Large Language Models for Research Data Management?! 2025 (LLMs4RDM 2025)

Magnus Bender[1,2], Sylvia Melzer[3,4], Ralf Möller[3] and Stefan Thiemann[5]

[1] *Aarhus University, Center for Contemporary Cultures of Text, Jens Chr. Skous Vej 4, 8000 Aarhus C, Denmark*
[2] *Aarhus University, Department of Management, Fuglesangs Allé 4, 8210 Aarhus V, Denmark*
[3] *University of Hamburg, Institute for Humanities-Centered AI (CHAI), Warburgstraße 28, 20354 Hamburg, Germany*
[4] *University of Hamburg, Centre for the Study of Manuscript Cultures (CSMC), Warburgstraße 26, 20354 Hamburg, Germany*
[5] *University of Hamburg, Center for Sustainable Research Data Management, Monetastraße 4, 20146 Hamburg, Germany*

Abstract

Research data management (RDM) has become an important discipline that enables researchers to effectively organise, preserve and share their research results. RDM is a new development that aims to prepare researchers for the future by building on the principles of open science. It utilises innovative approaches such as generative artificial intelligence (genAI), which is powered by large language models (LLMs), to complement traditional research methods. As data-driven research becomes increasingly complex, researchers often have to spend a lot of time learning how to manage, analyse and interpret large amounts of information. Traditional data literacy training can be time-consuming and doesn't always keep pace with evolving technologies and methods of analysis. Foundation models based on generative AI offer the potential to streamline this learning process. By automating data pre-processing, pattern recognition and even hypothesis generation, these models can lower the technical barriers to entry, allowing researchers to focus more on insights and discovery rather than spending excessive amounts of time mastering data skills. The objective of this workshop is an exchange of perspectives regarding the implementation of novel RDM approaches using LLMs or not, both past and prospective, in research and practice.

CEUR-WS.org/Vol-4140/preface.pdf

1. Introduction to the First Workshop on LLMs4RDM

The rapid rise of large language models (LLMs) has opened new possibilities for research across the humanities, transforming how scholars analyse texts, interpret complex datasets, and engage with cultural, linguistic and historical materials. While these technologies promise powerful new methods of discovery, they also bring conceptual, methodological and ethical challenges that humanities researchers are only beginning to explore.

The first Workshop on Large Language Models for the Humanities?! (LLMs4RDM) was created as a forum to address these emerging opportunities and questions. Bringing together researchers, practitioners and infrastructure specialists, the workshop aims to foster dialogue on how LLMs can be meaningfully integrated into research workflows, from data creation and curation to analysis, interpretation and communication. By examining both the potential and the limitations of these models, the workshop seeks to support a more informed, critical and creative engagement with LLMs.

Large Language Models for Research Data Management!? 2025 (LLMs4RDM 2025), Workshop at the INFORMATIK Festival 2025 (55th Annual Conference of the German Informatics Society), September 18, 2025, Potsdam, Germany
magnus@mgmt.au.dk (M. Bender); sylvia.melzer@uni-hamburg.de (S. Melzer); ralf.moeller@uni-hamburg.de (R. Möller); stefan.thiemannt@uni-hamburg.de (S. Thiemann)
https://person.au.dk/magnus@mgmt.au.dk (M. Bender); https://www.csmc.uni-hamburg.de/about/people/melzer.html (S. Melzer); https://www.philosophie.uni-hamburg.de/chai/personen/moeller.html (R. Möller); https://www.fdm.uni-hamburg.de/ueber-uns/mitarbeiter/thiemann.html (S. Thiemann)
0000-0002-1854-225X (M. Bender); 0000-0002-0144-5429 (S. Melzer); 0000-0002-3228-316X (R. Möller); 0000-0001-8300-2519 (S. Thiemann)
© 2025 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073
published 2025-12-22

1.1. Workshop Organisation

The LLMs4RDM 2025 Workshop was held as part the INFORMATIK Festival 2025 (55th Annual Conference of the German Informatics Society), September 18, 2025, Potsdam, Germany.

1.1.1. Organisers

- Magnus Bender, Aarhus University, Denmark
- Sylvia Melzer, University of Hamburg, Germany
- Ralf Möller, University of Hamburg, Germany
- Stefan Thiemann, University of Hamburg, Germany

1.2. Programme Committee of LLMs4RDM 2025

- Thomas Asselborn, University of Hamburg, Germany
- Magnus Bender, Aarhus University, Denmark
- Mahdi Jampour, University of Hamburg, Germany
- Sylvia Melzer, University of Hamburg, Germany
- Stefan Thiemann, University of Hamburg, Germany

1.3. Overview of papers

One keynote and five papers were presented at the workshop.

The keynote focuses on the crucial next step in Research Data Management (RDM), advocating for a transition from simple data immersion to structured scientific argumentation. The speaker presents historical examples, such as the 3D reconstruction of the theatre of Miletus, to illustrate how researchers formulate hypotheses about past human decisions, pointing out that these visualisations can be based either on empirical data or on imaginary 3D data generated via language models. The central argument is that RDM systems must evolve to support the formal representation of these scientific arguments so that they are machine-processable and the data used can be verifiably represented. The keynote calls for new RDM systems that not only host diverse data, but also enable researchers to use these resources directly for the development and validation of formal scientific hypotheses.

The first paper *Large Language Models in Labor Market Research Data Management: Potentials and Limitations* presents the application of LLMs within research data management (RDM), focusing specifically on tasks related to occupational data and labor market text interpretation. Through empirical studies, the researchers determined that LLMs struggle with the automated classification of job titles, often producing results that were less reliable and reproducible than those generated by traditional machine learning classifiers. The LLM tests using hermeneutical methods produced fundamentally inconsistent and unstable interpretations. The authors argue that LLMs are inadequate for tasks demanding methodological rigor or scientifically defensible classification due to their lack of consistency and interpretative depth. Therefore, LLMs should be used only as assistive tools for preliminary support functions.

The second paper *Challenges in Automatic Speech Recognition in the Research on Multilingualism* examines the significant challenges faced when applying Automatic Speech Recognition (ASR) technology, specifically the Whisper model, to complex spoken data collected for multilingualism research. The authors note that while commercial applications require clean, monolingual transcripts, linguistic studies require highly accurate recordings that capture every acoustic detail, including speech disorders and complex language switching between languages. Using Polish-German bilingual recordings from the LangGener corpus, the study identifies key shortcomings in ASR output, such as hallucinations and a problematic tendency towards code unification, which mis-transcribes or mis-translates embedded language elements.

The third paper *Improving Accessibility and Reproducibility by Guiding Large Language Models* presents proposes a combined method the general-purpose large language models (LLMs) and specialized research

data stored in Research Data Repositories (RDRs) by leveraging the expert knowledge of the data creators. The core innovation is the interpretation prompt, a field added during the data upload process that allows the expert data creator to provide specific instructions. When a user queries the RDR's LLM chatbot, this expert-generated prompt is prepended to the user's query, effectively guiding the LLM toward project-specific understanding. The authors demonstrate that these prompts result in more accurate, tailored responses by focusing the LLM's output and improving data accessibility and utility. Furthermore, the interpretation prompt facilitates automated reproducibility of research experiments by instructing the LLM to execute relevant algorithms or code associated with the data entry.

The forth paper *Talk to your Database: An open-source in-context Learning Approach to interact with Relational Databases through LLMs* presents an open-source large language model (LLM) framework designed to solve the Text-to-SQL problem through in-context learning. Researchers compared the performance of this method against a simpler default prompting technique using a PostgreSQL database. The results decisively show that in-context learning improved accuracy, boosting successful query execution rates from approximately 35% to over 85%.

The fifth paper *Verbalisation Process of a RAG-Based Chatbot to Support Tabular Data Evaluation for Humanities Researchers* presents the verbalization process of a RAG-based chatbot (ChatHA) engineered to support tabular data evaluation for humanities researchers. The core motivation for this research is enabling scholars to conduct free-form and semantic searches on structured data, moving beyond the limitations of simple string matching. To address the need for verbalizing database entries into natural language, the authors propose a hybrid verbalization method that minimizes the computational cost and risk of hallucination associated with LLMs.

2. Presentations

Abstracts and presentations are available at: https://doi.org/10.25592/uhhfdm.17955

2.1. Presentations

Magnus Bender
Aarhus University, Denmark
Welcome

Ralf Möller
University of Hamburg, Germany
Keynote

Jens Dörpinghaus[1,2,3], Michael Tiemann[1,2]
[1]*University of Koblenz, Germany;* [2]*Federal Institute for Vocational Education and Training (BIBB), Germany;* [3]*Linnaeus University, Sweden*
Large Language Models in Labor Market Research Data Management: Potentials and Limitations

Edyta Jurkiewicz-Rohrbacher[1,2], Thomas Asselborn[2]
[1]*University of Regensburg, Germany;* [2]*University of Hamburg, Germany*
Challenges in Automatic Speech Recognition in the Research on Multilingualism

Florian Marwitz, Marcel Gehrke
University of Hamburg, Germany
Improving Accessibility and Reproducibility by Guiding Large Language Models

Maximilian Plazotta, Meike Klettke
University of Regensburg, Germany
Talk to your database: An open-source in-context learning approach to interact with relational databases through LLMs

Thomas Asselborn[1], Magnus Bender[2], Florian Marwitz[1], Ralf Möller[1], Sylvia Melzer[1]
[1]*University of Hamburg, Germany;* [2]*Aarhus University, Denmark*
Verbalisation Process of a RAG-Based Chatbot to Support Tabular Data Evaluation for Humanities Researchers

Magnus Bender
Aarhus University, Denmark
Farewell

2.1.1. Acknowledgments

The organisers of the LLMs4RDM 2025 workshop would like to thank the organisers of the Informatik Festival conference in Potsdam for their excellent support. We also would like to thank the members of the programme committee for their help in carefully evaluating and selecting the submitted papers and all participants of the workshop for their contributions. We wish that new inspirations and collaborations between the contributing disciplines will emerge from this workshop.

Funding Information

This contribution was partially funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany´s Excellence Strategy – EXC 2176 'Understanding Written Artefacts: Material, Interaction and Transmission in Manuscript Cultures', project no. 390893796. The research was mainly conducted within the scope of the Centre for the Study of Manuscript Cultures (CSMC) at University of Hamburg.

This contribution was partially funded by the Danish National Research Foundation (DNRF193) through TEXT: Centre for Contemporary Cultures of Text at Aarhus University.

Declaration on Generative AI

During the preparation of this work, the authors used DeepL in order to: Grammar and spelling check. After using these tool(s)/service(s), the authors reviewed and edited the content as needed and take(s) full responsibility for the publication's content.