DAGSTUHL
REPORTS

Volume 14, Issue 2, February 2024

Are Knowledge Graphs Ready for the Real World? Challenges and Perspective (Dagstuhl Seminar 24061)
David Chaves-Fraga, Oscar Corcho, Anastasia Dimou, Maria-Esther Vidal, Ana Iglesias-Molina, and Dylan Van Assche 1

Beyond-Planar Graphs: Models, Structures and Geometric Representations (Dagstuhl Seminar 24062)
Vida Dujmović, Seok-Hee Hong, Michael Kaufmann, János Pach, and Henry Förster 71

RSafety Assurance for Autonomous Mobility (Dagstuhl Seminar 24071)
Jyotirmoy Deshmukh, Bettina Könighofer, Dejan Ničković, and Filip Cano 95

Triangulations in Geometry and Topology (Dagstuhl Seminar 24072)
Maike Buchin, Jean Cardinal, Arnaud de Mesmay, Jonathan Spreer, and Alex He 120

Computational Approaches to Strategy and Tactics in Sports (Dagstuhl Seminar 24081)
Ulf Brefeld, Jesse Davis, Laura de Jong, and Stephanie Kovalchik 164

AI for Social Good (Dagstuhl Seminar 24082)
Claudia Clopath, Ruben De Winne, Mohammad Emtiyaz Khan, and Jacopo Margutti 182

Reflections on Pandemic Visualization (Dagstuhl Seminar 24091)
Daniel Archambault, Fintan McGee, Nicolás Reinoso-Schiller, Tatiana von Landesberger, and Simone Scheithauer 191

Applied and Combinatorial Topology (Dagstuhl Seminar 24092)
Paweł Dłotko, Dmitry Feichtner-Kozlov, Anastasios Stefanou, Yusu Wang, and Jan F Senge 206

ISSN 2192-5283

Published online and open access by
Schloss Dagstuhl – Leibniz-Zentrum für Informatik GmbH, Dagstuhl Publishing, Saarbrücken/Wadern, Germany. Online available at
https://www.dagstuhl.de/dagpub/2192-5283

Publication date
July, 2024

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at `https://dnb.d-nb.de`.

Digital Object Identifier: 10.4230/DagRep.14.2.i

Aims and Scope
The periodical *Dagstuhl Reports* documents the program and the results of Dagstuhl Seminars and Dagstuhl Perspectives Workshops.
In principal, for each Dagstuhl Seminar or Dagstuhl Perspectives Workshop a report is published that contains the following:
- an executive summary of the seminar program and the fundamental results,
- an overview of the talks given during the seminar (summarized as talk abstracts), and
- summaries from working groups (if applicable).

This basic framework can be extended by suitable contributions that are related to the program of the seminar, e. g. summaries from panel discussions or open problem sessions.

Contact
Schloss Dagstuhl – Leibniz-Zentrum für Informatik
Dagstuhl Reports, Editorial Office
Oktavie-Allee, 66687 Wadern, Germany
reports@dagstuhl.de
https://www.dagstuhl.de/dagrep

Report from Dagstuhl Seminar 24061

Are Knowledge Graphs Ready for the Real World? Challenges and Perspective

David Chaves-Fraga*[1], Oscar Corcho*[2], Anastasia Dimou*[3], Maria-Esther Vidal*[4], Ana Iglesias-Molina†[5], and Dylan Van Assche†[6]

1 **Universidade de Santiago de Compostela, ES.** `david.chaves@usc.es`
2 **Universidad Politécnica de Madrid, ES.** `ocorcho@fi.upm.es`
3 **KU Leuven, BE.** `anastasia.dimou@kuleuven.be`
4 **TIB – Hannover, DE.** `vidal@l3s.de`
5 **Universidad Politécnica de Madrid, ES.** `ana.iglesiasm@upm.es`
6 **IDLab - Ghent University - imec, BE.** `Dylan.VanAssche@ugent.be`

Abstract

This report documents the program and results of the Dagstuhl Seminar 24061 "Are Knowledge Graphs Ready for the Real World? Challenges and Perspectives". The seminar focused on gaining a better understanding of the open challenges required for the development of Knowledge Graph ecosystems. The seminar focused on four different topics: access control and privacy in decentralized knowledge graphs, knowledge graph construction lifecycle, software methods for improving KG implementation, and a new wave of knowledge engineers and their expected skills. By focusing on these relevant research topics, the seminar aimed to reflect on KGs from a more fundamental computer science perspective. It brought together interdisciplinary researchers from academia and industry to discuss foundations, concepts, and implementations that will pave the way for the next generation of KGs ready for real-world use.

Seminar February 4–9, 2024 – https://www.dagstuhl.de/24061

2012 ACM Subject Classification Information systems → Resource Description Framework (RDF); Information systems → Data exchange; Information systems → Graph-based database models; Information systems; Information systems → Data model extensions; Information systems → Query languages; Information systems → Data structures; Information systems → Information integration; Information systems → Semantic web description languages; Information systems → Web applications; Social and professional topics → Computing education; Security and privacy → Security services; Theory of computation → Semantics and reasoning; Information systems → Information retrieval

Keywords and phrases access control and privacy, federated query processing, intelligent knowledge graph management, programming paradigms for knowledge graphs, semantic data integration

Digital Object Identifier 10.4230/DagRep.14.2.1

* Editor / Organizer
† Editorial Assistant / Collector

Are Knowledge Graphs Ready for the Real World? Challenges and Perspective, *Dagstuhl Reports*, Vol. 14, Issue 2, pp. 1–70
Editors: David Chaves-Fraga, Oscar Corcho, Anastasia Dimou, and Maria-Esther Vidal
Dagstuhl Reports
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

1 Executive Summary

David Chaves-Fraga (Universidade de Santiago de Compostela, ES, david.chaves@usc.es)
Oscar Corcho (Universidad Politécnica de Madrid, ES, oscar.corcho@upm.es)
Anastasia Dimou (KU Leuven, BE, anastasia.dimou@kuleuven.be)
Maria-Esther Vidal (TIB – Hannover, DE, vidal@l3s.de)

Graphs and knowledge bases have been around for many decades, and research results have had a tremendous impact on areas such as mathematics, artificial intelligence, and databases. However, although the term has been coined by the scientific community, technological developments and astronomical data growth have made knowledge graph (KG) management a fundamental topic in various areas of computer science today. The scientific and industrial communities have responded to the emerging field of knowledge management. As a result, formal frameworks for defining and representing KGs, as well as methods for creating, exploring, and analyzing KGs, have flourished to make KGs a reality. However, despite the tangible results, sustainability is still compromised by the lack of transparent and accountable management of CCs. The real-world application of KGs requires programming paradigms for KG management, transparent data integration and quality assessment techniques, and methods for maintaining access control and privacy. In addition to technological advances, societal adjustments can have a tremendous impact on the management of KGs. The seminar addressed these socio-technical challenges with a mix of invited talks, lightning talks, and small group workshops as follows:

The Incremental Creation of Knowledge Graphs. Creating a Knowledge Graph (KG) involves several open research challenges, such as data extraction, data quality, data integration, and data security. It also requires attention to architectural aspects such as scalability and interoperability. A working group was formed to discuss and focus on two main topics: the definition of a general pipeline for KG construction and its relationship to data quality. The main outcome is a standard formalization of the KG construction lifecycle and its associated components. This definition is accompanied by quality measures and provenance tracking of all steps.

Support of Knowledge Graph Implementation. Software engineering and programming languages have created approaches and techniques that support complex tasks during software development such as software dependencies, error identification, testing, syntactic validation, software lifecycle, etc. We look into these proposals to determine a set of requirements in software lifecycle management for knowledge graphs. They will improve and facilitate the implementation of knowledge graphs in industrial and complex environments, taking into account the relationships and dependencies between all the artifacts used (ontologies, shapes, mappings, tests, etc.) as well as their evolution and versioning. To achieve this goal, we believe that it is necessary to have a better understanding and general overview of how knowledge graphs are implemented. Therefore, a workshop on this topic has been proposed at ISWC2024[1]. After its celebration, the next step will be to create a community around this topic with researchers and industry stakeholders to standardize and implement the identified challenges/requirements.

[1] `https://w3id.org/soflim4kg`

Access Control in Decentralized Knowledge Graphs. Exploring access control in decentralized Knowledge Graphs has been a relatively underexplored area. Specifically, mechanisms for restricting access to knowledge to safeguard confidential information and personal data, as well as establishing consent models for the processing of personal data, have not received substantial attention within the realm of Knowledge Graph management. Additionally, ensuring compliance with usage policies has been inadequately addressed, particularly in the context of decentralized Knowledge Graphs. During the seminar, a dedicated group convened to deliberate on approaches for managing Knowledge Graphs across a federation of decentralized instances.

A New Generation of Knowledge Engineers. Improving the utilization and management of knowledge graphs requires educating a diverse audience about both the social and technical aspects of knowledge work. To address this need, a dedicated working group was established. This group conducted an analysis to identify existing educational resources and gaps in knowledge, exploring how consensus could be fostered among various stakeholders in the field. Moreover, the group investigated the specific educational requirements tailored to different audiences, including professional students, undergraduates, and postgraduates. By thoroughly examining these aspects, the working group aimed to formulate strategies for enhancing education and understanding in the domain of knowledge graph utilization and management.

2 Table of Contents

Executive Summary
David Chaves-Fraga, Oscar Corcho, Anastasia Dimou and Maria-Esther Vidal . . . 2

Invited Talks

Semantic Data Integration
Maurizio Lenzerini . . . 7

Access Control, Policies and Constraints
Sabrina Kirrane . . . 9

Programming Languages
Martin Giese . . . 10

Industry Perspectives
Souripriya Das . . . 12

Lightning Talks

From linked data for regulatory reporting towards an interoperable data layer for the rail sector: The ERA KG
Marina Aguado . . . 15

Knowledge Graphs for the Circular Economy
Eva Blomqvist . . . 16

Personal Data Protection in Knowledge Graphs (and KR in general)
Piero Andrea Bonatti . . . 17

Private Knowledge Graph Construction
Carlos Buil-Aranda . . . 18

Challenges in Procedural Knowledge governance in Industry 5.0
Irene Celino . . . 18

Are Knowledge Graphs ready for the World... Wide Web?
Pierre-Antoine Champin . . . 19

On the Need for Project Management for Knowledge Graph Construction and Usage Projects
Oscar Corcho and David Chaves-Fraga . . . 20

Grounding KGs in Natural Language
Christophe Debruyne . . . 20

KG4SE & SE4KG: Exploring the Intersection of Knowledge Graph and Software Engineering Research
Coen De Roover . . . 21

Challenges on Knowledge Graph Data Management Quality
Anastasia Dimou . . . 22

Semantic Knowledge Graphs as a Foundation for Advanced AI Applications
Michel Dumontier . . . 24

“Real World” Considered Harmful
George Fletcher . . . 25

Reliable Knowledge Graphs Need Reliable Processes
Paul Groth 25

Challenges for Knowledge Graphs (KGs) in an Era of Machines that Learn
Claudio Gutierrez 26

Knowledge Graphs in the Real World with metaphactory
Peter Haase 27

Improving UX/DX in Querying Federations of KGs
Olaf Hartig 28

Can Large Language Models improve the Usability of Knowledge Graphs?
Aidan Hogan 28

No Intelligence without Knowledge
Katja Hose 29

Lowering the Barriers for Declarative Knowledge Graph Construction Adoption
Ana Iglesias-Molina 29

Knowledge Graph Adoption: Unveiling User Perspective and Challenges
Samaneh Jozashoori 30

Semantic Reflection
Eduard Kamburjan 30

Exploiting Semantics for Integrating Data on Critical Minerals
Craig A. Knoblock 31

Knowledge-enhanced Representation Learning to Accelerate Scientific Discovery
Vanessa López 31

Using LLMs to Augment KG Construction
Paco Nathan 32

Queries over Evolving Knowledge Graphs
Edelmira Pasarella 33

From Traditional Data Quality to Knowledge Graph Quality
Anisa Rula 33

The Future of Knowledge is Social
Juan F. Sequeda 34

Towards Resource-efficient and Hybrid KG Construction Approaches
Dylan Van Assche 35

MisLED: Linked Enterprise Data Left without Research Attention
Ivo Velitchkov 35

Data and Knowledge Management in Knowledge Graphs
Maria-Esther Vidal 36

Breakout Groups

Access and Usage Control for Federations of Knowledge Graphs
Piero A. Bonatti, Pierre-Anoine Champin, Anastasia Dimou, Peter Haase, Olaf Hartig, Katja Hose, Sabrina Kirrane, Edelmira Pasarella 37

Quality-aware Knowledge Graph Construction
Cinzia Cappiello, Maria-Esther Vidal, Sandra Geisler, Ana Iglesias-Molina, Dylan Van Assche, David Chaves-Fraga, Anastasia Dimou, Irene Celino, Anisa Rula, Maurizio Lenzerini 47

Knowledge Engineering and Education
Juan F. Sequeda, Aidan Hogan, Axel Polleres, Christophe Debruyne, Eva Blomqvist, Samaneh Jozashoori, Paul Groth, Claudio Gutiérrez, George Fletcher, Ivo Velitchkov 59

Project Management for Knowledge Graph Construction and Usage
Oscar Corcho, Coen De Roover, Paco Nathan, Souripriya Das, Eduard Kamburjan, Martin Giese, Marina Aguado, David Chaves-Fraga, Carlos Buil-Aranda, Wouter Beek 65

Emerging Discussions

Knowledge Management and Knowledge Graphs
Axel Polleres, Irene Celino 68

Conclusions 69

Participants 70

4.7 On the Need for Project Management for Knowledge Graph Construction and Usage Projects

Oscar Corcho (Universidad Politécnica de Madrid, ES) and David Chaves-Fraga (Universidade de Santiago de Compostela, ES)

Knowledge graph construction projects today require creating a diverse set of artefacts (OWL ontologies, SHACL shapes, declarative mappings, SPARQL queries, etc.). Most of these projects may look now like an art. However, they should become a proper engineering activity, where all artefacts are well controlled and maintained, all processes are well understood and systematised, and, in general, we can be sure that they can be easily maintained and replicated. Let's work on this and normalise how these projects are done in the future.

4.8 Grounding KGs in Natural Language

Christophe Debruyne (University of Liège, BE)

I commence my talk by arguing that the first question of this Dagstuhl Seminar should be rephrased. First, we should discuss "languages" instead of "programming languages" as the latter is too narrow. Secondly, I believe that "using knowledge graphs" should be replaced with "engaging with knowledge graphs" as the former is unilateral, and the latter implies a bilateral interaction between agents (human and computer-based) and knowledge graphs. The question thus becomes: *"What are the key requirements for language paradigms for modeling, representing, storing, engaging with, and managing KGs in the real world?"* One of the requirements is to include humans by adopting their language.

Knowledge graphs are designed for machines, not for humans. Humans engage with each other using natural language, evidenced by the popularity of generative AI to engage with information. Humans must be kept in the loop of constructing, maintaining, and using knowledge graphs, but we cannot expect them to become "KG-literate." My call to arms is to critically reflect on the role of humans in a KG "ecosystem," as reducing them to "users" would be a disservice to them. There have been initiatives in the past where people used controlled natural languages such as NIAM [1] and RIDL* [2]) for knowledge engineering and querying. These initiatives can be applied to knowledge graphs, so our community should consider learning from the past. De Leenheer et al. [3], for instance, adopted these principles for a knowledge engineering method where knowledge is declared and used separately (a principle called double articulation), which allows for knowledge to be used in different interrelated contexts, much like humans perceive things from different angles depending on their activity or task. This talk briefly mentions these principles to open the room for discussion.

References

1 Nijssen, G. & Halpin, T. Conceptual schema and relational database design – a fact oriented approach. (Prentice Hall,1989)
2 Troyer, O. RIDL*: A Tool for the Computer-Assisted Engineering of Large Databases in the Presence of Integrity Constraints. *Proceedings Of The 1989 ACM SIGMOD International Conference On Management Of Data, Portland, Oregon, USA, May 31 – June 2, 1989.* pp. 418-429 (1989), `https://doi.org/10.1145/67544.66965`

3 Leenheer, P., Moor, A. & Meersman, R. Context Dependency Management in Ontology Engineering: A Formal Approach. *J. Data Semant.*. **8** pp. 26-56 (2007), `https://doi.org/10.1007/978-3-540-70664-9%5C_2`

4.9 KG4SE & SE4KG: Exploring the Intersection of Knowledge Graph and Software Engineering Research

Coen De Roover (VU – Brussels, BE)

Treating source code as data is common in software engineering research. Moreover, many software engineering success stories bear resemblance to knowledge graphs. For instance, logic rules over a database of program facts have proven popular for implementing program analyses [3] as well as program queries [8, 2] that identify code of interest (e.g., design patterns [5, 6], defective code to be repaired [4], meta-level code making assumptions about base-level code [7], etc ...) in a project. The same goes for graph query languages over various graph-based representations. In my talk, I will showcase the latter through a graph-based representation of the control and data flow within Ansible Infrastructure-as-Code scripts [10], which enables detecting design and security smells [11] through straightforward graph queries. Another notable success story in software engineering is the creation and sharing of datasets through mining software repositories. Examples include datasets of Helm Kubernetes charts [9], of build and test results [1], of StackOverflow posts [12], ... The goal of my talk is to raise the question "What if true knowledge graphs were used in all these success stories (KG4SE)?", and also "What software engineering needs exist among knowledge graph engineers (SE4KG)?". This with the aim of stimulating discussion and fostering new insights at the intersection of software engineering and knowledge graph research.

References

1 Beller, M., Gousios, G. & Zaidman, A. TravisTorrent: Synthesizing Travis CI and GitHub for Full-Stack Research on Continuous Integration. *Proceedings Of The 14th International Conference On Mining Software Repositories (MSR 2017)*. pp. 447-450 (2017,5)
2 De Roover, C., Noguera, C., Kellens, A. & Jonckers, V. The SOUL Tool Suite for Querying Programs in Symbiosis with Eclipse. *Proceedings Of The 9th International Conference On The Principles And Practice Of Programming In Java (PPPJ11)*. (2011)
3 Bravenboer, M. & Smaragdakis, Y. Strictly declarative specification of sophisticated points-to analyses. *Proceedings Of The 24th ACM SIGPLAN Conference On Object-Oriented Programming, Systems, Languages, And Applications (OOPSLA09*. **44**, 243-262 (2009)
4 De Roover, C. & Inoue, K. The Ekeko/X Program Transformation Tool. *Proceedings Of 14th IEEE International Working Conference On Source Code Analysis And Manipulation, Tool Demo Track (SCAM14)*. (2014)
5 De Roover, C. A Logic Meta-Programming Foundation for Example-Driven Pattern Detection in Object-Oriented Programs. *Proceedings Of The 27th IEEE International Conference On Software Maintenance (ICSM11)*. (2011)
6 Molderez, T. & De Roover, C. Search-Based Generalization and Refinement of Code Templates. *Proceedings Of The 8th International Symposium On Search-Based Software Engineering (SSBSE16)*. (2016)