Data Management for (Knowledge) Graphs

Stefania Dumbrava
ENSIIE, SAMOVAR, Télécom SudParis, France
IRIF, INRIA Paris, France

George Fletcher
Eindhoven University of Technology, The Netherlands

Matteo Lissandrini
University of Verona, Italy

Riccardo Tommasini
INSA Lyon, France
CNRS LIRIS, Lyon, France

Abstract

In this Special Issue of Transactions on Graph Data and Knowledge, entitled "Data Management for (Knowledge) Graphs", we present eight articles: one position paper, four research papers, and three resource papers. These span graph data querying, transformation, integration, evolution, and evaluation. Their applications include legislative knowledge management, geospatial data access, and knowledge graph construction from heterogeneous sources with large language models.

2012 ACM Subject Classification Information systems → Graph-based database models; Information systems → Resource Description Framework (RDF); Information systems → Query languages; Computing methodologies → Knowledge representation and reasoning

Keywords and phrases graph data, property graphs, knowledge graphs, RDF, data management

Digital Object Identifier 10.4230/TGDK.4.2.1

Category Preface

Acknowledgements We thank Dagstuhl Publishing and the TGDK Advisory and Editorial Boards for their continued support, our reviewers for their timely and constructive feedback, and Aidan Hogan for his guidance in preparing this special issue.

Published 2026-09-03

Editors Stefania Dumbrava, George Fletcher, Matteo Lissandrini, and Riccardo Tommasini

Special Issue Data Management for (Knowledge) Graphs

1 Special Issue Articles

We are pleased to present the TGDK Special Issue "Data Management for (Knowledge) Graphs", which collects eight articles: a position paper, four research papers, and three resource papers. We discuss them briefly below.

1.1 Position Paper

The issue opens with a position paper. In "How Human-Centric Are Our Graph Data Abstractions?", A. Bonifati et al. argue that expressive power alone does not make an abstraction usable, and that human accessibility should be a first-class design goal alongside expressiveness. Examining graph abstractions through four dimensions and use cases from Wikidata to clinical and biological data, they set out a research roadmap toward making graph abstractions easier to study, use, and learn.

1.2 Research Papers

"Transforming Shape Schemas with Composable Property-Graph Queries", by P. Seifer et al., addresses schema inference for composable property graph queries, deriving the schema of a query's output from that of its input without reference to any concrete instance. A family of mappings translates property graphs, their schemas, and their queries into RDF-based counterparts, isolating reification in a dedicated layer, which lets the authors reuse the correspondence between SHACL and description logics. They establish soundness and semantic equivalence.

"Multi-Layered Legislative Knowledge Management with Property Graphs" by A. Colombo et al. models legislative acts and their references as a property graph. Instead of storing redundant versions, a query reconstructs the text of a law as it applied on any given date, and Cypher triggers fire when a repeal removes the legal basis of active laws. An association rule operator exposes recurring overhauls of predecessor procurement legislation.

L. Asprino and E. Daga consolidate Façade-X, a uniform SPARQL access method for heterogeneous formats such as CSV, XML, and relational data. Because its meta-model captures only a fraction of RDF, not every basic graph pattern is satisfiable over a Façade-X source. "Towards a Theory of Façade-X Data Access: Satisfiability of SPARQL Basic Graph Patterns" addresses this by providing a sound, complete algorithm to statically detect unsatisfiable patterns.

In "OM4OV: Leveraging Ontology Matching for Ontology Versioning", Z. Qiang et al. formalise a framework repurposing ontology matching systems for versioning. They implement it as Agent-OV, extending their own Agent-OM, and evaluate it on a synthetic testbed from OAEI datasets and on Brick Schema version changes. The evaluation exposes weaknesses in naive reuse, which a cross-reference mechanism addresses by narrowing candidates via prior alignments.

1.3 Resource Papers

"Logica-TGD: Transforming Graph Databases Logically", by E. Skvortsov et al., expresses graph transformation as logic programming, compiling declarative rules to run over standard engines such as PostgreSQL, DuckDB, and BigQuery. It handles demanding scenarios including time-varying graphs and Win-Move games under the three-valued well-founded semantics, and is benchmarked against the Soufflé Datalog engine and SQL/PGQ on DuckPGQ.

"GeoSPARQL and SPARQL Benchmarking with GeoRDFBench Framework", by T. Ioannidis et al., contributes reusable infrastructure for benchmarking geospatial semantic stores. They specify benchmark components declaratively as JSON, served over a configurable REST interface, and manage store lifecycles for cold-cache measurements. Stores including RDF4J, GraphDB, Stardog, Strabon, Virtuoso, and Jena GeoSPARQL are provided as reusable modules.

Closing the issue, "BLINKG: A Benchmark for LLM-Integrated Knowledge Graph Generation" by C. Castedo et al. targets the mapping step of knowledge graph construction, whose manual effort can exceed six person months. Rather than assessing opaque end-to-end systems, they evaluate large language models on generating declarative, and therefore traceable, mappings across three scenarios of graded difficulty with gold standards, reporting promising results on simpler cases and clear limitations as complexity grows.

2 Data Management for (Knowledge) Graphs

This Special Issue spans the querying, transformation, integration, evolution, and evaluation of graph-structured data and knowledge, from foundational perspectives to reusable artifacts. We thank the authors and reviewers whose work made it possible, and we invite the reader working with graph data to consider TGDK as a venue for their own contributions.