ONTO COMMONS
ONTOLOGY-DRIVEN DATA DOCUMENTATION FOR INDUSTRY COMMONS

OntoCommons Project

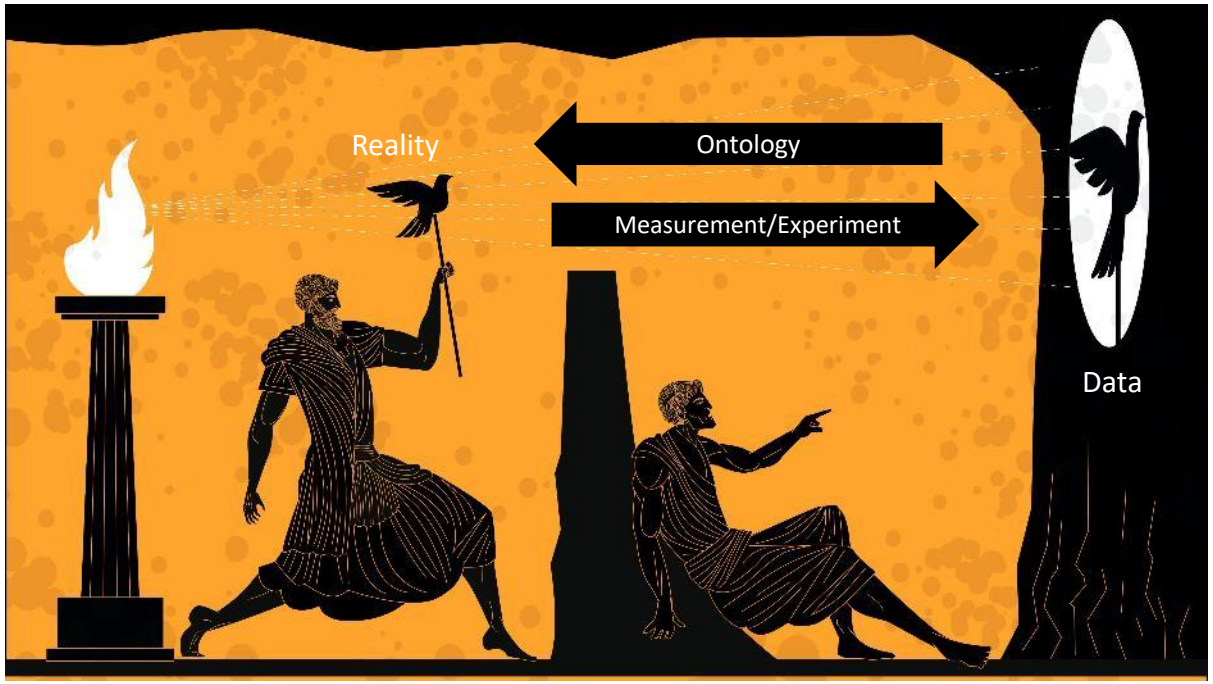
defining best practices for ontologies in EOSC

Arkopaul Sarkar (ENIT)

OntoCommons Member and WP3 Lead

OntoComm ons “Ontology-driven data documentation for Industry Commons” has received funding from the European Union’s Horizon Programme call H2020 -NMBP-TO-IND-2020-singlestage, Grant Agreement number 958371

Open Science Practice: The Generation of Value for Digitized Research Data



Ontology as the formal science of what is*?*
can play a key role in most of these steps

OntoCommons – a snapshot

- Consortium
  - 19 Partners from 10 member states
  - 15 RTDs and 4 companies
  - 22 use cases
- Timeline: Started > November 1st, 2020 (36 months)
- Overarching Goals
  - Overcoming interoperability bottlenecks & facilitating data sharing and valorization.
- Coordination and Support Action -CSA
  - Bringing together and coordinating activities of the most relevant EU and international stakeholders.
- Development of an Ontology Commons EcoSystem -OCES
  - as a foundation for data documentation.

SRIA-EOSC: Status and Gap for Ontology and Metadata (V1.0, 21 June, 2021)

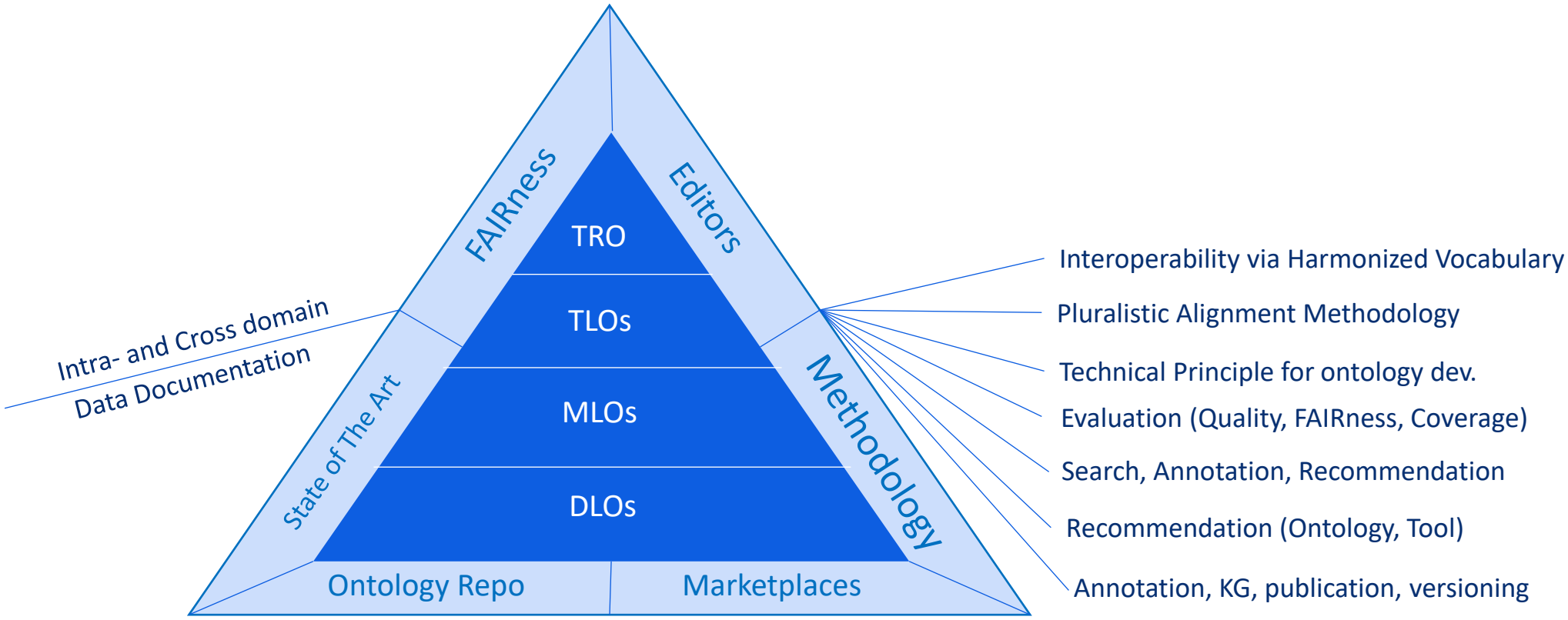
- The **adoption varies between research disciplines** and is, for example, strong in the life sciences or astronomy.
- Integration of discipline-specific metadata across communities and the **aggregation of metadata derived from different metadata schemas and ontologies is still lagging.**
- Work on developing, improving and applying metadata schemas and ontologies is **not well coordinated, mal-aligned, conflicting metadata**.
- Information about existing ontologies is scattered across organisations and services, making **it hard for users to find the relevant information.**
- **no standard way to describe or discover these existing crosswalks**, not maintained.
- Community specific metadata and lack of crosswalk **leads to silos of metadata**.
- **User-friendly tools to apply and maintain metadata for all types of research objects** are not easy to find or are not available.

Expected impacts of OntoCommons

- **Standardised** and operational **intra- and cross-domain data documentation** that meets the FAIR data principles
- Enable a mechanism to allow **practical and user-friendly re-usability of data** across domains and industrial sectors
- Enable a **maintained and continuously developed ontology and data documentation** to ensure long-term relevance and implementation
- Facilitate **uptake of new project results** by Making tools available that ensure practicality and user friendliness
- Improve ability to build **interoperable software solutions** in materials, process and manufacturing

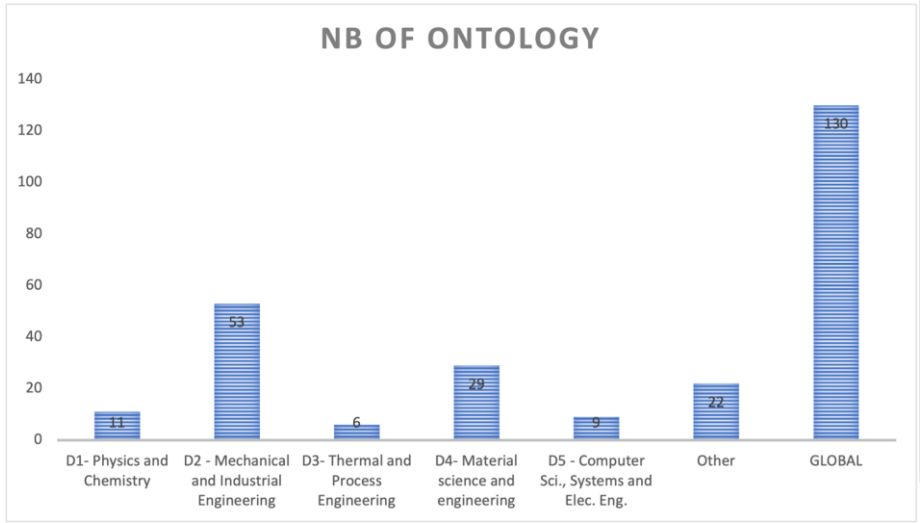
The most tangible outcome – OntoCommons EcoSystem

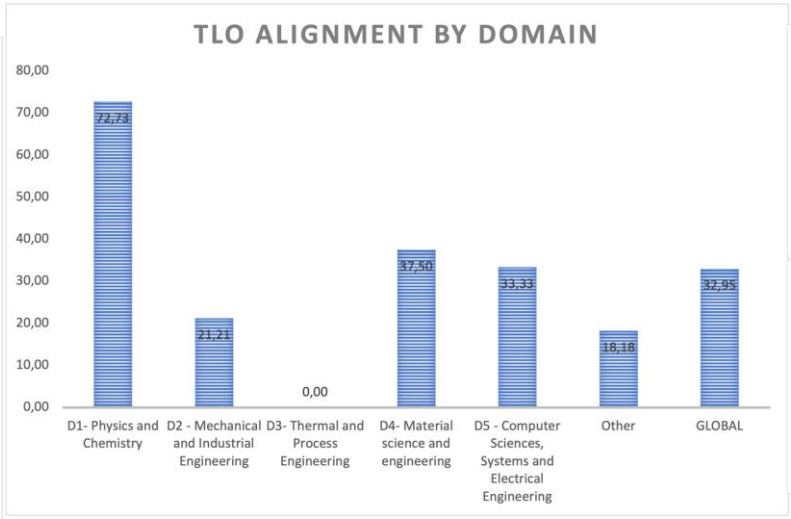
- OCES is a combination of fully harmonized ontology artifacts (from top to domain) and associated tools and methodologies for building upon existing and creating future ontologies. The complementary components of OCES therefore provide a complete solution for data documentation in the NMBP domains.

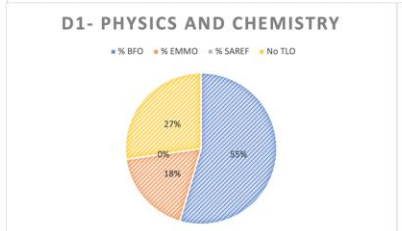


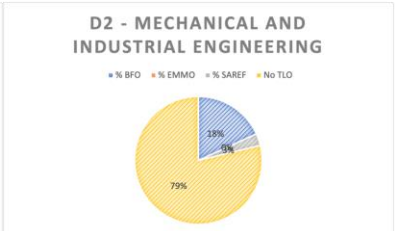
Current State of the Art

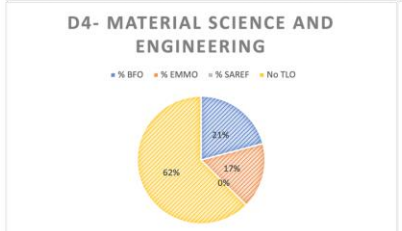
- OntoCommons landscape survey (D3.2 - 2022 https://zenodo.org/record/6504553)
  - 150 total ontologies, 90 machine readable in materials and manufacturing
  - Performed classification by topics, TLO usage, topology
  - FAIRness, coverage, overlap, semantic gaps, usage, maturity analysis
  - Sources stored in IndustryPortal and OntoCommons Registry.
- EUOS ontology survey (Stand-ICT) – ongoing, over 160 ontologies collected

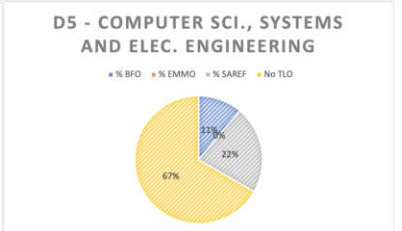






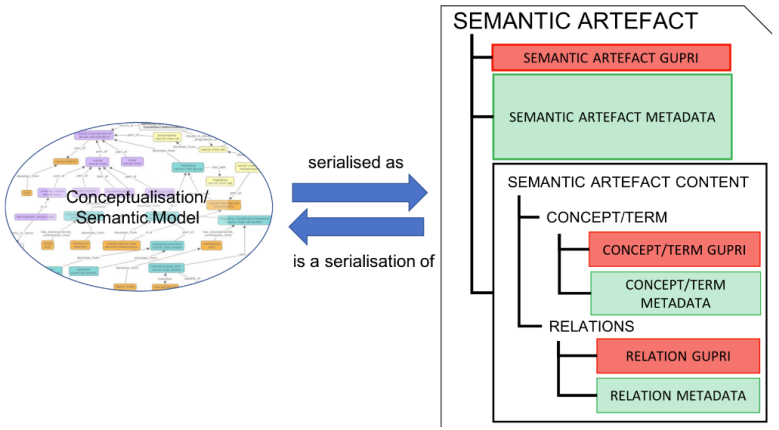






FAIR Ontology vs Ontology for FAIRness

- Ontologies help in interoperability, but they are not interoperable themselves
  - also, not findable
  - also, not accessible
  - also, not reusable
- SoTA says not FAIR: only 6 ontologies over 44 have a score that is more or equal to 50%



- Ontology is the “I” of the FAIR
  - the “I” (interoperability) of FAIR is only possible with the support of information structures that are ontologically consistent and that make explicit the ontological commitments that they inevitably make.
  - Domain and application ontology need to adopt Top and Mid level ontology
    - Alignment (“Not ontologies but Ontology”)
    - Common vocabulary and metadata for data and service
    - In turn, facilitate Findability, Accessibility, and Reusability of data and services

EOSC Priorities regarding Ontology and Metadata

- Develop **governance structures for coordinating the work on metadata and ontologies within EOSC**, both for specific disciplinary communities and for overall coordination.
- **Provide or embrace/stimulate existing registries of metadata schemas and ontologies, defining clear protocols for federation/harvesting, crosswalks and tools for metadata management.**
- Engage with the research information community in order to maximise the re-uptake of the information already residing in research information systems that **communicate in a semantically interoperable manner based on standards.**
- **Develop EOSC guidelines for a minimum metadata description based on existing metadata schemas** and tools to allow data discovery and metadata exchange across federated repositories and scientific communities.
- **Develop services that build on metadata registries and can facilitate the diffusion of metadata schemas across communities**, sharing and community maintenance of crosswalks, measurement of metadata resources uptake across communities, validation of data sources against metadata schemas, etc.

OntoCommons Roadmap and EOSC

- OntoCommons can support the EOSC governance structures by providing insights and recommendations on metadata, including metadata description for data discovery and exchange based on existing or emerging metadata schemas and tools.
- OntoCommons can provide a set of agreed ontologies, from top level to application level, and a set of tools for data documentation that enable data sharing.
- Ontocommons roadmapping and standardisation can help populate the part dedicated to metadata & ontologies in the next version of the EOSC SRIA

Industrial Needs

- Data integration and sharing.
- Standardisation
- **Various domain Perspectives**
- **Interface domain ontologies with TLOs**
- Link domain experts to Ontologists

Gaps in Domain Ontology

- Models granularity
- **Lack of Generic and Application-specific Ontologies**
- **Lack of standardised methodology and tools**
- Ontology as a conceptualization of reality vs information model
- **Ontology Sustainability**
- Lack of Standardised Method for Domain Ontology Evaluation

Recommended Actions

- Standardization of the ontology engineering steps
- **TLO-MLO Alignment**
- **Balance of Theory and Practice**
- **FAIRness**
- **Follow Domain related standards**
- Classify domains
- Bridging the gap between domain experts and ontologists

ONTO COMMONS
ONTOLOGY-DRIVEN DATA DOCUMENTATION FOR INDUSTRY COMMONS

OntoCommons Outputs

ONTO COMMONS TOOLS & SOLUTIONS

OntoCommons Ontology EcoSystem (OCES)

- A hierarchy of ontologies
- Toolkits
- Specifications
- OntoCommons Top Reference Ontology (TRO)
- Top Level Ontology (TLO)
- Domain Level Ontology (DLO)
- Application Level Ontology (ALO)
- Blueprinting reference implementation Toolkit
- OntoCommons Ontology Repository
- Ontology ecosystem knowledge graph

METHODOLOGICAL FRAMEWORK & ECOSYSTEM

- Methodological framework for ontology development and documentation
- Ontology ecosystem structure and reference implementation

REPORTS

- Data Management Plan
- Communities interested in domain-specific semantics
- Domains ontology requirements and specifications
- Feedback loops of cross domain ontologies interoperability
- The finalized Review of Domain Interoperability (RoDI)
- Dissemination, communication & stakeholder's engagement strategy & plan
- Exploitation & Sustainability
- Landscape of ontology development methodologies and platforms
- OntoCommons Standardisation Impact Report

EVENTS

2 DOMAIN ONTOLOGIES
2 HORIZONTAL WORKSHOPS
8 FOCUSED WORKSHOPS
2 EXPERT GROUP MEETINGS
3 EXTERNAL ADVISORY BOARD
6 SUPPORT WEBINARS

COMMUNITY

AN AUTHORITATIVE & ACTIVE EXTERNAL ADVISORY BOARD (EAB)

2,000 ENGAGED COMMUNITY MEMBERS FROM ALL STAKEHOLDER GROUPS & GLOBAL COVERAGE

PRESENCE AT **>30** 3RD PARTY EVENTS

DEMONSTRATORS

Use of Ontologies

- Airbus, Materials
- Bosch, Manufacturing of Microchips
- Aibel, Material, automated reasoning
- Teckniker, material, search and decision
- BASF, Material
- OAS, PSS on logistic and manufacturing, decision making
- IFAM, Material, quality management
- Manufacturing or chemical industry
- Holonix, Product life cycle management, manufacturing
- IRES, Nanosafety, manufacturing, decision making
- Adige SpA, Manufacturing, remote maintenance process

ONTOCOMMONS ROADMAP

ONTO COMMONS
ONTOLOGY-DRIVEN DATA DOCUMENTATION FOR INDUSTRY COMMONS

Thanks

Questions?

FOLLOW US ON

Contact

www.ontocommons.eu

Arkopaul Sarkar, asarkar@enit.fr

Hedi Karray, mkarray@enit.fr (Technical Coordinator)

OntoComm ons “Ontology-driven data documentation for Industry Commons” has received funding from the European Union’s Horizon Programme call H2020 -NMBP-TO-IND-2020-singlestage, Grant Agreement number 958371