LITHUANIAN COMPUTER SOCIETY
VILNIUS UNIVERSITY, INSTITUTE OF DATA SCIENCE AND DIGITAL TECHNOLOGIES
LITHUANIAN ACADEMY OF SCIENCES

16th Conference on

DATA ANALYSIS METHODS
for Software Systems

November 27–29, 2025

Druskininkai, Lithuania, Hotel "Europa Royale"
https://www.mii.lt/DAMSS

VILNIUS UNIVERSITY PRESS
Vilnius, 2025

Co-Chairs:

Dr. **Saulius Maskeliūnas** (Lithuanian Computer Society)
Prof. **Gintautas Dzemyda** (Vilnius University, Lithuanian Academy of Sciences)

Programme Committee:

Dr. **Jolita Bernatavičienė** (Lithuania)
Prof. **Juris Borzovs** (Latvia)
Prof. **Janusz Kacprzyk** (Poland)
Prof. **Ignacy Kaliszewski** (Poland)
Prof. **Bożena Kostek** (Poland)
Prof. **Tomas Krilavičius** (Lithuania)
Prof. **Olga Kurasova** (Lithuania)
Assoc. Prof. **Tatiana Tchemisova** (Portugal)
Assoc. Prof. **Gintautas Tamulevičius** (Lithuania)
Prof. **Julius Žilinskas** (Lithuania)

Organizing Committee:

Dr. Jolita Bernatavičienė
Prof. Olga Kurasova
Assoc. Prof. Viktor Medvedev
Laima Paliulionienė
Assoc. Prof. Martynas Sabaliauskas
Prof. Povilas Treigys

Contacts:

Dr. Jolita Bernatavičienė
jolita.bernataviciene@mif.vu.lt
Prof. Olga Kurasova
olga.kurasova@mif.vu.lt
Tel. (+370 5) 2109 315

https://doi.org/10.15388/DAMSS.16.2025
ISBN 978-609-07-1200-9 (digital PDF)

Enhancing Requirements Engineering QA with Retrieval-Augmented Generation and Knowledge Graphs

Darius Sabaliauskas, Jolanta Miliauskaitė

Institute of Data Science and Digital Technologies
Vilnius University

darius.sabaliauskas@mif.stud.vu.lt

Automated question answering (QA) systems in requirements engineering (RE) can greatly speed up the processes of specification analysis, validation, and decision-making. While traditional QA models, such as those based on BERT architectures, excel at extracting specific spans of text, they often encounter difficulties with long, fragmented, and domain-specific requirement texts [1-2]. Retrieval-Augmented Generation (RAG) improves the precision of answers by using external knowledge during the inference stage [3]. In this research, we assess five RAG strategies for RE QA like sparse lexical retrieval with BM25 [4], dense vector-based retrieval using semantic embeddings [2], a hybrid semantic reranking process utilizing cross-encoders [1], and graph-enhanced retrieval leveraging concept-based knowledge expansion [5]. Furthermore, we implement a multi-hop retrieval extension that incorporates entity-level reasoning to uncover contextual evidence spanning multiple segments [6]. These methods are designed to alleviate information overload by pinpointing potentially relevant parts of requirement documents, potentially enhancing answer precision and interpretability in subsequent reasoning activities. We assess the models using a domain-specific RE dataset with four standard QA metrics: Exact Match (EM) and F1 score for span-level accuracy, ROUGE-L for lexical similarity, and BERTScore for semantic alignment. The experimental outcomes indicate that retrieval-based approaches do not consistently surpass the non-retrieval BERT+LSTM baseline in all scenarios. Although some configurations, especially dense and multi-hop retrieval, show localized improvements in semantic alignment and contextual relevance, other RAG variants perform similarly or worse.

These results imply that while retrieval augmentation can boost performance in particular cases, its efficacy is largely contingent on the retrieval strategy and the quality of context selection, underscoring the necessity for meticulous method design in RE-focused QA systems.

References

[1] Fu, H., Yao, Y., Lin, J., & Liu, Z. (2023). S-BERT: A semantic information selecting approach for open-domain question answering. arXiv:2305.17413.

[2] Karpukhin, V., et al. (2020). Dense Passage Retrieval for Open-Domain Question Answering. ACL.

[3] Lewis, P., et al. (2020). Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. NeurIPS.

[4] Robertson, S., & Zaragoza, H. (2009). The probabilistic relevance framework: BM25 and beyond. Foundations and Trends in IR, 3(4), 333–389.

[5] Xu, X., et al. (2024). GraphTrace: A modular retrieval framework combining knowledge graphs and LLMs for multi-hop question answering. NeurIPS.

[6] Zhao, K., et al. (2024). Advancing Large Language Models with Enhanced RAG: Evidence from Biological UAV Swarm Control. Aerospace Science and Technology.

.