

Open Information Linking for Environmental Research Infrastructures (OIL-E)

Paul Martin, Zhiming Zhao,
Abraham Nieva De La Hidalga, Alex Hardisty, Markus Stocker et al.



H2020 Project
Number: 654182



Landscape

- Many and more research infrastructures (RIs)
- Each a complex system
- Each with a focus, in a scientific domain
- Most cover a region (country/continent)
- Heterogeneity across RIs, in
 - RI type (single-site, distributed, virtual)
 - Handled research data
 - Operational services
 - Adopted standards and technologies
- Heterogeneity hinders interoperability



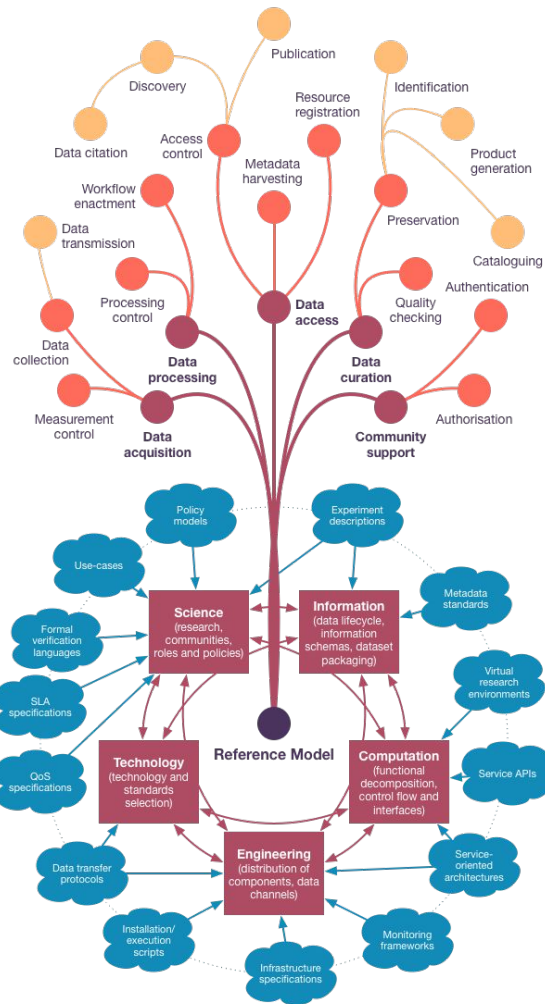
Problem

- System science needs *interoperable* RIs
- Global data in a domain
- Integrated data across domains



Approach: ENVRI Reference Model

- Design framework and common vocabulary for RI *specifications*
- Standardized descriptions of RIs from five viewpoints
 - Science: Actors and their behaviours
 - Information: Data objects and their state transitions
 - Computational: Services needed to support operations
 - Engineering: Distribution of services onto underlying resources
 - Technology: Standards and technologies used
- Structured along the phases of a research data lifecycle
 - Data acquisition
 - Data curation
 - Data publishing
 - Data processing
 - Data use



Approach: OIL-E

- Formal upper ontology for RI specifications
- Uses ENVRI Reference Model vocabulary
- Implemented with the Web Ontology Language
- OIL-E supports and enables
 - Linking different RI vocabularies by alignment with OIL-E
 - Interlinking RIs' semantic contexts: used standards, vocabularies, metadata schemes
 - Constructing an expert system for RI managers
 - Management of RI specifications in RDF databases

Discussion

OIL-E as legend for the Map of the MoLs

Does OIL-E describe the types of activities being undertaken in the MoLs?

- In OIL-E the unit of study is the RI, not a MoL
- OIL-E can describe **activities being undertaken in RI**
 - Specifically, actor behaviours in Science Viewpoint descriptions
 - Examples: Data collection, Data quality checking, Data identification
- It is not designed for MoL activities, landscaping RIs
 - MoL activities are surely different from RI activities
- However, database with RI descriptions in OIL-E may support MoL activities
 - Example: What RI do exist, who are their managers, what are their services, ...?

OIL-E as legend for the Map of the MoLs

Does OIL-E provide the semantic linking requirements of research data infrastructures?

- OIL-E is an upper ontology for RI specifications
 - It supports linking RI vocabularies by their alignment to upper ontology
 - Useful especially if RI vocabularies are different
 - Enables overcoming semantic gaps
- With OIL-E a RI can perhaps describe its semantic linking requirements