



Editorial

Christoph Gröger¹ · Holger Schwarz² · Theo Härder³

Angenommen: 27. Juni 2023 / Online publiziert: 19. Juli 2023
© Gesellschaft für Informatik e.V. and Springer-Verlag GmbH Germany, part of Springer Nature 2023

1 Schwerpunktthema “Managing Data and Metadata in Complex Enterprise Landscapes”

The digital transformation generates huge amounts of heterogeneous data, across the entire lifecycle of products and services and across all kinds of businesses. Comprehensive analytics efforts and vast amounts of data have made enterprise data landscapes far more complex revealing globally distributed, federated and hybrid deployed structures of analytical and operational data systems. This poses new challenges to both data management and metadata management: on the one hand, new kinds of platforms and techniques for managing data and metadata have emerged, e.g., around data lakes, data catalogs and data marketplaces; on the other hand, data governance and data strategy approaches are developed to ensure the purposeful use of data from the business and the legal point of view.

In this special issue of *Datenbank-Spektrum*, we explore and present current trends on managing data and metadata in complex enterprise landscapes. The idea for this topic originated from the Industrial Data Lab¹, a strategic cooperation of Bosch and the University of Stuttgart for talent development and innovation between research and industry

practice. After peer-review, we accepted an interdisciplinary set of in total seven papers. Five of them focus on technical and application-oriented aspects of complex enterprise landscapes covering components like data catalogs and data lakes as well as self-service capabilities.

- The contribution *Data Catalogs in the Enterprise: Applications and Integration* by Nils Jahnke and Boris Otto focuses on data catalogs as metadata-driven platforms for data transparency and data governance. Based on both a literature study and an investigation of different data catalog tools, the authors develop a typology for different kinds of data catalog applications ranging from enterprise data catalogs to ecosystem data marketplaces. Moreover, they investigate how different types are technically integrated into enterprise IT system landscapes discussing, e.g., on-premises and cloud deployments as well as metadata federation approaches. The authors conclude with highlighting future directions, e.g., regarding decentral metadata management.
- The second contribution *Metadata Extraction from User Queries for Self-Service Data Lake Exploration* by Jonas Gunklach, Sven Michalczyk, Mario Nadj and Alexander Mädche also focuses on data catalogs. The authors present a self-service system that automatically extracts metadata from data lakes and enables business analysts to explore the metadata through an easy-to-use interface. This bottom-up approach allows to tackle the complexity of implementing data catalogs top-down.
- The contribution *Metadata Management and Asset Exchange in the Agricultural Data Ecosystem of the Project Agri-Gaia* by Tobias Wamhof, Ansgar Bernardi, Daniel Martini, Martin Leinberger, Arka Sinha, Heiko Tapken, Andreas Schliebitz and Henri Graf takes a closer look at managing data and metadata in the agricultural domain. Based on the specific requirements for data exchange in agriculture, they present a federated architecture for data sharing called Agri-Gaia. Taking the generic GAIA-X concepts as a starting point, they develop ontology-based metadata models and data exchange processes.

¹ Industrial Data Lab, <https://www.industrial-data-lab.de>.

Christoph Gröger
christoph.groeger@de.bosch.com

Holger Schwarz
holger.schwarz@ipvs.uni-stuttgart.de

✉ Theo Härder
haerder@informatik.uni-kl.de

¹ Robert Bosch GmbH, 70442 Stuttgart, Deutschland

² IPVS, Universität Stuttgart,
Universitätsstraße 38, 70569 Stuttgart,
Deutschland

³ AG DBIS, RPTU Kaiserslautern-Landau,
67663 Kaiserslautern, Deutschland

ses suited for the agriculture domain and showcase a corresponding prototypical implementation.

- Corinna Giebler and Eva Hoos address data consumption in their contribution *A Systematic Approach to consuming Data in Complex Data Management Landscapes using Data Consumption Patterns*. They present a pattern hierarchy for data consumption that covers various categories as well as implementation aspects. The authors also demonstrate the usage of such patterns in a heterogeneous data management and analytics landscape and discuss the benefits of applying data consumption patterns.
- In their contribution *A Recommender Approach to Enable Effective and Efficient Self-Service Analytics in Data Lakes*, the authors Christoph Stach, Rebecca Eichler and Simone Schmidt focus on self-service data preparation for domain experts. They develop a recommender approach on top of an existing self-service data preparation system with an ontological knowledge base. The goal is to suggest suitable data preparation steps and data sources for self-service data preparation in data lakes to foster data reuse and boost data refinement. The detailed evaluation in a real-world IoT use case underlines the increased effectiveness and efficiency of self-service data preparation enabled by the recommender system.

Two further papers in this special issue focus on legal aspects and legal challenges of managing data in databases as well of data-driven business models.

- In the contribution *Steuerrechtliche Herausforderungen datengetriebener Geschäftsmodelle am Beispiel des Connected-Car-Geschäftsmodells*, Carsten Gröger discusses taxation issues in the context of data-driven business models offered by internationally operating companies and exemplifies them by means of services based on connected cars. The author highlights the importance of considering these issues when implementing new data-driven business models and points out shortcomings of international law in this field.
- Michael Beurskens and Stefanie Scherzinger focus on legal aspects of building database instances. In the contribution *Datenbankherstellerrecht und Datenbankforschung*, they discuss to what extent the various components of an application and in particular the data used by the application are protected by law with a specific focus on the situation in the EU. For the research community, this is of particular interest as more and more research artefacts are managed in database instances to which public access is provided.

2 Community-Beiträge

In einem mit Bildern illustrierten Beitrag *Datenbank-Community vernetzt sich in Dresden* berichtet Wolfgang Lehner ausführlich über die 20. GI-Fachtagung „Datenbanksysteme für Business, Technologie und Web“, die vom 6. bis 10. März 2023 an der TU Dresden stattfand. Über 300 Teilnehmer besuchten Workshops, Tutorien, Sessions mit wissenschaftlichen und Industrie-Beiträgen sowie ein Studierendenprogramm. Dazu gab es ein reichhaltiges Demo-Programm mit Vorstellungen von Forschungsprototypen.

Die Rubrik „Community“ berichtet unter „News“ über aktuelle Informationen, welche die DBIS-Gemeinde betreffen.

3 Künftige Schwerpunktthemen

3.1 Best Workshop Papers of BTW 2023

This special issue of the “Datenbank-Spektrum” is dedicated to the Best Papers of the Workshops, Demos and Data Science Challenge running at the BTW 2023 at the TU Dresden. The selected Workshop contributions should be extended to match the format of regular DASP papers.

Paper format: 8–10 pages, double-column

Selection of the Best Papers by the Workshop chairs and the guest editor: April 1st, 2023

Deadline for submissions: June 1st, 2023

Issue delivery: DASP-3-2023 (Nov. 2023)

Guest editor:

Uta Störl, FernUni Hagen

uta.stoerl@fernuni-hagen.de

3.2 Data Management on Quantum Hardware

With the recent availability of cloud-hosted quantum hardware, the potential of this technology for the field of data management is starting to be explored by a growing community of researchers. The topics of interest include machine learning on quantum computers, as well as core tasks of the database management system, such as query optimization and transaction scheduling. This enumeration is by no means final, since research in this area is at a very early stage, and the potential availability of future hardware, such as QRAM, may offer completely new opportunities.

However, quantum software engineering for data management, as well as dealing with a completely new family of hardware, holds its challenges: Learning how to pro-

gram quantum computers requires a fresh mindset and has a steep learning curve. A further challenge is dealing with the limitations of today's early prototype hardware.

In this special issue of *Datenbank-Spektrum*, we will provide a dedicated forum for exploring and presenting current trends in the intersection between data management research, quantum software engineering, and architecting database systems using quantum hardware.

We welcome original contributions, including technical papers, application-oriented papers, case studies, survey papers, and position papers. In particular, we also welcome contributions from industry.

Topics of interest include, but are not limited to:

- Machine learning and quantum computing
- Natural language processing and quantum computing
- Data management applications and quantum computing
- Database systems technology and quantum computing
- Database systems architecture and quantum hardware
- Teaching quantum computing in the context of data management

Paper format: 8–10 pages, double-column (cf. author guidelines at www.springer.com/13222). We welcome contributions in both German and English.

Deadline for submissions: Oct. 1st, 2023

Publication of special issue: DASP-1-2024 (March 2024)

Guest editors:

Stefanie Scherzinger, University of Passau

stefanie.scherzinger@uni-passau.de

Uta Störl, University of Hagen

uta.stoerl@fernuni-hagen.de

3.3 Data Management for Research Data Infrastructures

Research data plays a crucial role in advancing scientific knowledge and driving innovation across various fields. Effective data management approaches are essential for ensuring the integrity, accessibility, and usability of research data. The FAIR data principles (Findability, Accessibility, Interoperability, and Reuse) are generally accepted as a goal

in this context, but their comprehensive implementation remains a challenge. Hence, innovative data management solutions for Research Data Infrastructures (RDIs) are needed to unlock the full potential of research data and enable collaboration, reproducibility, and knowledge exchange.

In Germany, the German National Research Data Infrastructure (NFDI) started in 2020 and comprises 26 consortia addressing a broad variety of disciplines. Currently, research data infrastructure projects inside and outside the NFDI develop data management solutions. In this special issue of *Datenbank-Spektrum*, we will provide a forum for exploring and presenting current trends and solutions in the field of data management for research data infrastructures.

We welcome original contributions, including technical papers, application-oriented papers, case studies, survey papers, and position papers addressing for example the following challenges in data management for research data infrastructures:

- Data Volume and Variety
- Data Integration and Interoperability
- Data Quality and Integrity
- Data Security and Privacy
- Long-Term Preservation
- Ethical and Legal Considerations
- Data Discoverability and Access

Paper format: 8–10 pages, double-column (cf. author guidelines at <http://www.springer.com/13222>). We welcome contributions in both German and English.

Deadline for submissions: February 1st, 2024

Publication of special issue: DASP-2-2024 (July 2024)

Guest editors:

Andreas Henrich, University of Bamberg

andreas.henrich@uni-bamberg.de

Wolfgang Müller, Heidelberg Institute for Theoretical Studies

Wolfgang.Mueller@h-its.org

Bernhard Seeger, Philipps-Universität Marburg

seeger@mathematik.uni-marburg.de