



UNIVERSITÀ
di CAMPORNO



University of Applied Sciences and Arts
Northwestern Switzerland

A Hybrid AI Approach for Recommending Collaborators in Research Projects

FHNW Supervisor

Prof. Emanuele Laurenzi

Candidate

Piermichele Rosati

UNICAM Supervisor

Prof. Michela Quadrini

MSc in Computer Science and Business Information Systems

April 7, 2025

Consortia

A **consortium** is a group of organizations or individuals that collaborate to achieve a common objective [1].

Successful research proposals require a **consortium** experienced in the topics of the call.

Consortia advantages:

- ✓ Shared resources and infrastructure
- ✓ Access to complementary expertise
- ✓ Cost sharing and risk reduction
- ✓ Improved innovation and research outcomes

Related Work

- ▶ Traditional **recommender systems** such as Content-Based Filtering (CBF) and Collaborative Filtering (CF) suggest relevant items based on user preferences and item features [2, 3, 4]
- ▶ **Deep learning-based models** get high accuracy but do not offer explainable results [5]
- ▶ **Large Language Models (LLMs)**, especially when enhanced with **Retrieval Augmented Generation (RAG)**, offer explainability and improved contextual understanding [6]
- ▶ A promising direction combining **Knowledge Graphs (KGs)** and **LLMs** is explored in [7, 8]

Problem Statement

Forming a consortium is challenging and requires finding suitable research collaborators.

- ▶ Disciplinary and cognitive differences between researchers make it challenging to find suitable collaborators
- ▶ Traditional methods rely solely on social networks and, or the number of author citations, which proved to be limited in efficacy.

My thesis in a nutshell:

An approach that integrates LLMs with KGs could offer a promising alternative to existing methods.

Research Questions

Main Research Question:

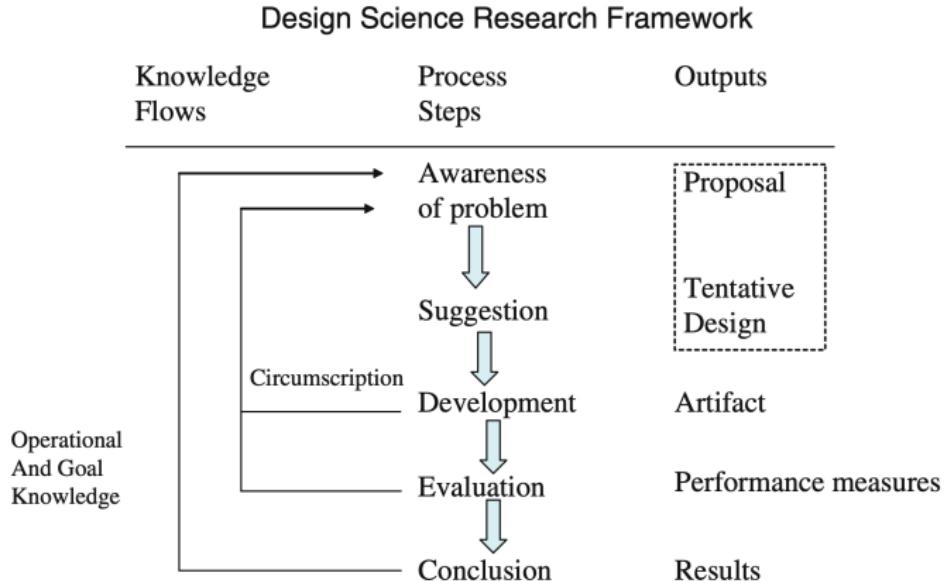
How can a KG and LLM-based approach enhance the process of suggesting collaborators for research projects?

Sub-Research Questions (SRQs):

- ▶ **SRQ1:** *How can research-related data (e.g., projects, people, affiliations) be modeled into a KG for research collaborator recommendation?*
- ▶ **SRQ2:** *How can a KG and LLM-based system be designed to efficiently retrieve relevant information from large, heterogeneous data sources to support personalized recommendations?*
- ▶ **SRQ3:** *How can the system generate human-readable explanations for its research collaborator recommendations using the KG and LLM outputs?*

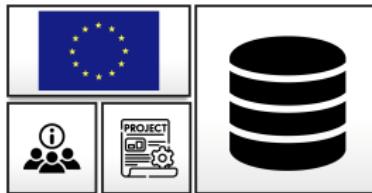
Research Methodology

Design Science Research (Hevner and Chatterjee [9] adapted from Vaishnavi [10])



Dataset

Community Research and Development Information Service (CORDIS): the EU Commission's main public source for EU-funded research projects [11].



The dataset includes projects information and details from the **FP7** [12] and **H2020** [13] programmes.

Project Information

BIMERR

Grant agreement ID: 820621

[Project website](#) ↗

DOI

[10.3030/820621](#) ↗

Project closed

EC signature date

3 August 2018

Start date

1 January 2019

End date

30 September 2022

Funded under

INDUSTRIAL LEADERSHIP - Leadership in enabling and industrial technologies - Advanced manufacturing and processing

Total cost

€ 6 933 320,01

EU contribution

€ 6 933 320,00



Coordinated by

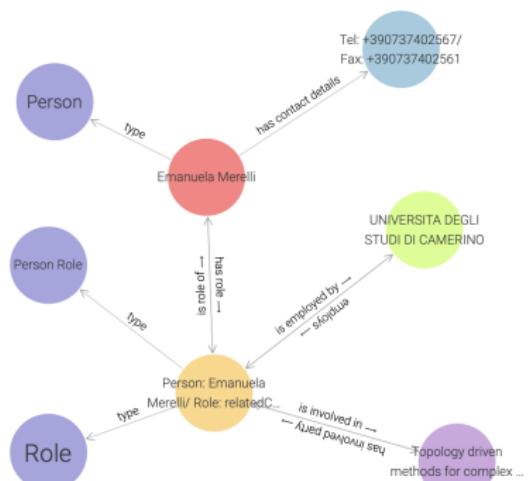
FRAUNHOFER GESELLSCHAFT ZUR FORDERUNG DER ANGEWANDTEN FORSCHUNG EV

Germany

Ontology Selection

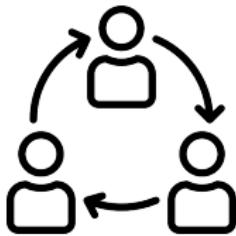
The European Research Information Ontology (EURIO): data model that formalizes and makes available structured, machine-readable data on EU-funded research projects.

- ▶ Semantic model of research information, e.g., projects, calls, funding schemes, organizations, people ...
- ▶ **EURIO KG [14]:** an RDF KG built from CORDIS data



Recommendation Types

The recommendation strategy leverages structured data properties and relationships in the EURIO KG to generate context-aware suggestions.



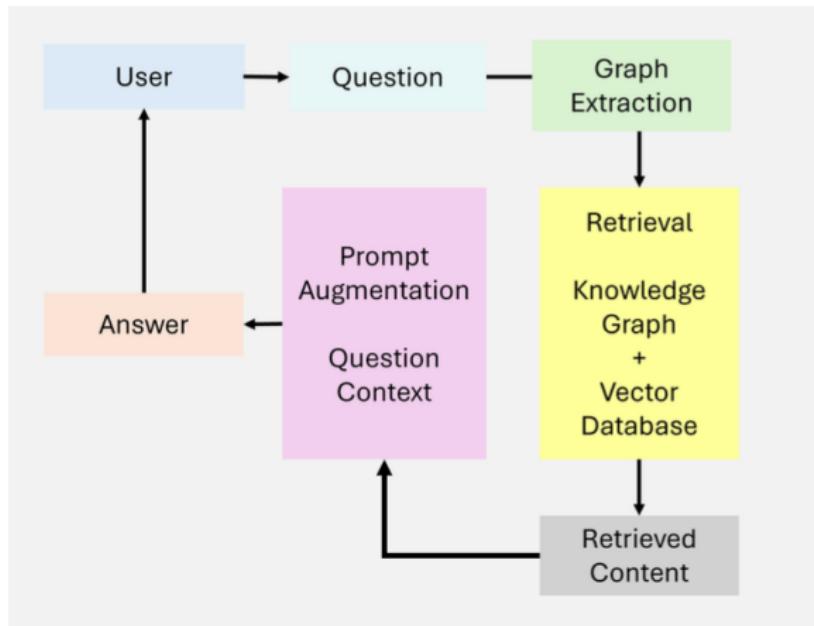
**Research collaborators
recommendation**



**Consortium organisations
recommendation**

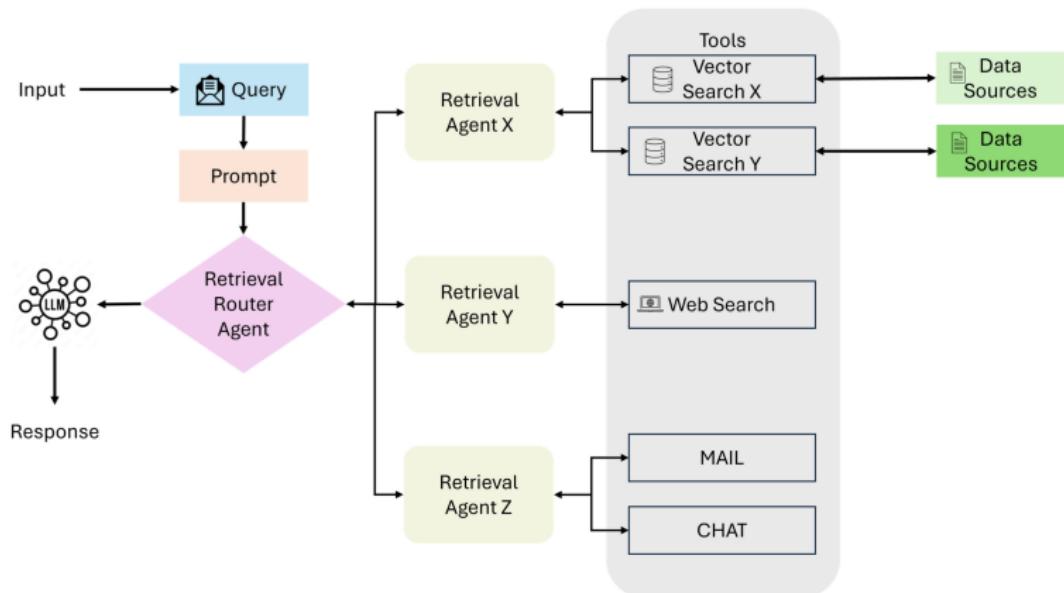
GraphRAG Architecture

(from Singh et al. [15])

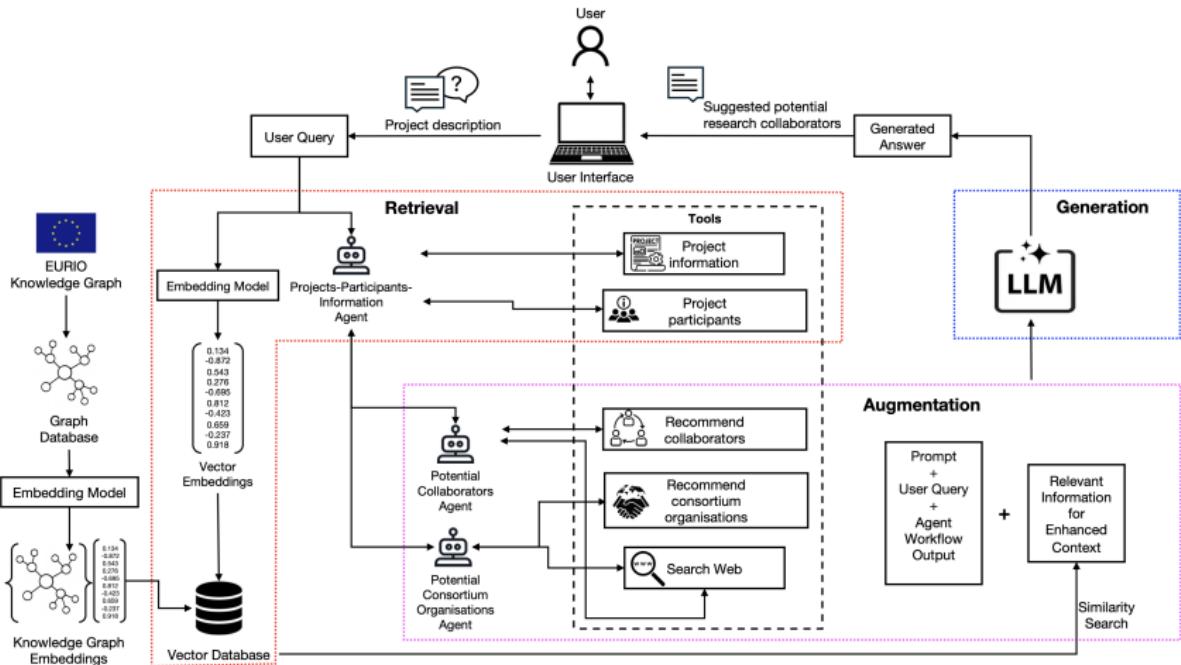


AgenticRAG Architecture

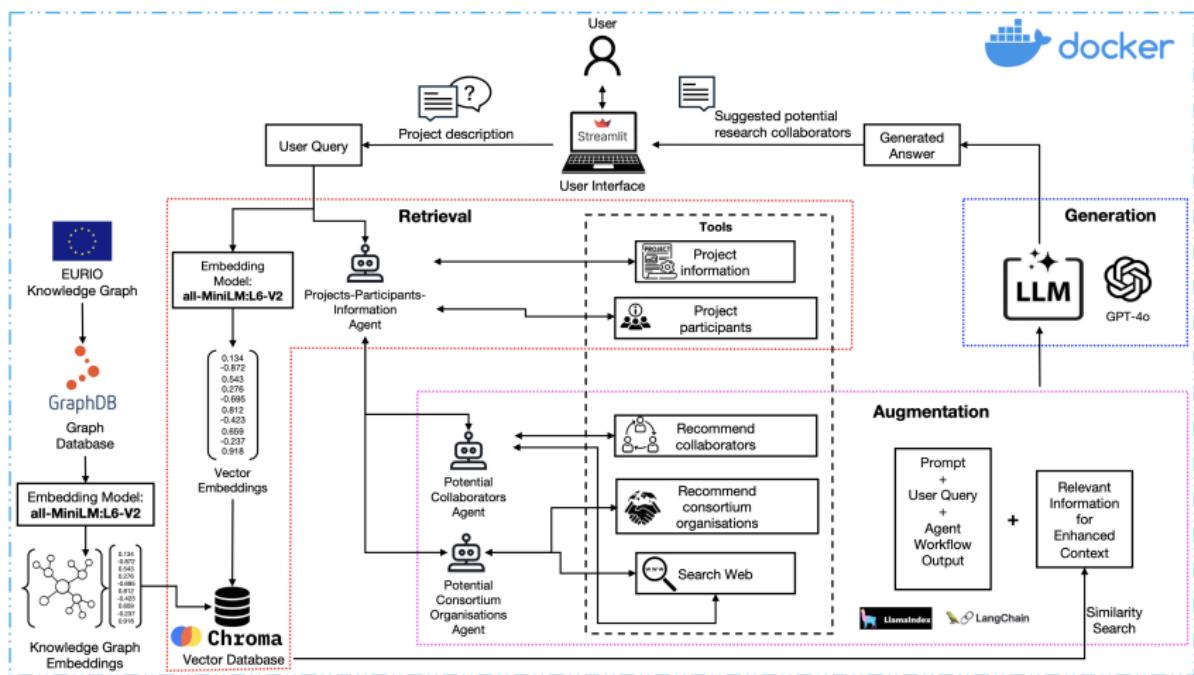
(from Singh et al. [15])



Artifact Design



Technology Stack



Evaluation

Goal: Evaluate the performance of the Agentic Graph RAG pipeline for collaborator and consortia recommendations.

Evaluation Framework: Retrieval-Augmented Generation Assessment (RAGAs) [16]

- ▶ Assesses both **retriever** and **generator** components
- ▶ Quantitative evaluation using automatic LLM-based metrics

Retriever Metrics:

- ▶ **Context Precision (CP)**: relevant chunks at top ranks
- ▶ **Context Recall (CR)**: how well context supports the answer
- ▶ **Context Entity Recall (CER)**: entity overlap between ground truth and retrieved context

Evaluation

Generator Metrics:

- ▶ **Faithfulness**: factual consistency with context
- ▶ **Answer Relevancy (AR)**: alignment with user prompt
- ▶ **Answer Semantic Similarity (SS)**: similarity to ground truth answer
- ▶ **Answer Correctness (AC)**: combined semantic and factual accuracy

Evaluation Datasets:

- ▶ Two manually annotated sets:
 - ▷ potential collaborator recommendations
 - ▷ potential consortia recommendations
- ▶ 21 queries each, derived from real EU project descriptions

Evaluation: Query Example

Project Description:

'The project aims to advance Artificial Intelligence (AI) for Connected, Cooperative, and Automated Mobility (CCAM) by enhancing situational awareness, predictive decision-making, and safety in time-critical scenarios. The initiative will develop AI-driven solutions that integrate seamlessly with active safety systems while ensuring trustworthy, explainable, and human-centric AI to increase public acceptance and usability. The project will focus on predictive system state awareness, moving beyond current reactive and adaptive AI approaches to anticipatory AI-driven automation.

...

Project Objectives:

- OBJ1: Develop AI-based situational awareness, prediction, and decision-making models to improve safety-critical CCAM applications.
- OBJ2: Enhance AI training and validation methodologies by leveraging real and synthetic traffic event datasets, ensuring unbiased and ethical AI development.

...

Which organisations could be potential collaborators to form a consortium for the specified project description and objectives?

Results

Values closer to 1 denote better performance; near 0 indicate worse.

Collaborators Recommendation Evaluation Results

Faithfulness	AR	CP	CR	CER	SS	AC
0.140	0.249	0.630	0.464	0.299	0.791	0.753

Consortium Organisations Recommendation Evaluation Results

Faithfulness	AR	CP	CR	CER	SS	AC
0.436	0.525	0.733	0.604	0.457	0.783	0.736

- ▶ **AR:** Answer Relevancy
- ▶ **CP:** Context Precision
- ▶ **CR:** Context Recall
- ▶ **CER:** Context Entity Recall
- ▶ **SS:** Answer Semantic Similarity
- ▶ **AC:** Answer Correctness

Contribution

Main Research Question:

How can a KG and LLM-based approach enhance the process of suggesting collaborators for research projects?

Research contributions:

- ▶ Investigated the limitations of existing approaches for suggesting collaborators.
- ▶ Proposed an approach that leverages KGs and LLMs to enhance the research collaboration process.

Application contributions:

- ▶ Designed and developed a user-friendly system for recommending research collaborators using real-world data
- ▶ Demonstrated the system's ability to retrieve and explain relevant collaborator recommendations using semantic enrichment

Future Directions

- ▶ **Automatic updating of the KG** (including information about Horizon Europe projects) to keep up with emerging research topics and thus keep the proposed approach relevant over time
- ▶ The **integration of academic KGs** could be a useful resource to enrich the EURIO knowledge base and thus add context regarding the recommendation of papers and researchers.
- ▶ The construction of **more detailed evaluation datasets** provided by experts could be future work that could help improve the artifact evaluation

Conclusion

Agentic Graph RAG was introduced to deliver contextual and explainable recommendations for research collaborators by combining **KGs** and **LLMs**.

Evaluation Highlights:

- ▶ High-quality and contextual reasoning, reduced hallucinations
- ▶ Areas needing improvement: *consistency* and *context retrieval*

- ▶ This thesis work has been submitted to the **Society 5.0 conference**
- ▶ A journal version is in preparation for submission to the **Semantic-Web Journal**

A Hybrid AI Approach for Recommending Collaborators in Research Projects

Piermichele Rosati¹(ORCID: 0009-0000-2038-8630), Emanuele Laemmeli¹(ORCID: 0009-0001-3042-7400), and Michela Quadrini²(ORCID: 0009-0010-0530-0200)

¹ FHNW - University of Applied Sciences and Arts Northwestern Switzerland
emanuele.laemmeli@fhnw.ch, piermichele.rosati@students.fhnw.ch
² School of Sciences and Technology, University of Camerino, Italy
michela.quadrini@unicam.it, piermichele.rosati@students.unicam.it

Abstract. The success of research project proposals heavily depends on the connections, which should be experienced and knowledgeable in the topics outlined in the corresponding calls, e.g., those in the EU's research and innovation programme Horizon Europe. Yet, one of the main chal-

The End



References (1)

- [1] J.Sylvan Katz and Ben R. Martin. "What is research collaboration?" In: *Research Policy* 26.1 (1997), pp. 1–18. ISSN: 0048-7333. DOI: [10.1016/S0048-7333\(96\)00917-1](https://doi.org/10.1016/S0048-7333(96)00917-1).
- [2] Dimitris Plexousakis and Themistoklis Kutsuras. "Alleviating the Sparsity Problem of Collaborative Filtering Using Trust Inferences". In: vol. 3477. May 2005, pp. 224–239. ISBN: 978-3-540-26042-4. DOI: [10.1007/11429760_16](https://doi.org/10.1007/11429760_16).
- [3] Jian Wei et al. "Collaborative filtering and deep learning based recommendation system for cold start items". In: *Expert Systems with Applications* 69 (2017), pp. 29–39. ISSN: 0957-4174. DOI: [10.1016/j.eswa.2016.09.040](https://doi.org/10.1016/j.eswa.2016.09.040).
- [4] Linyuan Lü et al. *Recommender systems*. Oct. 2012. DOI: [10.1016/j.physrep.2012.02.006](https://doi.org/10.1016/j.physrep.2012.02.006).
- [5] Zihuai Zhao et al. " Recommender Systems in the Era of Large Language Models (LLMs) ". In: *IEEE Transactions on Knowledge & Data Engineering* 36.11 (Nov. 2024), pp. 6889–6907. ISSN: 1558-2191. DOI: [10.1109/TKDE.2024.3392335](https://doi.org/10.1109/TKDE.2024.3392335).
- [6] Yashar Deldjoo et al. "A Review of Modern Recommender Systems Using Generative Models (Gen-RecSys)". In: *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. Association for Computing Machinery, Aug. 2024, pp. 6448–6458. ISBN: 9798400704901. DOI: [10.1145/3637528.3671474](https://doi.org/10.1145/3637528.3671474).

References (2)

- [7] Linyao Yang et al. "Give us the Facts: Enhancing Large Language Models With Knowledge Graphs for Fact-Aware Language Modeling". In: *IEEE Transactions on Knowledge and Data Engineering* 36.7 (2024), pp. 3091–3110. DOI: [10.1109/TKDE.2024.3360454](https://doi.org/10.1109/TKDE.2024.3360454)
- [8] Shirui Pan et al. "Unifying Large Language Models and Knowledge Graphs: A Roadmap". In: *IEEE Transactions on Knowledge and Data Engineering* 36.7 (2024), pp. 3580–3599. DOI: [10.1109/TKDE.2024.3352100](https://doi.org/10.1109/TKDE.2024.3352100).
- [9] Alan Hevner and Samir Chatterjee. "Design Science Research in Information Systems". In: *Design Research in Information Systems: Theory and Practice*. Boston, MA: Springer US, 2010, pp. 9–22. ISBN: 978-1-4419-5653-8. DOI: [10.1007/978-1-4419-5653-8_2](https://doi.org/10.1007/978-1-4419-5653-8_2).
- [10] V. K. Vaishnavi. *Design Science Research Methods and Patterns: Innovating Information and Communication Technology*. 1st. Auerbach Publications, 2007. DOI: [10.1201/9781420059335](https://doi.org/10.1201/9781420059335).
- [11] Publications Office of the European Union. *CORDIS reference data*. Data set. Originally published in 2015. 2018. URL: <http://data.europa.eu/88u/dataset/cordisref-data>.
- [12] Publications Office of the European Union. *CORDIS - EU research projects under FP7 (2007-2013)*. Data set. 2015. DOI: [10.2906/112117098108/11](https://doi.org/10.2906/112117098108/11).
- [13] Publications Office. *CORDIS - EU research projects under Horizon 2020 (2014-2020)*. Data set. 2015. DOI: [10.2906/112117098108/12](https://doi.org/10.2906/112117098108/12).

References (3)

- [14] Publications Office of the European Union. *CORDIS - EU research projects data from the EURIO Knowledge Graph (in the form of a database dump and named graphs)*. Accessed: 2025-02-10. 2022. URL: <http://data.europa.eu/88u/dataset/named-graphs-from-eurio-knowledge-graph>.
- [15] Aditi Singh et al. *Agentic Retrieval-Augmented Generation: A Survey on Agentic RAG*. 2025. arXiv: 2501.09136 [cs.AI]. URL: <https://arxiv.org/abs/2501.09136>.
- [16] Shahul Es et al. "RAGAs: Automated Evaluation of Retrieval Augmented Generation". In: *Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics: System Demonstrations*. Ed. by Nikolaos Aletras and Orphee De Clercq. St. Julians, Malta: Association for Computational Linguistics, Mar. 2024, pp. 150–158. URL: <https://aclanthology.org/2024.eacl-demo.16/>.