

Chatting with Papers

A Hybrid Approach Using LLMs and Knowledge Graphs

Slava Tykhonov [b], Han Yang [a], **Philipp Mayr** [a], Jetze Toubert [b] Andrea Scharnhorst [b]

[a] GESIS, Cologne, Germany [b] DANS-KNAW, The Hague, The Netherlands



Preprint: <https://arxiv.org/abs/2505.11633> Demo: <https://gesis.now.museum/>

Agenda

- Background and Motivation
- Related Work & Positioning
- Use Case: Method Data Analysis (mda) Journal
- System Architecture: EverythingData + GhostWriter
- Demo Walkthrough
- Outlook & Broader Applications

Science dynamics and AI

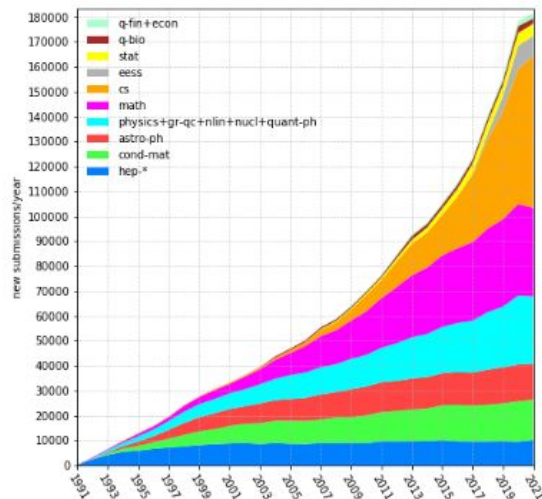
In the evolution of the sciences we find growth and increasing differentiation.

Leaving us with the problem to review, evaluate and select.

Precondition to any creating of (new) knowledge (in individual brains or for larger parts of academia) is **to find and understand**.

Machines (latest AI) have fostered growth, can they also support the knowledge production process?
This is a question of **Information Retrieval**.

Data for 1991 through 2021, updated 3 January 2022.



https://info.arxiv.org/help/stats/2021_by_area/index.html

Why Chatting with Papers? Background & Motivation

Problem

- Navigating large document collections is time-consuming
- Traditional search is keyword-based & lacks semantic depth

Two extremes

- LLMs = “Asking an expert” → flexible, but prone to hallucinations
- Knowledge Graphs or KOS = “Asking a librarian” → precise, but rigid

Goal: Combine both worlds:

- **Flexible natural language querying + structured semantic context**

Key Idea: A Hybrid AI Workflow

What is GhostWriter?

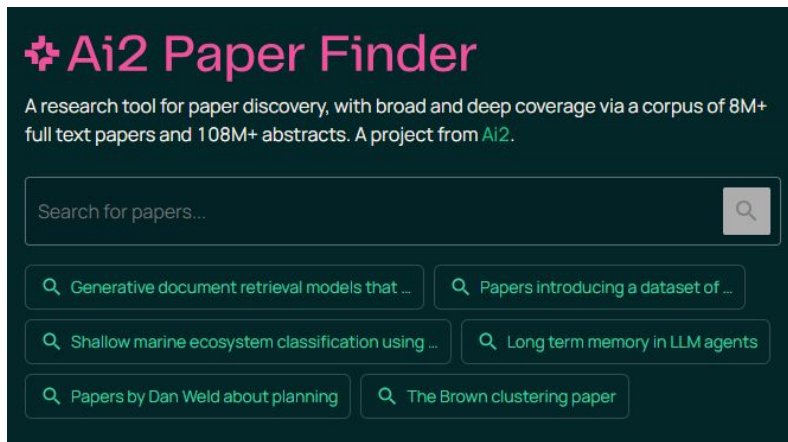
- A local chatbot that enables semantic interaction with paper collections
- Based on LLMs + Knowledge Graphs (semantic artifacts)

Key Features

- Natural language interface
- Iterative query refinement
- Contextual document retrieval
- Supports multilinguality

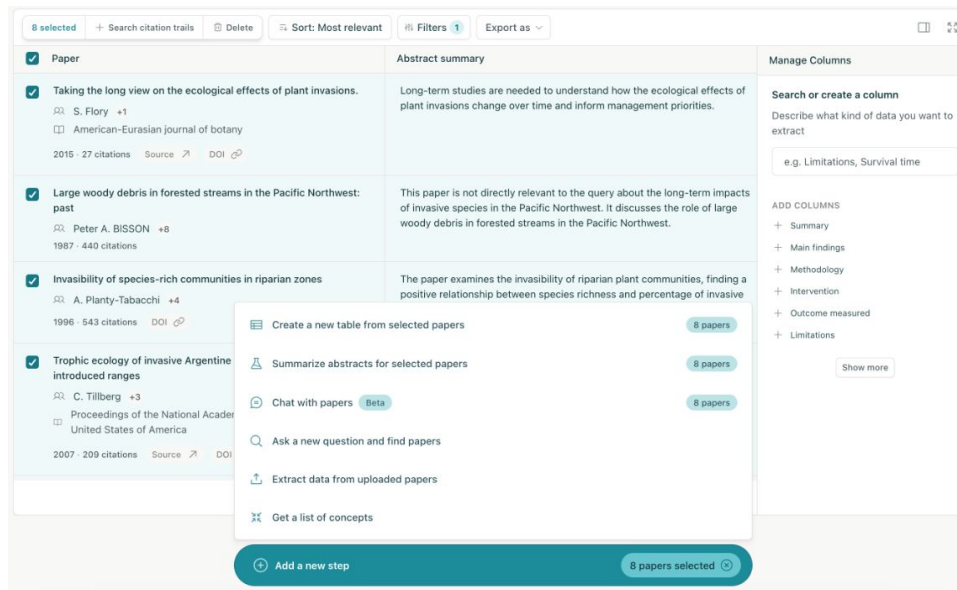
Goal: **Help users iteratively query and explore paper collections**

Related Work & Positioning



Search paper with semantic queries,
Result Summarization and exploration

Ai2 PaperFinder from allenai.org



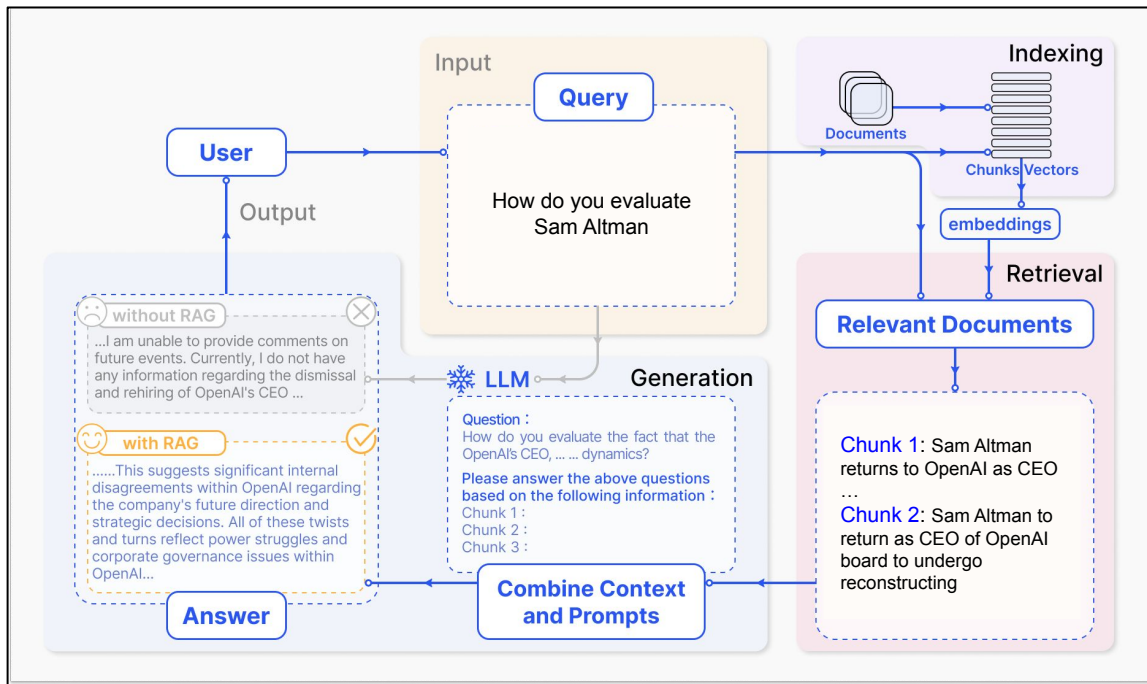
Chatting with papers

Elicit: the AI Research Assistant

<https://support.elicit.com/en/articles/1467969>

Related Work & Positioning

Retrieval-Augmented Generation (RAG)



Storing documents with vector database

Retrieval with semantic similarity

Generating answers with LLM in human language

Use Case: Method Data Analysis (mda) Journal

- Applied to 100 articles from methods, data, analyses (GESIS journal)
- Setup
 - Full-text PDFs processed with EverythingData
 - Live querying via GhostWriter interface

mda
methods, data, analyses

HOME ABOUT LOGIN REGISTER SEARCH ONLINE FIRST CURRENT ARCHIVES /

Home > Vol 18, No 2 (2024)

methods, data, analyses

methods, data, analyses (mda) publishes research on all questions important to quantitative methods in social special emphasis on survey methodology. All content is freely available and can be distributed without any rest ensuring the free flow of information that is crucial for scientific progress.

mda is an DIAMOND OPEN ACCESS Journal.

mda is indexed and abstracted in the Emerging Sources Citation Index.

mda is also listed in the Directory of Open Access Journals and SCOPUS.

Vol 18, No 2 (2024)

Full Issue

View or download the full issue

Research Reports

[Combining Information from Multiple Data Sources to Improve Sampling Efficiency](#)

Paul Burton, Sunghee Lee, Trivellore Raghunathan, Brady T. West

[How to Reduce Item Nonresponse in Face-to-Face Surveys? A Review and Evidence from the European Social Survey](#)

Malte Grönnemann

Search Search Scope All SEARCH ISSUES

Browse By Issue | By Author | By Title

mda
methods, data, analyses

HOME ABOUT LOGIN REGISTER SEARCH ONLINE FIRST CURRENT ARCHIVES /

Home > Vol 18, No 2 (2024) > Burton

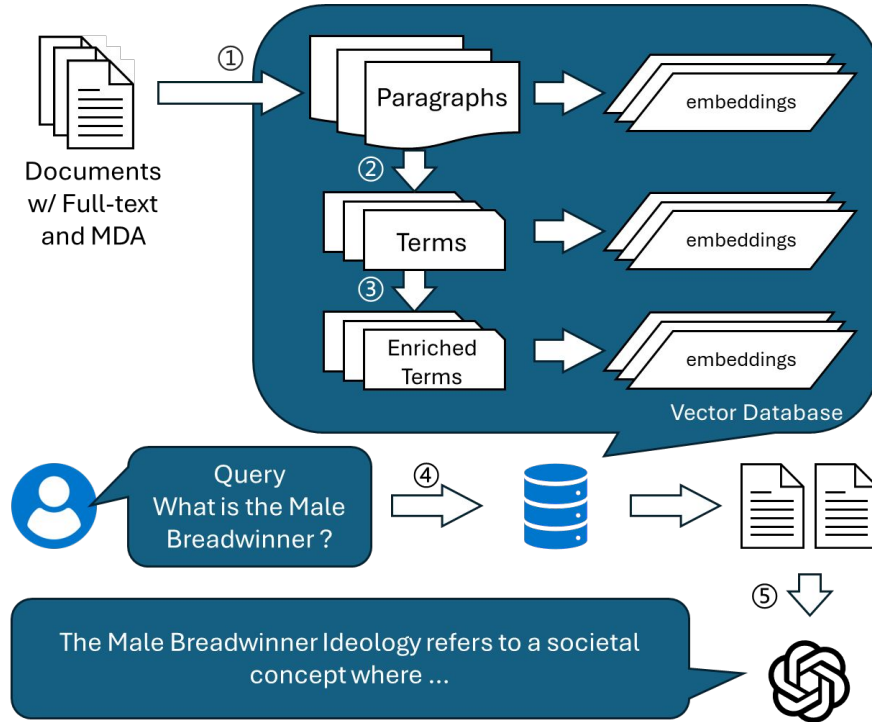
Download

methods, data, analyses | Vol. 18(2), 2024, pp. 143-164 DOI: 10.12758/mda.2024.

Combining Information from Multiple Data Sources to Improve Sampling Efficiency

Paul Burton, Sunghee Lee,
Trivellore Raghunathan & Brady T. West
University of Michigan, Survey Research Center (SRC)

System Architecture: EverythingData + GhostWriter



- Enhanced Retrieval Augmented Generation (RAG)
- Further augmented for matching targets
- Terms: extracted by NER enhanced by EverythingData & Wikidata
- User-interface: Ghostwriter

Demo Walkthrough

gesis

Leibniz-Institut
für Sozialwissenschaften

Add New Collection ▾

Available Collections

Select collection to explore

mda ▾

Collection

Vectors: 37,792

About Sources Manage Pages Login

🗨️ Ask Questions

Example: Explain the Rational Choice Theory

Enter your question:

🔍 Ask

📖 Add Page

Enter webpage URL or RSS feed:

Select source type:

☒ Single Webpage ☐ Website Crawler ☐ RSS Feed

📄 Add Content

Copyright © 2024. Ghostwriter created by Slava Tykhonov, [DANS-KNAW](#) and powered by [Now.Museum](#)

Demo is work in progress!

Demo Walkthrough

gesis
Leibniz-Institut
für Sozialwissenschaften

Add New Collection

Available Collections

Select collection to explore

mda

Collection

Vectors: 37,792

Ask Questions

Example: Explain the Rational Choice Theory

Enter your question:

explain male breadwinner model



Ask

Answer



Good



Bad

Share

Copy

Google

Based on the available information: Male Breadwinner Ideology and the Inclination to Establish Market Relationships: Model Development Using Data from Germany and a Mixed-Methods Research Strategy. Journal of Macromarketing, 36(2), 149–167. <https://doi.org/10.1177/0276146715576202> Hainmueller, J., Hangartner, D., & Yamamoto, T. (2015). Validating Vignette and Conjoint Survey Experiments against Real-world Behavior. Proceedings of the National Academy of Sciences, 112(8), 2395–2400.

Sources

	Chat	Title	Reference	Score
		The Past, Present and Future of Factorial Survey Experiments		

Based on the available information: Male Breadwinner Ideology and the Inclination to Establish Market Relationships: Model Development Using Data from Germany and a Mixed-Methods Research Strategy. Journal of Macromarketing, 36(2), 149–167. <https://doi.org/10.1177/0276146715576202> Hainmueller, J., Hangartner, D., & Yamamoto, T. (2015). Validating Vignette and Conjoint Survey Experiments against Real-world Behavior. Proceedings of the National Academy of Sciences, 112(8), 2395–2400.

Demo Walkthrough

Add New Collection

Available Collections

Select collection to explore

mda

Collection

Vectors: 37,792

Ask Questions

Example: Explain the Rational Choice Theory

Enter your question:

explain male breadwinner model

Ask

Answer

Good Bad Share Copy Google

Based on the available information: Male Breadwinner Ideology and the Inclination to Establish Market Relationships: Model Development Using Data from Germany and a Mixed-Methods Research Strategy. Journal of Macromarketing, 36(2), 149–167. <https://doi.org/10.1177/0276146715576202> Hainmueller, J., Hangartner, D., & Yamamoto, T. (2015). Validating Vignette and Conjoint Survey Experiments against Real-world Behavior. Proceedings of the National Academy of Sciences, 112(8), 2395–2400.

Based on the available information: Male Breadwinner Ideology and the Inclination to Establish Market Relationships: Model Development Using Data from Germany and a Mixed-Methods Research Strategy. Journal of Macromarketing, 36(2), 149–167. <https://doi.org/10.1177/0276146715576202> Hainmueller, J., Hangartner, D., & Yamamoto, T. (2015). Validating Vignette and Conjoint Survey Experiments against Real-world Behavior. Proceedings of the National Academy of Sciences, 112(8), 2395–2400.

Sources

Chat	Title
	The Past, Present and Future of Factorial Survey Experiments

Chat	Title	Reference	Score
0 Ask	The Past, Present and Future of Factorial Survey Experiments: A Review for the Social Sciences Treischl methods, data, analyses	https://mda.gesis.org/index.php/mda/article/download/2021.07/293	0.419529
1 Ask	Gender and Survey Participation: An Event History Analysis of the Gender Effects of Survey Participation in a Probability-based Multi-wave Panel Study with a Sequential Mixed-mode Design Becker methods, data, analyses	https://mda.gesis.org/index.php/mda/article/download/2021.08/299	0.417514

Demo Walkthrough (1)

gesis

Leibniz-Institut
für Sozialwissenschaften

Add New Collection ▾

Available Collections

Select collection to explore

mda ▾

Collection

Vectors: 37,792


[About](#) [Sources](#) [Manage Pages](#) [Login](#)

🙄 Ask Questions

Example: Explain the Rational Choice Theory

Enter your question:

what is relationship between interviewer experience and interview duration?


 Ask

📖 Add Page

Enter webpage URL or RSS feed:

Select source type:

☒ Single Webpage ☐ Website Crawler ☐ RSS Feed

 Add Content

what is relationship between interviewer experience and interview duration?



Ask Questions

Example: Explain the Rational Choice Theory

Enter your question:

what is relationship between interviewer experience and interview duration?



Ask



Answer



Good



Bad

Share

Based on the available information: We formulate this expectation even though overall interviewer experience has been found to affect interview duration in telephone survey interviews (Kirchner & Olson, 2017; Olson & Peterson, 2017). The possibility that both types of interviewer experience interact should also be considered. It can well be that interviewers with less overall experience will have a more pronounced learning curve as they gather more within-survey interviewer experience compared to more experienced interviewers which start working on a new survey study already with shorter interview durations (Kirchner & Olson, 2017; Kosyakova et al., 2021). In sum, the differentiation between overall and within-survey interviewer experience may be crucial, and thus call for different adjustments in training, feedback, and supervision. The Role of Other Factors on Interview Duration There are other influences on interview duration besides interviewer experience (Kosyakova et al., 2021; Loosveldt & Beullens, 2013b; Vandenplas et al., 2018). An alternative explanation for decreasing interview duration over the field phase is based on the changing composition of the sample over time (Kirchner & Olson, 2017). As the fieldwork progresses, harder-to-reach respondents become more common and the lower cooperativeness of the remaining sample could be what makes the interview duration shorter.



Sources

	Chat	Title	Reference	Score
0	Ask	Adjusting to the Survey: How Interviewer Experience Relates to Interview Duration Pirralha methods, data, analyses	https://mda.gesis.org/index.php/mda/article/download/2023.05/348	0.781093

what is relationship between interviewer experience and interview duration?

Experienced (survey-) interviewer has shorter interview duration



Ask Questions

Example: Explain the Rational Choice Theory

Enter your question:

what is relationship between interviewer experience and interview duration?



Ask



Answer



Good



Bad

Share

Based on the available information: We formulate this expectation even though overall interviewer experience has been found to affect interview duration in telephone survey interviews (Kirchner & Olson, 2017; Olson & Peck, 2017). The possibility that both types of interviewer experience interact should also be considered. It can well be that interviewers with less overall experience will have a more pronounced learning curve as they gather more within-survey interview experience. More experienced interviewers which start working on a new survey study already with shorter interview duration (2017; Kosyakova et al., 2021). In sum, the differentiation between overall and within-survey interview experience thus call for different adjustments in training, feedback, and supervision. The Role of Other Factors on Interview Duration: Other influences on interview duration besides interviewer experience (Kosyakova et al., 2021; Loosvold & Vandenplas et al., 2018). An alternative explanation for decreasing interview duration over the field period is the composition of the sample over time (Kirchner & Olson, 2017). As the fieldwork progresses, harder-to-recruit participants become common and the lower cooperativeness of the remaining sample could be what makes the interview duration longer.

what is relationship between interviewer experience and interview duration?

Experienced (survey-) interviewer has shorter interview duration

methods, data, analyses | Vol. 18(2), 2024, pp. 185-212

DOI: 10.12758/mda.2023.05

Adjusting to the Survey: How Interviewer Experience Relates to Interview Duration

André Pirralha, Christian Haag &
Jutta von Maurice

Center for Study Management (CSM), Leibniz Institute for Educational
Trajectories (LifBi)



Sources

	Chat	Title	
0	Ask	Adjusting to the Survey: How Interviewer Experience Relates to Interview Duration Pirralha methods, data, analyses	https://mda.gesis.org/index.php/mda/article/download/2023.05/348 0.781093

Demo Walkthrough (2)

gesis

Leibniz-Institut
für Sozialwissenschaften

Add New Collection

Available Collections

Select collection to explore

mda

Collection

Vectors: 37,792

🗨️ Ask Questions

Example: Explain the Rational Choice Theory

Enter your question:

why is research on young people political participation necessary?

Ask

📖 Add Page

Enter webpage URL or RSS feed:

Select source type:

☒ Single Webpage ☐ Website Crawler ☐ RSS Feed

Add Content

why is research on young people
political participation necessary?

16

Ask Questions

Example: Explain the Rational Choice Theory

Enter your question:

why is research on young people political participation necessary?



Ask

Answer

Good

Bad

Share

Copy

Based on the available information: In this connection, Quintelier has argued that “young people operate with a very narrow conception of politics that is restricted to formal politics only” (2007, 177; see also O’Toole et al., 2003, 52). If

314 methods, data, analyses | Vol. 12(2), 2018, pp. 309-334 we consider this limited and narrow conception of politics to conception and understanding of political participation as well, young people’s political participation may be less faceted and based on fewer modes of participation than that of adult people. In a similar manner, changes or delays in youth transition periods as highlighted by previous studies (cf. Soler-i-Martí & Ferrer-Fons, 2015, 96; García Albacete, 2011, 6) might also lead to varying structures of young people’s political participation over time. An empirical investigation of the underlying structure of political participation across age groups and over time therefore becomes indispensable in order to shed more light on the differences and similarities concerning the structure, levels and developments of young and adult people’s political participation. The ‘identity-equivalence Procedure’ for Political Participation For our empirical investigation, we make use of the so called ‘identity-equivalence procedure’ which has originally been introduced by Przeworski and Teune, (1966) in the context of cross-cultural research. The basic premise of this procedure is that, in order to be comparable, measurements of the same concepts do not have to be identical but rather equivalent (cf.

Sources

	Chat	Title	Reference	Score
0	Ask	Blaming the Young Misses the Point: Re-assessing Young People’s Political Participation over Time Using the ‘Identity-equivalence Procedure’ Schnaudt methods, data, analyses	https://mda.gesis.org/index.php/mda/article/download/2017.12/228	0.738237

why is research on young people political participation necessary?

To know about the differences and similarities of the structure, levels and developments of young and adult people’s political participation

Ask Questions

Example: Explain the Rational Choice Theory

Enter your question:

why is research on young people political participation necessary?



Ask

Answer

Good

Bad

Share

Copy

Based on the available information: In this connection, Quintelier has argued that “young people operate with a very narrow conception of politics that is restricted to formal politics only” (2007, 177; see also O’Toole et al., 2003, 52). If

314 methods, data, analyses | Vol. 12(2), 2018, pp. 309-334 we consider this limited and narrow conception of politics to conception and understanding of political participation as well, young people’s political participation may be less faceted and based on fewer modes of participation than that of adult people. In a similar manner, changes or delays in youth transition period are highlighted by previous studies (cf. Soler-i-Marti & Ferrer-Fons, 2015, 96; Garcia Albacete, 2011, 6) might also lead to variations of young people’s political participation over time. An empirical investigation of the underlying structure of political participation across age groups and over time therefore becomes indispensable in order to shed more light on the differences and similarities concerning the structure, levels and developments of young and adult people’s political participation. The ‘identity-equivalence procedure’ for Political Participation For our empirical investigation, we make use of the so called ‘identity-equivalence procedure’ originally been introduced by Przeworski and Teune, (1966) in the context of cross-cultural research. The basic premise of the procedure is that, in order to be comparable, measurements of the same concepts do not have to be identical but rather

why is research on young people political participation necessary?

To know about the differences and similarities of the structure, levels and developments of young and adult people’s political participation

methods, data, analyses | Vol. 12(2), 2018, pp. 309-334 DOI: 10.12758/mda.2017.12

Blaming the Young Misses the Point:
Re-assessing Young People’s Political
Participation over Time Using the
‘Identity-equivalence Procedure’

Christian Schnaudt¹ & Michael Weinhardt²

¹ GESIS - Leibniz Institute for the Social Sciences

² Bielefeld University

Sources

	Chat	Title	References
0	Ask	Blaming the Young Misses the Point: Re-assessing Young People’s Political Participation over Time Using the ‘Identity-equivalence Procedure’ Schnaudt methods, data, analyses	https://mda.gesis.org/index.php/mda/article/download/2017.12/228 0.738237

Strengths: Why This Approach Works

Combines Strengths of LLMs + KGs

- Natural language flexibility (LLMs)
- Structured, contextual enrichment (Knowledge Graphs)

Grounded, Transparent Answers

- Reduces hallucinations via local document grounding
- Includes source references and confidence scores

Multilingual Ready

- Uses multilingual KOS like Wikidata

Local, Adaptable Workflow

- Works on any document collection: PDFs, websites, Dataverse datasets
- Transparent, open-source tooling (EverythingData, GhostWriter)

Challenges & Limitations

Ranking Issues

- Relevance ranking may favor local terms over document-level meaning

Loss of Nuance

- Domain-specific meanings can be lost in paraphrasing

Preconditions

- High-quality metadata & full text
- Access to KOS services
- API interoperability

Outlook & Broader Applications

- Our method can be easily implemented for museums, libraries, archives and other cultural heritage institutions **with their local collections**
- We could optimize the ranking theory in document-level (information from same document aggregate together) to improve logical consistency
- We could also introduce multiple meta-data to enhance the matching and answer

Takeaways

- GhostWriter is a prototype enabling researchers to “chat” with scholarly documents
- Blends the flexibility of LLMs with structure and traceability of KGs
- Supports Open Science and reproducibility

Thank you for your attention

Contact

Dr. Philipp Mayr, philipp.mayr@gesis.org

Preprint: <https://arxiv.org/abs/2505.11633>

Demo: <https://gesis.now.museum/>

Credit: **Slava Tykhonov** developed the system. Frequent updates!