Organizing Scholarly Knowledge in the Open Research Knowledge Graph

An Open-Science Platform for FAIR Scholarly Knowledge

Sören Auer[1][2][https://orcid.org/0000-0002-0698-2864], Markus Stocker[1][2][https://orcid.org/0000-0001-5492-3212],
Oliver Karras[1][https://orcid.org/0000-0001-5336-6899],
Allard Oelen[1][https://orcid.org/0000-0001-9924-9153],
Jennifer D'Souza[1][https://orcid.org/0000-0002-6616-9509], and
Anna-Lena Lorenz[1][https://orcid.org/0000-0002-1660-1463]

[1]TIB Leibniz Information Centre for Science and Technology, Germany

[2]L3S Research Center, University of Hannover, Germany

Abstract: The Open Research Knowledge Graph (ORKG) is an Open Science digital infrastructure for the production, curation, publication, and reuse of machine-actionable scholarly knowledge. Built on top of the RDF data model and extensible ontologies, the ORKG provides a common vocabulary for researchers to describe their research contributions and data, improving the discoverability and reusability of scholarly knowledge and research data. The ORKG includes tools for visualizing the relationships between different entities, making it easier to understand the connections between different pieces of research and their findings. It facilitates collaboration between researchers by providing a collaborative platform for organizing and sharing scholarly knowledge and data, reducing duplication and enabling more efficient use of resources. As research becomes increasingly data-driven, tools like the ORKG will become essential for enabling efficient, transparent, and collaborative research.

Keywords: Scholarly Communication, Knowledge Graph, Ontologies, Open-Science

1 Introduction

While many domains have significantly transformed in the digital age (e.g. encyclopedias, mail-order catalogs, street maps) scholarly communication is still based on static and relatively unstructured documents. This results in severe problems, such as a publication flood, deterioration of peer review, and inadequacy of machine assistance. The Open Research Knowledge Graph (ORKG) is a novel open-science platform that allows researchers to share and access research knowledge and data in a more efficient, transparent, and collaborative way [1], [2]. In this essay, we will explore what the ORKG is, how it works, and its potential benefits for research data stewards.

2 What is the ORKG?

The ORKG is a knowledge graph that connects scientific research papers, datasets, and methods. It is an open-science platform that provides a structured way of organizing and sharing research knowledge and data. The ORKG is designed to improve the discoverability and reusability of research data, enabling researchers to build on each other's work and collaborate more effectively. The core feature of the ORKG is to describe research contributions (e.g. published in an article) in a semantic manner. Based on these structured and semantic descriptions, the ORKG provides comparisons (i.e. overviews) on research contributions addressing the same research challenge. Based on the structured descriptions of research contributions a number of other features are provided. This includes visualizations, literature lists, and review articles comprising the structured content.
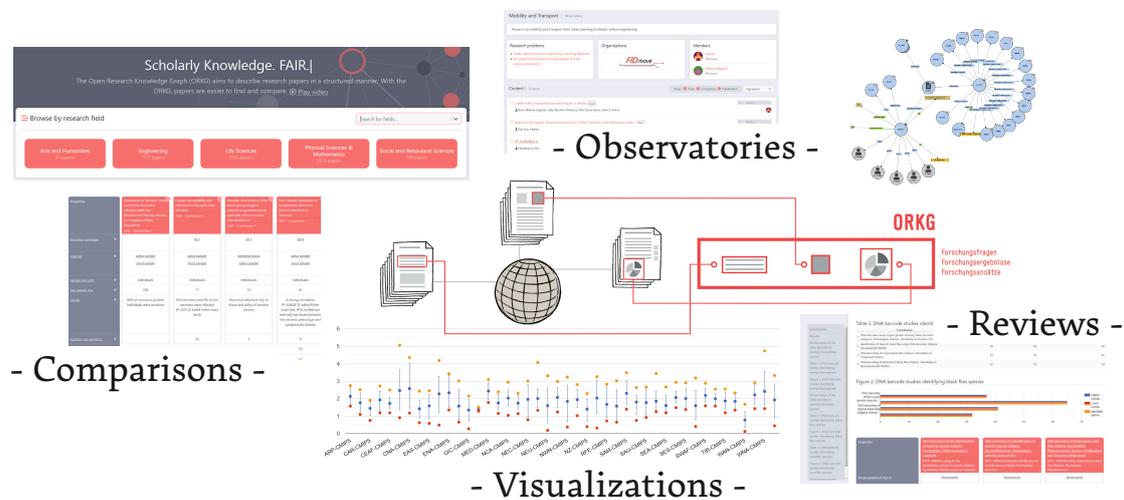

Figure 1. Semantic descriptions of research contributions can be arranged in comparisons and organized in domain-specific observatories, which in turn build the basis for visualizations, reviews, and other structured elements.

3 How does the ORKG work?

The ORKG is built on top of the Resource Description Framework (RDF) data model. Based on this data model contributors can add, curate, and organize descriptions of scientific contributions in a crowd-sourcing manner. New properties and resources for describing papers, datasets, and methods can be created on the fly, thus enabling a domain-specific and evolving ontological description of the relationships between different research entities. Pre-defined templates allow the structuring of typical, reusable knowledge patterns. The emerging domain ontologies provide a common vocabulary for researchers to describe their data, making it easier to find and understand.

The ORKG is also designed to be modular and extensible, allowing researchers to add and link to their own ontologies and customize the platform to suit their specific research needs. Automated extraction methods leveraging Natural Language Processing assist users in creating structured representations. The ORKG provides a user-friendly interface that allows researchers to create, edit, and search for research contributions.

The platform also includes tools for visualizing the relationships between different entities, making it easier to understand the connections between different pieces of research.

An example of an ORKG comparison of 25 studies regarding scenarios for the transition to renewable energy in Germany is shown in Figure 2 and available online at https://orkg.org/comparison/R153801/. The structured representation of the findings and scenarios from these studies enables novel ways of analysis e.g. about the installed capacity of renewable energy sources.
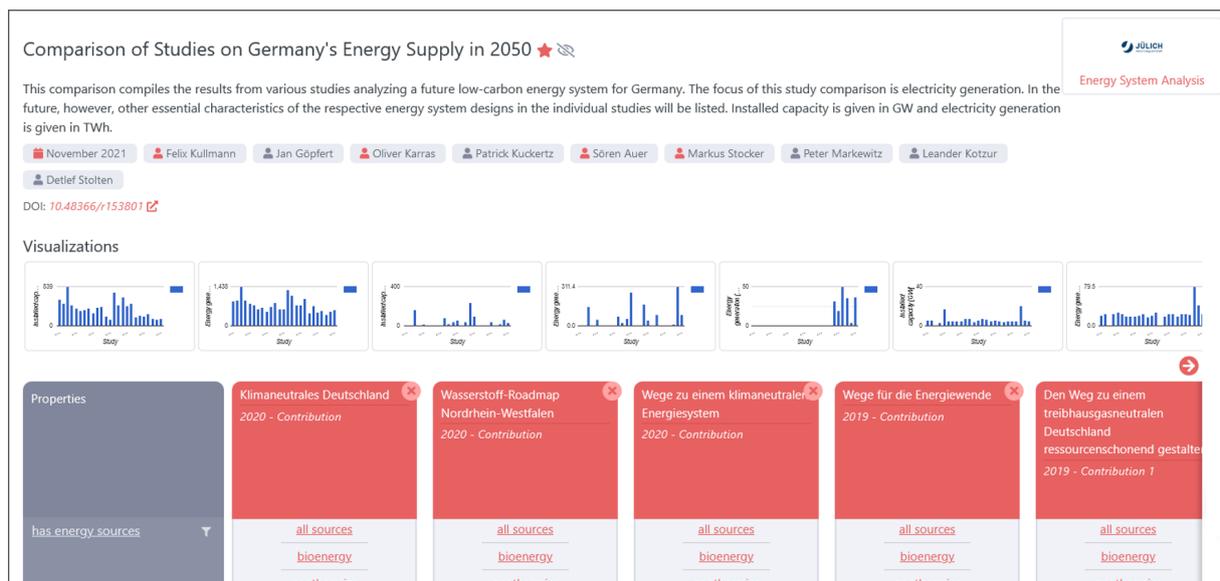

Figure 2. ORKG comparison of 25 studies regarding scenarios for the transition to renewable energy in Germany.

4 Benefits of the ORKG

The ORKG has several potential benefits for researchers. Firstly, it can improve the *discoverability* of research. By organizing research contributions into a structured format and providing a common vocabulary, the ORKG makes it easier for researchers to obtain comparative overviews of research approaches addressing a common research problem and to find relevant data for their work. This can save time and effort, as researchers no longer need to sift through large amounts of unstructured data.

Secondly, the ORKG can increase the *reproducibility* and transparency of research. By providing a clear description of the relationships between different entities and artifacts, the ORKG makes it easier for researchers to understand how data was generated and how it can be used. This can improve the reproducibility of research, as other researchers can more easily replicate experiments and verify results.

Thirdly, the ORKG can facilitate *collaboration* between researchers. By providing a shared platform for organizing and sharing research conceptualizations and data, the ORKG can enable researchers to work together more effectively (e.g. simplifying the creation of meta-analyses). This can lead to new insights and discoveries that would not have been possible otherwise.

Finally, the ORKG can help to address the issue of data silos in research. Many researchers collect and analyze data in isolation, leading to duplicate efforts and wasted resources. The ORKG provides a common platform for researchers to share their con-

tributions, data, and methods, thus reducing duplication and enabling more efficient use of resources.

5 Conclusion

The Open Research Knowledge Graph (ORKG) is an innovative novel tool for organizing and sharing research data. Built on top of the RDF data model and a set of ontologies, the ORKG provides a structured way of describing the relationships between different research entities. The ORKG has several potential benefits for research data enthusiasts, including improved discoverability, increased transparency, facilitated collaboration, and reduced data silos. As research becomes increasingly data-driven, tools like the ORKG will become essential for enabling efficient, transparent, and collaborative research.

References

[1] S. Auer, A. Oelen, M. Haris, *et al.*, "Improving access to scientific literature with knowledge graphs," *Bibliothek Forschung und Praxis*, vol. 44, no. 3, pp. 516–529, 2020. DOI: 10.1515/bfp-2020-2042.

[2] M. Stocker, A. Oelen, M. Y. Jaradeh, *et al.*, in *FAIR Connect*, B. Magagna, Ed., vol. 1, IOS Press, 2023, pp. 19–21. DOI: 10.3233/fc-221513.