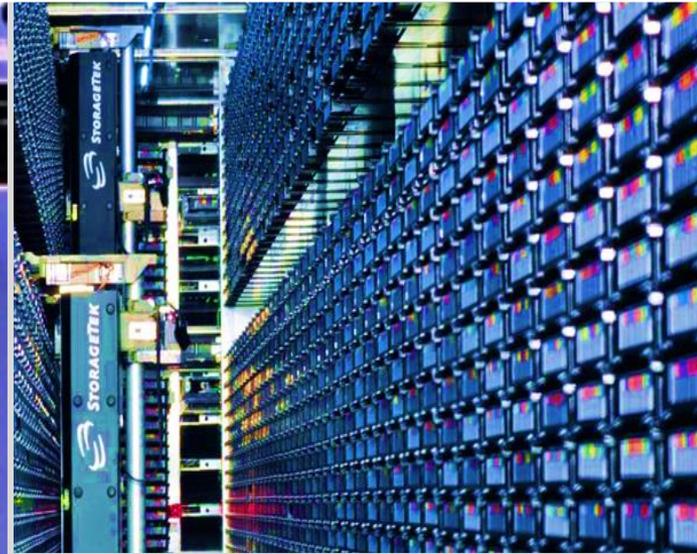


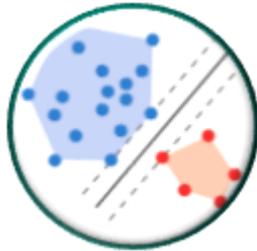
Concept of Data Life Cycle Labs (DLCLs)

Achim Streit

Steinbuch Centre for Computing



Enabling Data-Intensive Computing



Data Analysis Visualization

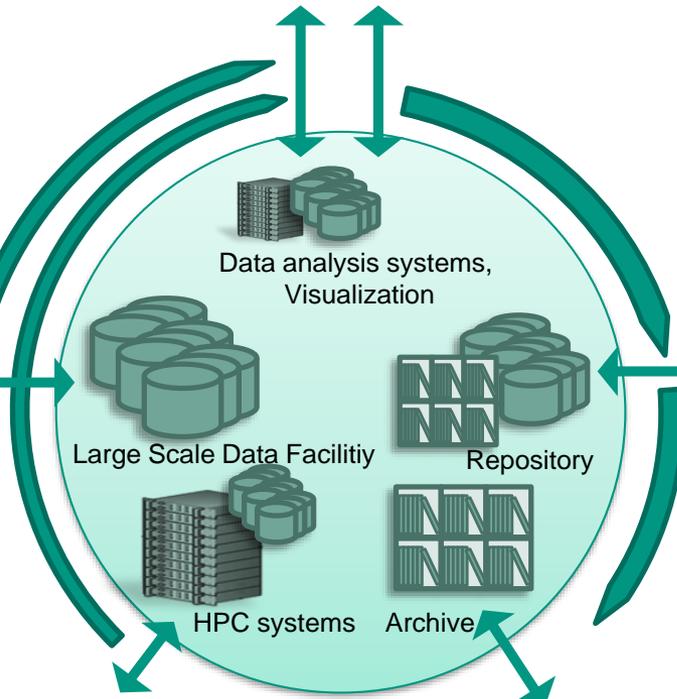


Experiment

Data Creation



Simulation



Publications



Archive



Enabling Data-Intensive Computing

■ Operation of **GridKa**

- German Tier-1 in WLCG for an international community



■ Operation of the **Large-Scale Data Facility**

- Multi-disciplinary data centre for climate research, systems biology, energy research, etc. in BaWü



■ **Joint R&D** with scientific communities

- Generic data management research
- Data Life Cycle Labs in Helmholtz Programme SBD



■ **Innovation driver** for SMEs

■ **Active role** in large projects & initiatives



Human Brain Project



INDIGO - DataCloud



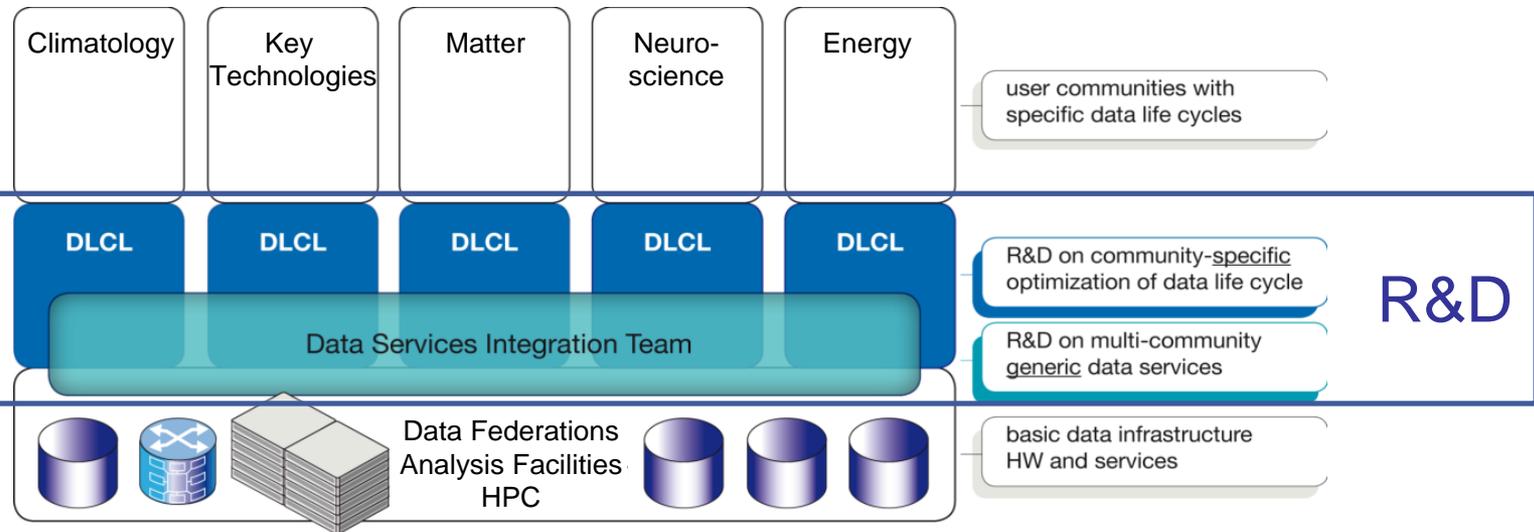
BWDATADISS



RESEARCH DATA ALLIANCE



Concept of Data Life Cycle Labs (DLCL)



Data Life Cycle Labs

Joint R&D with communities

- Optimizing the data life cycle
- Specific data analysis tools and services

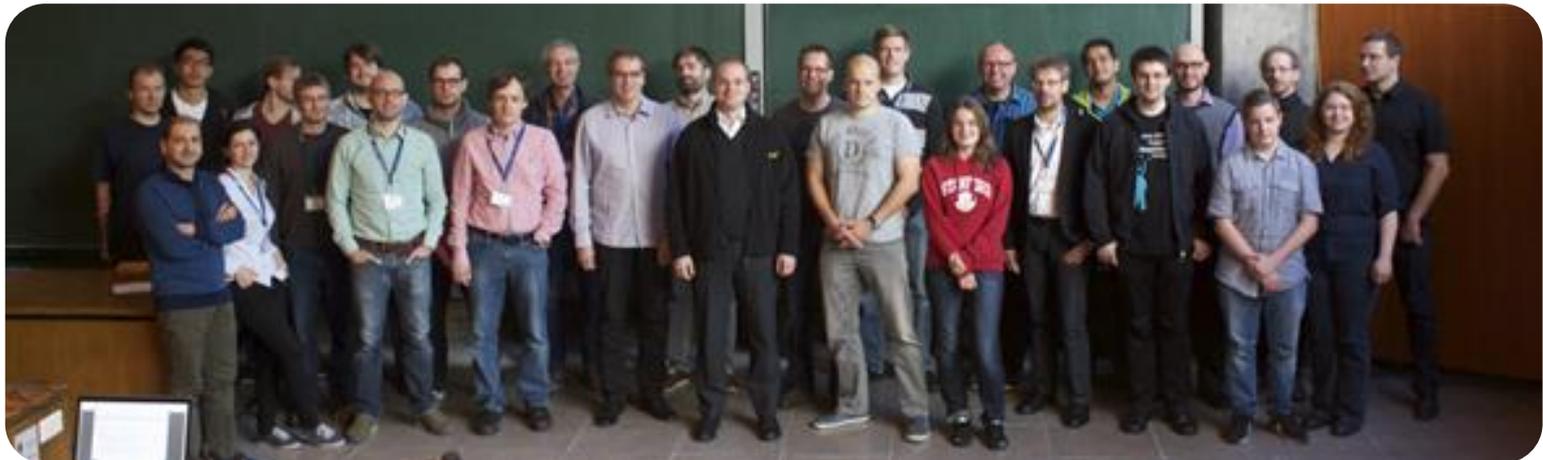
Data Services Integration Team

Generic, multi-community R&D

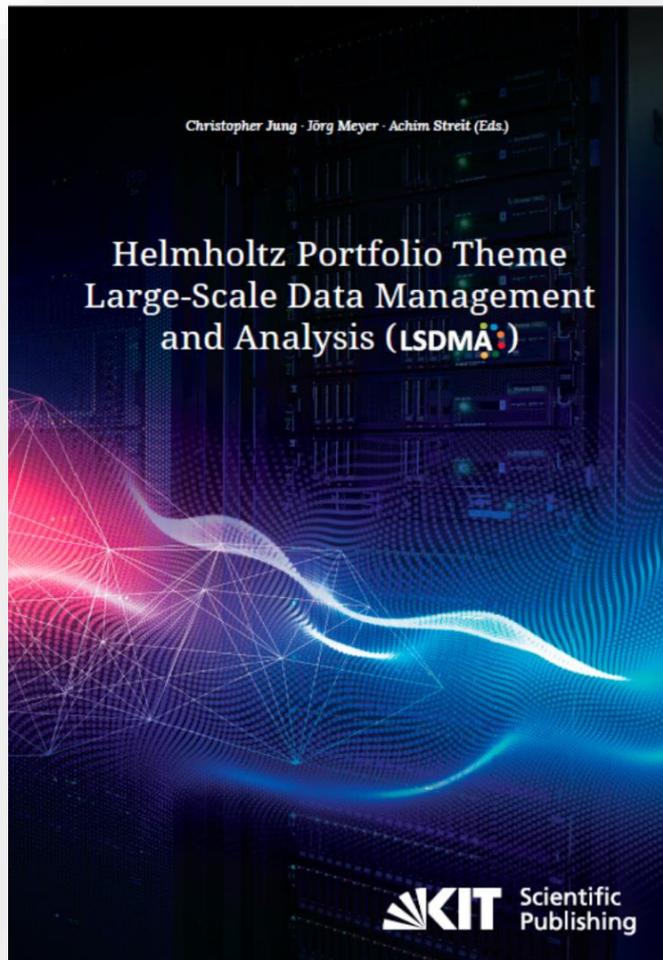
- Interface between federated data infrastructures and DLCLs resp. Communities
- Integration of data services in scientific working process

Project LSDMA

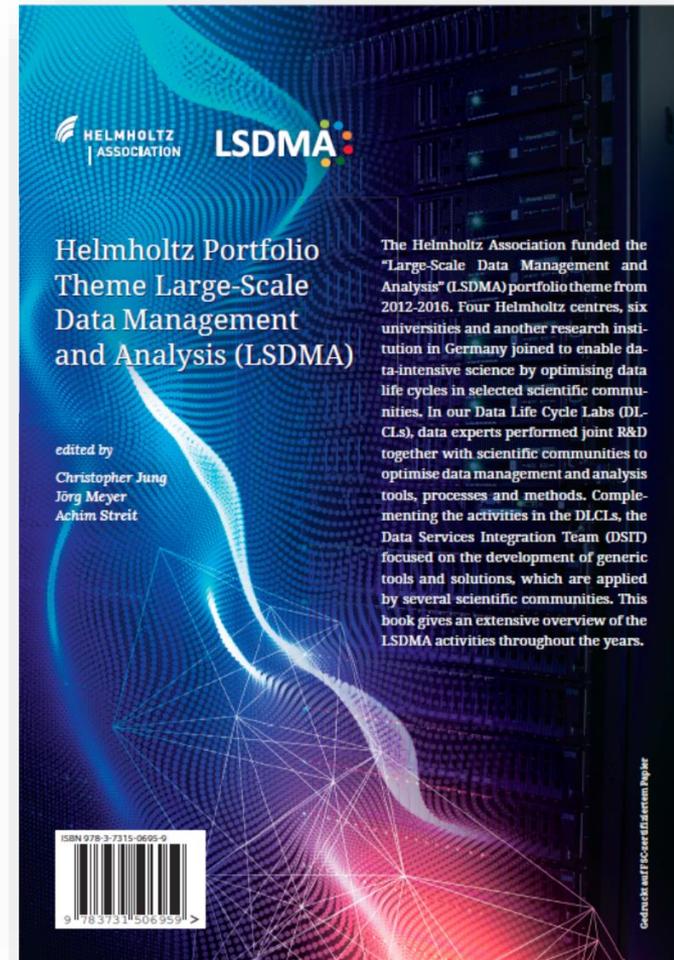
- Helmholtz portfolio theme of the 2nd round
- Initial Runtime 2012-2016
- From 2017 integrated in Helmholtz programmatic funding



LSDMA Book



DOI: 10.5445/KSP/1000071931

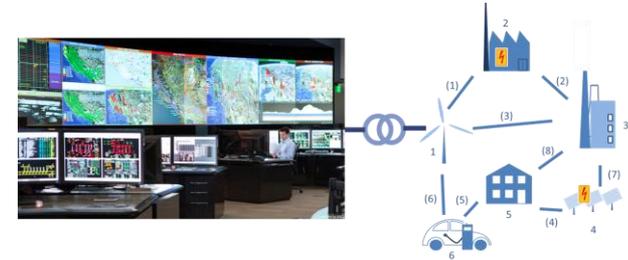


ISBN 978-3-7315-0695-9

Highlights: Data Life Cycle Labs (DLCL)

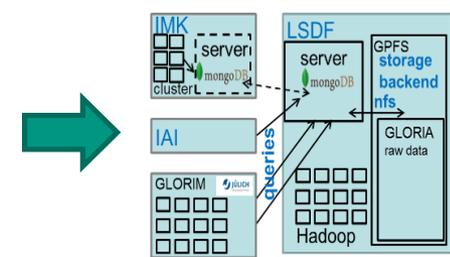
■ Data Life Cycle Lab – Energy

- Generic services for data access
- Data Management for Energy Lab 2.0
- User-oriented Management of Energy Data
- Joint R&D with Helmholtz programs on Energy



■ Data Life Cycle Lab – Climatology

- Management of heterogeneous meteorological datasets
- Distributed, scalable analysis framework for earth-observing data
- Lossy and lossless compression of environmental data using temporal and spatial information
- Joint R&D with Helmholtz programs on Climate & Environment



Highlights: Generic Data Management (DSIT)

■ AAI / federated Identities

- Integration of FedID (eduGAIN) and OpenID in various services, SAML on the commandline, Interoperability of AAI and Levels of Assurance

■ Archive technologies

- International harmonization of Quality of Services in Storage (QoS) via RDA and SNIA/CDMI
- Quantifying the reliability of bit preservation architectures

■ Provenance extensions for UNICORE and KIT Data Manager

■ MetaStore: Generic metadata management framework

- Flexible integration of metadata schemas
- Metadata storage via interoperable interfaces
- Integration of scientific workflow provenance

■ Successful new EU and DFG projects from DSIT collaboration

- EC projects EUDAT2020, AARC, Indigo-DataCloud
- DFG Metadata Management for Applied Sciences (MASi)

European e-Infrastructures

- Enable Access to distributed infrastructures for distributed scientists
 - ➔ Globally accepted **policies**, **standards** and **architectures**





Questions?

