines. It includes a section for misbehaving bots, advertising-related bots (Mediapartners-Google), crawlers that are not recommended (UbiCrawler, DDC, Zao), and bots known to be trouble (sitecheck.internetseer.com, Zealbot, MSIECrawler). Each bot is listed with its User-agent and a Disallow: / directive.

```
#####  
# Section for misbehaving bots  
# The following directives to block specific robots were borrowed from Wikipedia's robots.txt  
#####  
  
# advertising-related bots:  
User-agent: Mediapartners-Google*  
Disallow: /  
  
# Crawlers that are kind enough to obey, but which we'd rather not have  
# unless they're feeding search engines.  
User-agent: UbiCrawler  
Disallow: /  
  
User-agent: DDC  
Disallow: /  
  
User-agent: Zao  
Disallow: /  
  
# Some bots are known to be trouble, particularly those designed to copy  
# entire sites. Please obey robots.txt.  
User-agent: sitecheck.internetseer.com  
Disallow: /  
  
User-agent: Zealbot  
Disallow: /  
  
User-agent: MSIECrawler  
Disallow: /
```



da Veiga, A. (2025). Ethical guidelines for the use of generative artificial intelligence and artificial intelligence-assisted tools in scholarly publishing: a thematic analysis. *Science Editing*, 12(1), 28–34.
<https://doi.org/10.6087/kcse.352>

Editors' statement on the responsible use of generative artificial intelligence technologies in scholarly journal publishing

Kaebnick, G. E., Magnus, D. C., Kao, A., Hosseini, M., Resnik, D., Dubljević, V., Rentmeester, C., Gordijn, B., Cherry, M. J., Maschke, K. J., Rasmussen, L. M., Haupt, L., Schüklenk, U., Chadwick, R., & Diniz, D. (2023). Editors' statement on the responsible use of generative artificial intelligence technologies in scholarly journal publishing. *Developing World Bioethics*, 23(4), 296–299. <https://doi.org/10.1111/dewb.12424>

Written by a group of journal editors: “We believe that generative AI may pose a threat to the goals that animate our work but could also be valuable for achieving those goals”.

Recommendations include the human in the loop defense:

- Editors and reviewers should not rely solely on generative AI to review submitted paper - believe that a complete substitution should not take place and urge that publishers retain humans as the final arbiters in the review process.
- Editors retain final responsibility in selecting reviewers and should exercise active oversight of that task

AI Generated Metadata

Cocciolo, A. (2025). Oral History Metadata and AI: A Study from an LGBTQ+ Archival Context. *Preservation, Digital Technology & Culture*, 54(1), 27-33. <https://doi.org/10.1515/pdtc-2024-0054>

The following prompt was given to ChatGPT 4: “Hello ChatGPT. I would like you to summarize an interview so that it could be used as a description field in a Dublin core metadata field,” and then an export of the transcript available as text was provided.

Research participants preferred the AI-generated description of oral histories documenting LGBTQ + persons compared to that of the description created by a human.

Potential erosion of these skills amongst library workers, and the need for human-in-the-loop for due to bias, hallucination, etc.

AI in Peer Review

Singh Chawla, D. (2024). Is ChatGPT corrupting peer review? Telltale words hint at AI use. *Nature*, 628(8008), 483–484.
<https://doi.org/10.1038/d41586-024-01051-2>

Research study that examined the extent to which AI chatbots could have modified the peer reviews of conference proceedings submitted to four major computer-science meetings, since the release of ChatGPT



Identified buzzword adjectives that could be hallmarks of AI-written text

Their analysis suggests that up to 17% of the peer-review reports have been substantially modified by chatbots - although it's unclear whether researchers used the tools to construct reviews from scratch or just to edit and improve written drafts

"It seems like when people have a lack of time, they tend to use ChatGPT,"

Copyright implications when giving the tools access to confidential, unpublished material

Case Study: Automatic Plagiarism Detection

Journal of Public Knowledge

[← Back to Submissions](#)

Submit an Article

1. Start

2. Upload Submission

3. Enter Metadata

4. Confirmation

5. Next Steps

You must read and accept the [iThenticate EULA](#) before proceeding to finalise the submission.

☐ **I have read the end user license agreement and accept it.**

Confirm and Continue

Cancel

Case Study: Automatic Plagiarism Detection

27 / Author1 / ANOTHER AUTOMATIC SUBMISSION





Activity LogLibrary

WorkflowPublication

SubmissionReviewCopyeditingProductionHelp

Submission Files

Q SearchUpload File

▶		52	Submission_Dataset.docx	<div>SCREENED BY  iThenticate Professional Registered Provider</div>	100%	October 1, 2024	Data Set
▶		53	Submission_New Submission.docx	<div>SCREENED BY  iThenticate Professional Registered Provider</div>	100%	October 1, 2024	Article Text

Download All Files

Pre-Review DiscussionsAdd discussion

Send to Review

Accept and Skip Review

Decline Submission

ParticipantsAssign

Journal editor

▶ Daniel Barnes

Section editor

Case Study: Automatic Plagiarism Detection



Rejected paper Draft

BY TIM RIBARIC

COMPARING ACADEMIC FREEDOM

14

2024). It is also possible to make generated responses cite documents in the corpus that were foundational in the construction of the response. This analysis was performed using a platform

calli

Publications

Full Source View

ena

Bo T. Christensen, Linden J. Ball, Kim Halskov. "Analysing Design Thinking: St...

Inst

the removal of very short, no-content turns (e.g., "um", "yeah"), so no additional work is required to deal with the problem of finding content in such short turns. 2.3.3 Building the topic **model** **The final step is to** build **the** topic **model**. LDA requires **the** user to specify the number of topics to estimate upfront. Thus, this final step yields our fourth and final free parameter to select: how many topics should the model learn ? There is no hard-and-fast

share is that the *temperature* parameter was adjusted to 0.0 for all testing (Zhu et. al., 2023) By

doing this a consistent repeatable state is created with the **model**. **The final step is to prompt the**

primed **model** with **the** following question: *What is the purpose of Academic Freedom?* Also

worth noting here is that the unadulterated full-text of each publication was fed into the RAG system; preprocessed text is not required for LLM use.

Case Study: Automatic Plagiarism Detection



Rejected paper Draft

BY TIM RIBARIC

COMPARING ACADEMIC FREEDOM

14

2024). It is also possible to make generated responses cite documents in the corpus that were foundational in the construction of the response. This analysis was performed using a platform

calli

ances

Rejected paper Draft

BY TIM RIBARIC

Quotes Excluded
Bibliography Excluded

3%
SIMILAR

build the topic model. LDA requires the user to specify the number of topics to estimate upfront. Thus, this final step yields our fourth and final free parameter to select: how many topics should the model learn ? There is no hard-and-fast

ary to

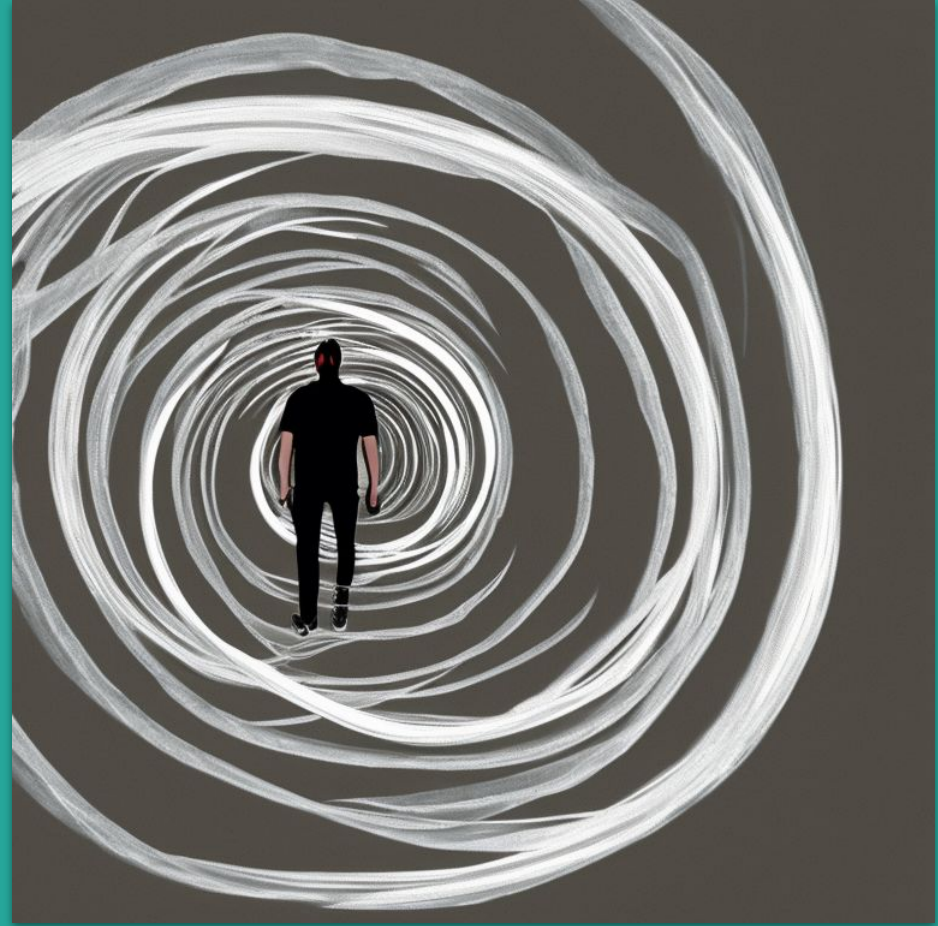
share is that the temperature parameter was adjusted to 0.0 for all testing (Zhu et. al., 2023) By

doing this a consistent repeatable state is created with the model. The final step is to prompt the

primed model with the following question: *What is the purpose of Academic Freedom?* Also

worth noting here is that the unadulterated full-text of each publication was fed into the RAG system; preprocessed text is not required for LLM use.

Is Human-in-the-loop
Enough?



Critiquing the ‘Human-in-the-loop’ defense

<https://pluralistic.net/2024/10/30/a-neck-in-a-noose/#is-also-a-human-in-the-loop>

Cory Doctorow argues that people (including experts) are highly susceptible to "automation bias."

Experts who are put in charge of overseeing an automated system get out of practice, because they no longer engage in the routine steps that lead up to the conclusion - cognitive offloading

Leads to the creation of what economist Dan Davies calls ‘Accountability Sinks’ - structures that absorb or obscure the consequences of a decision such that no one can be held directly accountable for it.

So...what do we do about
this? If anything?

—

Mock Language

“The University agrees that because neither traditional AI nor generative AI is a person, it cannot independently perform the duties and responsibilities of a librarian, as articulated in the collective agreement.”

– inspired by the Writers Guild of America contract language

“The parties acknowledge the importance of human performance in librarianship and will not consider synthetic performers as equivalents.”

– inspired by the Screen Actors Guild MOA

Questions
And
Thanks

