ONTO COMMONS
ONTOLOGY-DRIVEN DATA DOCUMENTATION FOR INDUSTRY COMMONS

LGP
PRODUCTION ENGINEERING LABORATORY

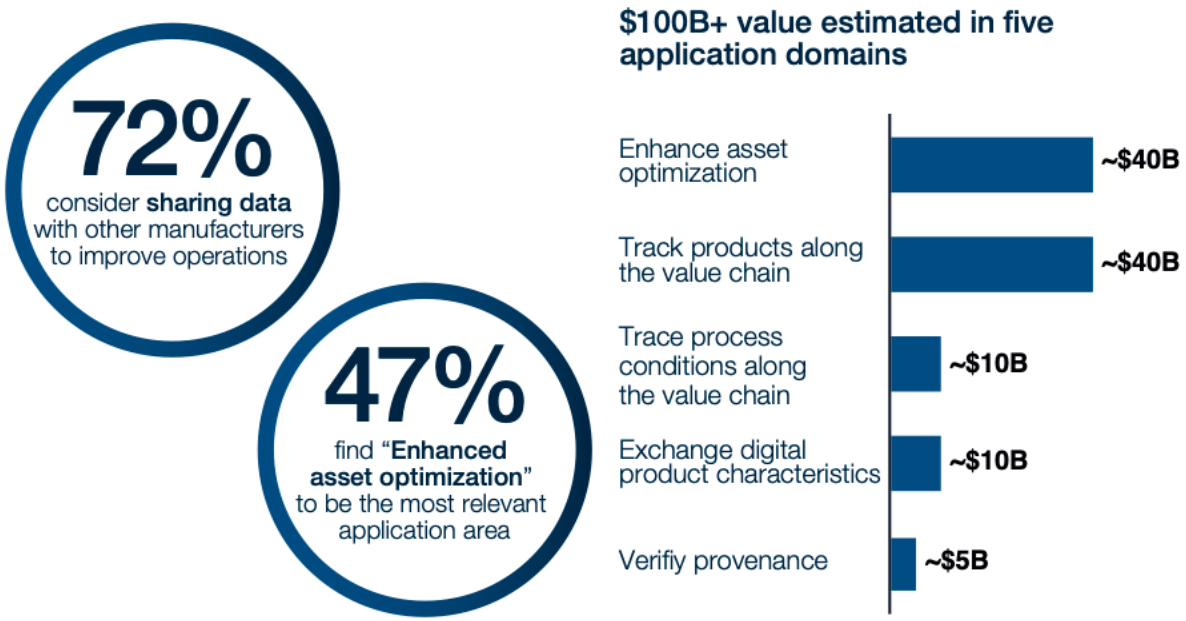
OntoCommons: Standardising materials and manufacturing data documentation to support data sharing in a Cross domain data spaces

Hedi Karray (ENIT)

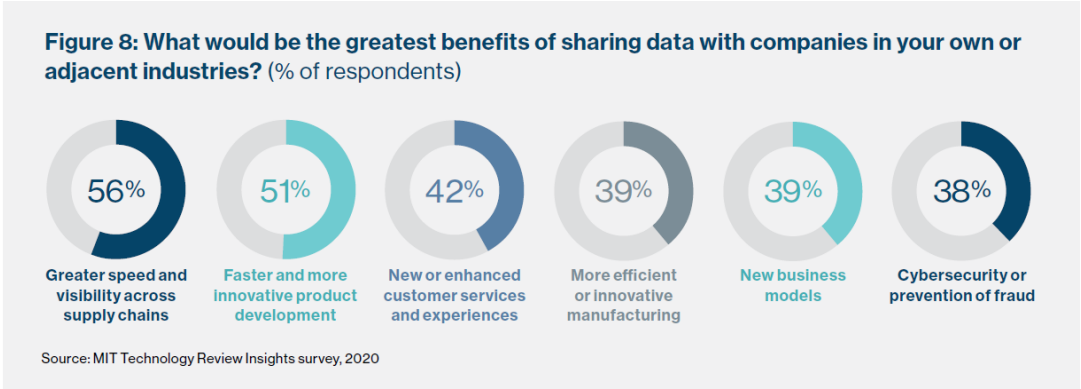
OntoCommons Technical coordinator

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

Share to Gain



Source: Share to Gain: Unlocking Data Value in Manufacturing (WEForum 2020)



Seeing its **data** as a business **asset** helps a company to place a **value** on it and take appropriate actions to **protect**, **share** or **sell** it.

Every manufacturer has an opportunity to immediately start unlocking value through data sharing.
To make it happen, leaders in manufacturing must establish a clear vision, develop the right value proposition and select the right set of partners by building trust within its ecosystem. (WEForum 2020)

Barriers for the Data Value Generation through sharing

70% of production generated data are not used ! (McKensey 2019)

Managerial barriers

Ignorance
Lack of Trust
Lack of strategy and vision
Lack of means (IT experts, budget, etc..)
Resistance to change

- 20% of companies do not engage in data sharing because they fear that their data be made available to third parties.
- 80% of companies reported problems in data access.
- 43.5% of those companies signalled the 'impossibility to find data of the relevant quality'
- 73% of the 129 companies surveyed indicated that poor or insufficient data quality hampers data sharing
- EUR 10.2 billion per year of cost generated by not having research data compliant with standards developed by the FAIR initiatives.

Source: Data Governance act SWD(2020) 295 final

Technical Barriers

Isolated Data Silos (Interoperability)

- Local (company divisions)
- External Data (customer, subcontractor…)
- Data context (operating data, technical data…)
- Data type (sensors, images, text…)

Data Accessibility

- Limited search functionalities
- Complicated Access path
- No integrated view of Data
- Different technologies (and old) for storage and access

Inefficient Workflows

- Long delay from need to access
- Qualified IT experts
- Complex data control

Data Quality

- Outdated
- Noise
- Duplicated
- Incorrect
- Contradictory
- Voluminous

Data spaces in all of that

The role of common European data spaces

Source: Data Governance act SWD(2020) 295 final

The European Strategy for Data proposes to establish sector- or domain-specific data spaces, as the concrete arrangements in which data sharing and/or data pooling can happen beyond one single Member State. A common European data space will be composed of a secure IT environment for processing of data by an open number of organisations, and a set of rules of legislative, administrative and contractual nature that determine the rights of access to and processing of the data. Data will be made available on a voluntary basis and can be reused against remuneration or for free, depending on the data holder's decision.

"A dataspace system manages the large-scale heterogeneous collection of data distributed over various data sources in different formats. It addresses the structured, semi-structured, and unstructured data in coordinated manner without presuming the semantic integration among them."
(Singh 2013)

Semantics is becoming increasingly important within Data Spaces and potentially becomes an area of competitive advantage in European data space and data market projects where semantics plays the role of federator and ecosystem facilitator. In order to facilitate the use of data platforms, they should be able to semantically annotate and enhance the data without imposing extra effort on data owners and data producers. The semantically enhanced data will improve various data processes and unlock data silos using interoperability standards and efficient technologies of the

Source: Common European Data Spaces: Challenges and Opportunities, 2022

Is Ontology still a thing?

- Ontology is not a data model but a model for relating data to the “Reality”.
- Ontology helps us in having a shared view of the “Reality”, based on:
  - Consensus (including scientific findings)
  - Common sense
  - Experiences
  - Metaphysics

- Ontology is a key enabler **to interpret and understand the meaning of data => Data Documentation have to be ontology - based.**
- Ontology may help in making data interoperable, but the ontologies are not themselves interoperable.
- Ontology in Semantic web (or data) and Ontology in philosophy has overlap in their scope and purpose but do not identical.
- Ontology often perceived as “difficult”, “academic”, “for experts only”.
  - Different contexts, No unified methodology, Lack of tools

What about ontologies initiatives in industrial domains !!

- **Physics** & **Chemistry** : EMMO-Crystallography, MDO, CIF, TribAIn, VIMMP, NPO, CHEBI, ChemInf, RXNO, ENM
- ***Mechanical and industrial engineering*** *:*OFM, IOF-Core, ManuService, MSDL, MASON, SAREF4INMA, BWMD Domain Ontology, SCONTO, PRONTO, SIMPOM , ExtruOnt, VAR, PSS, LIDON, SOM, SCOR, QU4LITY-RMPFQ, ,Di-Con, BIMERR, QU4LITY-AIRBUS ontology, IOF core, IOF-Maintenance, IMAMO, ROMAIN, Z-BRE4K, QU4LITY-GFMS ontology, BOOST4.0-GFMS ontology
- **Thermal Engineering/Process Engineering** : BWMD Domain Ontology, EEPSA, REACT, RESPOND
- **Materials Science and Engineering** : EMMO-Mechanical Testing, BWMD Domain Ontology, DEB, MatOnto, MMFO, NanoMine, Material properties ontology, TribAIn, BattInfo, EMMO-Microstructure, BWMD Domain Ontology, DEB, MatOnto, MMFO, TribAIn, NPO, ENM, VIMMP
- **Computer science, systems and electrical engineering:** ADACO, SOSA, SSN, SAREF, iiRDS , MDO, VIMMP, EMMO-Atomistic, MMFO, IEEE Robotics

See OntoCommons landscape survey (D3.2 - 2022 https://zenodo.org/record/6504553)

Most of those ontologies are not FAIR and not linked to standards!!

Typical reasons for ontology failure

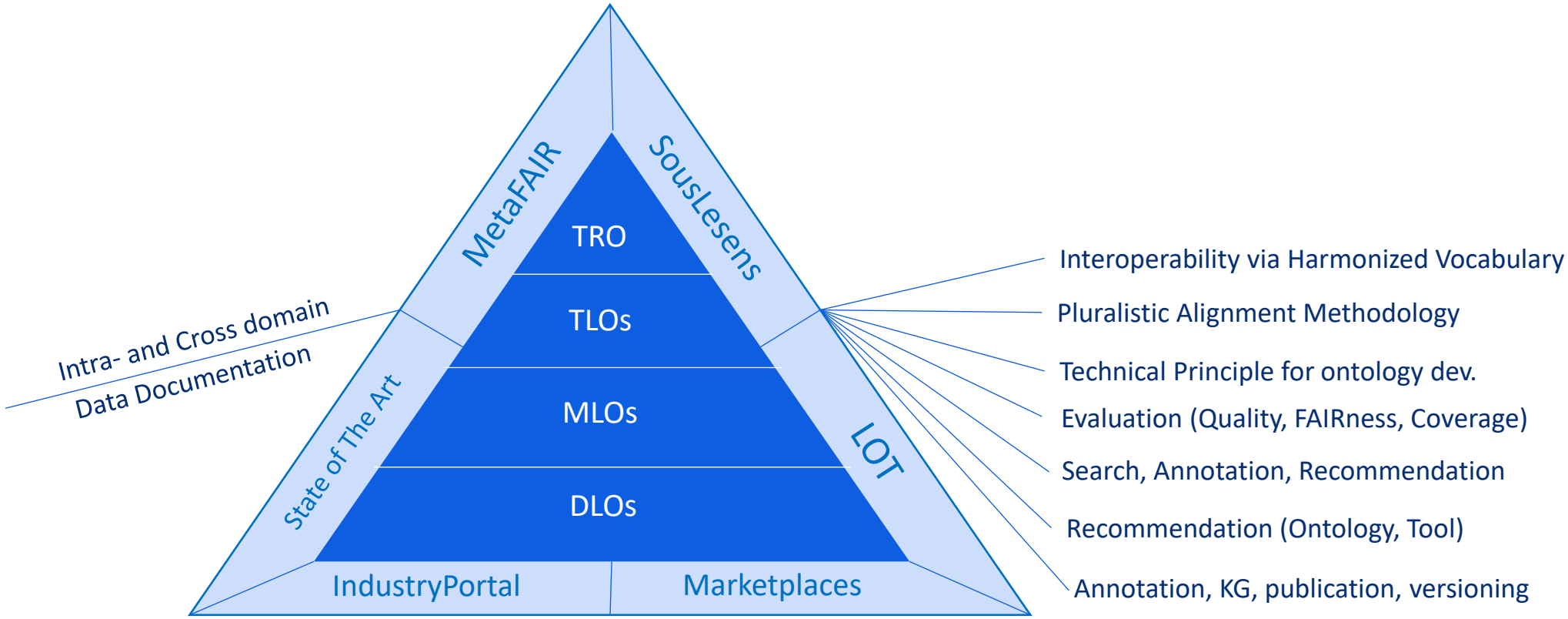
- Too **many** ontologies (everybody wants one; everybody thinks they are easy to build)
- They are built in **ad hoc** ways
- No common **methodology** (50% of ontology projects did not follow a particular ontology engineering methodology and 60% of ontologies are built from scratch)
- Wrong **reuse** strategy (ontology reuse was predominantly interpreted as usage of arbitrary information sources of the relevant cases.)
- Domain related standards are ignored
- No commonly accepted **quality** control standards
- Different philosophical **paradigms** of thinking (Conceptual, Realism, 3D, 4D, etc.)
- Poor **training**
- Poor **documentation**
- Poor **fairness compliance**
- **Short** half life (often EU funded)
- etc.

OntoCommons – a snapshot

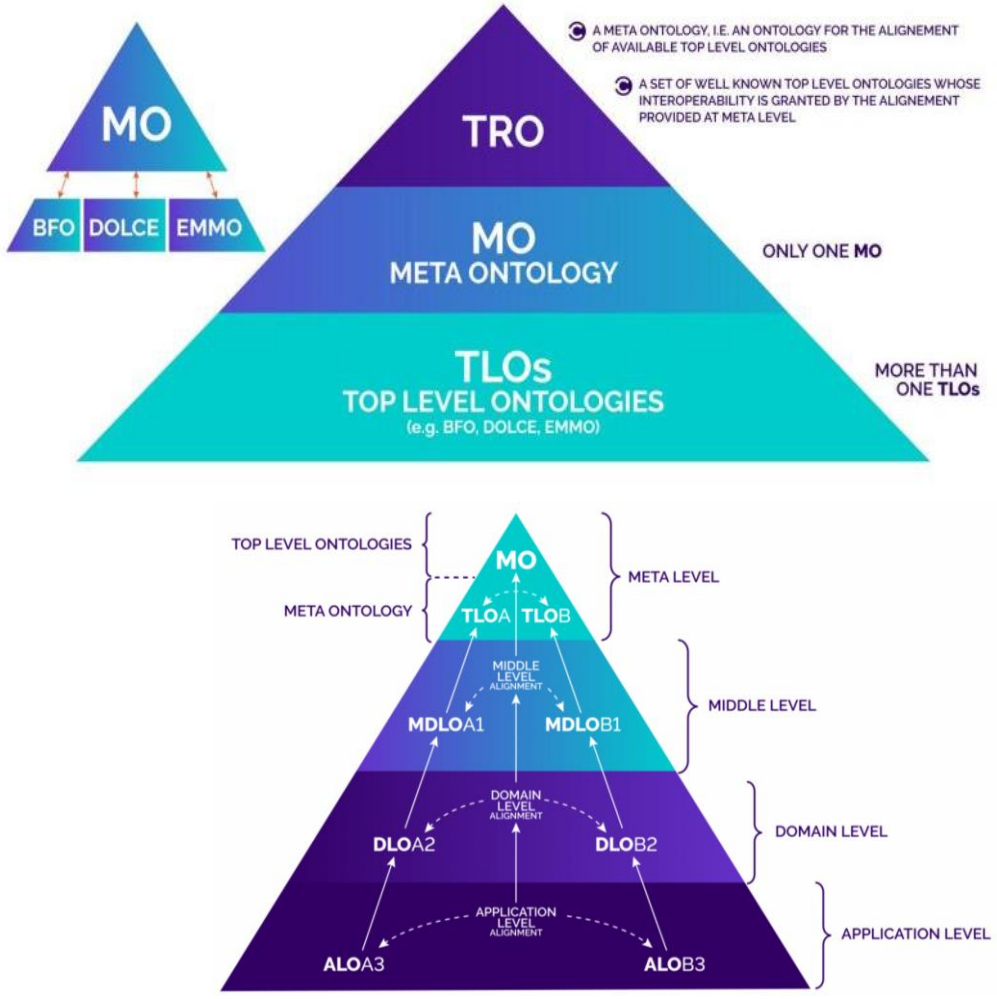
- Consortium
  - 19 Partners from 10 EU countries
  - 15 RTDs and 4 companies
- Timeline: Started > November 1st, 2020 (36 months)
- Overarching Goals
  - Overcoming interoperability bottlenecks & facilitating data sharing and valorization.
- Coordination and Support Action -CSA
  - Bringing together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

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.



Ontologies harmonisation

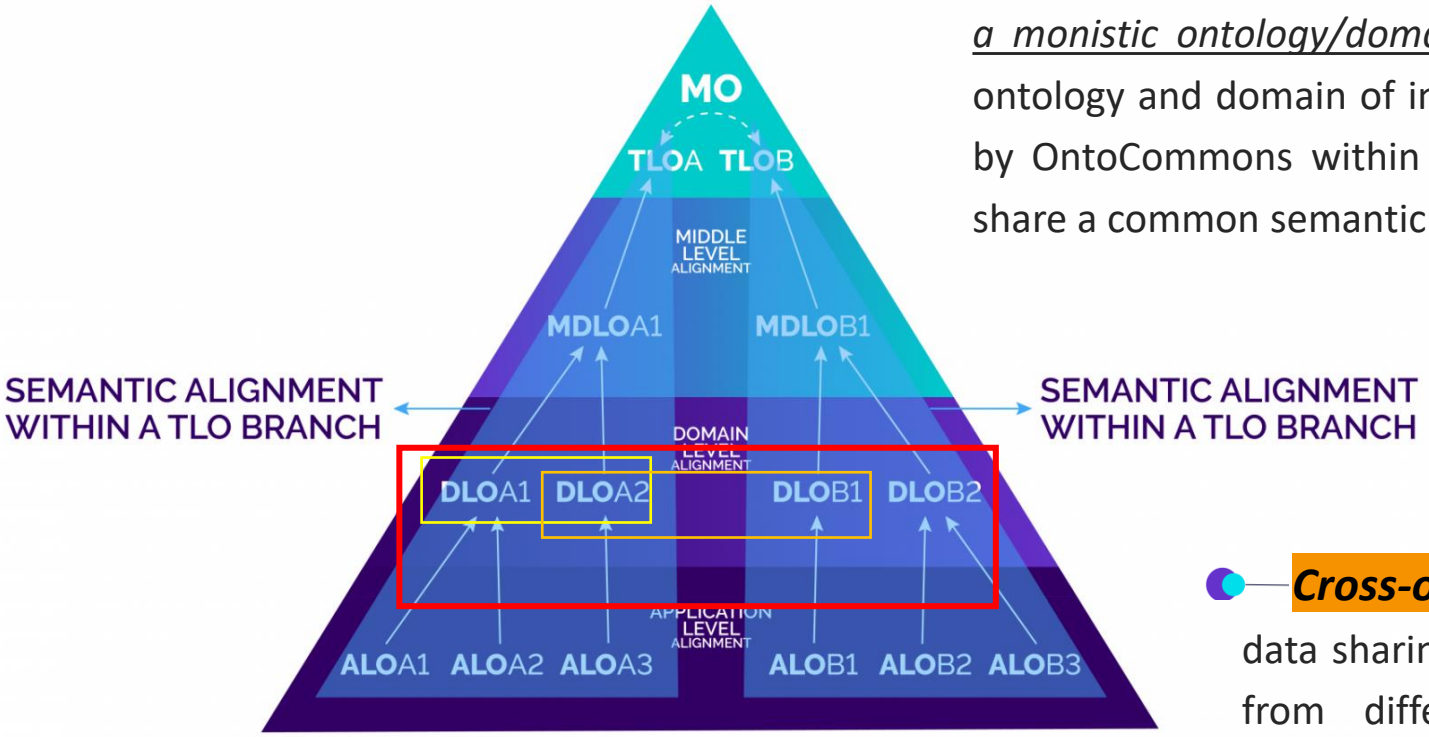


OntoCommons will provide harmonisation between ontologies, through Top Reference Ontology through a multilevel alignement:

- ***Syntactic*** alignment (OWL, FOL, etc.) for all the ontologies that will be part of the OES.
- **Terminological** alignment enabling a minimum taxonomical interoperability between ontologies, by pasting a sub-branch of one ontology under another ontology.
- **Semantic** alignment will be targeted primarily by OntoCommons only within TLO branches,.
- **Formatting** alignment including e.g. labelling of classes, the definition of terms and the annotations.

The OCES will adopt a pluralist approach for the ontological representation of a domain of interest, meaning that more than one ontology for the same domain may be hosted.

Intra and Cross-ontology interoperability



- ***Intra-ontology interoperability*** **:** The capability to enable data sharing between a single semantic representation of data from TLO to ALO coming from *a monistic ontology/domain approach* (one-to-one exclusive relation between ontology and domain of interest). This type of interoperability will be addressed by OntoCommons within a TLO ontology branch whose lower ontology levels share a common semantic framework.

- ***Cross-ontology interoperability*** the capability to enable data sharing between different semantic representations of data from different TLOs branches coming from a *pluralistic ontology/domain approach*.

Initial mapping of all EU and International standardization initiatives

SDOs

Initiatives & Associations

Projects

Domain ontologies in materials and manufacturing

- Ontology landscape survey (AIOTI Report - 2021: http://tinyurl.com/y86s82ac)
- 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 130 ontologies collected

- OntoCommons Roadmap

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**

How OntoCommons endorses the EU single data market Direction

- THE IMPORTANCE OF OPENNESS
  - Interoperable meta-models for assuring F.A.I.R-ness of Open Source, Open Data, Open Standards, Open Hardware. (catalogue, taxonomy, registry)
  - Standardized data documentation for both intra- and cross-domain technology and know-how transfer.

- BLEND AND EVOLVE THE RELEVANT EUROPEAN COMMUNITIES
  - OntoCommons, being a CSA project, is actively engaging SDOs, European and International initiatives and project to create synergy.
  - Create roadmap from stakeholder's input (Gap analysis)

- REFINE AND GROUND THE VISION FOR A COMPUTING CONTINUUM
  - Repository for sharing semantically annotated, FAIR results of academic and industrial research results and best practices produced across the overall AMI2030 ecosystem.
  - Sustainability of Materials and Manufacturing collaboration based on standardized knowledge exchange.

OCES will contribute to data spaces semantisation to allow cross domain data spaces interoperability and then data Market places !

ONTO COMMONS
ONTOLOGY-DRIVEN DATA DOCUMENTATION FOR INDUSTRY COMMONS

Thanks

Questions?

FOLLOW US ON

Contact

www.ontocommons.eu

Hedi Karray, mkarray@enit.fr

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