Preface
First International Workshop on Scholarly Information Access at ECIR 2025

1. Introduction

These are the proceedings of the First International Workshop on Scholarly Information Access (SCOLIA 2025)[1]. SCOLIA 2025 was held as a half-day in-person event at the European Conference on Information Retrieval (ECIR) in Lucca, Italy. The aim of the SCOLIA workshop, following the successful and long-established BIR workshop series [1, 2], is to bring together researchers and practitioners from Information Retrieval (IR), Natural Language Processing (NLP), and Scientometrics/Bibliometrics who are working on the analysis of scientific/scholarly documents. Bibliometrics and Scientometrics are concerned with all quantitative aspects of information and academic literature, which naturally make them interesting for IR research, as well as domains in which citations play a central role, for example, legal and patent retrieval. The newly established SCOLIA workshop has a broader scope than the previous BIR workshop series. This allows us to cater for recent developments in NLP and Generative AI, which are becoming more important for SCOLIA's core focus on scholarly information access and recommendation.

2. Overview of the papers

The following eight research papers are included in the proceedings. All articles were peer-reviewed by at least two experts in the field. The works were presented in two sessions during the workshop. All presentation slides are available online[2].

Session 1: Publications and applications

- Birger Larsen and Roman Jurowetzki:
  Using reasoning for citation context classification
- Dirk Tunger and Philipp Schaer:
  On the Alignment of Post-Publication Reviews & Bibliometric and Altmetric Impact
- Fabian Richter, Benjamin Schäfer and Klemens Böhm:
  A review of query systems for temporal n-gram corpora
- Fabian Karl and Ansgar Scherp:
  CRAWLDoc: A Dataset for Robust Ranking of Bibliographic Documents

Session 2: RAG and domain specificity

- Pierre Achkar, Tim Gollub and Martin Potthast:
  Ask, Retrieve, Summarize: A Modular Pipeline for Scientific Literature Summarization
- Maël Lesavourey and Gilles Hubert:
  Reshaping Biomedical Scientific Literature in a RAG Pipeline for Question Answering

---

[1]https://sites.google.com/view/bir-ws/scolia-2025
[2]https://figshare.com/collections/SCOLIA_2025_ECIR_-_Presentations/7778048

- Stavroula Svolou, Fotis Aisopos, Anastasios Nentidis, Anastasia Krithara and Georgios Paliouras:
  A framework for disease-specific information extraction from biomedical literature and open databases, aiming at drug re-purposing
- Shufan Jiang, Mary Ann Tan and Harald Sack:
  Towards Disambiguation of Mathematical Terms based on Semantic Representations

3. Programme Committee

We thank our reviewers.

- Akiko Aizawa; National Institute of Informatics, Japan
- Birger Larsen; Aalborg University, Denmark
- Cyril Labbé; Université Paris Cité, France
- Dietmar Wolfram; University of Wisconsin-Milwaukee, USA
- Gilles Hubert; IRIT, France
- Haiko Lietz; GESIS - Leibniz Institute for the Social Sciences, Germany
- Hermann Kroll; TU Braunschweig, Institute for Information Systems, Germany
- Kanishka Silva; University of Wolverhampton, UK
- Katriina Byström; Oslo Metropolitan University, Norway
- Lucy Lu Wang; University of Washington, Allen Institute for AI, USA
- Marc Bertin; Université Claude Bernard Lyon 1, France
- Norbert Fuhr; University of Duisburg-Essen, Germany
- Patrice Bellot; Aix-Marseille Université - CNRS (LIS), France
- Petr Knoth; The Open University, UK
- Philipp Schaer; TH Köln (University of Applied Sciences), Germany
- Rajesh Piryani; IRIT - Equipe MELODI Universite Paul Sabatier, France
- Ralf Schenkel; Trier University, Germany
- Rob Koopman; European Space Agency, France
- Roman Kern; TU Graz, Austria
- Wolf-Tilo Balke; TU Braunschweig, Institute for Information Systems, Germany

Acknowledgments

Declaration on Generative AI

The authors have not employed any Generative AI tools.

The SCOLIA 2025 Organisers

Ingo Frommholz, University of Wolverhampton, UK[3]
Philipp Mayr, GESIS – Leibniz-Institute for the Social Sciences, Cologne, Germany
Guillaume Cabanac, University of Toulouse, France
Suzan Verberne, Leiden University, The Netherlands
Christin Katharina Kreutz, TH Mittelhessen – University of Applied Sciences & Herder Institute, Germany

---

[3]New affiliation: Modul University Vienna, Austria

References

[1] P. Mayr, A. Scharnhorst, B. Larsen, P. Schaer, P. Mutschke, Bibliometric-enhanced information retrieval, in: M. de Rijke, T. Kenter, A. P. de Vries, C. Zhai, F. de Jong, K. Radinsky, K. Hofmann (Eds.), Advances in Information Retrieval - 36th European Conference on IR Research, ECIR 2014, Amsterdam, The Netherlands, April 13-16, 2014. Proceedings, volume 8416 of *Lecture Notes in Computer Science*, Springer, 2014, pp. 798–801. URL: https://doi.org/10.1007/978-3-319-06028-6_99. doi:10.1007/978-3-319-06028-6_99.

[2] I. Frommholz, P. Mayr, G. Cabanac, S. Verberne, Bibliometric-enhanced information retrieval: 14th international BIR workshop (BIR 2024), in: N. Goharian, N. Tonellotto, Y. He, A. Lipani, G. McDonald, C. Macdonald, I. Ounis (Eds.), Advances in Information Retrieval - 46th European Conference on Information Retrieval, ECIR 2024, Glasgow, UK, March 24-28, 2024, Proceedings, Part V, volume 14612 of *Lecture Notes in Computer Science*, Springer, 2024, pp. 442–446. URL: https://doi.org/10.1007/978-3-031-56069-9_61. doi:10.1007/978-3-031-56069-9_61.