GeRDI - GENERIC RESEARCH DATA INFRASTRUCTURE

Atif Latif & Fidan Limani

ZBW - Leibniz Information Centre for Economics, Kiel, Germany

GO FAIR Workshop on

"Semantic Interoperability of Metadata for Cross-Domain Research of the Future"

11th November 2019, ZBW - Leibniz Information Centre for Economics, Hamburg

GeRDI at a glance

- Research Data (RD) Management Across Disciplines
- Targets long tail RD produced by universities and non-university research centers
- Funded by DFG (~€ 3 M)
- 5 partners + associated community partners
- Project runs Nov. 1st, 2016 - Jan. 31st, 2020
- GeRDI II – proposal submitted and under review

MAIN PURPOSE AND OBJECTIVES

GeRDI aims at multidisciplinary and FAIR RD management

- Provides generic, sustainable and open software for RDM
- Federated-like approach over existing RD repositories
- Integrated access to RD
- Services to support RD Lifecycle

GeRDI offers knowledge and technology in

- Community requirements gathering
- Harvestering of multidisciplinary RD metadata
- Central index
- Architectural design: Self-Contained Systems
- Metadata schema

Semantic Interoperability: Challenges

Long tail RD: High diversity

- Communities (large and small)
- Research disciplines
- RD management practices (including the “no RDM practice” option)

Semantic formalization status

- No common Knowledge Organization Systems (KOS), such as vocabularies, ontologies, etc.
- No common definitions of terms within and across communities
- No common reference “registries”

No silver-bullet solution

- Metadata standard(s) to describe them all
- Cover all the communities suit their (disciplinary) needs
- Provide the corresponding metadata mappings for the communities

Challenges (Cont.)

Metadata Standards

- Generic and Disciplinary standards available
- Different metadata breadth vs. depth capability
- The lowest common metadata denominator

Metadata Mapping

- Mapping consistency
- Metadata (and corresponding mapping) recommendation when there is nothing to map to

FAIR Principles

- F: Missing identifiers and (rich) metadata descriptions
- A: An issue that stems from missing identifiers
- I and R generally harder to achieve
  - (I2) Missing information on KOS
  - (R1.2) Missing provenance tracking

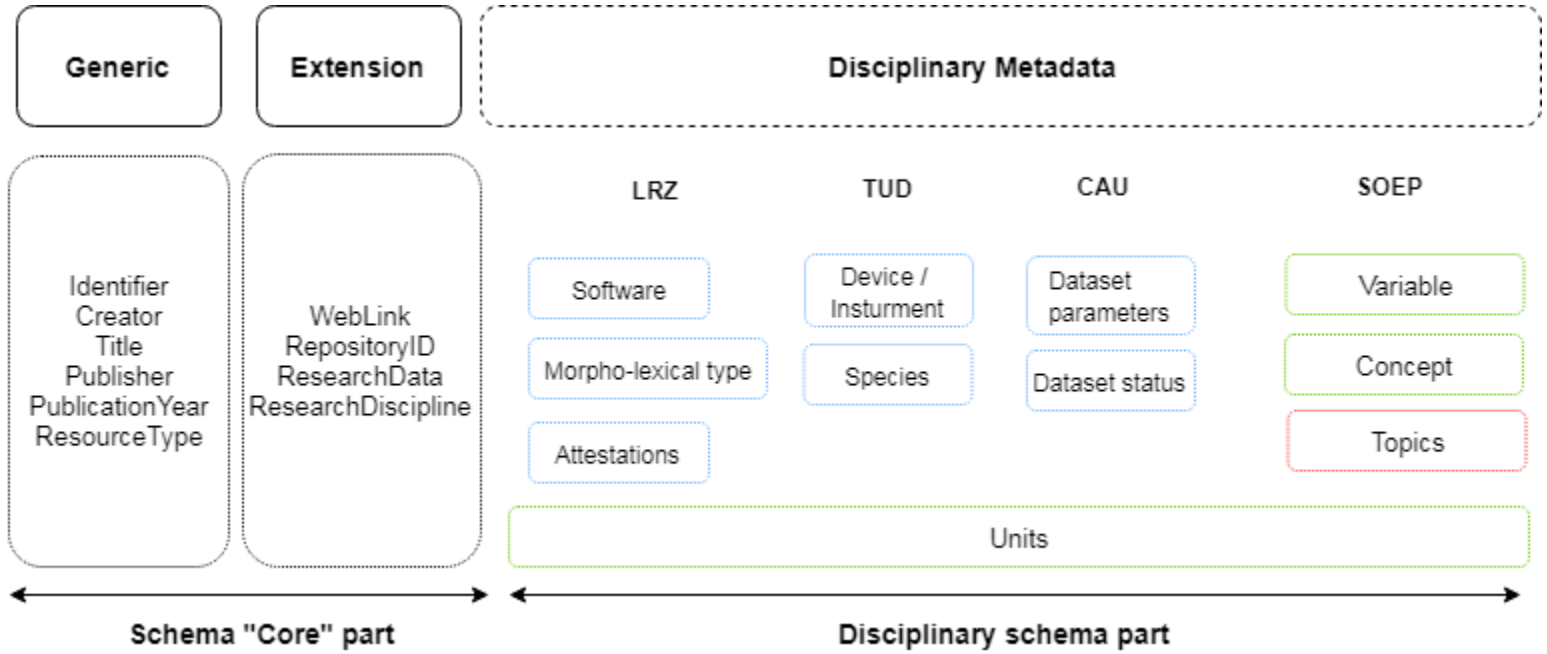
Semantic Interoperability: Ideas, Methods, Solutions

Metadata Standards

- Hybrid approach: GeRDI Schema
- GeRDI Schema parts: a generic, and extension, and a disciplinary metadata parts

Metadata mapping

- Improved consistency by introducing disciplinary metadata and more informed mapping



Semantic Interoperability: Ideas, Methods, Solutions (Cont.)

Start small (Interdisciplinary aspects)

- Adopt common KOS for certain metadata elements
- Harmonize schema elements b/w projects: DFG's subject area classification for "research discipline" description

Explore alternative metadata solutions

- Component-based metadata solution
- More flexible and (more) reusable-by-design

FAIR Principles

- Start with the "low-hanging fruits": F and A
- F: All MD records have ID's
- A: Access to the MD enabled
- I & R:
  - (R) Relative rich metadata descriptions available
  - (R) License information available (repositories where we harvest metadata collections)
  - (R) Domain-relevant standards are adopted to some extent (disciplinary part in GeRDI Schema)

GO FAIR

Thanks & Questions

Atif Latif (A.Latif@zbw.eu)
Fidan Limani (F.Limani@zbw.eu)
ZBW – Leibniz Information Centre for Economics,
Kiel, Germany