

Exploring AI Capabilities in Indexing Swedish LGBTQ+ Fiction

Koraljka Golub, Daniel Ihrmark

NKOS Workshop at TPDL, Tampere, 23 September 2025



The Role of AI in Library Cataloging

Growing Importance of Automation

- Efforts to automate cataloging have increased over past decades
- AI and Large Language Models (LLMs) like ChatGPT offer new possibilities
- AI-assisted subject indexing is a growing research focus

Challenges in Fiction Indexing

- Traditional cataloging emphasizes genre, time, and place, missing specific themes
- Users seek more detailed indexing, especially for niche topics
- Automated tools could help address these challenges



LGBTQ+ Fiction & Challenges in Automated Indexing

- **Difficulties in LGBTQ+ Fiction Indexing**
 - LGBTQ+ themes are often nuanced, requiring in-depth reading
 - Existing subject headings are often too broad or outdated
 - Older works may not be explicitly labeled as LGBTQ+
- **Limitations of AI in Fiction Indexing**
 - Struggles with connotation, hidden themes, and metaphors
 - Potential for bias in training data, leading to inaccurate indexing
 - Risk of hallucinations (false but convincing information)
 - Lack of transparency in AI decision-making ("black box" issue)



Research Aim & Questions

- **Exploring AI in Automated Indexing**
 - Investigating four LLM tools' ability to generate subject index terms for LGBTQ+ fiction
- **Key Research Questions**
 1. How effectively can AI tools identify main LGBTQ+ themes in fiction?
 2. How effectively can AI tools identify minor LGBTQ+ hemes in fiction?
- **Evaluation Method**
 - Comparing AI-generated index terms with those assigned by professional catalogers in the Queerlit database

The Queerlit Database

- **Scope of Queerlit**
 - LGBTQ+ fiction in Swedish and minority languages
 - 2420 catalog records (January 2025)
- **Indexing Languages**
 - QLIT thesaurus (900+ LGBTQ+ terms, adapted from Homosaurus)
 - Also employs Swedish controlled vocabularies (SAO, Barnämnesordslistan, SAOGF, Barngf)



Queerlit Indexing Policy

- **Indexing Practices**

- Main subjects: $\geq 20\%$ of the work's content
- Minor subjects for LGBTQ+ themes
- Uses the most specific subject term available for precise retrieval

- **Quality Control**

- Four librarians collaborate for consistency.
- Scope notes ensure uniform term usage.

Game over Jenny Milewski
2023
Ungdomsböcker Rysare
Lättlästa böcker
Barn- och ungdomslitteratur

Queera ungdomar – perifert : Hbtqi-ungdomar – perifert : Döden och döende (hbtqi) – perifert :

Ungdomar : Överlevnad : Monster : Övernaturligt : Klasskompisar :

Skräck lockar många unga till tv-serier, dataspel och filmer. Den sista skolresan lockar även in dem mellan bokens pärmar. Här sparas det inte på blod, splatter och fantasifula fasor. Det här är...

En av personerna i boken, Freja, har haft en flickvän som dött. Detta omnämns inte i denna del, men framgår i den första delen i serien.

Hemlösa hbtqi-ungdomar

Anvisning Används för skildringar av hbtqi-ungdomar som lever i hemlöshet. Hemlöshet innebär att personen saknar en fast bostad i en kultur där fast bostad är norm.

Identifierare <https://queerlit.dh.gu.se/qlit/v1/iq97le56>
❗ Identifieraren har formen av en URL, men används inte främst som en webbadress. Formen är mest ett sätt att precisera att definitionen är en del av Queerlits tesauros.

Sök i Queerlit på *Hemlösa hbtqi-ungdomar* :

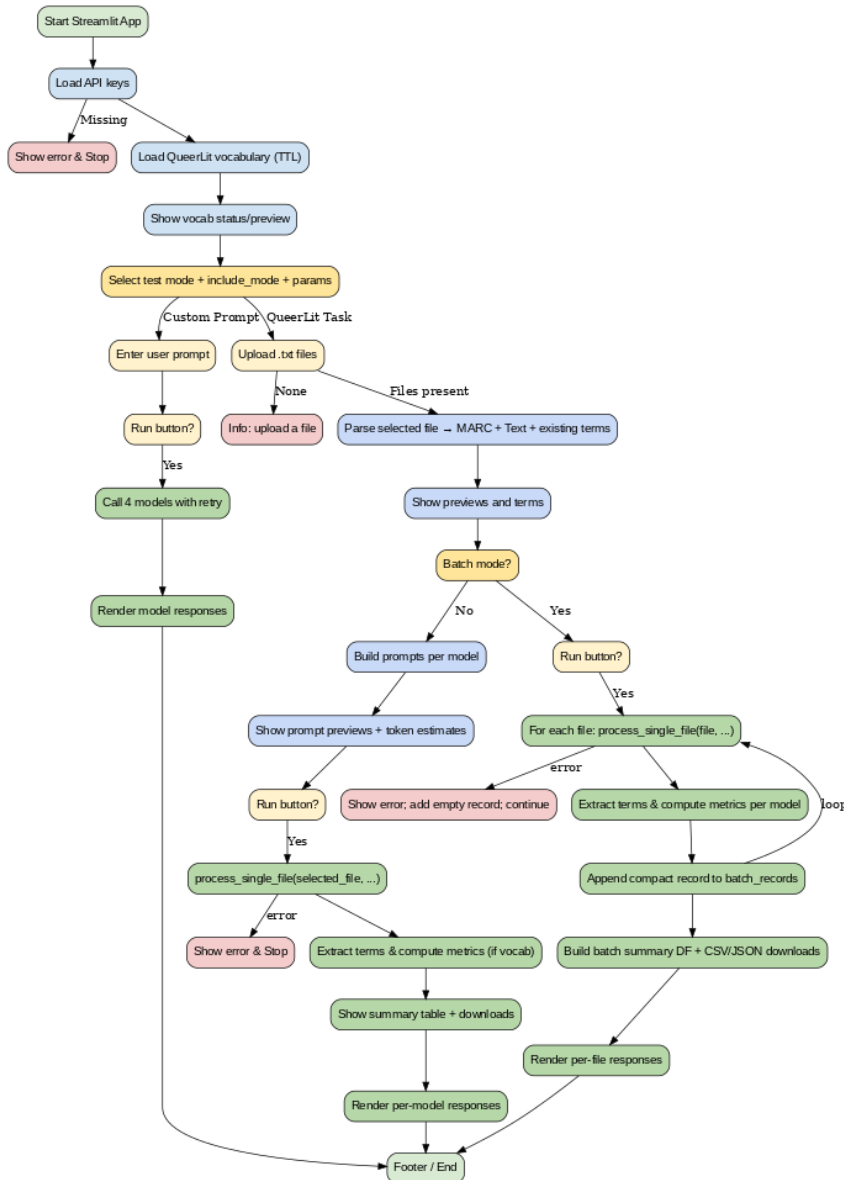


Study Sample & Selection

- **Dataset**
 - 78 metadata records & full-text works from Queerlit
 - Freely available on Litteraturbanken (full-text literary fiction database)
- **AI Tools**
 - ChatGPT (GPT-4 o)
 - Claude (Sonnet 4)
 - DeepSeek (Chat model)
 - Gemini (1.5-flash)

Platform: Streamlit Dashboard

- Two modes: Prompt testing and Classification task
- Prompt testing sends the prompt directly via the API
- Classification task facilitates the indexing process
- **Zero-shot classification** performed based on Queerlit vocabulary
- Indexing can be done in **batch or for single files**
- **Precision, Recall and F1** are calculated based on model responses with MARC record as ground truth



AI Indexing Process

- **Source documents**
 - Full text .txt file for each work preceded with MARC record
 - Queerlit Vocabulary with 913 terms
- **Streamlit Dashboard**
 - Different token cut off (ChatGPT cca 5000)
 - Temperature setting: 0.7

AI Model Comparator for Subject Indexing

✓ Loaded 913 terms from 913 TTL files

▼  QueerLit Vocabulary Preview (913 terms)

AFAB, AMAB, Aboriginer (hbtqi), Adopterade (hbtqi), Adoption (hbtqi), Adoptivfamiljer (hbtqi), Adoptivföräldrar (hbtqi) (hbtqi), Anala samlag, Analkulor, Analsex

... and 893 more terms

▼  Model Information

Models being compared:

- **Claude:** claude-sonnet-4 (Anthropic)
- **ChatGPT:** gpt-4o (OpenAI)
- **DeepSeek:** deepseek-chat (DeepSeek)
- **Gemini:** gemini-1.5-flash (Google)

All models use temperature=0.7 as default for consistent comparison.

Pick models to run

☐ Claude ×

🔥 ChatGPT (...) ×

 DeepSeek ×

💎 Gemini ×

Select Mode:

AI Prompting

Full text

“You are a subject indexer specializing in LGBTQI+ literature analysis. Your task is to analyze the provided literary work and suggest relevant subject terms from the QueerLit controlled vocabulary.

Please analyze ONLY the literary text provided and:

- 1. Identify specific LGBTQI+ themes, characters, relationships, or content in the text*
- 2. Suggest appropriate QueerLit vocabulary terms that would apply (use exact terms from the vocabulary when possible)*
- 3. Provide brief justification for each suggested term based on textual evidence*
- 4. If no exact vocabulary match exists, suggest the closest appropriate terms or describe what terms might be needed*

Base your analysis solely on the provided text, not on external knowledge about the author or work.

”

+ *Queerlit terms*
+ *Full text*

Metadata

“You are a subject indexer specializing in LGBTQI+ literature analysis. Your task is to analyze the provided literary work and suggest relevant subject terms from the QueerLit controlled vocabulary.

Please analyze ONLY the MARC metadata provided and:

- 1. Identify specific LGBTQI+ themes, characters, relationships, or content in the text*
- 2. Suggest appropriate QueerLit vocabulary terms that would apply (use exact terms from the vocabulary when possible)*
- 3. Provide brief justification for each suggested term based on textual evidence*
- 4. If no exact vocabulary match exists, suggest the closest appropriate terms or describe what terms might be needed*

Base your analysis solely on the provided text, not on external knowledge about the author or work.”

+ *Queerlit terms*
+ *Metadata*

Findings: Major Terms (16 works)

Metadata

Model	Precision	Recall	F1	Terms_Found	Correct
claude	0.25	0.94	0.39	156	39
chatgpt	0.39	0.80	0.49	97	36
deepseek	0.22	0.89	0.35	156	36
gemini	0.37	0.94	0.50	110	39
<i>all models average</i>	0.31	0.89	0.43	129.75	37.5

Full text

Model	Precision	Recall	F1	Terms_Found	Correct
claude	0.09	0.45	0.14	178	18
chatgpt	0.06	0.27	0.09	125	8
deepseek	0.09	0.42	0.14	178	15
gemini	0.07	0.30	0.11	113	9
<i>all models average</i>	0.08	0.36	0.12	148.5	12.5



Findings: Minor Terms (50 works)

Metadata

Model	Precision	Recall	F1	Terms_Found	Correct
claude	0.23	0.95	0.36	443	93
chatgpt	0.33	0.97	0.46	307	94
deepseek	0.20	0.97	0.32	488	94
gemini	0.33	0.97	0.46	328	93
<i>all models average</i>	0.27	0.97	0.40	391.5	93.5

Full text

Model	Precision	Recall	F1	Terms_Found	Correct
claude	0.02	0.13	0.03	590	15
chatgpt	0.02	0.09	0.03	334	7
deepseek	0.02	0.13	0.03	558	13
gemini	0.00	0.02	0.01	218	2
<i>all models average</i>	0.02	0.09	0.03	425	9.25



Findings: Major and Minor Terms (12 works)

Metadata

Model	Precision	Recall	F1	Terms_Found	Correct
claude	0.34	0.72	0.45	142	50
chatgpt	0.46	0.72	0.54	107	48
deepseek	0.29	0.70	0.40	166	48
gemini	0.41	0.73	0.51	118	47
<i>all models average</i>	0.37	0.72	0.48	133.25	48.25

Full text

Model	Precision	Recall	F1	Terms_Found	Correct
claude	0.08	0.21	0.11	185	14
chatgpt	0.03	0.07	0.04	99	4
deepseek	0.04	0.10	0.06	178	8
gemini	0.04	0.06	0.04	80	4
<i>all models average</i>	0.05	0.11	0.06	135.5	7.5

Overall Findings

- AI-generated index terms are largely **unhelpful**
 - F1 up to 54% when given solutions
 - F1 up to 14% when full text
- No winner as all four tools underperforming, although variations across the data sets
- Lacks contextual awareness of LGBTQ+ nuances
- Generates incorrect, irrelevant or overly broad index terms



Future Research & Recommendations

- **Challenges in AI Training**
 - Older LGBTQ+ themes may be too implicit for AI to detect
 - Copyright restrictions limit AI training on newer LGBTQ+ fiction
- **Expert Evaluation**
 - Entire metadata set
 - Involve subject librarians to assess AI-generated terms
 - Evaluate for general subject headings, too
- **Need for Critical AI Adoption:**
 - AI is often overhyped—users should critically assess its limitations & biases





Lnu.se