KG-EmpiRE: A Community-Maintainable Knowledge Graph of Empirical Research in Requirements Engineering

Oliver Karras\textsuperscript{1}, Felix Wernlein\textsuperscript{2}, Jil Klünder\textsuperscript{2}, Sören Auer\textsuperscript{1,2}

\textbf{Abstract:} In this summary, we present our paper “Divide and Conquer the EmpiRE: A Community-Maintainable Knowledge Graph of Empirical Research in Requirements Engineering” [Ka23b] which received the best paper award of the 17th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement, 2023. Several papers address the topic of “Empirical research in requirements engineering (RE)” and provide a snapshot of its “current” state and evolution using literature reviews. However, these papers have never built on or updated earlier ones, as the underlying data is not available. Researchers need technical infrastructures to conduct sustainable literature reviews. We examine the use of the Open Research Knowledge Graph (ORKG) as such an infrastructure to build and publish a Knowledge Graph of Empirical research in RE (KG-EmpiRE) whose data is openly available. Our long-term goal is to continuously maintain KG-EmpiRE with the research community to synthesize a comprehensive, up-to-date, and long-term available overview. We conduct a literature review using the ORKG to build and publish KG-EmpiRE which we evaluate against competency questions. From 570 papers of the IEEE International Requirements Engineering Conference (2000 – 2022), we extract and analyze data on the reported empirical research and answer 16 out of 77 competency questions. These answers show a positive development of applied empirical research in RE, but also the need for future improvements. The ORKG is a ready-to-use and advanced infrastructure for organizing data from literature reviews as knowledge graphs, making the data openly available and maintainable by research communities and thus enabling sustainable literature reviews.

\textbf{Keywords:} Knowledge graph, empirical research, infrastructure, sustainability, literature review

\section{Introduction}

Empirical research in requirements engineering (RE) is a constantly evolving topic. Several papers have already pursued the same goal of synthesizing a comprehensive, up-to-date, and long-term available overview of the state and evolution of this topic. Despite having the same goal, using similar research methods, and even examining overlapping periods, venues, and themes, they have not collaborated to build on and update earlier works. The underlying problem is the unavailability of the extracted and analyzed data, as unavailable data complicates collaboration and updating literature reviews. While recent research mainly provides social and economic decision support and guidance for updating literature reviews, researchers need support in the form of technical infrastructures and services to conduct sustainable literature reviews so that all data is openly available in the long term.

\textsuperscript{1} TIB - Leibniz Information Centre for Science and Technology, Welfengarten 1B, 30167 Hannover, Deutschland, oliver.karras@tib.eu, soeren.auer@tib.eu

\textsuperscript{2} Leibniz University Hannover, Welfengarten 1, 30167 Hannover, Deutschland, felix.wernlein@stud.uni-hannover.de, jil.kluender@inf.uni-hannover.de, auer@l3s.de

doi:10.18420/sw2024_11
2 Research Approach

In our paper [Ka23b], we investigate the Open Research Knowledge Graph (ORKG) [St22] as a technical infrastructure by building, publishing, and evaluating an initial Knowledge Graph of Empirical research in RE (KG-EmpiRE). For this purpose, we collect all 751 papers from the research track of the IEEE International Requirements Engineering Conference (1993 – 2023) and extract (so far) data from 570 papers (2000 – 2022) on the reported empirical research, i.a., data collection, data analysis, and research design. In this way, we show how we can consistently organize scientific data on empirical research in RE using the ORKG to build and publish the initial KG-EmpiRE that the research community can constantly maintain, (re-)use, update, and expand. We evaluate KG-EmpiRE by analyzing its data to provide initial insights into the state and evolution of empirical research in RE. In particular, we answer 16 out of 77 competency questions derived from the published vision of how researchers should conduct empirical research in all fields of SE in 2020 – 2025.

3 Conclusion

Our initial findings show a positive development of the state and evolution of empirical research in RE towards the vision for 2020 – 2025, but also the need for future improvements. We provide all answers to the 16 competency questions with the associated analyses, visualizations, and explanations as an interactive Jupyter Notebook [Ka23a]. Based on our findings, we conclude the use of the ORKG and RKGs, in general, is a step in the right direction to allow researchers to build on and update earlier literature reviews, enabling sustainable literature reviews to ensure the quality, reliability, and timeliness of research results for successful long-term collaboration among researchers. Comprehensive, up-to-date, and long-term available overviews of the state and evolution of broad topics such as empirical research in RE are major research challenges that we as a research community can only conquer by dividing the efforts, true to the principle: Divide et Impera.

Data Availability

All data from this study is openly available in the ORKG. We also provide all materials, including our interactive Jupyter Notebook on GitHub and Zenodo [Ka23a].

Bibliography


[St22] Stocker, Markus; Heger, Tina; Schweidtmann, Artur; Ćwiek-Kupczyńska, Hanna; Penev, Lyubomir et al.: SKG4EOSC - Scholarly Knowledge Graphs for EOSC: Establishing a Backbone of Knowledge Graphs for FAIR Scholarly Information in EOSC. Research Ideas and Outcomes, 8, 2022.