ONTO COMMONS
ONTOLOGY-DRIVEN DATA DOCUMENTATION FOR INDUSTRY COMMONS

OntoCommons Roadmap v1

30 April 2022

OntoCommons - Ontology-driven data documentation for Industry Commons, has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement no. 958371.

Glossary of terms

Item	Description
BFO	Basic Formal Ontology
CNR	National Research Council
NSB	National Standardisation Body
SDO	Standard Developing Organisation
AUWP	Annual Union WP for European Standardisation
ESO	European Standardisation Organisation
TLO	Top-Level Ontology
MLO	Mid-Level Ontology
TRO	Top Reference Ontology
BFO	Basic Formal Ontology
DOLCE	Descriptive Ontology for Linguistic and Cognitive Engineering
EMMO	Elementary Multiperspective Material Ontology
HPC	High Performance Computing
AI	Artificial Intelligence
DLO	Domain Level Ontology
ALO	Application Level Ontology
OWL	Web Ontology Language
CAD	Computer-Aided Design
JSON	JavaScript Object Notation
UML	Unified Modeling Language
FIBO	Financial Industry Business Ontology
FAIR	Findability, Accessibility, Interoperability and Reusability
EOSC	European Open Science Cloud
LOT	Linked Open Terms
OCES	Ontology Commons EcoSystem
API	Application Programming Interface

Keywords

Standardised data documentation, ontology, Ontology Commons EcoSystem, Industry Commons, OCES Toolkit, interoperability of data, FAIR data, metadata, FAIR principles, Knowledge Management Translator, demonstrators on use of ontologies, networking and cooperation, coordination and support actions, knowledge-based AI, Top Reference Ontology, Top-level ontology, Mid-Level Ontology, Domain-level Ontology, Application-Level Ontology

Disclaimer

OntoCommons.eu has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement no. 958371.

The information and views set out in this report are those of the authors and do not necessarily reflect the official opinion of the European Commission, which cannot be held responsible for the use which may be made of the information contained therein.

Executive Summary

Data sharing within and across domains can offer enormous opportunities for innovation and for overcoming various bottlenecks in industry. It can help organisations to achieve the general objectives of both green and digital transition, but also more specifically, to improve resilience and bring safe and sustainable materials and products to market more quickly. However, capitalising on the unprecedented opportunities for innovation based on sharing of common assets requires a structured, systemic approach. Rather than assuming ad-hoc discovery based on a "data-soup", an Industry Commons ecosystem based on horizontal enablers is needed. This includes an ontology-driven approach which can harmonise data documentation, support data sharing across application domains, and has the potential to stimulate and support sustainable cross-domain industrial innovation. The development, implementation and widespread uptake of an Ontology Commons EcoSystem (OCES) is an important pillar of the Industry Commons ecosystem. In particular, high-level needs for OCES include the following:

- Effective and meaningful data documentation to enable the harvesting of data resources that are generated by European industry and that are currently not fully exploited.
- Covering the gap between human and digital tools (Industry 5.0): The need for expressing human knowledge through a widely applicable and general methodology that is both human understandable and at the same time easily processable by digital devices.
- The need to provide machines with existing knowledge and thus enable the exploitation of the computational power of modern HPC facilities and software.
- Knowledge-based reasoning as opposed to generic AI processing by formalising the knowledge already available from existing scientific disciplines and to provide knowledge structures that can be used as reference points for the analysis and interpretation of machine data and deep learning AI findings.

In this context, the OntoCommons Roadmap considers Needs, State of the Art, Gaps, Definition of Success and Recommended Actions for a number of topics contributing to an Ontology Commons Ecosystem for ontology-based data documentation grouped into:

1. Ontology Foundations: Top Reference, Middle, Domain and Application Levels
2. Integrated Development Environment (Tools) and Infrastructures
3. Industrial Impact including Marketplaces, Standardisation, Education and Human Resources

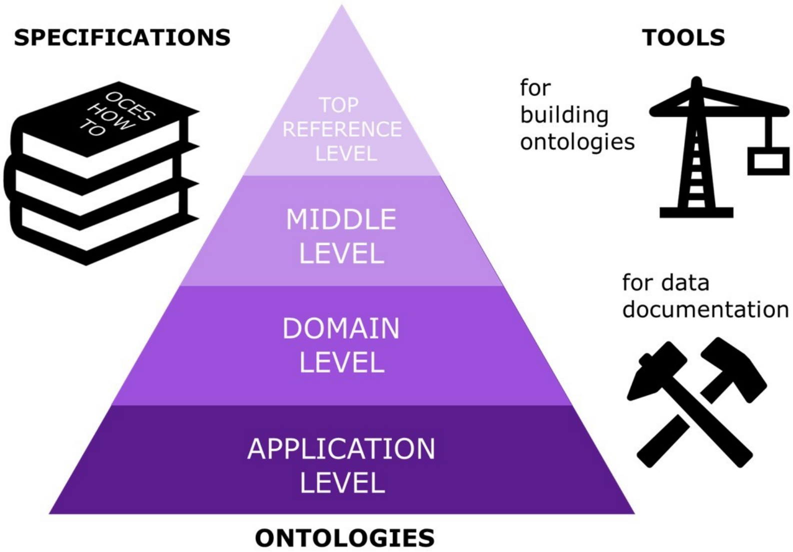

Figure 1 - Ontology Commons Ecosystem Overview

Ontology Foundations

The objective of OntoCommons is to create a system of interoperable ontologies based on widely accepted and used Top, Middle and Domain Level ontologies (TLO/MLO/DLO). OntoCommons takes a plurastic approach, recognising that there are different Top-Level Ontologies, with different axiomatic commitments already in widespread use in materials, manufacturing and related fields. Hence, OntoCommons introduced the concept of a Top Reference Ontology (TRO), defining a common foundation for data interoperability to enable knowledge sharing across TLOs.

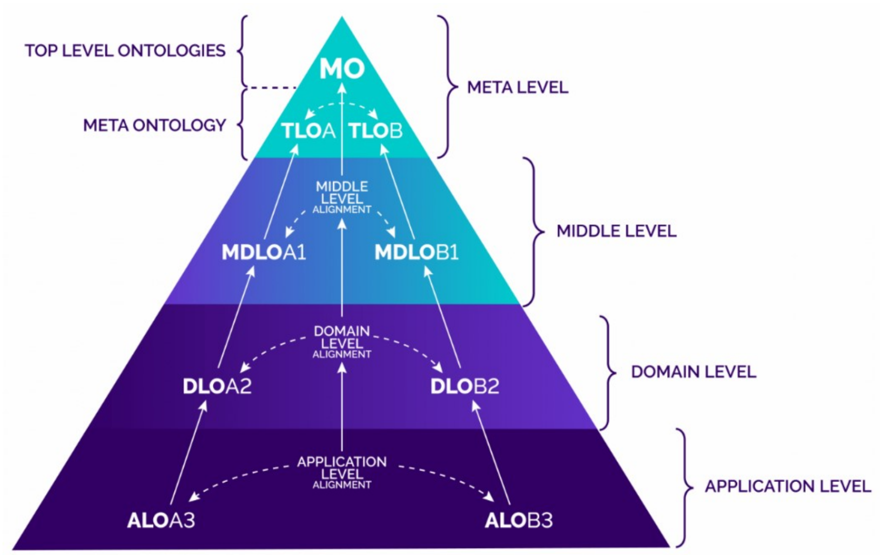

Figure 2 - Different Ontology Levels in OCES and their alignment

The industrial needs and current gaps regarding Ontology Foundations discussed in this Roadmap include:

- Overheads of using TLOs as knowledge representation foundations;
- Cross-disciplinary barriers as found between philosophy-based TLOs and their applications;
- Lack of inter-TLO connections creating silos in terms of communities and data;
- Coherence between MLOs/DLOs and existing definitions or standards provided by each discipline;
- Lack of DLOs and a lack of standardised methods for their evaluation; and
- Ontology sustainability.

Overall, the industrial success of a TLO/MLO/DLO framework can be defined as its capability to enable sharing of data or data re-use, where the same data is being used by different companies for different materials, products, and processes. To achieve that, actions are needed to:

- Establish the OntoCommons TRO and stimulate TLO communities to work together for adoption;
- Address the TLO/MLO/DLO gaps above, and in particular, work towards a set of widely community adopted, continuously maintained industrial domain ontologies utilising a collaborative framework of developing, standardising, sharing, maintaining, updating, and documenting ontologies;
- Improved TLO/MLO scalability regarding tools with different expressivity levels used in applications; and
- Establish interoperability by means of an effective methodology for mapping between relevant TLOs and the mapping of relevant MLOs with more than one TLO.
- Overcome the cultural and disciplinary gap between the top and the application levels, enabling successful development by non-TLO/-MLO experts of application ontologies

Integrated Development Environment

For community and industry to develop, maintain and use ontologies, a powerful, well-supported integrated environment is required. An environment that makes it easy and efficient to map existing data sources to ontological concepts would speed up the development of ontologies and increase their sustainability and impact. The Ontology Commons EcoSystem (OCES) Toolkit is intended to provide a specification for both methodological and tool support for industry-oriented ontology development, maintenance, and usage. The Toolkit consists of state-of-the-art methodologies as well as references and specifications for tools.

The current gaps and industrial need regarding the Integrated Development Environment (OCES Toolkit) discussed in this Roadmap includes issues such as the facts that:

- Many ontology development, engineering and usage tasks are poorly supported by current tools;
- Existing tools are poorly integrated or not integrated at all, and lack support for collaborative development;
- The usability and user experience of ontology engineering tools is not satisfactory, and domain experts in particular find it challenging to understand and work with existing tools.

An Integrated Development Environment (OCES Toolkit) is needed that can:

- Offer a number of functional services covering all phases and common activities of ontology engineering and use; and
- Provide a set of non-functional services, ensuring flexibility, openness, ability to integrate with other tools, an interoperable interface, collaborative development opportunities and FAIR semantics.

The impact of an Integrated Development Environment would be to significantly broaden the community of industrial researchers using ontology reference documentation and hence contribute to optimisation and innovation. The OCES Toolkit will be a catalyst for overcoming scepticism by creating an ecosystem within which industries can benefit from a turnkey solution with a set of good practice recommendations that will support them in integrating ontologies successfully and effectively into their operations.

The Roadmap recommends a number of key actions to achieve these outcomes including:

- Raise awareness about the need for tools in ontology development;
- Refine methodologies and tools, especially for ontology validation;
- Leverage the Industrial Ontology Portal to support ontology adoption;
- Provide a user-friendly tool chain and reference implementation.

Infrastructure

In the domains addressed by OntoCommons, there is a clear need for an infrastructure on which to base the functionality of the ecosystem of ontology engineering tools, for (standardised) data documentation tools and processes, to support secure communication between stakeholders that collaborate, for secure data sharing or integration, etc. This needs to be combined with intelligent data infrastructure (e.g., provisioning, exploring, transforming) that will allow any data to be exposed as semantically enriched information (e.g., Knowledge Graphs), allowing for a transparency layer that enables users and down the line (semi-)automatic processes to seamlessly collaborate and run experiments that may involve several data sources and several computation modules. Current gaps include the low maturity level of available infrastructure components including the lack of infrastructure services such as data management and curation, security and privacy, as well as appropriate legal and financial frameworks. Actions are required to develop low-level ontologies, provide secure platforms for ontology data creation, provisioning, and exchange, and a Virtual Research and Innovation Environment blueprint for the domains covered by OntoCommons.

Industrial Impact

The adoption of ontology-based data documentation and knowledge management practices in industry is still at a low level, with some noteworthy exceptions. To reach wider adoption and impact, the ontologies, tools and infrastructure, as well as human resources, need to be developed as discussed in this Roadmap. The industrial needs that can be addressed by such an ecosystem are clear, and are also discussed in this Roadmap and in some detail in the OntoCommons Demonstrators. However, there are current barriers such as the high cost of ontology development, learning barriers, as well as the specific barriers discussed regarding TLOs/MLOs/DLOs, tools and infrastructure. On the other hand, with the adoption of ontology-based systems, industry can expect to reap substantial benefits including:

- Standardised data documentation and FAIR data within and across organisations;
- Improved communication within a company (between personnel) and with external partners;
- Time and cost saving;
- Increased innovation capacity; and
- Optimised product quality and environmental footprint.

In addition to the benefits for individual companies, there are huge untapped opportunities of data sharing in an "Industry Commons". A system of digital marketplaces utilising ontology-based data documentation can support needs such as data integration and interoperability, as well as improving the transfer of data between industries and marketplaces.

Achieving industrial impact requires that the Ontology Foundations, Integrated Development Environment, and Infrastructure be established, as discussed above. In addition, there is a need for Standards as well as Human Resources, skills and training as outlined below. Support actions that are recommended include:

- Networking for people to share their experience with ontology adoption in industrial settings;
- Success stories showing savings in both time and cost;
- Close cooperation with FAIR communities and demonstration of FAIR benefits;
- Actions on marketplaces such as well-defined demonstration for marketplaces and linking marketplaces via alignment and mapping of their ontologies.

The role of standards

The field of materials and manufacturing requires widely agreed standards for documenting their data, reaching from vocabularies/terminologies to taxonomies and ontologies. Ontologies can further help to clarify the meaning of terms in certain contexts and provide a means of establishing machine-readable standards, increasing interoperability and automation. Improving the inclusivity and interoperability of standards will strengthen the competitive position of Europe's worldwide innovative technologies.

In particular, there is a need to:

- Focus on the use of ontologies to contribute to standards inclusivity, harmonisation and interoperability, offering better categorisation of information and process efficiency;
- Better formalise standards using a unambiguous set of definitions for concepts and their relations in a clear way which should be formal but also easily understandable by human agents;
- Integrate existing terminologies and definitions in current standards into OntoCommons domain ontologies and vice versa, utilising OCES to provide machine readable ontologies;
- Engage with SDOs on the role of OCES in materials and manufacturing standards.

Human resources, skills and training

The success and industrial impact of ontology-based data documentation and knowledge management strongly depends on establishing a knowledge culture based on training, continuous education and dedicated human resources, including the new role of a Knowledge Management Translator for Industry Commons.

The difficulty of finding people trained in ontologies and knowledge engineering is a gap that needs to be addressed. Training on ontology usage and development issues is hence very important, and this education must be adaptable to the needs and competencies of various stakeholders. A curriculum needs to be developed, comprising literature, training, forums, etc., to provide continuous professional development opportunities. Relevant policies, programmes, training courses and supporting infrastructure to upskill capabilities across industry must be developed.

In order to bridge the gap between business functions, domain experts and ontologists, a 'Translator' role is required similar to those established in Materials Modelling and Data Analytics. Such a Knowledge Management Translator for Industry Commons role needs to be established and supported by training programmes and certification from relevant professional associations.

Table of Contents

List of Figures

List of Tables

1. About OntoCommons and this Roadmap

1.1 Overview

The OntoCommons project brings together and coordinates the activities of the most relevant EU stakeholders for the development of an Ontology Commons EcoSystem (OCES), consisting of ontologies and tools following specific standardisation rules, that can be effectively used as a foundation for data documentation in the industrial domain, in order to facilitate data sharing and utilisation/exploitation, and overcome existing interoperability bottlenecks.

The OCES will provide a way to harmonise data documentation through ontologies and taxonomies, making the data FAIR, and enabling intra- and cross-domain interoperability in a range of domains including materials and manufacturing. The effectiveness of the OCES in accelerating data-driven innovation will be proven at the end of the project through the delivery of demonstration cases, covering several application domains.

OntoCommons will achieve this overall objective through activities that are consistent with its nature as a Coordination and Support Action (CSA), strongly relying on communication, networking and coordination between EU and international stakeholders, and by making use of the available state-of-the-art tools and solutions as much as possible. OntoCommons brings together a wide range of communities/stakeholders, including subject-matter experts (e.g., material scientists), ontologists (e.g., philosophers, semantic web experts), implementers (e.g., software developers, database experts), industrial stakeholders (e.g., manufacturers), and end users.

The OntoCommons Roadmap includes a number of recommendations for policy instruments towards Data Sharing for the European Single Market. OntoCommons will support industry to the opportunities of digitalisation, 21st century societal challenges and the Green Deal.

This document will be updated at the end of the project. The consortium will use a variety of methodologies to harness stakeholder inputs, contributions and foresights, including online open consultation(s), targeted interviews, and open feedback sessions with specific groups of stakeholders and users.

1.1.1 Overall Methodological Approach

The OntoCommons project has been consistently built around a develop-test-validate-agree methodology, and follows the steps as organised in following phases:

- Stakeholder Networking: reach the widest set of stakeholders, via stakeholder engagement and cooperation with international organisations in order to provide widely accepted guidelines and good practices for the implementation of project activities
- Top Reference Ontology Development and Agreement: build the foundations of the OCES by developing the Top Reference Ontology (TRO) and mid-level ontologies (MLOs) following extensive stakeholder networking and feedback about design choices

- Domain level ontologies harmonisation and new development: development of new domain ontologies following stakeholder consultation and demonstrators' requirements, and reuse through harmonisation with the TRO of those already-available domain ontologies
- Demonstration: using the OCES approach, implement the data documentation process for various predefined and community-proposed application case studies.

Figure 3 - OntoCommons phases

1.2 Stakeholders

OntoCommons engages high-level domain experts, ontologists and ontology implementers, as depicted in Figure 4, through active participation in focused stakeholder workshops, which will help define the requirements for the OCES. Industrial stakeholders will be engaged through demonstrators and through awareness and capacity building in dedicated workshops and webinars. Beyond the key industrial users, *OntoCommons* aims to encourage the development of innovation based on newly enabled cross-domain interoperability, and will engage business ecosystem stakeholders in collaborative global workshops to identify novel cross-domain application opportunities in view of creating an *OntoCommons*-driven multiplier effect, feed emerging applications markets, and identify future research and innovation directions for Horizon Europe. In the long term, by engaging the community listed in Figure 4, the project will also impact and engage indirect beneficiaries such as citizens and consumers who benefit from products created by industrial developers and the services created by business entrepreneurs utilising OntoCommons.

Figure 4 - OntoCommons stakeholders

1.2.1 Cooperation and Stakeholder Engagement

OntoCommons will leverage relevant ongoing efforts and create synergies by close cooperation with all relevant bodies and initiatives, in particular European and international standardisation bodies, work done under European and international umbrella organisations and initiatives such as the European Open Science Cloud (EOSC), the Research Data Alliance (RDA), the European Materials

Modelling Council (EMMC), the Industrial Ontologies Foundry (IOF), as well as relevant EU projects, in particular the DT-NMBP-40-2020 funded marketplace.

OntoCommons facilitates the interests of its stakeholders regarding standardised documentation based on ontologies, and disseminates the CSA's results into these initiatives (including, e.g., recommendations, roadmaps and specific actions). Also, vice versa, key outcomes, recommendations and requirements from these initiatives are taken account in OntoCommons (roadmaps, recommendations, reference designs and blueprints, etc).

1.3 Vision

OntoCommons is designed to greatly strengthen the interconnectedness and interoperability of the industrial technologies ecosystem by means of establishing and enhancing a common semantic basis on which digital interactions, collaborations, co-innovation, re-use and valorisation of data can take place. *OntoCommons* aims at providing an **Ontology Commons EcoSystem (OCES)** that facilitates the development of domain ontologies through a set of specifications, an underlying top-level ontology, and a set of tools that will enable end users to harvest the potential of ontologies at different levels (see Figure 1), according to their specific needs, from simple metadata schema generation to more complex integration with RDF databases. The OCES will enable users to adopt ontologies for data documentation, facilitating their data management under FAIR principles.

It is widely accepted that a standardised data documentation and semantic interoperability across NMBP will have a major impact on the digitalisation of European industry while at the same time strengthening its competitiveness and growth opportunities.

1.4 Disclaimer on the approach

The OntoCommons Roadmap recognises the importance of making advances in standardised data documentation and semantic interoperability across NMBP domains in order to support the digitalisation and competitiveness of European industry. It strives to identify shortcomings and barriers in current approaches and proposes actions to address them based on the input of experts and many discussions between various stakeholder communities. The OntoCommons CSA has organised and facilitated many of these interactions with and amongst stakeholders who are broadly identified as: domain experts, ontologists, ontology developers, industrial stakeholders, business ecosystems stakeholders, etc. However, the OntoCommons CSA cannot be held responsible or liable in any form for statements made by stakeholders, or potential socio-economic consequences of implementation and use of any of the roadmap's recommendations. All recommendations are based on stakeholder feedback and are presented in good faith. They do however represent the OntoCommons position on the best possible approach for ensuring economic or societal impact in the NBMP industrial technologies' ecosystem as regards data interoperability and its potential.

2. Introduction

OntoCommons CSA coordinates and supports an Industry Commons in the field of NMBP, in particular to facilitate and enhance the use of the same underlying data to support the development of numerous new products, services or manufacturing processes, and to enable data valorisation by any business or public entity based on the same data in different data-sharing collaborations to accelerate data-driven innovation and spill-over into new areas of the economy.

The development, implementation and widespread uptake of the Ontology Commons EcoSystem (OCES) is an important pillar of the Industry Commons ecosystem. In particular, high-level needs for OCES include the following:

- Effective and meaningful data documentation to enable the harvesting of data resources that are generated by European industry and that currently are not fully exploited.
- Covering the gap between human and digital tools (Industry 5.0): The need for expressing human knowledge through a widely applicable and general methodology that is both human understandable and at the same time easily processable by digital devices.
- The need for machines to be able to ingest existing knowledge and therefore enable the exploitation of the computational power of modern HPC facilities and software.
- Knowledge-based reasoning as opposed to generic AI by formalising the knowledge already available from existing scientific disciplines and to provide knowledge structures that can be used as references for the analysis and interpretation of machine and deep learning AI findings (XAI).

In this context, the OntoCommons Roadmap considers Needs, State of the Art, Gaps, Definition of Success, and Recommended Actions for a number of topics contributing to an Ontology Commons Ecosystem for ontology-based data documentation, grouped into:

1. Ontology Foundations: Top Reference, Middle, Domain and Application Levels
2. Integrated Development Environment (Tools) and Infrastructures
3. Industrial Impact including marketplaces, standardisation, education and human resources

This Roadmap summarises the outcomes from a number of events organised by the project in 2020-2022:

from the Global Workshop: Ontology Commons addressing challenges of the Industry 5.0 transition (02.-05.11.2021), Workshop on Industry Commons Marketplaces (29.04.2021), Workshop Top-Level and Mid-Level Ontologies Multi-Disciplinary Workshop (25+29.03+08.04.2021), Workshops: Creating a Knowledge Exchange Space for data management and documentation – KExS (01.07.2021+ 22.10.2021), Workshops: DORIC-MM 2021 (15.03.2021 +07.06.2021), Workshop on Tools for Ontology Engineering (19.03.2021), Workshop Demonstrators at work to deliver an Industry Commons Marketplace (09.-10.03.2021), Joint Workshop within Open Science Fair "Applying FAIR principles to open science and industry to drive innovation: challenges and opportunities" (22.09.2021), Webinar "What can ontologies do for standardisation? OntoCommons.eu project in the standardisation ecosystem" (30.09.2021), Webinar on "An industrial ontology journey stopping at Standardisation", FAIR Data & Innovation (23.02.2021) etc.

and provides the industry needs, current gaps and policy recommendations.

The materials and manufacturing domains community can continue to benefit from the insights of the "OntoCommons ontology and data documentation development Roadmap" created with the expertise, feedback and guidelines shared by OntoCommons' stakeholders during the project's lifespan.

3. Ontology Foundations

3.1 TOP Reference Ontology

A top-level ontology (TLO), or foundation ontology, is a formal ontology which consists of general concepts, expressed by terms such as "object", "property", "relation", that are common across all application domains. A mid-level ontology (MLO) is an ontology primarily intended to include concepts that are general in the context of one or more specific disciplines (e.g., manufacturing, materials science, chemistry) with the aim of providing the core of a shared vocabulary to which lower-level ontologies may refer to. An MLO will provide a higher level of detail than a TLO and can be seen as an extension of the taxonomical structure of a TLO into more detailed subclasses.

TLOs provide a sound theoretical foundation of consistent and well-defined logical theories and interpretation frameworks that fits the needs of a large community of users sharing the same knowledge representation requirements and epistemological approach. The adoption of an already developed TLO/MLO system, acting as a common knowledge representation framework for a community, facilitates mutual understanding and information sharing between its members. It saves development and training costs, those being the specific skills requirements for ontology development (e.g., conceptual and ontology engineering, logics) which are not always available within each specific community. Moreover, ontologies provide an expression of knowledge through machine-understandable formal languages that can enable the transfer of knowledge between humans and machines, and vice versa.

Several TLOs and MLOs exist, each one expressing a different representational perspective according to the community whose needs the ontology is aimed to fulfil, based on a specific terminology, philosophical commitments, or epistemological approaches. A list of relevant resources (i.e., researchers, ontologies, initiatives) in TLO and MLO fields is available. The different approaches used by TLO/MLOs may lead to information silos in which the same knowledge about the world's entities may be represented in different and non-compatible ways. A single ontology might be a philosophical ideal, but it is almost certainly not a practical reality, and the use of many different TLOs in practice is a reality that cannot be neglected, nor removed.

A Top Reference Ontology (TRO) is a concept introduced by the OntoCommons project, defining a common foundation for data interoperability to enable knowledge sharing across TLOs, and that aims to add a pluralistic perspective in the choice of a TLO. A TRO can allow the comparison and interoperability of TLOs by providing a foundation for harmonisation and data sharing between ontologies based on different TLOs.

A TRO will consist of i) a set of TLOs (i.e., BFO, DOLCE, EMMO) and ii) a set of mappings between TLOs. The mappings will allow us to establish semantic correspondences between entities belonging to different ontologies (the so-called semantic alignments) and include the alignment links (matching relations) expressed in the form of bridging axioms, as well as annotations and carefully chosen examples and counterexamples. In this way, the TRO will account for the mutual relations, similarities, differences, design rationales of the selected TLOs and, when this is not possible, will explain the reasons of disagreement in order to guide developers and users of domain and application ontologies and so that they can better understand and characterise their ontological commitment.

The choice of a TRO as foundation for interoperability between communities (e.g., scientific, industrial, economic communities) instead of committing to a single TLO ontology, has several important benefits in terms of inclusivity and reuse of existing solutions:

- avoiding the risk of alienating existing communities already committed to a specific TLO, by providing a framework that would expand the user target of their ontologies instead of restricting it,
- enabling reuse of available domain ontologies developed with different TLO ontologies,
- allowing for multiple representations for the same problems (pluralism) depending on the perspective of specific sets of users (e.g., the same problem could be represented following a physics or chemistry approach),
- facilitating the choice of ontologies that adhere to the users' own language and conceptualisation, while at the same time providing alignments to bridge with other representations for the same domain, minimising the risk of clashing against i) already in use or de-facto standards or ii) important international actors that may use their global positioning to promote another standard in the future (e.g., Google).

3.1.1 Industrial Need

The TRO will meet the following industrial needs:

Need #	Need Description
1	**Covering the gap between human and digital tools (Industry 5.0)** The need for expressing human knowledge through a widely applicable and general methodology that is both human understandable and at the same time easily processable by digital devices. The need for machines to be able to take in existing knowledge and thus enable the exploitation of the computational power of modern HPC facilities and software.
2	**Interoperability between standards, vocabularies, data, and software tools** The need to facilitate industrial innovation which nowadays requires more and more multi-disciplinary interactions between experts, resources, tools, and data coming from different disciplines with different languages, standards, and formats. The need to enhance interoperability at several levels, from the human level to the machine and data level.
3	**Better formats for standards and vocabularies**

	The need to formalise standards using a unambiguous set of definitions for concepts and providing relations between concepts in a clear way which should be formal but also easily understandable by human agents (e.g., how a product relates to a manufacturing process, how a simulation is related to a material, how a measurement device relates with how a characterisation process is structured).
4	**Effective data documentation** The need for a generic data documentation methodology that makes it possible to easily find, assess, interpret, and retrieve data generated in different sectors is required to enable the harvesting of the data resources that are generated by European industry and that are not fully exploited.
5	**Knowledge based reasoning as opposed to generic AI** The need to provide strong foundations for knowledge-based AI, by formalising as much as possible the knowledge already available from existing scientific disciplines. At the same time, to provide knowledge structures that can be used as reference for the analysis and interpretation of machine and deep learning AI findings.

3.1.2 State of the Art

A landscape analysis of the currently active TLOs and MLOs that have a potential applicability to the domains defined by the H2020 NMBP Work Programme has been provided by [Ref 3.1.1], focusing on BFO, EMMO, DOLCE, BORO, ISO 15926, OPM and SUMO. Another general survey of TLOs [Ref 3.1.2] has been recently provided by the Construction Innovation Hub [Ref 3.1.3]

TLOs are expressed at many different levels of expressivity/computability (e.g., from RDF to HOL). Most existing TLOs provide an axiomatisation based on semantic web technologies, enabling the use of tools that have been developed in the past two or so decades and that are nowadays fully available (e.g., RDF databases, reasoners) at production level.

TLOs are actually used in several projects and initiatives tha aim to facilitate the development of a framework of homogeneous and interoperable MLOs and DLOs [Ref 3.1.1]. The interoperability between frameworks based on different TLOs can be achieved by creating mappings between the concepts expressed by each TLO. However, mappings between TLOs are actually not available for practical usage and need to be further developed.

Wider usage of TLOs is prevented by various barriers, such as different vocabularies and disciplinary barriers between top level and domain level ontologists, or by the overhead of effort in the initial phase. For this reason, end users usually develop *ad hoc* domain or application-level ontologies focused on their particular domains, without committing to a more general TLO framework, thus creating ontological silos for their knowledge representations.

3.1.3 Gaps

GAP #	GAP DESCRIPTION

1	**TRO and TLO establishment** Ontology developers are not investing in the overhead of introducing a TLO as a knowledge representation foundation. Little to no TLO adoption could be the fastest and cheapest approach for a company to go for, putting interoperability as a mid- or long-term objective posing unmotivated cost in the short term. A TRO and TLO are however recognised as advantageous when the scope is beyond a single project.
2	**Cross-disciplinary gap** There is a gap in the adoption of TLOs, based on the lack of cross-disciplinary understanding. TLO terms and concepts come from the ontology or philosophy fields that are often unknown to people working in a specific application field. Adoption of TLOs is also hampered by gaps in knowledge regarding TLOs approaches and architectures, and this is in many cases the reason why they are not used. MLOs and DLOs are easier to understand by non-ontology experts, who usually refer directly to them instead of investigating more high-level general approaches provided by TLOs.
3	**Lack of inter-TLOs connections** There are many TLOs, expressions from different communities, often with conflicting commitments, so that there is not a one-size fits all approach. The adoption of TLOs is always done on the basis of the community to which the user belongs, mainly because the TLO adopts ontological perspectives that are easy to understand by the community, but also for historical and practical reasons (e.g., easy compatibility within existing resources). This has created silos in terms of both communities and data.
4	**Lack of high-level tools for TLO/MLO engineering** There is a lack of professional tools for ontology engineering. These include tools for reasoning, validation (not only verification), code generators, complexity management tools, as well as tools that help with ontology reuse, including leveraging ontology patterns and anti-patterns.
5	**Lack of computability and computation power** There is a lack of development of new algorithms and HPC solutions for symbolic computations (e.g., provers, reasoners) and theorems in expressive logical languages (e.g., OWL 2 DL, FOL), that would enable the exploitation of the TLOs' semantic richness. The few existing tools often impose constraints as to the choice of ontology expressivity (i.e., the language used to represent the logical system, such as FOL, RDFS, OWL) in favour of computability. Both tractability and expressivity are needed: a heterogeneous approach, where the same ontology is provided at different levels of expressivity (e.g., from HOL to XML and RDBs), is to be developed.
6	**Lack of multidisciplinary competences (from domain expert to ontology)** There is a lack of multi-disciplinary competences within all communities involved in ontology development and usage. Potential users, scientific communities and industry are

	often affected by a competence bottleneck when reaching for formal representation of knowledge. There is a lack of people who can write good and reusable ontologies.

3.1.4 Definition of Success

The applicability and usability of TLOs and MLOs can be measured according to their capability i) of being used in different knowledge domain environments and ii) of being the foundations for mutual interoperable lower-level ontologies. This success is a measure of the ability of the ontology to capture the general aspects that are common to all domains (for TLOs) or to a selected set of expert domains (for MLOs).

The scalability of a TLO/MLO can be measured by its capacity to be applied in environments that makes use of tools based on different expressivity levels. Ontology scalability enables users to adapt their knowledge representation and reasoning to the computational resources that are actually available. In relevant industrial application fields, this would likely require RDFS or OWL2 formalised ontologies (enabling the usage of RDF triplestores), while at research level, it would require FOL/HOL formalised ontologies (enabling the usage of higher-level theorem provers).

The ability to enable cross-ontology interoperability is demonstrated with the establishment of an effective methodology for the mapping between relevant TLOs and the mapping of relevant MLOs with more than one TLO.

The ability of a TLO/MLO to overcome the cultural and disciplinary gap between the top level and the application level is demonstrated with the successful development by non-TLO/MLO experts of application ontologies based on a TLO/MLO framework, without the support of high-level ontology experts.

The **industrial level success** of this TLO/MLO framework can be defined as its capability to enable sharing of data or data re-use, where the same data is being used by different companies for different materials, products, and processes.

3.1.5 Recommended Action

The TLO terms and concepts should be made accessible and understandable to domain and application ontology developers. End users, the ones more interested into data documentation, should be facilitated to find, select, understand, and use MLOs/DLOs.

ACTION #	ACTION DESCRIPTION
1	**Push for adoption of TRO/TLOs/MLOs** Background: Referring to a common TRO within an organisation, community or project or a framework of mapped TLOs for the development of a domain specific ontology (MLO, DLO or ALO) facilitates its development and greatly enhances interoperability between ontologies. However, the overhead of introducing a TRO may not be justified in all application cases (e.g., small projects) and partial or no TLO adoption could be the

	fastest and cheapest approach. Demonstrators may play an important role in showing the capabilities and limits of TRO. Action: A well-defined and accurate adoption route (e.g., dedicated training courses and educational paths), together with good documentation should be the priority to overcome some of these barriers.
2	**Stimulate TLO communities to work together and adopt a TRO** Background: There may be multiple TLOs suitable to represent any application. For this reason, it is not expected, nor is it reasonably feasible to assume that a single specific TLO will be chosen by all the communities. A bridging between TLOs is needed to retain a community-preferred approach and at the same time provide ways to share knowledge with other communities. Tightening the TLO communities by providing common grounds for discussion and by taking the effort to understand similarities and differences can be a way to overcome this bottleneck. This will also strengthen their overall position towards policy makers. Action: Clear communication based on an understanding of common principles, differences, and capabilities between TLO/MLOs communities, as well as concrete mappings, should be developed to prevent creating ontological silos. Promotion and funding of initiatives to further develop the OntoCommons TRO (which is less likely to be coherently supported by industry stakeholders). Consider in the evaluation of future EU projects the mandatory inclusion of their ontologies in the OntoCommons ecosystem.
3	**Ontology scalability (expressivity vs computability)** Background: TLO authors are often focused on expressivity, philosophical commitments, and generality of the approach, while application domain experts focus on “getting the things done” (tractability/being easy to handle). Expressivity (conceptual models) and computational efficiency (implementational models) are both important and necessary. We should find a good balance between them, and fortunately there is still room for manoeuvre in ontologies that are focussed on expressiveness to achieve further efficiency. Reasoning based on knowledge is a step that should precede Artificial Intelligence, and this is certainly one route that is possible for scenarios with low data volume. Action 1: Promote a scalable approach to ontology development and usage by proposing funding of multi-disciplinary projects that puts together philosophers, formal ontologists, computer scientists, applied scientists, engineers, and industrial end users. Action 2: Promote projects dedicated to the development of HPC tools for symbolic based reasoning (which is part of the AI field), to be proposed as an alternative/complementary method to machine and deep learning approaches.
4	**Tools** Background: Awareness about the need for tools for ontology development is falling short with policy makers and scientists. Good tooling combinations and ecosystems that

	make it easy and efficient to map existing data sources to ontological concepts would speed up development. Action: Raise awareness about the need of tools for ontology development.
5	**Education and Training** Action 1: Clear communication, such as which problem classes can be solved with ontologies and which problems cannot be solved with ontologies, is to be targeted. Action 2: Education and training, together with incentives for using ontologies in the scientific community and in industry are to be developed.
6	**Expert** Background: MLO definition faces the challenges of being coherent with existing definitions or standards provided by each discipline of interest (e.g., chemistry, metrology), and at the same time of being able to provide constraints on the formal structure of an ontological representation of the discipline (e.g., taxonomy class levels). A clear separation between disciplines is not a realistic goal since each domain has already created its standard terminology and definitions. We are always dealing with already established domain-specific classificatory systems that often lack rigorous and consistent logical structures, and that introduce ambiguities in the definitions. It is very important that a MLO at least reuses already established terms and standards, instead of starting from scratch at the top, using a pragmatic approach in which we also propose changes to already established standards, to improve their consistency in an ontological environment. It is reasonable to believe that abstractions can help bridge disciplines, even if it is difficult for some disciplines to be formally defined.

3.1.6 References

Ref 3.1.1 Laura Ann Slaughter, & Jens Otten. (2022). OntoCommons - TLOMLO Landscape Analysis Report. Zenodo. https://doi.org/10.5281/zenodo.6504440

Ref 3.1.2 https://www.cdbb.cam.ac.uk/files/a_survey_of_top-level_ontologies_lowres.pdf

Ref 3.1.3 https://constructioninnovationhub.org.uk/

3.2 Industrial Domain Ontologies

A Domain-Level Ontology (DLO) can be seen as a specialised module of a Mid-Level Ontology (MLO), targeting the specific domain of applications (e.g., additive manufacturing, composite materials). A DLO is characterised by an increased level of detail concerning an MLO, a more pronounced horizontal extension, and a strong dependency on the domain of application, while still maintaining some neutrality as to the specific problem addressed.

According to the CSA nature of the OntoCommons project, the activities and the choice of technical solutions were guided by consultations with the stakeholder community, and until now made use as much as possible of the available state-of-the-art solutions, thereby limiting the development of domain-level ontology definitions, harmonisation, and demonstration.

The harmonisation and development of domain ontologies are towards the standardisation of ontologies following the agile development/test/validate/agree procedure. The harmonisation methodology should be a hybrid approach, based on a top-down alignment to a Top Reference Ontology (TRO), and a bottom-up focus on domain specification and requirements, as well cross-domain interactions that are grounded from consultation with domain experts and industrial stakeholders as well as feedback loops driven by real industrial use cases.

3.2.1 Industrial Need

Need #	Need Description
1	**Data Integration and sharing** There is a unanimous understanding by industrial stakeholders that they will benefit from improvements in data integration, sharing and format conversion, while 70% of the respondents to a survey conducted during the DORIC-MM Workshop [Ref *3.2.1*] responded that they have started or already adopted such standards in practice.
2	**Standardisation** Though there are many standards available for the domain of materials and manufacturing, there is a general lack of consensus among these standards. while in some of our domains we have standards at the level of ISO (as per 2/3 of respondents being in favour), in others we are very far from that (e.g., a CWA). Though there is no doubt that standards are key, it is very hard/impossible for them to be produced within the timescale of a typical EU project, unless the project is really about just producing the standard. In this context, de facto standards are also important, and for that, it is necessary to get standardisation organisations, universities, supporters of data spaces, and especially commercial partners (these could be manufacturers, software vendors) including large vendors (e.g., as in the case of Cape-Open [Ref *3.2.2*]) on board. It is also important for such standardisation efforts to be validated and showcased by several industrial demonstrators. Meaningful data sharing is becoming a more and more clear need, for which global semantic alignment is a requirement.
3	**Various domain perspectives** Regarding domain ontology development a major problem is how to combine various views and domains. According to industry, it is still an unsolved problem in engineering. In this regard, a dynamic visualisation may help. For example, one could think in terms of dynamic visualisation: it would be very good to have something like that for ontologies, where you can zoom in and out in detail as needed. Think of Google Maps: that is a type of visualisation that might be valuable for an ontology ecosystem. For a system of (domain)

	ontologies, a Unix-like philosophy could be used, like a periodic table of elements: many building blocks that can be combined, and there are rules to combine them. Different tiers of complexity are present and not all blocks need to be used at the same time.
4	**Interface domain ontologies with TLOs** The interface to a TLO is more relevant from the point of view of developers of domain ontologies, but not necessarily the intricacies of the TLO. In other words, we should "isolate" the domain ontologies from the TLO's theoretical and technical details. It also appeared from discussions with industrial partners that it is urgent for domain ontology development to include efforts in modelling some well-developed pieces of information and take them further and put them in the form of an ontology (e.g., as carried out in EMMO-Crystallography), as a domain reference model or MLO to help in harmonising the application-specific discrepancies among different communities.
5	**Link domain experts to ontologists** Domain ontology development needs input from both domain experts and ontologists. In other words, the domain ontology development needs to mix both top-down and bottom-up approaches. The domain experts and ontologists complement each other's roles where the former brings the domain level requirements and helps in characterising the ontology terms from a domain's point of view and the latter provides formalisation in the ontology model using theoretical grounding and ontology engineering best practices. In this effort, we need to have both domain experts and ontologists collaborating during the development of the ontology and on the final product to be accepted by domain users.

3.2.2 State of the Art

In a large-scale state-of-the-art survey, we collected 222 entries, out of which we initially identified a total number of 108 relevant ontologies, including 74 machine-readable ontologies vs 34 non machine-readable ontologies. This dataset has been then extended with 26 ontologies, including 11 non-machine readable ones. The dataset contains 134 ontologies (89 machine-readable ontologies, including 9 from MatPortal, and 45 non-machine-readable ontologies). This set of ontologies spans multiple domain areas of material and manufacturing: 10 from physics and chemistry, 51 from mechanical and industrial engineering, 14 from thermal and process engineering, 22 from material science, 11 from computer science, and the rest from other domains. This representation shows that the most prominent domain in terms of ontologies is the mechanical and industrial engineering domain. Below that, we find material science and engineering. For other domains, it is quite likely that very few resources are available.

3.2.2.1 Use of TLO by the DLO

It was discovered that among the corpus of ontologies considered, two main TLOs were used: the Basic Formal Ontology (BFO) (40%) and EMMO (18%). In the overall list of ontologies, without any domain consideration, we found that 36.5% of the ontologies were aligned with TLOs, while the physics and chemistry domain had a maximum percentage (80%) of ontologies aligned to a TLO. Concerning the serialisation format chosen for these ontologies, the main format used to publish machine-readable ontologies was RDF/XML followed by Turtle and OWL/XML. This distribution is preserved across different domains considered except for D1 – Physics and Chemistry (see Figure 5) for which there are no ontologies using the OWL/XML syntax.

3.2.2.2 Topological analysis

A detailed topological analysis of the ontologies is performed as part of the landscape survey. It is found that the size of the ontologies (in terms of the number of axioms they define) varies a lot from ontology to ontology. The largest is CHEBI, which defines over 2.5 million axioms. The Universal Standard Product and Services Classification (USPSC) is another large ontology. However, unlike CHEBI which mainly defines classes, USPSC defines many individuals, which might be unexpected for a domain ontology. While looking at other aspects of ontology size, we see that the Coordinated Holistic Alignment of Manufacturing Processes, MatOnto and Allotrope ontologies all have a significant higher annotation count compared to others. When it comes to object properties and data properties, the Schema.org ontology is dominant. For data type count, it is harder to get a clear picture as the variation is relatively small. For Rbox-size, the Allotrope, Coordinated Holistic Alignment of Manufacturing Processes and ScorVoc ontologies all have a significantly larger size than others. If we look at relative measures, the Organization and Sensor Observation Sampling Actuator ontologies both have more than one annotation property per class. The LinkedDesign ontology has by far the largest Tbox (Terminology Box, i.e., classes) size per class, while the Standards ontology is dominating the Abox (Assertion Box, i.e., individuals) size per class and the average instance per class. We see that the average number of asserted sub- and superclasses is close to one for all ontologies in all domains. Just a few ontologies have a slightly higher average number of sub- and superclasses, including CHEBI and EMMO Mechanical Testing. The average number of superclasses with more than one subclass per class is about 10% overall, and also on average this holds for the domains. However, within the domains, there are some variations. For example, EMMO Atomistic stands out with an average number of superclasses with more than one subclass of 50%. Multi-inheritance is used in 35% of the analysed ontologies and is most used in the NanoParticleOntology. It is most frequently used in the Physics & Chemistry Material Science and Engineering domains. However, for the ontologies within Mechanical and Industrial Engineering that use multi-inheritance, it is rather intensively used (in terms of multi-inheritance per class) in ontologies like Gist Upper Enterprise Ontology, Product Life Cycle Engine and Scheduling Reference Ontology. The overall average number of axioms per class is 52%. A few ontologies, like ExtruOnt, use axioms more extensively, but there are also several ontologies with no complex axiomatisation.

3.2.2.3 FAIRness of DLOs

The FAIR analysis of a sample of ontologies revealed some disparities between domains. Indeed D1- Physics and Chemistry (see Figure 5) is the domain with the highest FAIR score on average. This is because ontologies listed as part of the domain are following up on some of the OBO Foundry recommendations. The "FAIRest" ontology is the Allotrope Ontology which is the only ontology tracking provenance using PROV. In most cases, considering the global FAIR score improves the situation with respect to FAIRness. However, no ontologies passed the threshold of minimally FAIR score (i.e., compliant with the mandatory FAIR recommendations).

3.2.2.4 Overall state of affair

In the domain of physics and chemistry, one domain-related TLO is available, EMMO (Elementary Multiperspective Material Ontology), where multiple mentions have been found. There are also some vocabularies and metadata in this domain that are at a pre-standard level. These standards may be harmonised by using ontologies like EMMO. There is also a need for EU-wide standard vocabularies and metadata in this domain. More emphasis on trust, demonstrations and tool-based development will allow domain experts to easily map concepts to ontologies in this domain.

Although lots of work has been done for mechanical and industrial engineering, they are largely at a low Technology Readiness Level - TRL (i.e., research work), and very little of this has direct economic commercial value. Therefore, a framework, within which mechanical and industrial engineering domain ontologies can be developed and extended in such a way that able to meet commercial needs, is required to be developed. Some supports were also found for highly expressive ontologies using Common-Logic, besides that also available from OWL, in this domain.

For process and thermal engineering, alignment among processes from the different industrial sectors is needed, along with modularity, reasoning, and usability. Some sets of ontologies have been available for about twenty years and need to be consolidated. They include variants of ISO-15926 and OntoCape. The process sector and the oil-and-gas industry have funded initiatives such as READI and DEXPI which needs to be further harmonised. There is a need to make ontologies in this domain usable by domain experts.

For the material science domain, various ontologies are adopted within the field. Some of them need to be linked to "biology" ontologies. There is also a need for: building a common entry point to find ontologies; a common vocabulary for linking experiments and modelling; standards designed to enable interoperability across different scales, especially in model validation; and modelling and characterisation for molecular and nanomaterials. Different file formats for experimental data or simulation outputs used in this area need to incorporate metadata. Needs were expressed for metadata standards for multiscale simulation workflows, mechanical testing, and crystallography. There is difficulty in linking metadata standards with ontologies (DICOM).

3.2.3 Gaps

GAP #	GAP DESCRIPTION
1	**Models granularity** In the material science domain, some of the important focus areas are material properties, crystallography, statistical methods, characterisations, molecular materials, and chemical kinetics. Although some ontologies are identified for these domains, participants of the workshops emphasised the need for more extensive and granular models addressing these areas. Areas, such as tribology, corrosion, and powder materials are identified by participants as an immediate focus area for ontology development.

2	**Lack of generic and application-specific ontologies** Some of the areas in which no ontology has been found yet are the roles of chemicals and materials, materials for electronics, and extensive physical properties, to name but a few. There is also a general lack of ontologies that covers fundamental and application-specific physics and chemistry-related topics as evident from the domain distribution statistics in Figure 5.
3	**Lack of standardised methodology and tools** Although a number of existing ontology development methodology and tools are available as surveyed by the OntoCommons "Report on Landscape Analysis of Ontology Engineering Tools" [Ref 3.2.3], no such methodology or tools have been standardised with a wide agreement from the community. Furthermore, no significant methodology or tool specific to harmonising ontologies is available. It is to be noted that harmonising ontologies includes integrating the domain ontologies in an integrated framework where mapping is available not only between intra- and cross-domain ontologies but also between DLOs and upper level ontologies (e.g., TLO, MLO).
4	**Ontology as a conceptualisation of reality vs information model:** The key need for an ontology here is to be able to formalise the terms used by engineers in the manufacturing field. Engineers often find it difficult to change their perspective because they find it difficult to connect their domain-specific view to a global point of view. For example, some ontologies address CAD, but they just classify the elements (point, line, surface, body) in some formal language, e.g., OWL, but these ontologies fail to capture the cognitive understanding that CAD models capture. We are not interested in the ontology being another representational scheme of some existing representational scheme, but rather as a way to model the concept that the existing scheme represents. In this sense, ontologies should not be used as another data structure but instead as the conceptualisation of domain knowledge. As a counterargument, a trade-off approach at the intersection between a foundational view on ontology engineering and an application view is possible as well. Indeed, the Semantic Web community provides a lot of technical support to handle linked data and knowledge-based representation and reasoning systems where ontologies do play a fundamental role.
5	**Ontology sustainability** During the landscape survey, some gaps were identified in the overall state of ontologies in the domain of material and manufacturing. For example, many good quality ontologies are lost due to a lack of maintenance and have not found wider adoption. A permanent repository needs to be established for long term persistence of ontologies. At the same time, this repository needs to be specific to industry so that the ontologies are not hard to find/crowded out by other ontologies from nearby domains such as the biological or agricultural sectors. It can also be difficult to search for a suitable ontology when they are not tagged with the reference domain they cover. Ontologies are also built using different languages, tools with different nomenclature and metadata, suggesting the usual lack of a

	common development framework. Also, a lot of ontologies are developed by academia that sometimes does not capture the real needs of industry. A lack of a sustainable strategy also hinders the development and maintenance of an ontology and ultimately its quality. Because of this lack of quality, some ontologies lose trust amongst industrial users. One of the difficulties in reusing existing ontologies is the difference in the complexity level they take under their scope. If we want to reuse an existing ontology, first determining at which level this ontology falls (domain reference, domain, or application) can be a complicated effort.
6	**Lack of Standardised Method for Domain Ontology Evaluation** The quality and coverage of DLOs need to be evaluated by formal methods. One method is to compare the concepts covered in different domain ontologies with a Gold Standard catalogue; several metrics related to coverage, overlap, and gaps are described in the literature in this area. However, selecting suitable Gold Standard catalogues for each domain of focus is challenging and requires consensus among domain experts. Along with formal methods, more inputs from the community and internal reviews need be conducted to understand the usage and maturity of these ontologies. To quantify the discrepancies between each subdomain and to evaluate potential overlaps, a more detailed analysis of high-level domain and subdomain coverage is required. However, the mapping of ontologies to domains and sub-domains currently can only be done by experts inspecting the ontologies one by one, which is very time consuming and not sustainable. For a more automated and scalable analysis of domain applicability, at least some of the 'Domain' metadata would be required for ontologies, which many existing ontologies lack.

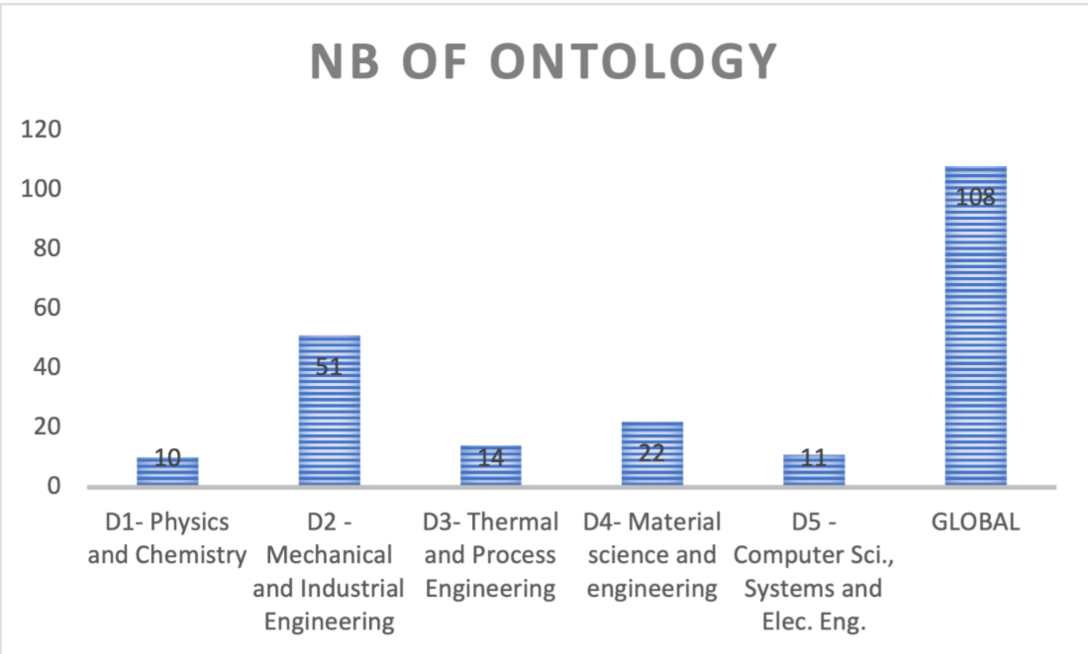

Figure 5 - Distribution of ontologies by domains

3.2.4 Definition of Success

The spectrum of ontologies overlaps across different communities from different areas such as philosophy, Semantic Web, logic, programming languages, and engineering domains. As a branch of philosophy, "ontology" is the science of what is, of the kinds and structures of the objects, properties, and relations in every area of reality. In simple terms, it seeks the classification of entities.

We have noticed that in computer science or philosophy, the perspective of ontologies is the same. It is about the representation of entities, ideas, and events, along with their properties and relations, according to a system of categories. However, the focus in both domains is different. In the computer science and engineering area, scientists are more focused on establishing fixed, controlled vocabularies, while in philosophy, the focus is more on the perception and representation of the world.

In the computer science and engineering area, we are more focused on the formats of the vocabularies (OWL, JSON, UML, etc.) and the capacities to process them to answer specific applications. This focus should not make us miss an important part, which is the semantic disambiguation of the vocabulary. We are convinced that we should care about the quality of the produced ontology. For that, we must consider also the philosophical focus (the nature of things) when building an ontology so as to define any semantics in the controlled vocabulary logically.

High-quality domain ontologies are essential for their successful adoption in any domain. However, the field of ontologies is nascent and therefore, somewhat unstable. In practice, many new ontologists (ontology modellers) begin as software developers who are accustomed to thinking strictly in terms of data and not of the entities or associated concepts to which such data relates. The result is that many ontologies are little more than data dictionaries. This orientation needs to change in the development of ontologies. We also acknowledge that such improper orientation has led to problems in many ontology development projects and has often led to a poor reputation for the notion of an ontology itself. The quality of an ontology should be checked, not only according to some metrics and/or if the ontology answers the competency questions of a specific application, but also it should be carried out according to a philosophical basis of common understanding of different perspectives within the same domain.

In biomedicine, by contrast, ontologies have made significant inroads as valuable tools for achieving interoperability between data systems whose contents are derived from widely heterogeneous sources. The OBO (Open Biological and Biomedical Ontologies) Foundry was designed to be a community resource consisting of public domain ontologies in the health and biological sciences space. It is a collaborative experiment based on the voluntary acceptance by its participants of an evolving set of principles designed to maximise the degree to which ontologies can support the needs of working clinical and biological scientists. The OBO Foundry has been a successful venture in the bioinformatics domain, and its approach is now being copied by others such as the Industrial Ontologies Foundry. Barry Smith provided a demonstration of the utility of the Foundry methodology in the neurophysiological, neuroanatomical, and biomedical domains. The United Nations Environmental Program (UNEP) approach to achieving interoperability is modelled on the OBO

Foundry, and OBO Foundry principles for ontology development are also now in use by ontology developers in other areas, including manufacturing, geology, transport, and security.

The Financial Industry Business Ontology (FIBO) developed as an open standard by the EDM Council is another success story of domain ontologies. For FIBO to become a widely accepted standard to be implemented in many systems across their industry, the Council is testing and demonstrating FIBO ontologies as they move through a rigorous process and become available to industry.

ISO 15926-8:2018 specifies an ontology integration of life-cycle data for process plants including oil and gas production facilities. Since its inception and standardisation under ISO in 13 parts, including rigorous upper level semantic, oil industries (such as the Norwegian Oil Industry Association) and many other capital-intensive projects have adopted ISO 15926. Several initiatives to extend ISO 15926 has also been funded by the oil-and-gas industry such as READI and DEXPI. One of its successes lies in the fact that the ontology (part 8) is based on a prior industrial standard. This also helped ISO 15926 to be quickly adopted by industry.

From these examples, we can argue that the success of domain ontology critically depends on its adoption by the community. To reach an agreement on the ontology, the ontology must demonstrate that it can bring the desired interoperability in the industrial value chain.

The success of an ontology needs to be ensured by the right choice of development process. Apart from proper requirement engineering and technical development, the ontology needs to be continuously maintained with periodic release cycles to address progress in the industry. A framework to focus collaborative efforts on developing, standardising, sharing, maintaining, updating, and documenting industrial ontologies needs to be implemented to ensure the success of domain ontology development. An ontology, being a tool for industrial data interoperability and organisation, needs to be viewed as an asset for the organisation. Following other asset life cycle approaches, all the ontologies used by an organisation need to be periodically reviewed and maintained, and, if necessary, new ontologies may need to be developed, or existing ones upgraded, re-organised, and re-purposed.

To bring in open collaboration among industries or between industries and academia, a comprehensive set of fully open-source, stable ontologies for different aspects of science, technology, and business need to be established. These ontologies need to be fully harmonised by using unambiguous and formalised foundational concepts and well-documented for their use. In this regard, the adoption of FAIR metadata for annotation, standardised by initiatives like FAIRsFAIR and EOSC, may help in ensuring easy discovery, accessibility, and reusability of these ontologies by the community.

Finally, domain ontologies need close collaboration between ontologists and domain experts. To ensure interoperability across domains, disciplines and industries, domain ontologies need semantically unambiguous and formalised definitions for their content. For this purpose, every domain ontology needs to be built with a coherent top-level ontology. At the same time, the scope and purpose of the ontologies need to be derived from content contributed and monitored by domain specialists and need to take the domain expert's viewpoint into account during modelling. As already mentioned, a successful ontology needs to be understandable to the domain practitioners and be able to capture the interpretation of application data as intended by industry.

3.2.5 Recommended Action

Action #	Action Description
1	**Standardisation of the ontology engineering steps** Standardise every facet of the steps for domain ontology engineering. Some of the recommendations for standardising ontology engineering methods is to adopt one of the formal methodologies such as LOT, including the use of well-defined competency questions for requirement engineering and validation of the ontology using well-defined completion criteria. A detailed technical principle needs to be agreed upon by the development team, and rigorously followed regarding the choice of language, expressivity level, editing tools, metadata, naming conventions, and most importantly a continuous development-integration-release strategy.
2	**TLO-MLO alignment** Adopt a coherent top-level ontology and a set of mid-level ontologies to ensure interoperability across domains in the domain ontology model. The domain ontology should refrain from adding foundational concepts (primitives) and reuse existing terms from the upper-level ontologies. Special attention must also be given to identify overlaps with other domain ontologies, and if possible a set of mappings to those ontologies need to be included.
3	**Balance of theory and practice** Adopt a hybrid approach for the definitions of terms in the domain ontology by making a balance between utility and deep ontological (philosophical grounding) analyses on the conceptualisations and formalisations.
4	**FAIRNESS** Make domain ontologies FAIR by storing the ontology in a permanent ontology repository specific to the industry (industryportal), adopting FAIR metadata for annotation and documentation. At the same time, the current proposals for FAIR metadata require enhancement to support domain ontology alignment and FAIRification at the content level (classes and relationships).
5	**Adopt existing standards** While building an ontology for a certain domain area, existing standards covering that topic need to be identified and ontologised as much as possible. As the nomenclature of these standards is already well accepted in the community, they need to be directly adopted in the ontology.
6	**Domain classification** Classify all existing, under development, and future ontologies (domain level) as per their target domain. For this purpose, a standardised domain classification needs to be globally implemented. At the same time, a gold standard vocabulary (benchmark) needs to be

	established for each domain so that existing and new ontologies can be evaluated for coverage for that domain.
7	**Appropriate training** Bridge the gap between domain experts and ontologists by supporting educational, training and professional development needs and in particular supporting a 'Translator' role, able to bridge gaps in the stakeholder value chain from ontology design to exploitation for data documentation.

3.2.6 References

Ref 3.2.1 OntoCommons (2021). Domain Ontologies for Research Data Management in Industry Commons of Materials and Manufacturing (DORIC-MM 2021). Online Focused Workshop, 15.03.2021. https://ontocommons.eu/news-events/events/ontocommons-kick-doric-mm-2021-0

Ref 3.2.2 https://www.colan.org/

Ref 3.2.3 Martin G. Skjæveland, Laura Ann Slaughter, & Christian Kindermann. (2022). OntoCommons - Report on Landscape Analysis of Ontology Engineering Tools. Zenodo. https://doi.org/10.5281/zenodo.6504670

4. Integrated Development Environment

4.1 Ontology Commons EcoSystem Toolkit

The Ontology Commons EcoSystem (OCES) Toolkit is intended to provide a specification for both methodology and tool support for industry-oriented ontology development, maintenance, and usage. To this end, existing methodologies and tools for supporting the life cycle of ontologies are firstly reviewed with experts/practitioners from industry and research. As a result, industrial needs and associated gaps/challenges are identified. Later, what is termed the OCES Toolkit – that consists of the most state-of-the-art methodologies, references and specifications for tools, and guidelines for implementation – is framed.

To prepare this section of the Roadmap, we have engaged with relevant stakeholders, such as ontology experts, practitioners, and users, as well as the developers of tools to understand their expectations for ontology engineering and use.

4.1.1 Industrial Need

Consulting ontology practitioners, tool developers, ontology users and tool users has given us a better understanding of how the tool ecosystem should be used in terms of industrial needs, and later what gaps they see a required ecosystem addressing for the different stages of ontology development and usage.

Therefore, in consultation, we have obtained the following industrial needs for the tool ecosystem:

NEED #	NEED DESCRIPTION
1	**Improved delivery of various functional services** The functional services need is that a tool ecosystem should support all the phases and common activities in the process of ontology engineering and use. To be specific, it should integrate a set of components that can perform specific tasks in ontology engineering and use. These functional requirements need to be further elaborated in detail.
2	**Supporting non-functional services** First, the architecture should ensure flexibility, openness, and ability to integrate with other tools. Second, an interoperable interface is crucial for improving user experience and allowing domain experts to participate in the development of ontologies. A third need is that the tool ecosystem should be collaborative, meaning that multiple users can contribute to each phase of the process of ontology engineering and use. Last but not least, the tool ecosystem needs to consider the FAIR principles (Findable, Accessible, Interoperable, and Reusable) for ontology publication.

This needs gathering has been carried out at the Focused Workshop (OntoCommons Workshop on Tools for Ontology Engineering) [Ref 4.1.1] and the First Global Workshop [Ref 4.1.2], both in 2021. The Focused Workshop organised in March collected feedback from ontology practitioners, tool developers and users of ontologies about the current state of ontology engineering tool. The Global Workshop organised in November 2021 included a session on "User Experience on Ontology Engineering Tools" and a session on "Ontology Adoption". These sessions aim to identify ways to better engage users with ontology engineering tools and ontologies. Ultimately, the development of an OCES Toolkit will be based on input and learnings from experts, including the requirements of demonstrators to identify the components for the OCES Toolkit, along with other (mostly non-functional) requirements.

4.1.2 State of the Art

An in-depth review of existing ontology engineering methodologies and tools is essential as the OCES Toolkit seeks to support industry-targeted ontology engineering and use. Therefore, this section will briefly discuss:

- the state of the art of ontology engineering and use processes, i.e., methodology, and
- the tools available for supporting the process.

With regards to methodologies, we distinguish between ontology engineering (i.e., creating and maintaining ontologies) and ontology use. In other words, ontology engineering activities are primarily concerned with the ontology development process. Moreover, ontologies as an engineering artefact and an end product can be utilised for a variety of purposes, which is known as ontology use and exploitation. To identify and understand the existing methodologies, our review is organised into two parts:

- The first part covers the current ontology engineering methodologies;
- The second part sums up the most important uses of ontologies.

Apart from methodologies, we also surveyed the existing tools for ontology engineering and use. This section will also provide a summary of the state of the art of ontology engineering tools based on the results from a tool survey and the key learnings from the workshops.

4.1.2.1 Existing methodologies for ontology engineering

There exists many ontology engineering methodologies. Here, we briefly review five primary traditional ontology engineering methods as methodological references, namely:

- Grüninger and Fox (Grüninger and Fox, 1995), [Ref 4.1.3]
- METHONTOLOGY (Fernández-López *et al.*, 1997), [Ref 4.1.4]
- On-To-Knowledge (Sure et al., 2004), [Ref 4.1.5]
- DILIGENT (Pinto et al., 2009), [Ref 4.1.6] and
- NeOn (Suárez-Figueroa et al., 2012). [Ref 4.1.7]

On the basis of these existing five methodologies, we analysed the common workflow or common steps for ontology development. In addition, the OCES Toolkit reuses many of the core components. The following is a brief summary of the results of the analysis.

Initially, Grüninger and Fox [Ref 4.1.3] proposed a procedure for ontology design and evaluation in 1995. As part of the methodology, requirements in the form of questions that the ontology should answer, i.e., its competency, are defined. A second step involves defining the ontology's terminology. The third step involves defining the ontology's definitions and constraints. Finally, the competency of the ontology is tested by proving completeness theorems for its competency questions.

A second existing methodology is METHONTOLOGY [Ref 4.1.4], which describes a set of life cycle models and a development process to provide an overview of how an ontology should be developed. There were three life cycle models that were proposed: a waterfall model, an incremental model (which ensures compatibility between all successive versions) and an evolved prototype model (similar to agile software development).

The On-To-Knowledge methodology [Ref 4.1.5] focuses on the application-oriented development of ontologies. It includes five phases, namely: (a) feasibility study, (b) ontology kick-off, (c) refinement, (d) evaluation, and (e) maintenance. As part of the ontology kick-off phase, ontology engineers should collect user requirements as competency questions to provide an overview of possible queries to the system that can indicate the scope and content of the domain ontology. In the evaluation phase, ontology engineers verify that the target ontology complies with the ontology requirements specification document and that it supports or addresses the competency questions analysed in the kick-off phase of the project. The evaluation phase often results in new requirements that need to be addressed by the ontology.

According to Pinto *et al.*, [Ref 4.1.6], the DILIGENT methodology supported ontology development in a distributed environment. As a result, the actors involved in the development of the same ontology have different complementary skills, including ontology users and ontology developers. This methodology proposes five general activities, including (1) build, (2) local adaptation, (3) analysis, (4) revision, and (5) local update.

The NeON methodology [Ref 4.1.7] supports the collaborative development of ontologies and offers concrete guidelines for reusing and reengineering knowledge sources. For this purpose, this methodology identifies nine scenarios for creating ontology networks, which are the most common ones during ontology network development.

Currently, lightweight and agile methodologies are becoming more prevalent in ontology engineering (e.g., Peroni, 2016 [Ref 4.1.8]; Hristozova and Sterling, 2002 [Ref 4.1.9]; Presutti et al., 2009 [Ref 4.1.10]). These methodologies represent the main workflow in sprints and iterations. The agile methodology also emphasised better communication between ontology developers and domain experts, as well as continuous assessment of the project status and the ability to respond to changes quickly. (It was also noted in the Ontology Adoption session of the Global Workshop that domain experts sometimes struggle with top-level ontology definitions/perspectives, and external advice may be required in such cases.)

4.1.2.2 Linked Open Terms: A methodology for ontology engineering in OCES Toolkit

Linked Open Terms (LOT) methodology combines the previously mentioned existing methodologies, current trends in ontology engineering, and new needs arising from the Linked Data paradigm. Methodologies prior to this did not consider nor orient the development of ontologies to their online publication. As a result of analysing the existing ontology engineering methodologies, the OCES

Toolkit adopts and adapts when needed the LOT methodology, which divides the common activities involved in the ontology development process into four phases:

- ontology requirements specification,
- ontology implementation,
- ontology publication, and
- ontology maintenance.

These phases along with ontology use completes the OCES methodology framework. Next, we focus on the usage of ontologies.

4.1.2.3 Main usage of ontologies

A common use of ontologies is to transform data from an unstructured format into a structured format. Shapes can also be generated using ontologies. Data shapes define a set of restrictions that data must meet, which can be related to the data model (e.g., cardinality) and to the values of the data (e.g., string patterns). Additionally, APIs can be generated from ontologies to make it easy for developers to consume data.

4.1.2.4 Existing tools for ontology engineering and use

According to the results of the tool survey and workshops, we conducted a landscape analysis on the existing tools for ontology engineering and use. The most familiar and frequently used tool for authoring and editing ontologies is Protégé [Ref 4.1.11]. Tools such as GraphDB [Ref 4.1.12], Apache Jena [Ref 4.1.13], OWL API [Ref 4.1.14], Virtuoso [Ref 4.1.15], and rdf4j [Ref 4.1.16], and RDFLib [Ref 4.1.17] are available for storing and querying ontologies. Ontology drafting, concept identification, and requirement elicitation consists of only a few tools, namely Magic Draw Cameo [Ref 4.1.18] and OWBO - Ontology White Board [Ref 4.1.19]. For ontology validation, OOPS! [Ref 4.1.20] is commonly used. Among the results of our survey, WIDOCO [Ref 4.1.21] was the preferred tool for publication and deployment, however it is interesting to see that GitHub [Ref 4.1.22], a general issue tracking and version control system, is used by many practitioners. Except for Linked Open Vocabularies [Ref 4.1.23], very few ontology repositories exist for publication and reuse. Common reasoning engines include HermiT OWL Reasoner [Ref 4.1.24] and Pellet Reasoner [Ref 4.1.25]. Tools like WebVOWL [Ref 4.1.26] are sufficiently known by the community for visualisation and visual editing. OntoPortal [Ref 4.1.27] and Ontology Lookup Service [Ref 4.1.28] emerged as useful terminology services.

4.1.3 Gaps

From the feedback from the survey, workshops [Ref 4.1.29], [Ref 4.1.30], and the review of the state of the art, we found the following gaps in the creating an methodological and tool support ecosystem for ontology engineering and use.

Considering the needs and the state of the art, we compiled the following gaps for the OCES Toolkit:

GAP #	GAP DESCRIPTION
1	**Gap in ontology validation** A typical ontology engineering process does not include the generation of linked data, APIs or SHACL [Ref 4.1.31] shapes for data validation. Yet, it is often necessary to evaluate different aspects of the obtained ontological data format, for example, checking for literals that do not fit in the data type ranges, determining whether the data set contains redundant objects, determining whether the data set uses existing established ontologies to represent its entities, or determining whether the data set provides possibilities for obtaining the desired information (for example, in terms of SPARQL queries). Despite the fact that tools based on SHACL are well established, ontology validation remains a challenge.
2	**Limited tool support** From the survey, we identified a few issues with existing tool support. Currently, domain experts from industrial domains find it challenging to understand and work with existing tools. In addition, there is limited tool support for concept identification, constraint specification, test specification, visual drafting, and navigating ontologies efficiently. Only a small number of tools are available for ontology visualisation.
3	**In need of integration among existing tools or a tool chain** Many criticisms pointed out that the tools were not sufficiently integrated with each other, offered poor user experience, and did not enable enough expressiveness. There is also a lack of integration among ontology engineering tools. Ontology engineering requires better support for collaborative development and validation, in particular, in the context of collaborative work with domain experts.
4	**Unsatisfactory and unsustainable solutions** The usability and user experience offered by ontology engineering tools are not satisfactory. Ontology engineering tools also presented some cost-related challenges. These tools are not sustainable or robust.

4.1.4 Definition of Success

As a result of the state-of-the-art review and learnings from the survey and workshops, a set of components for the OCES Toolkit was developed as follows:

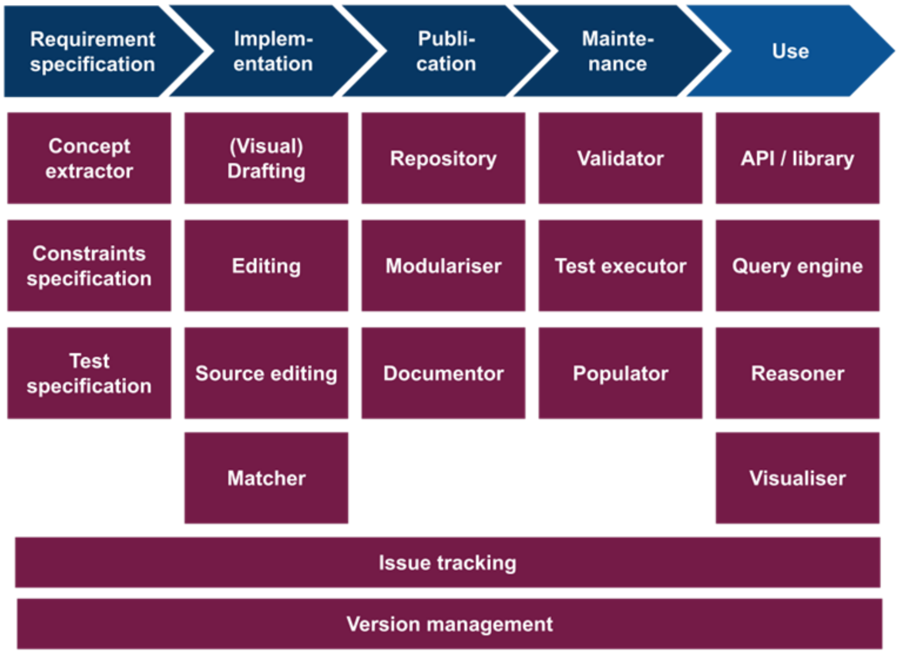

Figure 6 - Components of the ontology ecosystem toolkit

A critical factor for the OCES Toolkit becoming a successful methodology and tool support is that its components should fulfil the functional requirements for ontology engineering and use, and meet the non-functional industrial needs. As shown in Figure 6, the components of the OCES Toolkit are organised into the four phases of the LOT methodology, plus the additional activity of using ontologies. It should be noted that the OCES Toolkit components could play multiple roles throughout the process of engineering and implementing ontologies. Hence, a component is placed under each phase in order to structure the figure based on the phase to which it contributes most. As such, every component is regarded as a service, i.e., a piece of software or broader tool that performs a specific function within the process of ontology engineering and use. Therefore, the structure of the OCES Toolkit ensures the services can interact with one another, and some dependencies may exist.

Another success criterion is whether the OCES Toolkit can support the entire lifecycle of ontologies. In fact, it has been designed with the following methods and tools in mind. Taking into account METHONTOLOGY, both technical and documentation activities are addressed in the OCES Toolkit. Based on On-To-Knowledge, it includes refinement cycles in different phases. The Toolkit contains technical considerations regarding the distribution of ontologies. When appropriate, it also considers the reuse of ontologies and non-ontological resources utilising the NeOn methodology.

Looking at the success of this Toolkit from a practical point of view, using the OCES Toolkit, the experts should be able to create and document ontologies based on recommendations tailored to meet the present challenges. They can benefit from guidelines and best practice derived from existing ontology development methodologies and from success stories of tools that have been adopted. They can also benefit from the reference implementation, which supports the development phases of the methodology and realises its guidelines. The methodological framework and its references will be incorporated into a knowledge graph that will be both machine-readable and human-readable, which will be valuable to both experts and users.

In a coordinated effort, the OCES Toolkit is expected to significantly broaden the core community of industrial researchers and developers using ontology reference documentation, contribute to optimisation, evolve methodologies for manufacturing and services, and create methods for innovation within existing industrial systems.

With regard to industrial domains, the OCES Toolkit will be a catalyst for overcoming scepticism in many circles by creating an ecosystem within which industries can benefit from a turnkey solution with a set of good practice recommendations that will support them in integrating ontologies successfully and effectively into their operations.

4.1.5 Recommended Action

To avoid uncertainty around data formats and access, as well as to address any known limitations with tools, their availability/functionality and integration, the future work should enable the following:

ACTION #	ACTION DESCRIPTION
1	**Refine methodologies and tools especially for ontology validation** Background: Shneiderman (1996) [Ref 4.1.32], in a seminal human-computer interface research paper, identified seven tasks that should be supported in order to improve modern ontology engineering environments. In addition, there are issues around a lack of some tools or at least of sufficiently advanced tools (e.g., in terms of specification, navigation visualisation/visual editing, etc.) for carrying out certain tasks, along with the problems of tools not being sufficiently well integrated together. Action 1: Taking these tasks into account, as well as considering the industrial needs, the current state of ontology engineering ecosystem, and the gaps, the future work should provide recommendations on principles, best practice, and methods for the creation and maintenance of ontologies. Action 2: Better coordination amongst tool providers (with input from tool users) is required to identify those missing aspects/integrations of tools in the components (and phases) of the OCES Toolkit. This is key given the aforementioned gaps around data and ontology validation, evaluation and querying.
2	**Leverage industrial ontology portal and ontology adoption** Background: In the Focused Workshop [Ref 4.1.1] as well as the First Global Workshop [Ref 4.1.2], external expert Clement Jonquet presented best practices for FAIR adoption in portals similar to that being deployed by the OntoCommons project for the materials and industry portals. This also reduces the issues around data/ontology validation, querying, etc. mentioned previously. In addition, the best way to influence industry's willingness to adopt ontologies is to understand what kinds of operations it needs to run

	with the data schema: the better it understands this, the better it can build an ontology that is useful. As the term "ontology" is still used in different ways, e.g., knowledge structures that range from terminology lists, to thesauri, to taxonomies, and to ontologies, the key to successful ontology adoption is to identify a correct purpose and use cases by using the OCES Toolkit for picking the most appropriate technical platform or system. Action 1: Establish and promote Industrial Ontology Portal in order to make ontology adoption easier and meaningful and integrate it in the ecosystem (especially the FAIR principles, recommendations regarding ontology languages, expressivity levels, and so forth). Action 2: Define a clear Roadmap during the first steps of developing an ontology adoption strategy, i.e., how will the ontology be used, and partner with important stakeholders to demonstrate how ontology adoption works.
3	**User-friendly tool chain and reference implementation** Background: At the First Global Workshop [Ref 4.1.2], external expert Gianmaria Bullegas described the fact that he eventually had to build his own tools to develop/maintain ontologies for the aircraft domain. He also highlighted the Common Core Ontology [Ref 4.1.33] for unified semantics in an enterprise, along with prior work from the Industrial Ontologies Foundry [Ref 4.1.34]. Therefore, an advanced tool chain must facilitate the development and use of ontologies by coordinating actions on tools, guided by stakeholder input and demonstrator requirements. Tools should be better integrated and more usable (with a better user experience). Action: A methodological framework and reference implementation toolkit should be developed that offers a practical and user-friendly method for re-using data across domains and industries.

4.1.6 References

Ref 4.1.1 OntoCommons (2021). OntoCommons Workshop on Tools for Ontology Engineering. Online Focused Workshop, 19.03.2021. https://ontocommons.eu/news-events/events/ontocommons-workshop-tools-ontology-engineering

Ref 4.1.2 OntoCommons (2021). Global Workshop: Ontology Commons addressing challenges of the Industry 5.0 transition. Online Global Workshop, 02.-05.11.2021. https://ontocommons.eu/news-events/events/global-workshop-ontology-commons-addressing-challenges-industry-50-transition

Ref 4.1.3 Gruninger M. and Fox M.S. (1995). Methodology for the design and evaluation of ontologies. In: Proceedings of the Workshop on Basic Ontological Issues in Knowledge Sharing, IJCAI-95, Montreal.

Ref 4.1.4 Fernández-López M., Gómez-Pérez A., and Juristo N. (1997). METHONTOLOGY: From Ontological Art Towards Ontological Engineering. In: Proceedings of the AAAI-97 Spring Symposium Series, AAAI-97, 24-26 March 1997, Stanford University.

Ref 4.1.5 Sure Y., Staab S., and Studer R. (2004). On-To-Knowledge Methodology (OTKM). In: Staab S., Studer R. (eds) Handbook on Ontologies. International Handbooks on Information Systems. Springer, Berlin, Heidelberg.

Ref 4.1.6 Pinto H. S., Tempich C., and Staab S. (2009). Ontology engineering and evolution in a distributed world using DILIGENT. In Handbook on ontologies (pp. 153-176). Springer, Berlin, Heidelberg.

Ref 4.1.7 Suárez-Figueroa M.C., Gómez-Pérez A., and Fernández-López M. (2012). The NeOn Methodology for Ontology Engineering. In: Suárez-Figueroa M., Gómez-Pérez A., Motta E., Gangemi A. (eds) Ontology Engineering in a Networked World. Springer, Berlin, Heidelberg.

Ref 4.1.8 Peroni S. (2016). A simplified agile methodology for ontology development. In OWL: Experiences and Directions–Reasoner Evaluation (pp. 55-69). Springer, Cham.

Ref 4.1.9 Hristozova M. and Sterling L. (2002). An eXtreme method for developing lightweight ontologies. In: Proceedings of 1st International Joint Conference on Autonomous Agents and Multi-Agent Systems.

Ref 4.1.10 Presutti V., Daga E., Gangemi A., and Blomqvist, E. (2009). eXtreme Design with Content Ontology Design Patterns. In: Proceedings of the 2009 International Conference on Ontology Patterns, Washington DC.

Ref 4.1.11 https://protege.stanford.edu/

Ref 4.1.12 https://graphdb.ontotext.com/

Ref 4.1.13 https://jena.apache.org/

Ref 4.1.14 http://owlcs.github.io/owlapi/

Ref 4.1.15 https://virtuoso.openlinksw.com/

Ref 4.1.16 https://rdf4j.org/

Ref 4.1.17 https://rdflib.readthedocs.io/en/stable/

Ref 4.1.18 https://www.3ds.com/products-services/catia/products/no-magic/cameo-systems-modeler/

Ref 4.1.19 https://github.com/mdaquin/OWBO

Ref 4.1.20 http://oops.linkeddata.es/

Ref 4.1.21 http://dgarijo.github.io/Widoco/doc/tutorial/

Ref 4.1.22 https://github.com/

Ref 4.1.23 https://lov.linkeddata.es/

Ref 4.1.24 http://www.hermit-reasoner.com/

Ref 4.1.25 http://pellet.owldl.com/

Ref 4.1.26 http://vowl.visualdataweb.org/webvowl.html

Ref 4.1.27 https://ontoportal.org/

Ref 4.1.28 https://www.ebi.ac.uk/ols/index

Ref 4.1.29 OntoCommons (2021). User Experience on Ontology Engineering Tools. In: OntoCommons. Global Workshop: Ontology Commons addressing challenges of the Industry 5.0 transition. Online Global Workshop, 02.-05.11.2021. https://www.youtube.com/watch?v=w-8DBo35eKk&list=PL-cwgiwYXckPOatW5dGBK_wDr0ZoP72L3&index=3

Ref 4.1.30 OntoCommons (2021). Ontology Adoption. In: OntoCommons. Global Workshop: Ontology Commons addressing challenges of the Industry 5.0 transition. Online Global Workshop, 02.-05.11.2021. https://www.youtube.com/watch?v=mLWoFK6QDF4&list=PL-cwgiwYXckPOatW5dGBK_wDr0ZoP72L3&index=6

Ref 4.1.31 https://www.w3.org/TR/shacl/

Ref 4.1.32 Shneiderman B. (1996) The eyes have it: a task by data type taxonomy for information visualisations, In: Proceedings of the 1996 IEEE Symposium on Visual Languages, pp. 336-343.

Ref 4.1.33 https://www.cubrc.org/index.php/data-science-and-information-fusion/ontology

Ref 4.1.34 https://www.industrialontologies.org/

4.2 Infrastructure

In general, "infrastructure" is defined as basic and necessary systems and services facilitating the effective operation of countries, organisations, or companies (see for example [Ref 4.2.1]). Researchers have tried over the years to outline what makes an infrastructure in, e.g., the technological, societal, or material domains of our society (see [Ref 4.2.2], [Ref 4.2.3] and [Ref 4.2.4]). In the context of this Roadmap document, we limit the term "Infrastructure" to Research Infrastructures as defined by art. 2 of the EU Regulation 2021/695 [Ref 4.2.9] that establishes Horizon Europe. In this respect, infrastructures provide resources and services for communities to efficiently conduct their research and innovation development. We can further define the term "infrastructure", then, as a concept comprising of everything from hardware and physical networks, via a software stack, services and API definitions, to organisational aspects such as rules of participation, financial regulations, up to actual human resources to operate and maintain the infrastructure, with the additional requirement of long-term stability [Ref 4.2.5].

As such, this section covers mostly the aspects of actually making instances of the components listed in the preceding sections of this document available (or rather: providing an environment where these components can be made available), and serves as the basis for generating the impact addressed in the subsequent section.

In the domains addressed by OntoCommons, there is a clear need for an infrastructure on which to base the functionality of the ecosystem of ontology engineering tools, for (standardised) data documentation tools and processes, to support secure communication between stakeholders that collaborate, for secure data sharing or integration, etc. The long-term vision is to reach a level where using the ontologies, data documentations, etc., combined with intelligent data infrastructures (e.g., provisioning, exploring, transforming) will allow us to expose any data as semantically enriched information (e.g., Knowledge Graphs), allowing for a transparency layer that enables users and down the line (semi-)automatic processes to seamlessly collaborate and run experiments that may involve several data sources and several computation modules. This vision presumes, firstly, that low-level data and metadata representations are available, that workflows for data processing, integration, documentation – as described by the FAIR principles – are well described, and that basic services for managing this kind of information are available. Secondly, a layer of tools and services for data provisioning (e.g., standardisation, FAIRification, etc.) will be available such that computing infrastructures or Virtual Research Environments can be built on top of it, and including computational facilities as and where necessary. In our vision, data markets and tool spaces interoperate seamlessly to a user.

4.2.1 Industrial Need

As observed in the interactions between the OntoCommons project partners, internally and externally, either through focused expert workshops or global workshops (November 2021), to develop and test their products, industry stakeholders are doing complex computations that necessitate integration of data coming from conceptually different systems. Furthermore, complex simulations and models must be run in a timely manner, with requirements for accountability and quality of results.

Abstracting out from the industrial needs identified in sections 3 and 5 of this document, we pinpoint the following infrastructure related needs, which are closely related to those described in Section 4.1:

Need #	Need Description
1	**Secure collaborative tools for multiple stakeholders (inter-organisation)** • Need for secure interaction, exchange of data through common standards and data documentation, and tools for understanding (see section 5.2). • Data collection and cataloguing following data documentation and description standards, that are adopted by the industry. • Availability of a digital market (see section 5.1, 5.4).
2	**Interactive visualisation of data** • As an integral part of data analysis and understanding dataset characteristics, the choice of appropriate tools for aggregated, visual display are paramount for understanding the outputs of processes or grasping the characteristics of data that is to be used in other processes. • Identifying trends in the various industrial domains, grasping the (ontological) structure of the data as well as its content, so that it is easily transmitted with suitable visualisation methods.

	• These tools must allow for interaction with the data.
3	**Data quality assurance and analytics, data validation** • Data that is to be published must comply with commonly agreed on quality norms. When necessary, automatic processing of the data to detect points of non-compliance must be possible. Similarly, metrics for interoperability, re-usability and data documentation must be part of the quality assessment. FAIRification of the data will contribute to high levels of data-sharing between stakeholders, also ensuring its traceability through the use of dedicated ontologies.
4	**Trustworthy data repositories and trusted computation** • Access to data and computation resources must always be done in a secure way. Computation on data must be performed such that no third party has access neither the computation's inputs nor outputs. Ultimately, data residing on different trusted repositories and platforms must be brought together by means of shared metadata.

4.2.2 State of the Art

Research infrastructures that are dedicated to specific domains (renewable energies, neurosciences, scholarly communication, to name a few) have been developed over the last number of decades, not only in Europe but also worldwide ([Ref 4.2.6] [Ref 4.2.7] [Ref 4.2.8]). The European Union, through its Research Framework Programmes, has made important contributions to the development of research infrastructures. The European Strategy Forum on Research Infrastructure (ESFRI) provides a pan-European coherent approach to policy making and investment on Research Infrastructure (see EU Regulation 2021/695) [Ref 4.2.9].

The set of tools and services, for every existing infrastructure, are specific to the domain of research that is the subject of the infrastructure. A common thread in managing infrastructures is the existence of services for data provisioning, semantic enrichment of data and services, etc., even when this is not explicitly stated in the infrastructure documentation, such as in [Ref 4.2.10] or [Ref 4.2.11].

In the material modelling domain, there is a large number of individual building blocks created for specific tasks, with high technical readiness levels (TRLs). The individual building blocks have been identified and landscaped during focused meetings and workshops with OntoCommons stakeholders, and we will not reiterate them here; instead see [Ref 4.2.12] ,[Ref 4.2.13] and [Ref 4.2.14].

4.2.3 Gaps

As previously stated, many building blocks of the envisioned infrastructure exist, with various examples on all layers. These cover highly advanced concepts for ontology mapping and evaluation, with several powerful examples demonstrating the techniques' potential. Highly qualified experts are driving research and development, showing the potential of semantic technologies and their scalability across a range of disciplines. However, these activities are usually fragmented, serving isolated communities of experts familiar with their respective infrastructures, and are specifically lacking in sometimes trivial components (and thus not worthy of research funding) that would, however, enable a massive boost in the wide-scale deployment and take-up of (ontological) services across disciplines, and specifically for interdisciplinary work.

These components are also made to serve different parts of the research life cycle, ranging from data management via service discovery and the actual research and development being performed, to the publication, meta-analysis and re-use of results. The requirement of FAIRness, with the specific focus on machine-actionability for virtually all types of content and all layers of a society evolving in a world undergoing digital transformation, puts trustworthy services based on semantic technologies at the core of all such endeavours.

While a multitude of gaps at different levels of granularity exist, as already identified in the other chapters of this document, the following items are representative of the core components that need to be put in place in order to support the vision and to realise the potential envisaged.

Gap #	Gap Description
1	**Low maturity level of available infrastructure components** The components identified in the state-of-the-art sections in this document are usually operated as prototypes rather than as professionally deployed services, or operated in isolated environments, with interoperability and/or sharing of data being only possible with considerable additional effort.
2	**Missing fundamental low-level ontologies** A lot of work is dedicated to developing mid- and high-level ontologies that map between different spaces. While these are absolutely essential and addressed in-depth in Sec. 3.1 and 3.2, there is a lack of trusted, quality-assured and professionally maintained low-level ontologies capturing trivial but essential building blocks of knowledge such as measurement units, chemical elements etc. Many such ontologies exist, covering different sub-parts of such spaces (e.g., the range of ontologies covering various types of measurement units), but none of them covers the entire space that can be derived from the International System of Units (SI) and the International System of Quantities (ISQ), and alternative unit systems as well as the according transformations. However, such low-level ontologies are central to developing standards for the interoperability of data as described by higher-level ontologies. In several cases, (non-semantically structured) code is widely used to robustly and efficiently compensate for the lack of low-level ontologies, e.g., code to read the multitude of formats that time/date can be expressed in. In this example, ontologies that capture the wide range of calendar systems, the variations in expressing dates within them, and corresponding transformations are not readily available. Creation of these is a non-trivial, but predominantly effort-intensive activity that does not qualify as research, and yet would constitute essential building blocks for myriads of applications, and thus forming a core "Commons" for society as well as for industry (in an Industry Commons).
3	**Ontology data provisioning** The provisioning of the above-mentioned (low-level) ontologies as well as the services, APIs, and tools needed for creating, maintaining, and interacting with these ontologies (as

	outlined in Sec. 4.1) is not largely available outside the environments these tools were developed in. This makes them less likely to be shared and re-used as trustworthy, stable environments that are properly maintained, quality-assured, and that have business and succession plans. Allowing others to rely and thus integrate such services deep into their own applications is not yet a common practice in the material modelling domain.
4	**Federated, interconnected virtual research environments** These run on massively scalable and continuously evolving hardware platforms, providing essential infrastructure services such as data management and curation, security and privacy, as well as an appropriate legal and financial framework that are not readily available to the OntoCommons community. These environments need to address the specific requirements of research that mostly overlap with industrial requirements, but sometimes at different levels of prioritisation or scale, such as e.g., reproducibility and traceability of results, minimisation of lock-in risks associated with any specific standard or technology, and full transparency on the hardware, software, legal and other resource levels.
5	**Few pilot end-to-end application demonstrators (cross-walks)** There is a lack of end-to-end demonstrators (cross-walks) for reasonably complex settings across a representative range of domains that, combined with sandboxes, support exploration, adoption, testing service quality and integration. Services that apply the concepts underlying semantic technologies to the actual system, encompassing the entire stack from hardware, software, data, as well as the associated legal, financial, and human resources, are not readily available. Such services need to be themselves FAIR, i.e., exposed in an ontologically structured form, relying on the very concepts and standards established for the operation of the infrastructure, forming one massive Knowledge Graph that seamlessly merges the metadata about the infrastructure with the metadata and data held within the infrastructure and all associated services.
6	**No well-defined (and semantically described) pipelines** Pipelines with sufficient maturity and quality of service, to continuously develop, test, roll-out (eventually retire) and evaluate novel services, to allow an infrastructure to evolve and evaluate novel solutions on a continuous basis, do not exist. Such pipelines play an important role in defining a routine process for infrastructure evolution with the agility required to more easily adopt new trends and more quickly deploy advanced services.
7	**Insufficient support for transfer from R&D department and funding stream activities into infrastructure operations** Advanced concepts, tools and processes are usually being developed by R&D departments, funded via time-limited research grants, maturing to a certain TRL. Irrespective of the TRL achieved, it is usually other units within an organisation that need to operate any such tool or service. These need to on-board a service, integrate it into their daily operations/routines, and also need to ensure cost coverage over prolonged periods of time. Especially in these times of digital transformation, each additional service deployed will cause significant additive workload and responsibilities on the units responsible for continuous operations

	and the sustainability of such services. As these are different stakeholder groups with differing interests and operational incentives/pressures, specific types of projects and funding streams need to be devised that target these kinds of activities, with a highly skewed cost distribution, allowing for significant spending up-front to ensure take-over/on-boarding, while guaranteeing partial funding to cover operational costs for a prolonged period of time.

4.2.4 Definition of Success

From the many options that one could choose to define the success criteria for an infrastructure dedicated to the domains addressed by the OntoCommons project, we list the ones that contribute the most to addressing the industrial needs previously identified.

- **Number of quality-assured concept spaces**, represented exhaustively, by specific and integrated low-level ontologies, provided and maintained by trusted entities (repositories, organisations). These should be selected following a community-guided prioritisation level starting from e.g., measurement units and conversion rules to domain-specific but widely applicable base ontologies such as chemical element tables. The availability of such ontologies will contribute to a significant increase on the level of interoperability, as not-formalised implicit knowledge (e.g., a basic measurement ontology) is a source of errors.
- **Number of federated institutions providing the mentioned services:** A high number of institutions and organisations, with industrial stakeholders playing a lead role, that participate in providing services to the OntoCommons community, organisations that offer guarantees on the service level, financial sustainability and succession regulations, to ensure stability and resilience of the services thus provided as Commons.
- **Number of federated virtual (research) environments:** An important aspect of a successful infrastructure is the number of available virtual environments, with associated documentation of the ratio of users served by these infrastructures and key usage statistics, observing the trend of up-take of these infrastructures in comparison to dedicated, isolated non-federated specialised infrastructures.
- **Number and domain-coverage of end-to-end demonstrators**: Uptake of data documentation standards, metadata description methods, common ontologies is slow when there are too few cross-walks that have demonstrators, at different levels of complexity and across different domains. The rate of take-up and deployment of further pilot services based on the inspiration provided by an initial set of end-to-end demonstrators is expected to be high.
- Number of **services deployed at each maturity levels of the pipeline** including evaluation criteria and user feedback evaluation influencing the continuation, QoS-improvement or retirement of services.
- **Sufficient funding streams devised to support hand-over from R&D results into institutional operations**; number of services transferred to operations.

4.2.5 Recommended Action

The recommended actions in this section are part of a longer term vision that understands the life cycle of services to be deployed, as well as the need to complement R&D activities with purely development-oriented activities (and thus funding streams), that allow advanced services to find their way into trustworthy, sustainable, powerful infrastructures across all levels of infrastructure elements, ranging from hardware, via software services to human support as well as financial and legal regulations. These need to be operated in a redundant, and transparent manner with succession agreements and high quality-of-service levels. To achieve this, a tighter collaboration between the typical R&D departments as well as infrastructure provider/operations units needs to be encouraged and supported, addressing their distinct planning horizons and operational principles in terms of staffing, expertise and responsibilities.

Action #	Action Description
1	**Development of low-level ontologies** Fundamental low-level ontologies are a critical soft infrastructure component that will significantly contribute to the development of interoperability tools by eliminating the potential misinterpretation of implicit knowledge during knowledge engineering. These fundamental ontologies need to be developed and their take-up in knowledge engineering environments (current and future) must be strongly encouraged.
2	**Secure platforms for ontology data creation, provisioning, and exchange** Develop secure tools and software components that support coordination and collaborative work for knowledge engineering in the OntoCommons domains. The tools and processes must ensure that agreed-upon data documentation standards are followed, that any data that is engineered using these tools are compliant with the FAIR principles, that the information and data exchange is regulated.
3	**Virtual Research and Innovation Environments blueprint** Establish a Virtual Research and Innovation Environment *blueprint* for the domains covered by OntoCommons. The blueprint must describe the minimal hardware, software, and organisational requirements for an environment encompassing – among others – secure access to computing hardware, software tool availability and use, and secure data storage and provision.

4.2.6 References

Ref 4.2.1 See, for example, the Cambridge Dictionary entry for “Infrastructure”: https://dictionary.cambridge.org/dictionary/english-german/infrastructure

Ref 4.2.2 W. Buhr (2003) ‘What is infrastructure?’, Universität Siegen, Fakultät Wirtschaftswissenschaften, Wirtschaftsinformatik und Wirtschaftsrecht, 107–03.

Accessed: Feb. 27, 2022. [Online]. Available: https://ideas.repec.org/p/sie/siegen/107-03.html

Ref 4.2.3 L. Videka, J. A. Blackburn, and J. R. Moran (2008) 'Building Research Infrastructure in Schools of Social Work: A University Perspective', Social Work Research, vol. 32, no. 4, pp. 294–301, Dec. 2008, doi: 10.1093/swr/32.4.294.

Ref 4.2.4 R. Prud'homme (2004) 'Infrastructure and development', Paper Prepared for ABCDE (Annual Bank Conference on Development Economics): Washington DC, May 3-5, 2004.

Ref 4.2.5 https://ec.europa.eu/info/research-and-innovation/funding/funding-opportunities/funding-programmes-and-open-calls/horizon-europe/research-infrastructures_en (Accessed: Feb. 27, 2022.)

Ref 4.2.6 NSF 09-013, NSF-Supported Research Infrastructure: Enabling Discovery, Innovation and Learning. https://www.nsf.gov/news/nsf09013/index.jsp (Accessed 17.02.2022)

Ref 4.2.7 Rakeshnie Ramoutar-Prieschl, Sepo Hachigonta (2020), Management of Research Infrastructures: A South African Funding Perspective, Springer, Cham, https://doi.org/10.1007/978-3-030-37281-1

Ref 4.2.8 Mapping of Research Infrastructures in CELAC Countries https://celac.d2c2.gub.uy/en/home-page/ (Accessed 26.02.2022)

Ref 4.2.9 Regulation (EU) 2021/695 of the European Parliament and of the Council of 28 April 2021 establishing Horizon Europe – the Framework Programme for Research and Innovation, laying down its rules for participation and dissemination, and repealing Regulations (EU) No 1290/2013 and (EU) No 1291/2013 (Text with EEA relevance). (2021). *Official Journal*, L 170, 1-68. ELI: http://data.europa.eu/eli/reg/2021/695/oj[legislation]

Ref 4.2.10 P. Martin, B. Magagna, X. Liao, and Z. Zhao (2020), 'Semantic Linking of Research Infrastructure Metadata', in Towards Interoperable Research Infrastructures for Environmental and Earth Sciences: A Reference Model Guided Approach for Common Challenges, Z. Zhao and M. Hellström, Eds. Cham: Springer International Publishing, 2020, pp. 226–246. doi: 10.1007/978-3-030-52829-4_13.

Ref 4.2.11 Z. Zhao, K. Jeffery, M. Stocker, M. Atkinson, and A. Petzold (2020), 'Towards Operational Research Infrastructures with FAIR Data and Services', in Towards Interoperable Research Infrastructures for Environmental and Earth Sciences: A Reference Model Guided Approach for Common Challenges, Z. Zhao and M. Hellström, Eds. Cham: Springer International Publishing, 2020, pp. 360–372. doi: 10.1007/978-3-030-52829-4_20.

Ref 4.2.12 Mathieu d'Aquin. (2021). OntoCommons - Ontology Ecosystem Specification. Zenodo. https://doi.org/10.5281/zenodo.6504633

Ref 4.2.13 Martin G. Skjæveland, Laura Ann Slaughter, & Christian Kindermann. (2022). OntoCommons - Report on Landscape Analysis of Ontology Engineering Tools. Zenodo. https://doi.org/10.5281/zenodo.6504670

Ref 4.2.14 María Poveda-Villalón. (2022). OntoCommons - Report on OntoCommons ontology registry infrastructure. Zenodo. https://doi.org/10.5281/zenodo.6504709

5. Industrial Impact

5.1 Industrial Application

This section investigates the adoption of ontology technologies by practical applications. It determines the needs the industry has (both perceived by them and from the outside), bridging to relevant state of the art, and to the gaps, success definitions and recommended actions.

Among the needs that are determined from the outside are in particular, increasing the TRL of the ontology documentation adoption, the use of TLO, MLO, DO and tools, and the FAIRness level of the use case.

The OntoCommons project aims to evaluate its impact via a set of demonstrators. The project has started with 11 demonstrators [Ref 5.1.1] from NMBP domains. These demonstrators are

1. IRIS - IndustRIal codesign Support (Airbus Design and Manufacturing)
2. SeDIM - Semantic Data Integration for Manufacturing (BOSCH Manufacturing)
3. EngDemonstrator (Aibel Procurement)
4. Tribomat (Tekniker Materials' Tribological Characterisation)
5. EVMF - European Virtual Marketplace Framework (Digital Materials Marketplaces)
6. PSS - Product Service Systems (OAS Product Service System)
7. Feedstock Quality Assurance (Fraunhofer Quality)
8. NanoMaterials Characterisation (IRES Nano-Materials)
9. Ontology-based Maintenance (Adige BLM Group Ontology-based maintenance)
10. Cu/Al Data (Elvahalcor Metal Industry)
11. Complex Equipments (Siemens Digital Manufacturing)

The content of the following sections is based on various events and meetings in the form of workshops and interviews in which these 11 initial demonstrators were involved. These events are:

- **Focused Workshop "Demonstrators at work to deliver an Industry Commons Marketplace - Workshop on Demonstrators and Use Cases" [Ref 5.1.2]:** This focused demonstrator workshop enabled the OntoCommons demonstrators to present their use cases and initiate engagement with a larger audience. Harmonisation of requirements between initial use cases, ontologies and ontology tools was the main topic of the focused workshop. For this purpose, the requirements from the 11 initial use cases were collected and analysed concerning data interoperability and ontology use. Furthermore, the workshop was used to build a platform for project internal discussions between the use cases and the linked working groups on industrial domain ontologies in the OntoCommons EcoSystem. From these discussions in small groups consisting of two to four use cases and experts from the mentioned areas, various requirements could be collected, indicating what the industrial needs and hurdles are. A first overview on ontologies and development tools was provided and the public was informed about the initial use cases and their role in the project. In addition, networking of industrial use cases and wider stakeholders was supported.

- **First Global Workshop: Ontology Commons addressing challenges of the Industry 5.0 transition [Ref 5.1.3]:** The demonstrators were in the centre of a session in the OntoCommons global workshop, where developments from existing demonstrators were mixed in with the introduction of new demonstrators. Additionally, a comprehensive panel discussion regarding FAIRness was held.
- **One-to-one interviews with the demonstrators:** A series of one-to-one interviews were conducted with industry to understand their use cases in detail.

The events mentioned above provided a means for input to the remaining part of this section. The industrial needs focus on high-level industrial requirements. The state-of-the-art section provides a brief overview of how these needs are currently attempting to be addressed. Complementarily, gaps specify explicitly where the state of the art falls short in covering industrial needs. The last two sections provide a definition of success from the eyes of industrial stakeholders and the recommended actions.

The items in each section are grouped based on what they mainly relate to, namely people, data, processes and tools. There can be of course items belonging to multiple categories as well as interplay between different items. This will be investigated for further consolidation in future versions of the Roadmap.

5.1.1 Industrial Need

The following industrial needs manifested during our interaction with the initial demonstrators and external stakeholders:

NEED #	NEED DESCRIPTION
People	
1	**Ease of interoperability and communication between different stakeholders** The ontology development tools should allow different stakeholders to work simultaneously, and the ontologies should provide a “commons language” for this to happen.
2	**Best practices for data model governance as well as modelling tools** Industrial stakeholders need best practices about how to maintain data models and intuitive tool support. This is particularly important for bringing domain experts on board.
Data	
3	**Improved reusability of (meta-)data and processes** With little or no use of standard vocabularies and ontologies the reusability of (meta-)data is not very high.

4	**Easy to use and to understand** The industry needs ontologies that are easy to use and understand. They need to be applicable without much explanation. This points to the need for proper documentation and concrete examples of usage for ontologies.
Processes	
5	**Time savings in industrial processes** One of the main industrial needs is in saving resources, particularly time in industrial processes. Time savings is expected in terms of increased automation for tasks like decision making and interaction between different actors.
6	**Avoidance of physical testing** In many industrial processes, it is desired to avoid physical testing and create reliable simulations for resource and cost saving reasons. The need for simulation particularly manifests in manufacturing, for example in the aircraft industry.

5.1.2 State of the Art

Industrial applications currently address their needs (particularly in the scope of the needs mentioned above) to some extent already with state-of-the art ontolologies and tools. In the core of fulfilling many of these needs, ontologies play an important role. To that end, there is already a significant amount of (planned) ontology adoption, particularly of well-known ontologies like BFO, CheBI, DOLCE, EMMO, IOF-Core and SSN (see Sections 3.1 and 3.2). From the domain ontologies perspective, many stakeholders rely on in-house development. For the development of these ontologies, academically or industrially established methodologies are typically not used, however there is some usage of methodologies like the LOT methodology (see Section 4.1.2). From a tools perspective, we see that Protegé is the mostly adopted tool for ontology development. It is usually combined with reasoners like Hermit++ and Pellet. A wide variety of triplestores are used to store semantically described data, such as Stardog and Virtuoso, as well as property graph databases like Neo4j, and virtualisation solutions like OnTop on top of relational databases.

Several community efforts and W3C recommendations are also adopted as declarative languages. RML is adopted for mapping heterogeneous sources to RDF, and SHACL is used for defining data shapes for verification purposes.

Needs regarding interoperability and reusability are particularly related with FAIR principles. There is already a certain level of implementation of FAIR, however it was also observed that the adoption has reached a certain limit that prevents further adoption. Mostly, data privacy and proprietary data issues are cited as the reason for this situation.

5.1.3 Gaps

Considering the needs and the state of the art, the community proposed the following gaps for industrial applications:

GAP #	GAPS
People	
1	**High cost of ontology development** Related to the end, ontology development incurs high costs due to high learning barriers for non-ontologists. This gap hinders the fulfilment of reusability of data, metadata and processes.
2	**Learning barriers for semantic technology in the industry** The ontology development and its support tools should be made more intuitive for easy introduction of semantic technology in industry. *This is particularly important for bringing non-ontology experts on board.*
3	**Ontologies are difficult to maintain** The ontology shall be easy to maintain (e.g., adding lower-level terms, additional relations, etc.) from non-ontology experts (e.g., software engineers).
4	**Company internal/partner interaction should be optimised** Currently a major barrier across many industrial parties is being able to speak a common language during the development of industrial processes. Tooling frameworks and methodologies are not mature for enabling such communication (e.g., between domain experts and ontology developers).
Data	
5	**Lack of comprehensive domain ontologies in the NMBP domains** There are many domain ontologies scattered around, however there are not many reference domain ontologies that cover a large portion of their domain (NMBP) and that contain canonicalised definitions of concepts and their relationships.
6	**The ontologies are not well documented** The ontology documentation should define how the reuse and harmonisation of different ontologies could be achieved. This also includes the formal documentation of the ontology where the formal constraints and scope are clear.
7	**Arguments for using FAIR principles** It is not always clear for industrial stakeholders what are the concrete benefits arising from the application of FAIR principles. This contributes to natural barriers occurring ahead of further FAIR adoption.

8	**Dealing with content protected with IPR** Many industrial standards are protected with licenses that prevent publishing derivations of the work. This hinders the creation of semantic resources from those standards. From a data perspective, this also creates a hindrance for FAIR adoption.
9	**Interoperability between TLOs** There should be interoperability between TLOs to facilitate harmonisation of ontologies, allowing for interoperability among ontologies that are based on different top-level ontologies.
10	**The ontologies should follow higher-level ontologies** The aligned ontologies should follow top- or mid-level ontologies to allow a higher compatibility with other ontologies.
Processes	
11	**Lack of standards and guidelines** Although ontology usage is there to some extent, there are still challenges in terms of heterogeneity of ontologies and a lack of standards for alignment, as well as documentation. There is also a lack of a comprehensible methodology.
Tools	
12	**User interface** There are already tools like Protegé used for ontology development, however the user interfaces can be somewhat incomprehensible, particularly for non-ontology experts.
13	**Maturity of the (collaborative) ontology development tools** Many of the ontology development tools are not always intuitive and easy to use. One needs to already have some experience with ontologies, their structure, and what are the possibilities available, in order to be able to use the existing development tools. Many of them also have serious drawbacks in terms of collaborative development.
14	**Tools for ontology engineering are not complete** Tools should support visualisation, debugging, validation, and search of existing ontologies and importing same. Tools should be provided to support initial brainstorming and conceptualisation around models of concepts relevant for the domain and applications, to enhance the transition from initial ideas to standards.

5.1.4 Definition of Success

The following are the definitions of success compiled from the demonstrators and other industrial stakeholders with regards to the adoption of semantic technologies, particularly ontologies. These are mainly a reflection of concrete KPIs from industrial parties.

#	Definition of Success
People	
1	**Improved communication within company personnel and with external partners** Using a "common language", i.e., ontology and vocabularies, the communication between stakeholders will improve. This can be also seen as a consequence of achieving standardised data documentation from people's perspective.
Data	
2	**Achieving standardised data documentation** Achieving standardised data documentation, typically via ontologies, is seen as a sign of success for many industrial stakeholders. Such data documentation increases Findability, Interoperability and Reusability of data within and across organisations for different projects, and allow companies to increase their innovation capacity.
Processes	
3	**Time and cost saving** An important factor for all industrial customers is time saving. Saving costs can also be important for customers, but saving time is more globally comprehensible (e.g., ontology-enabled automation, optimised communication, and increased reuse).
4	**Gaining competitive advantage for small and large companies** Small and large companies can benefit from the use of ontologies. Large companies can benefit because they repeat a process very often. Small companies can benefit from the time improvement because they are faster than the competition.
5	**Optimised product quality and environmental footprint** Many industrial stakeholders provided a KPI for improving product quality and reducing environmental footprint e.g., in terms of CO_2 emission.

5.1.5 Recommended Action

Action #	Recommended Action
People	
1	**Knowledge engineering education** A major gap on ontology development and usage is the high cost of and struggles with finding trained people. Training on ontology usage and development issues is an important point, to allow early education on ontologies. This education must be adaptive to the needs and competencies of various stakeholders.
2	**Networking** Networking events where people share their experience with ontology adoption in industrial settings may be beneficial for a large audience and increase engagement.
3	**Demonstrate examples on saving time and cost** Examples and success stories should be shown on the topic of time and savings to increase awareness of the benefits.
4	**Highlight advantages of ontology usage** Demonstrate what the use of ontologies can do. This can be done by establishing a translator role in companies (see Section 5) and disseminating success scenarios with concrete improvements on specific KPIs (e.g., increased automation, time saved, reduced carbon footprint).
Data	
5	**Data sharing and standardisation** Several gaps are related to the reusability of (meta-)data and lack of standardisation. Ontologies make data sharing and data standardisation easier/possible. In general, standardisation is crucial (e.g., for legal requirements). At a minimum, ontologies must be aligned with industrial standards as much as possible.
6	**FAIR principles also for metadata** Implementation of all FAIR principles is hard, therefore implementing it for metadata is a good starting point.
7	**Close cooperation with FAIR communities** Close cooperation with communities, use/development of standardised tools for the implementation and the evaluation of FAIR principles. This will also help to clarify the misunderstandings about FAIR principles that prevent further adoption.

8	**Demonstration of FAIR benefits** Industrial stakeholders may need concrete examples around how adopting a specific or a set of principles will help them. The community should provide minimal examples to demonstrate the benefits.
Processes	
9	**Follow good ontology development practices and provide a comprehensible methodology** This would guarantee a high quality of ontology development. Best practices must be supported by comprehensive methodologies to enable sustainable development of ontologies.
Tools	
10	**Increase user-friendliness of tools** The most significant gap regarding tools is their usability. Tools should be user friendly, and complex details should be in the background. Tools must be developed more user centric with constant feedback regarding their usability. Research and Development (R&D) projects targeting higher TRLs can include usability testing of developed prototypes as a criterion.

5.1.6 References

Ref 5.1.1 https://ontocommons.eu/ontocommons-demonstrators

Ref 5.1.2 OntoCommons (2021). OntoCommons.eu: Demonstrators at work to deliver an Industry Commons Marketplace – Workshop on Demonstrators and Use Cases. Online Focused Workshop, 09.03.-10.03.2021. https://ontocommons.eu/workshop-demonstrators-and-use-cases

Ref 5.1.3 OntoCommons (2021). Global Workshop: Ontology Commons addressing challenges of the Industry 5.0 transition. Online Horizontal Workshop, 02.-05.11.2021. https://ontocommons.eu/news-events/events/global-workshop-ontology-commons-addressing-challenges-industry-50-transition

5.2 Standardisation

As mentioned in Chapter 5.1, the industrial impact of the implementation of ontologies contributes to making applications easy to use and understand, which can save time for companies. Ontologies can help make data sharing and data standardisation easier, allowing interoperability between standards and better data formatting, which in turn enables better data sharing between industrial companies. In this section, we explore the importance of standardisation and how ontologies can contribute, firstly, to the harmonisation of standards use in materials and manufacturing processes, and secondly, to building a common understanding between materials and manufacturing companies.

Despite their growing strategic importance worldwide, standards and standardisation are still not well integrated into Europe's journey towards strategic autonomy [Ref 5.2.1] The geopolitics of new technologies and advanced manufacturing require Europe to ensure the efficient and effective functioning of its standardisation systems. As the EU Strategy on Standardisation [Ref 5.2.2] states "Europe's competitiveness, technological sovereignty, ability to reduce dependencies and protection of EU values […] depends on how successful EU actors are in standardisation at an international level" This is true of the materials and manufacturing sectors. Europe must be a rule-maker and not a rule-taker in terms of international rules and standards – every European standard adopted at an international level brings a competitive advantage to European businesses. As the pace of digitalisation accelerates, European standards are essential to ensuring that Europe's digital space remains safe, secure, and cyber-proof. The slow approval of harmonised standards is weakening the coherence of the (Digital) Single Market. It is also sapping the competitiveness of the EU's digital players, where speed to market is critical and must be resolved urgently.

If Europe is to achieve its strategic objectives and retake its role as a global standards setter, EU political leaders must recognise that a well-functioning standardisation system is crucial to achieving strategic autonomy. This demands leadership and advocacy across the European Commission and strong coordination. Therefore, this first Roadmap report by OntoCommons is delivered at a crucial time when Thierry Breton, Commissioner for the Internal Market at the European Commission, announced at an EC Conference on the 2nd of February 2022, a shift in the standardisation strategy for Europe, and which is focused on the following actions:

1. Anticipate, prioritise and address standardisation needs in strategic areas;
2. Improve the governance and integrity of the European standardisation system;
3. Enhance European leadership in global standards;
4. Support innovation,
5. Enable the next generation of standardisation experts.

With this Roadmap, we would like to draw upon some relevant recommendations and help bridge EU policies and standardisation activities across materials and manufacturing, that are aligned with the Annual Union WP for European Standardisation (AUWP) [Ref 5.2.3].

The Ontocommons project finds itself in the advantageous position of tapping into a set of 11 demonstrators, with additional ones being onboarded throughout the project duration [Ref 5.2.4], with industrial relevance which responds well to Horizon Europe call requirements around standardisation efforts.

The work conducted with the OntoCommons demonstrators in the scope of the "OntoCommons Standardisation Impact Report", which analyses the standards used by each use case, can help the relevant standards Working Groups and Technical Committees to understand user requirements needed for the adoption of existing standards, and the tools and software that are necessary to exploit them: thanks to our demonstrators we can gain a better understanding of what standards the industrial communities are using, and what requirements they should have, for users in the NMBP domain to be able to exploit them. It has also highlighted the need for future proposals to clearly demonstrate the promotion of the relevance and benefits of ICT standardisation, for the improvement of data sharing between materials and manufacturing, especially in support of European industry competitiveness, driving sustainability sovereignty, the green deal, values and ethics. The second iteration of the "OntoCommons Standardisation Impact Report" will cover the standardisation progress and recommendations linked to the results of the demonstrators.

5.2.1 Industrial Need

Need #	Need Description
1	European industrial strategic autonomy through better integration of materials and manufacturing standards and standardisation.
2	Stronger integration of multi-domain stakeholder clusters with streamlined, digitally-supported workflows.
3	Agile and market responsive SMART standards.
4	Widely recognise standardisation as a channel of technology transfer from science to industry and a way to valorise those results.
5	Engineering software systems need to be reusable.

As industry is moving towards closer collaboration and optimisation across domains, standardisation processes require stronger integration of multi-domain stakeholder clusters with streamlined, digitally-supported workflows for greater efficiency. Standardisation should be promoted as a key enabler for industry which can also reinforce links between research, innovation and standardisation.

In principle, it is in industry's own interest to participate in and try to influence standardisation. As standards codify how new products or services entering the market are produced, innovative producers sometimes have an opportunity to shape the rules of the game and benefit from a first-mover advantage. Participation in standardisation is voluntary and not remunerated directly, so any additional free aid should incentivise engagement. However, projects such as StandICT2023 provides funding to individuals to contribute to ICT standardisation activities. Therefore, companies innovating in the ICT sector are now able to apply for funding to propose their technologies or results for the revision of existing standards or the creation of new standards.

Dedicated efforts on developing SMART standards that are agile and market responsive need to be tackled with end-users through dedicated interoperability test-bed frameworks. OntoCommons is identifying the right channels for this to happen and recommends that more attention should be dedicated in the future to the connection between National (NSBs) and International Standards Bodies (SDOs). Literature on the role of standards and standardisation in technology transfer, and more specifically in research initiatives, is not widespread, and it is evident that this is an effort that has been evolving recently [Ref 5.2.5]. Indeed, standardisation has not been traditionally and widely recognised as a channel of technology transfer from science to industry [Ref 5.2.6].

5.2.2 State of the Art

Europe's competitiveness in the domain of advanced manufacturing is strongly linked to its ability to embed digital capabilities in operations of competitive industries and services. Good performance alone is not sufficient: FAIRness, interoperability, trust, security and reliability in data sharing are crucial attributes that must be guaranteed in all phases of the production chain.

The European Standardisation System (ESS) supports the overall competitiveness of EU industries, and the Standardisation Strategy Roadmap [Ref 5.2.7] has identified coordination of European Standards and addressing bottlenecks within the standardisation system as some fundamental challenges, among others, to achieve an effective and harmonised standardisation.

This is where ontologies play a fundamental role, contributing to standards harmonisation and interoperability, offering better categorisation of information and process efficiency.

5.2.3 Gaps

The following gaps are aligned with the challenges and opportunities for standardisation in Europe as identified in the new EU Standardisation Strategy, addressing delays in the approval of new European standards, inclusiveness for environmental and other societal voices and interests, as well as transparency and legal certainty for European standards [Ref 5.2.7]. A better understanding of these gaps will help the materials and manufacturing communities by improving standards applicability in the manufacturing process.

GAP #	GAP DESCRIPTION
1	Coordination of European standards and addressing bottlenecks within the standardisation system.
2	Slow approval of harmonised standards.
3	Re-usability of engineering software systems.
4	Literature on the role of standards and standardisation in technology transfer and in research initiatives is recent and not widespread.
5	Different data representations making it difficult to reuse different systems.
6	Barriers to communication between devices in ICT.
7	Facilitating greater R&I contributions to the standardisation ecosystem.

Through the first year of its life, the OntoCommons project has worked on ensuring that efforts around standards harmonisation through ontologies are channelled in the same direction. In particular, the OntoCommons Standardisation Impact Report released in November 2021 has highlighted challenges and opportunities in the use of standards related to ICT, engineering properties and material information standards.

- One of the current gaps concerns the re-usability of engineering software systems: each of these systems might have different representations of data and different digital information, that, therefore, cannot be re-used by other systems, resulting in extra costs.
- Information describing the characteristics and properties of each specific domain within Materials 4.0 should be understandable by receiving systems, even when these become obsolete. Standardisation, and bringing the application of such standards to an ontology level, has great potential to provide new opportunities and new business models for the application and presentation of material information.
- In the ICT domain, the "network effect", which is the capacity of a device to communicate with others, and therefore the ability to interoperate, is a crucial aspect to determine the device's value. Accessibility, approval procedures, awareness, engagement and Intellectual Property rights are the main barriers preventing ICT standards from achieving a tangible impact towards full device interoperability.
- Standardisation experts in the ontologies field should support the contributions of R&I projects to EU standardisation activities through initiatives such as the EC's Standardisation Booster [Ref 5.2.9].

5.2.4 Definition of Success

The new knowledge resulting from publicly funded research and innovation programmes to industry related projects using standards can be included in new or improved standards, contributing both to the implementation of research and industrial innovation outcomes through the usage of standards, addressing key impacts. The latter is crucial to defining success in the standardisation ecosystem.

5.2.5 Recommended Action

ACTION #	ACTION DESCRIPTION
1	Europe to ensure the efficient and effective functioning of its standardisation system improving speed to market.
2	Focus on the achievement of a well-functioning standardisation system.
3	Demonstrating SMART standards with end-users through dedicated interoperability test-bed frameworks.
4	Promote standardisation as a key enabler for industry.
5	Reinforce links between research, innovation and standardisation.
6	Improve the connection between National (NSBs) and International Standards Bodies (standards development organization, or SDOs).
7	Embed digital capabilities in operations of competitive industries and services.

8	Focus on the use of ontologies to contribute to standards inclusivity, harmonisation and interoperability, offering better categorisation of information and process efficiency.

Moreover, improving the inclusivity and interoperability of standards will strengthen Europe's worldwide competitive position as regards innovative technologies. In fact, additional investments in the quality, quantity and speed of the standardisation effort will allow Europe to cope with the increasing speed of new technology development, which in return, will attract investment and technologies from research organisations and companies all over the world wishing to invest in countries where new technology regulation is more advanced.

To introduce a more inclusive level of participation to standardisation work, efforts need to improve to break down the following barriers to engage, which are:

- Lack of knowledge of standards/standardisation, especially for start-ups or SMEs,
- financial barriers,
- long-term nature of the standardisation process,
- lack of recognition of standardisation work for researchers,
- standards as a mutual outcome (where there are virtually no incentives or mention for authors). [source: RTD/2021/SC/005-Developing a code of practice on standardisation-scoping study for supporting the development of a code of practice for researchers on standardisation].

Some success factors that feature in overall standardisation developments, that may be considered as recommended actions to go forward, are taken from page 12 of Hermann, P., Blind, K., et al., 2020 on "Relevance of standards and standardisation for knowledge and technology transfer", and are as follows:

Success factors for the overall standardisation undertaking	
Thematic fit for individuals	• Thematic complementarities and synergies • High degree of individual domain expertise
Involvement of an industry partner	• Facilitates bringing other partners into the project • Increases the likelihood that the standard is later taken up in practice
Previous experience with standardisation work	• Successful collaboration in past initiatives provides a basis for further undertakings • Informal links win SDO committees may lead to future joint projects • An understanding of how standardisation processes work increases efficiency and effectiveness
Existence of a driving force	• "Lighthouses" who motivate others to participate and increase the likelihood of success

	• More intense contribution results from own interest in the standard (relevance, urgency)
"Right timing"	• Depending on the driver and technology choosing the right timing for initiating standardisation activities.

A recommendation going forward would be that the OntoCommons project representatives, together with its esteemed Expert Advisory Group, may contribute to the chapters around ontologies and semantic interoperability within future editions of the ICT Rolling Plan of Standardisation [Ref 5.2.10] as well as contributing to the SRIA editions of EOSC [Ref 5.2.11].

5.2.6 References

Ref 5.2.1 https://www.epc.eu/content/PDF/2020/EPE_JB_Europe_as_a_global_standard-setter.pdf

Ref 5.2.2 https://ec.europa.eu/docsroom/documents/48598

Ref 5.2.3 https://www.eesc.europa.eu/en/our-work/opinions-information-reports/opinions/annual-union-work-programme-european-standardisation-2020

Ref 5.2.4 https://www.ontocommons.eu/ontocommons-demonstrators

Ref 5.2.5 RTD/2021/SC/005 – Developing a code of practice on Standardisation – Input to stakeholder workshop, Dec 2021 (EFIS, IMC, FH KREMS, ECORYS)

Ref 5.2.6 Radauer, A. (2000) Driving from the fringe into the spotlight. The underrated role of standards and standardisation in RTDI policy and evaluation. Fteval Journal for Research & Technology Policy evaluation (51). pp.59-65. ISSN 1726-6629.

Ref 5.2.7 https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/13099-Standardisation-strategy_en

Ref 5.2.8 https://ecostandard.org/wp-content/uploads/2022/01/ECOS-Media-Briefing-Understanding-the-EU-Standardisation-Strategy-%E2%80%93-and-why-it-matters-for-the-climate.pdf

Ref 5.2.9 https://www.hsbooster.eu/

Ref 5.2.10 https://joinup.ec.europa.eu/collection/rolling-plan-ict-standardisation/rolling-plan-2021

Ref 5.2.11 https://www.eosc.eu/sites/default/files/EOSC-SRIA-V1.0_15Feb2021.pdf

5.3 Knowledge Management Translator for Industry Commons

5.3.1 Industrial Need

The Knowledge Management Translator is a key role in an Industry Commons that aims to advise companies on the adoption, use and benefits of ontology-based data documentation to support knowledge management and data sharing. These Translators can enable Industry Commons and the deployment of OCES, helping to bridge skills and capabilities gaps in the application of knowledge management technologies and supporting data sharing to support a wider leveraging of data value. Their aim is to remove key obstacles for industrial beneficiaries in utilising ontology-based data documentation and OCES knowledge engineering. In particular, a Knowledge Management Translator is an "ontologist" translator and coach, with expertise spanning across data management, semantic tools, ICT, analytical philosophy and the science/engineering domains.

Similar "Translation" roles have emerged in other fields, where there is also a gap between complex technology potential and industrial impact, including materials modelling [Ref 5.3.1], [Ref 5.3.2] and data analytics [Ref 5.3.3], [Ref 5.3.4].

An inventory of Industrial Needs has been made during two workshops organised by OntoCommons. [Ref 5.3.5], [Ref 5.3.6].

They fall into the following categories:

Need #	Need Description
1	**High quality, reliable advice on ontology-based knowledge management/engineering** Industry has a strong interest in advancing their data and knowledge management, supporting digitalisation, achieving more interconnected solutions, having information more easily and readily available for anyone who needs it, improving information exchange with suppliers and customers, etc. Ontology-based approaches are very much of interest in this context, but industry will need advice on what approach to take to data and knowledge management, especially given some past failed approaches. The perception is that it takes a huge amount of time to develop ontologies and that the technology may not be ready enough or scalable enough.
2	**Information on what to expect, best practice** Industry needs to know whether and how ontologies can help, and what the costs, efforts and returns on investment of different approaches would be. Best practice needs to avoid incorrect expectations and that knowledge engineering is misapplied.
3	**Skill and capability in industry** There are only a few people around in the world that possess both semantic and domain knowledge and thus can master the skills needed. Hence, more experts are required and

	there is the need to create the Knowledge Management Translator role and associated education to train new Translators.
4	**Remove obstacles** The path to implementing and successfully adopting ontology-based data and knowledge management may involve too high an up-front investment. Industry will be keen on Translators that can help them deliver early successes and an onboarding process that is fast and smooth.

5.3.2 State of the Art

The role of a Translator first emerged in the Materials Modelling community since the transfer of materials modelling was lacking an expert who could translate business requirements into potential modelling solutions, in particular in, small- and medium-sized companies. When finishing in 2019, the H2020 EMMC-CSA reached a milestone by defining this concept and supporting tools for Translation processes, which is documented in the EMMC Translators Guide, see [Ref 5.3.7]. In continuation of this work, an EMMC related group of scientists combined the Translator concept with Business Decision Support Systems and developments of supporting ontologies, published in [Ref 5.3.8]. From this point onwards, a continuous elaboration of the Translator role in novel, contemporary context began to happen. One of these re-interpretations is currently active in the EU H2020 project OntoTrans [Ref 5.3.9], where an "OntoTransLator" was introduced who needs to know about simulation ontologies, to keep their subject matter expertise up-to-scratch and to research what global use a product may have. Furthermore, the EMMC is continuing this line with a Benefits Analysis and Method Comparison tool. A necessary skill will also be how a translator can interface with other roles as they will have to work with many stakeholders to make innovation FAIR on all its many facets. Hence, as this evolving is happening, we are confident that the state of the art is a superb basis to launch the "Industry Commons Translator" as a new job role. [Ref 5.3.10].

Similarly, in data analytics, the role of an Analytics Translator is very much reminiscent of the one of a Materials Modelling Translator, as they are expected to bridge the technical expertise of data engineers and data scientists with the operational expertise of marketing, supply chain, manufacturing, risk, and other frontline managers, i.e., close the "language gap" between industrial stakeholders and data experts. AIANDUS [Ref 5.3.4] praises Analytics Translators as one of the new trendiest jobs of the 21st century, and attempts to build a community around personas who can aid with building Data Science solutions for getting business value.

In the field of ontology-based data documentation, Knowledge Management (KM) [Ref 5.3.11] and Knowledge Engineering (KE) [Ref 5.3.12], the current status is to aim for building the organisational structures that feed an AI system [Ref 5.3.14]. It is well understood that well-structured information is needed to make AI work correctly as intended. Ontologies can be seen as vital knowledge representations [Ref 5.3.15] but it is still an ongoing effort to bring them to organisations and find the staff to do this [Ref 5.3.16]

5.3.3 Gaps

In order for a community of semantic Translators to grow and thrive, we need to address a range of gaps as listed below. These include gaps related to the external environment, including a certain degree of readiness on the side of industry, the tools available to Translators, as well as skills, training and recognition of the role itself.

GAP #	GAP DESCRIPTION
1	The maturity of organisations with respect to their quality of information exchange is too low to embrace ontologies.
2	Knowledge is trapped in persons' minds and not captured externally.
3	Workflow tools for Translators such as Knowledge Graphs and Transformation Systems are not provided.
4	There is a lack of a coherent framework/technology stack (as is available, e.g., in AI) that makes the work of the Translator and team more difficult.
5	Upfront investment can be too high and/or there is a lack of clarity of investments required.
6	The role of a Knowledge Management Translator has to be defined in such a fashion that there are clear job specifications, and the candidates can have identifiable career progression in this role. Also, an educational curriculum is needed to integrate semantic and domain knowledge, either for self-study or as part of university courses.

5.3.4 Definition of Success

Success will be attained when:

- There is a wide agreement on the role of the Knowledge Management Translator and best practice, supported and ideally certified by relevant professional associations;
- Enough experts possess semantic as well as applied science skills so that they can act as Industry Commons Translators both within companies and as independent consultants;
- There is a network of consultancies offering "Industry Commons Translation" services;
- Industry embraces the role of an Industry Commons Translator for improving knowledge management and data sharing;
- Organisations are willing to upskill their staff; there is continuous professional development in ontology-based knowledge management/engineering (KE) and a culture of KE CPD in industry;
- A stack of high quality, easy to use knowledge engineering tools is available to Industry Commons Translator teams.

5.3.5 Recommended Action

OntoCommons, by coordinating with external experts, will define the role of an Industry Commons Translator using a bottom-up approach. A community of practice will be established and a manifesto similar to the Agile Manifesto of Software Development will be agreed.

ACTION #	ACTION DESCRIPTION
1	**Role**: Clear definition of the role(s) of Knowledge Management Translators for Industry Commons, achieving wide agreement regarding the skills and tasks performed by the Translator (or team of translators). Delineation of the Translator role as an advisory/coaching/consulting role to more technical data science and knowledge engineering roles. (Industrial Need 1+3, Gap 2+6)
2	**Education**: Educate applied scientists and semantic experts. Use a curriculum developed within OntoCommons comprising literature, training, forums, etc., to provide self-training. Relevant policies, programmes, training courses and supporting infrastructure to upskill capabilities across industry must be developed. Education and CPD programmes that support establishing data and knowledge sharing via connected digital twins of the built environment provide a useful guide as to what is required also for Industry Commons Translators. They include: • A Skills and Competency Framework (SCF) [Ref 5.3.17] • A Capability Enhancement Programme (CEP) [Ref 5.3.18] The SCF framework identifies priority skills and competencies required across a range of roles and can act as a baseline for industry and individual organisations in the assessment of their current capabilities, to identify gaps and plan how they may be addressed. As a part of the SCF, targeted role-based training plans need to be developed. The CEP programme identifies the steps necessary to bring organisations and individuals up to the level of expertise required, and equip organisations with tools, guidance and materials to understand and cultivate the skills and knowledge they need. This is to provide guidance and resources to drive the development of the right skills, at the right level, to achieve the goal of enhancing industry capabilities and deliver on the objectives of Industry Commons data sharing. Further approaches can include: • Establishing OntoCarpentry, similar to Software Carpentry [Ref 5.3.19], [Ref 5.3.20], possibly within Data Carpentry [Ref 5.3.21]. • Marie Skłodowska-Curie Actions [Ref 5.3.22] that enable domain researchers to visit experts in ontologies and semantics and gain education, and vice versa. (Industrial Need 1+2+3, Gap 1+2+3+5+6)

3	Establish "Translator Tools" for the comparison of different data processing technologies. With such a tool the industry could compare ML knowledge graphs (and other data analytic methodologies) and ontologies that enable companies to envisage their value streams and know where and why to invest. Such a tool would point out the strong points and benefits of ontologies to organisations. The translators will be provided with a best practice guide to make their work transparent and FAIR, interoperable with existing standards, and trackable. They also will have access to the OCES Toolkit and the training surrounding it. We suggest that translators need to identify a particular data analytics problem where they could provide a quick answer and then gradually progress to long-lasting solutions: i.e., ontologies. The demonstrator cases can aid with identifying suitable data analytics problems for a variety of industry sectors (Industry Needs 1+2+3+4, Gaps 2+3+4+5)
4	Formation of a community of practice (CoP), e.g., within the Research Data Alliance (RDA) [Ref 5.3.23]. In RDA, CoPs are composed of experts from that community that have an interest in the discipline/research domain, and are committed to directly or indirectly enabling data sharing, exchange, and/or interoperability. CoPs serve as platforms for communication and coordination among individuals, building bridges outside and within the RDA, with shared interests. [Ref 5.3.24] (Industrial Need 1+2, Gap 2+6)
5	Establish a directory of Knowledge Management Translators. OntoCommons should facilitate a means so that industry can identify individuals and organisations that can provide translation services. (Industrial Need 1+4, Gap 1+5)

5.3.6 References

Ref 5.3.1 EMMC Translators Guide: https://doi.org/10.5281/zenodo.3552260

Ref 5.3.2 Translation in Materials Modelling – Process and Progress: https://zenodo.org/record/4729918

Ref 5.3.3 https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/analytics-translator

Ref 5.3.4 https://www.aiandus.com/ats/

Ref 5.3.5 OntoCommons (2021). Industry Commons Translator. In: OntoCommons. Global Workshop: Ontology Commons addressing challenges of the Industry 5.0 transition. Online Global Workshop, 02.-05.11.2021. https://www.youtube.com/watch?v=vQrmnsLjVRM&list=PL-cwgiwYXckPOatW5dGBK_wDr0ZoP72L3&index=16

Ref 5.3.6 OntoCommons (2022). The Industry Commons Translator - 1st Expert Meeting. Online. Focussed Workshop, 21.2.2022 https://www.ontocommons.eu/news-events/events/industry-commons-translator-1st-expert-meeting

Ref 5.3.7 Hristova-Bogaerds, Denka, Asinari, Pietro, Konchakova, Natalia, Bergamasco, Luca, Marcos Ramos, Alicia, Goldbeck, Gerhard, Hoeche, Daniel, Swang, Ole, & Schmitz, Georg J. (2019). EMMC Translators Guide. Zenodo. https://doi.org/10.5281/zenodo.3552260

Ref 5.3.8 Dykeman, Donna, Hashibon, Adham, Klein, Peter, & Belouettar, Salim. (2020). Guideline Business Decision Support Systems (BDSS) for Materials Modelling. https://doi.org/10.5281/zenodo.4054009

Ref 5.3.9 https://cordis.europa.eu/project/id/862136, https://www.ontotrans.eu

Ref 5.3.10 https://www.earley.com/insights/who-owns-business-ontology-staffing-ontology-development

Ref 5.3.11 Knowledge management (KM) is the collection of methods relating to creating, sharing, using and managing the knowledge and information of an organisation. It refers to a multidisciplinary approach to achieve organisational objectives by making the best use of knowledge.

Ref 5.3.12 https://en.wikipedia.org/wiki/Knowledge_management#:~:text= Knowledge%20management%20(KM)%20is%20the,the%20best%20use%20of%20knowledge

Ref 5.3.13 Knowledge engineering (KE) refers to all technical, scientific and social aspects involved in building, maintaining and using knowledge-based systems. https://en.wikipedia.org/wiki/Knowledge_engineering

Ref 5.3.14 Earley, 2017, https://www.earley.com/insights/knowledge-managements-rebirth-knowledge-engineering-artificial-intelligence)

Ref 5.3.15 Earley, 2016, There is no AI without IA. IT Pro May/June, 58-64.

Ref 5.3.16 Earley, 2020, https://www.earley.com/insights/who-owns-business-ontology-staffing-ontology-development

Ref 5.3.17 https://www.cdbb.cam.ac.uk/files/010321cdbb_skills_capability_framework_vfinal.pdf

Ref 5.3.18 https://www.cdbb.cam.ac.uk/files/120421_cdbb_capability_enhancement_vfinal.pdf

Ref 5.3.19 https://software-carpentry.org/

Ref 5.3.20 Simperler, A.; Wilson, G.; Software Carpentry get more done in less time, https://arxiv.org/abs/1506.02575

Ref 5.3.21 https://datacarpentry.org/

Ref 5.3.22 https://marie-sklodowska-curie-actions.ec.europa.eu/about-marie-sklodowska-curie-actions

Ref 5.3.23 https://www.rd-alliance.org/rda-announces-new-group-category-communities-practice

Ref 5.3.24 https://www.rd-alliance.org/sites/default/files/attachment/RDA_Communities_of_Practice_Framework_v0.2_Dec2020.pdf

5.4 Ontology-based digital-marketplaces cooperation

Digital Marketplaces are multisided collaborative and trading platforms that facilitate materials innovation by easing access to otherwise disparate sources and deployments of information, expertise, software applications and data. There are multiple marketplace projects running at the moment, with little interaction between each other; this is often driven by beneficiaries who take part in more than one project. MarketPlace [Ref 5.4.1] and ViMMP [Ref 5.4.2] were funded under NMBP-25-2017 – "Next generation system integrating tangible and intangible materials model components to support innovation in industry", and DOME 4.0 [Ref 5.4.3] is funded under DT-NMBP-40-2020 – "Creating an open market place for industrial data (RIA)". We also are in contact with beneficiaries from MARKET 4.0 [Ref 5.4.4] and WeldGalaxy [Ref 5.4.5], both of which were funded under DT-NMBP-20-2018 – "A digital 'plug and produce' online equipment platform for manufacturing (IA)".

5.4.1 Industrial Need

The ontology-based digital-marketplaces will meet the following industrial needs:

Need #	Need Description
1	**Integrating data generated by simulation and experiments** The amount of data generated by simulation and experiments are continuously increasing and therefore, integrating and interconnecting these scattered repositories and data is essential. Such integrated approach will also support the generation of datasets that can leverage the data-driven methodologies.
2	**Interoperability based on common standards** From the industrial perspective, it is required to have more modelling workflows with a focus of interoperability. The interoperability should be based on common ontologies and open standards that enable the description of data in a unified fashion, capturing the meaning of data in an explicit and sharable manner understood by both humans and machines. The adoption of common standards increases the data's FAIR characteristics.
3	**Better user-friendly platforms** Marketplaces should provide a functional and user-friendly platform for industrial partners and experts to run different modelling workflows, and provide the functionality of linking various models and tools. In order to reach this goal, tutorials and use cases are required.
4	**Effective data exchange between simulators and databases**

	The seamless exchange of information between simulators and databases is another important aspect, in addition to the ability to create executable and adequate simulation and modelling workflows. This provides another useful benefit such as enabling tools/data generated in one platform to be easily used/post-processed in another platform.
5	**Improving transferring data between industries and marketplaces** It is necessary to use a single platform to store databases and repositories in many industrial use cases. Users should be able to search as well as find data and information, including how to exchange it and use it through such a platform. Currently there is no such platform. We are developing different digital marketplaces based on different APIs. Transferring data between industry and marketplaces should be improved. In this sense, having a single user-friendly platform for data transfer that an industrial partner can access to connect and use the available marketplaces would be beneficial. In this regard, marketplaces are required to have data accessible from multiple sources and deliver it to the customer, so there has to be trusted, aggregated and ready-to-use data, that also provides data provenance and IP/sovereignty.
6	**Access possibility** Marketplaces must provide a platform that provides the user with the required access rights (through adequate authentication services) to commercial and academic software as well other tools of interest.

5.4.2 State of the Art

Similarly, amongst other stakeholders, ontologies are recognised as being the key for interoperability and for paving the way to using AI in the most efficient way. EU marketplaces are using ontologies for their services and operations, i.e., Connect, Search, Test/Simulate, Tender/Bid, Compare. Feedback. Ontologies are also useful for them to be able to interact with external entities, such as sources of information, experts, software, matchmaking, interoperability with other marketplaces, etc. The latter does require them to widely agree on the ontologies used. The two materials modelling marketplaces, MarketPlace and ViMMP also need ontologies to power semantic interoperability within modelling workflows offered on their marketplace.

- OntoCommons and the marketplaces are actively working on finding agreements on the key concepts and taxonomies of the ontologies that are used. The main idea is to build a knowledge graph for one marketplace so that another marketplace can understand it.
- After a knowledge graph is created, it can be used for data analysis, business intelligence, etc.
- The figure below adapted from the MARKET4.0 project shows how a global ontology framework for marketplaces could look like. Furthermore, the European Virtual Marketplace Framework (EVMF) has shared fundamental concepts and a small (mid-level) ontology that provides connections (EVMPO). EVMPO could, in a way, be the basis for such a knowledge graph for marketplace projects and potentially could be extended. It is an ontology that includes common concepts we all need and how they are connected (and so is a good starting point and nucleus).

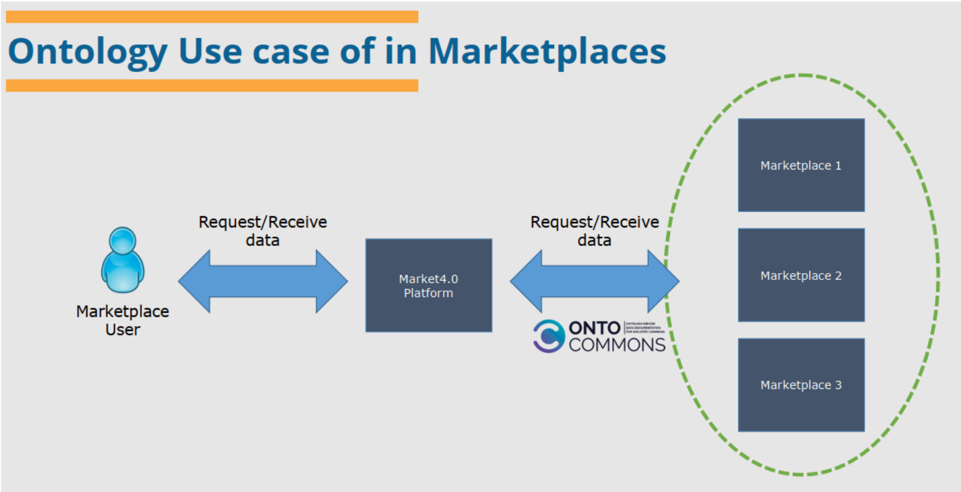

Figure 7- Distribution of ontologies by domains

5.4.3 Gaps

GAP #	GAP DESCRIPTION
1	**Lack of communication between marketplaces to develop a common ontology** The necessary efforts and quantification of benefits of a common ontology for marketplaces has been addressed in the OntoCommons Focused Workshop on Industry Commons marketplaces. Given the complexity and novelty of the efforts, they have been somewhat hard to quantify. However, it was clear that a collaborative action among the marketplaces is of utmost importance in order to achieve OntoCommons' ambitious goal: providing a common place for developing a common ontology for marketplaces. No ontology has been used by MARKET4.0 marketplace. As represented in the report of the OntoCommons Focused Workshop on Industry Commons marketplaces [Ref 5.4.6], the MarketPlace, VIMMP and DOME 4.0 marketplaces use the Elementary Multiperspective Materials Ontology (EMMO) as a TLO. The MARKET4.0 marketplace does not use TLOs in its semantic framework. However, the use of a TLO in MARKET4.0 is being considered for future work. The MARKET4.0 consortium has highlighted the wish and need to create a common ontology based on the similarities between the individual marketplaces which can be further extended with the additional requirements specific to each marketplace. Therefore, OntoCommons would not want marketplaces to start from scratch when they want to develop ontologies for their use cases, and should offer an ecosystem with methods and tools to develop interoperable ontologies.
2	**Lack of tools and methodology** The lack of industry-ready tools and methodologies for the alignment of various DLO and TLOs have been discussed during the OntoCommons Focused Workshop on Industry

	Commons marketplaces. The OntoCommons consortium should identify the most adequate methodologies and tools for an effective alignment.
3	**Lack of demonstrators** Get more marketplace demonstrators to show the benefit of guidelines.
4	**Lack of user-friendly Graphical User Interfaces** Have better GUIs and user-friendly interfaces for marketplaces. Searching for information in these marketplaces should be as easy as possible.
5	**Better communication between EMMO and the marketplaces** According to observations from a preparation meeting before the OntoCommons global workshop, it was mentioned that communication between the EMMO core group and a marketplace consortiums is quite essential and currently is missing. It is therefore required to have more discussions on the expectations of marketplaces from EMMO.

5.4.4 Definition of Success

- Try to find the synergies (and 'common points') between marketplaces and then find out the needs, issues and gaps in this respect from the marketplaces, and finally what are the needs for further collaborative developments.
- Organising another workshop for marketplaces. This workshop should be in-person/on-site to have the possibility for better discussions.
- Create a common space for sharing updates from each marketplace project, discussing the issues and gaps, and sharing the development status of ontologies.
- According to discussions in the OntoCommons global workshop [Ref 5.4.7], one of the important experiences from H2020 projects is that there should be a balance between new ontology developments that can take a long development time when compared to the timescale/duration of a European project, and use existing resources and available ontologies like FOAF. Therefore, the reuse of ontologies, integration/merging with other marketplace application ontologies, are very important.
- There were some solutions suggested during the OntoCommons global workshop on how to provide and establish links between marketplaces. One of the solutions is to share documents and technologies which are open amongst the marketplaces. In terms of semantic interoperability, a convergence of underlying ontologies/taxonomies is proposed and will be further developed in OntoCommons.
- It was also suggested during the OntoCommons global workshop to share fundamental concepts and small (mid-level) ontologies that provide connections between marketplaces (e.g., the EVMPO one that connects VIMMP ontologies to EMMO and other TLO concepts).

5.4.5 Recommended Action

Action #	Action Description
1	**Well-defined demonstration for marketplaces** Background: Often each marketplace is used for a specific purpose, for instance, database storage which is often called marketplace for data or execution of specific modelling workflows. There should be a well-defined demonstrator in which marketplaces share their specific purpose along with adopted methodologies, with a change in mindset/practices to develop and collaborate further with each other. Action: Define a demonstrator for marketplaces and create a shared space to foment the collaboration and development in a harmonised manner.
2	**Developing a common "global" ontology framework for the marketplaces.** Background: One of the goals of OntoCommons relies on facilitating interoperability across various marketplaces by enabling the seamless exchange of data across different databases, simulation engines and tools available at the marketplaces. Such an integrated approach could be achieved by using a unified schema provided via an ontology that can be understood across the marketplaces. Action: Resume the EVMPO development and create a collaborative environment that joins members from the various marketplaces in a collaborative effort towards a unified ontology. Establishment of an EMMC task group for the purpose.
3	**Establishing link between marketplaces** An OntoCommons demonstrator will work on how to establish connections between marketplaces and provide the OntoCommons EcoSystem with prototypical needs from digital marketplaces and similar NMBP platforms. Expected benefits include supporting federated queries on high-level categories and data ingestion. For example, we could look for an "agent" or "infrastructure" on multiple platforms.

5.4.6 References

Ref 5.4.1 https://cordis.europa.eu/project/id/760173;https://www.the-marketplace-project.eu/

Ref 5.4.2 https://cordis.europa.eu/project/id/760907;https://www.vimmp.eu/

Ref 5.4.3 https://cordis.europa.eu/project/id/953163;https://dome40.eu/

Ref 5.4.4 https://cordis.europa.eu/project/id/822064;http://market40.eu/

Ref 5.4.5 https://cordis.europa.eu/project/id/822106;https://www.weldgalaxy.eu/

Ref 5.4.6 OntoCommons (2021). OntoCommons Workshop Industry Commons Marketplaces. Online Focused Workshop, 29.04.2021. https://ontocommons.eu/news-events/events/ontocommons-workshop-industry-commons-marketplaces

Ref 5.4.7 OntoCommons (2021). Global Workshop: Ontology Commons addressing challenges of the Industry 5.0 transition. Online Global Workshop, 02.-05.11.2021. https://ontocommons.eu/news-events/events/global-workshop-ontology-commons-addressing-challenges-industry-50-transition

5.5 Innovation and perspectives

Horizontal connectivity and the ability to replicate results across domains stimulates increased information flows and knowledge exchanges between industry verticals, contributing to a blurring of the traditional vertical boundaries and resulting in something resembling a "data soup" (Magas 2017a). Within the context of Open Innovation 2.0 [Ref 5.5.1], this "new primordial soup" has been identified as an opportunity for all actors to quickly prototype transformative solutions using agile methods (Curley and Salmelin, 2017). In order to capitalise on the unprecedented opportunities for innovation, however, exploitation of common assets requires a more structured systemic approach. The "data soup" relies on ad-hoc discovery of opportunities, unmanageable at scale, and unreliable as an investment of proprietary assets. The resulting disruptive business models occur in the "gaps", outside of the core competence business focus areas, with the potential to threaten existing businesses, while missing the valuable supply networks and distribution channels which would allow them to scale. To solve these challenges, innovation requires structural integration with the horizontal cross-domain industry value network. For this reason, the Industry Commons model has been launched with the ambition to provide structure and support to cross-domain innovation with a series of horizontal enablers (Magas, 2017b).

5.5.1 Types of innovation and associated challenges

According to Pisano, incremental innovation for existing competencies and business models is *routine*; *radical* innovation builds new competencies for an existing business model; *disruptive* innovation proposes a new business model for existing competencies; and *architectural* innovation introduces both new competencies and business models (Pisano, 2019). Since significant breakthroughs may fracture established organisational learning patterns (Dodgson and Gann, 2010), the established tendency is to avoid radical innovation which may destabilise existing capabilities, or disruptive innovation which may disengage existing customers and the related secure income streams. The progress from breakthrough innovation occurring in the "gaps" between verticals to one which is fully integrated in the cross-domain value network entails a shift in organisational learning patterns and knowledge exchange practices.

5.5.2 Opportunities for innovation

Piloting cross-domain data-driven applications in experimental labs with state-of-the-art technology toolkits has demonstrated a steep rise in innovative solutions which combine data sets from two or more domains, and have attracted engagement from across industry sectors (Magas, 2016). Novel business models have already emerged around 10 years ago (e.g., Airbnb and Uber), built from combinations of data from two or more domains using Application Programming Interfaces (APIs). Within the context of experimental labs, interfaces such as APIs, GUIs and TUIs [Ref 5.5.2] have been key enablers for cross-domain innovation. While agile, these interfaces are high-maintenance because of continuous version updates, and therefore present challenges for long-term implementation. A top-level reference documentation system has the potential of solving these challenges. Geographical mapping conventions offer a useful analogy: working with APIs is similar to negotiating multiple geographical symbols and graphic styles, such as diverse road signage in different countries, with conventions which are constantly being updated. A unified top-level reference means that instead, from a high level, all geographical features are described with the same symbol and therefore intelligible to all, regardless of regional differences of the features on the ground. For this reason, an ecosystemic ontology-driven approach, which aims to harmonise top-level reference data documentation across data-driven application domains, has the potential to stimulate and support sustainable cross-domain industrial innovation in the Industry Commons.

5.5.3 The role of Industry Commons

Industry Commons enables a series of breakthrough innovation scenarios by fostering mechanisms and standards for shareable and reusable knowledge across industrial domains, including the enabling of data sharing, cross-domain data-driven hybrid applications, interoperability among involved software systems, identification of business value in the junctions between the verticals, and testing of early adoption and emerging market scenarios. The holistic approach integrates a series of horizontal data-driven enablers, essential for sustainable industrial innovation.

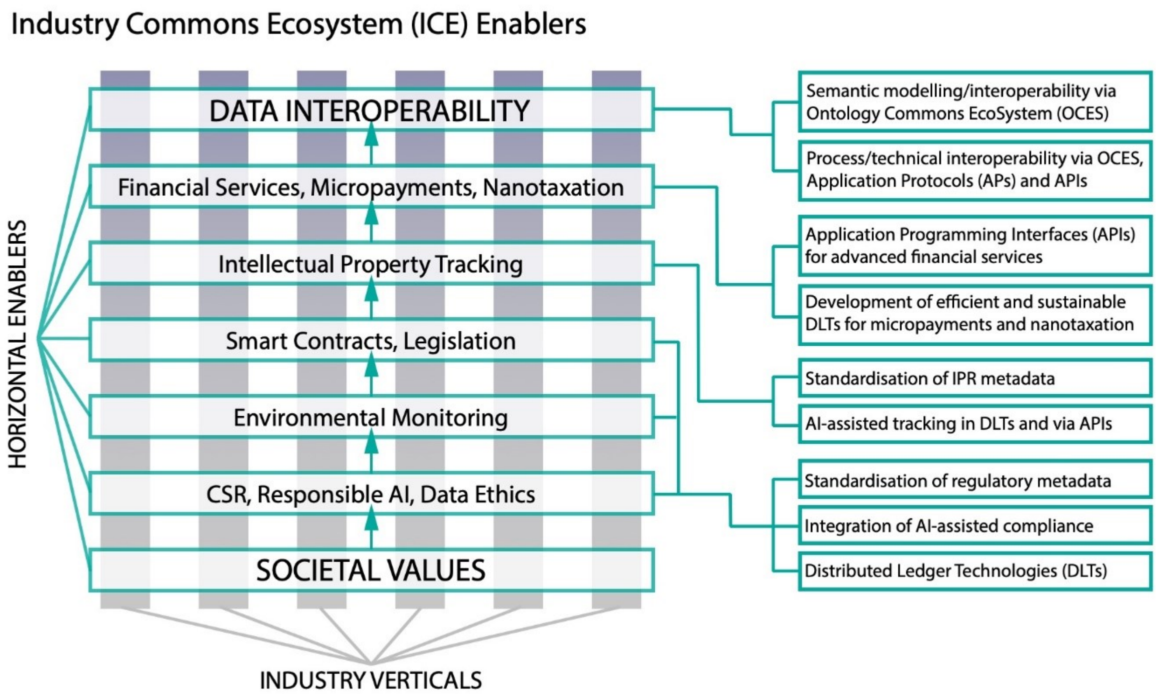

Figure 8 - Industry Commons Ecosystem horizontal enablers across industry verticals

A holistic view of the proposed Industry Commons Ecosystem (ICE) is shown in Figure 8. Starting with society as the foundation, horizontal enablers for cross-domain data exchanges build upwards conditional on the layers below. Data interoperability is conditional on ethical, cultural, regulatory, environmental and societal parameters, not only on technical, semantic and FAIR (findable, accessible, interoperable and reusable) aspects. ICE expands beyond the supply network's regulation and industrial agreements, to an ecosystemic approach of the Commons as a sustainable societal resource. The system therefore connects industrial domains and enables a variety of exchanges linked to an interdependent series of parameters which may include, but are not limited to, financial services/micropayments, Intellectual Property (IP) tracking [Ref *5.5.3*], legislation and regulation, environmental parameters [Ref *5.5.4*], Corporate Social Responsibility (CSR) considerations (e.g., responsible AI and ethics), and societal values.

The model for the Industry Commons Ecosystem (ICE) builds on the work of Enterprise Modelling and Enterprise Integration by adopting the ecosystem view across domains. Enterprise Modelling (EM) is "the process of producing models, i.e., abstract representations, that can be used by humans or machines to support understanding, analysis, (re)design, reasoning, control and even learning about various aspects of interest of an enterprise" (Vernadat, 2020). It covers the needs of software and information systems engineers, manufacturing and industrial engineers, business analysis and organisation experts by determining business drivers and guiding principles which can be used to elicit technological and organisational requirements. According to Vernadat it is "an art, meaning an engineering discipline and not a scientific discipline" [Ref 5.5.5]. Dietz frames Enterprise Modelling primarily within a social system of transactions emphasising human agency and advocating for an Enterprise Ontology to capture the essence of an enterprise network (Land and Dietz, 2012; Dietz, 2006). Enterprise Integration is the process of ensuring the interaction between enterprise entities necessary to achieve domain objectives (EN/ISO I9439, 2003; Bousdekis and Mentzas, 2021, [Ref 5.5.6]). It relies on Enterprise Modelling in order to optimise connectivity, communication and operations between people, processes, systems, and technologies. It enables successful operation, "in a world of continuous and largely unpredictable change, of a single manufacturing company or an ever-changing set of extended (or "virtual") enterprises – by enabling quick and accurate decisions and adaptation of operations to respond to emerging threats and opportunities" (Brosey, 2002).

The ICE model adopts an ecosystem view of Enterprise Modelling uniting well-informed decision-making, technological harmonisation, and socio-environmental responsibility in a common, multi-enterprise, multi-actor and multi-domain ecosystem. Extending Enterprise Integration through shareable and reusable knowledge across industrial domains, supported by tracking of provenance and attribution, and conditional on a series of regulatory, social and environmental parameters, contributes to enterprise resilience and behavioural and cultural adaptation, which results in novel affordances, emerging behaviours and business opportunities. The ICE model builds on the assumption that sustainable cross-domain industrial innovation can be achieved when all aspects of Enterprise Integration are: i) sufficiently transparent to allow all involved actors to be proactive in their decision-making workflows; ii) technologically harmonised to allow interoperability between involved actors' technological components; and iii) effectively supported by responsible societal and environmental parameters embedded in the system.

5.5.4 Emerging framework conditions for cross-domain ecosystems

Table 5.1 - Properties of the Cross-Domain Ecosystem, building on Weichhart, Panetto and Molina, 2021.

Property	System	System-of-Systems	Cross-Domain Ecosystem
Autonomy	No autonomy of parts; only autonomy of the system.	Autonomy is exercised by constituent systems in order to fulfil the purpose of the SoS.	**Autonomy** allows for an increase in dynamic states for greater modularity and adaptability.
Belonging	Parts are akin to family members; they did not choose themselves but came from parents. Belonging of parts is in their nature.	Constituent systems choose to belong on a cost/benefits basis; also in order to cause greater fulfilment of their own purposes, and because of belief in the SoS supra purpose.	**Belonging** is decentralised but closely monitored and tracked across the ecosystem. Market competitiveness is balanced by the ecosystem's supra purpose, encoded in the social dimensions.
Connectivity	Prescient design, along with parts, with high connectivity hidden in elements, and minimum connectivity among major subsystems.	Dynamically supplied by constituent systems with every possibility of myriad connections between constituent systems, possibly via a net-centric architecture, to enhance SoS capability.	**Connectivity** is considered to be all-pervasive rather than a series of nodes and synapses. The value networks operate simultaneously in several dimensions creating value ecosystems.
Diversity	Managed, i.e., reduced or minimised by modular hierarchy; parts' diversity encapsulated to create a known discrete module whose nature is to project simplicity into the next level of the hierarchy.	Increased diversity in SoS capability achieved by released autonomy, committed belonging, and open connectivity.	**Diversity** of capabilities is key to enabling innovation breakthroughs and therefore encouraged. Diverse ecosystem-oriented modules are open and ecosystem-facing, for modelling and coupling on the fly.
Emergence	Foreseen, both good and bad behaviour, and designed or tested as appropriate.	Enhanced by deliberately not being foreseen, though its crucial importance is, and by creating a potential for emergence that will support early detection	**Emergence** is closely monitored and trackable throughout the ecosystem allowing for detection of breakthrough innovation, leading to informed decision-making, investment and resource allocation.

		and elimination of bad behaviours.	

Building on Weichhart, Panetto and Molina (2021), Table 5.1 shows the value proposition of a Cross-Domain Ecosystem (CDE) with respect to those of a System and System-of-Systems (SOS) approach, following the five key properties of systems as identified by Boardman and Sauser (2006).

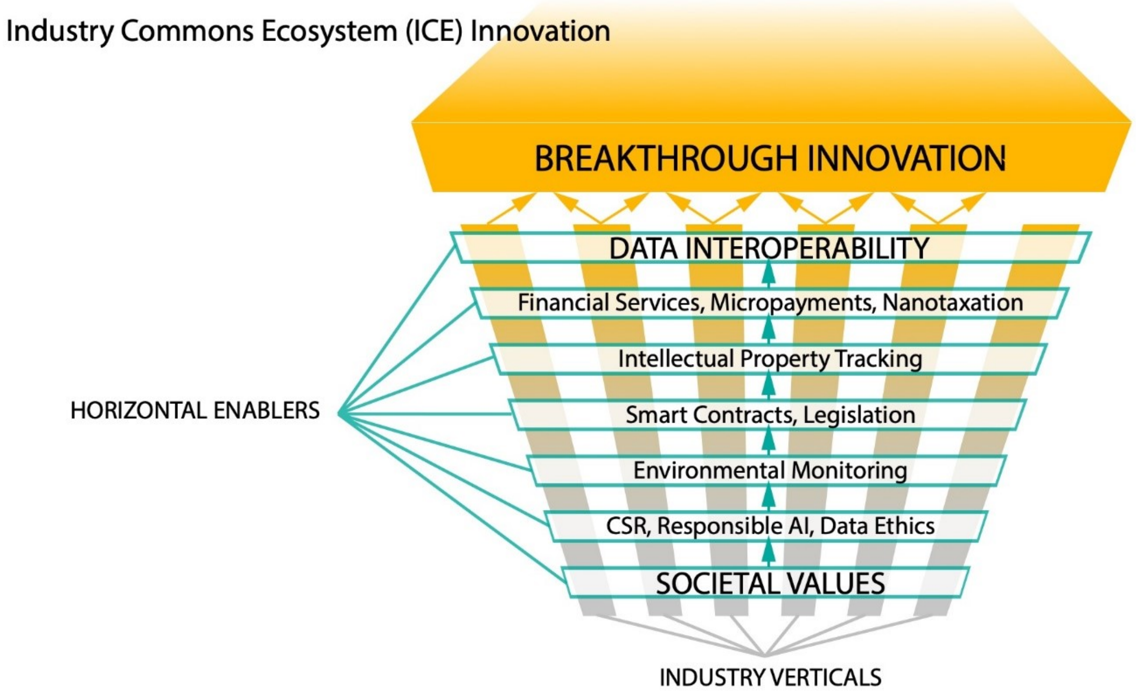

Figure 9 - The Industry Commons cross-domain innovation layer

The value-added layer of the ICE horizontal system is shown in Figure 9. The breakthrough innovation layer which builds and sits directly on top of Cross-Domain Ecosystem Interoperability (CDEI) is able to combine assets from various domains into radical, disruptive or architectural innovation solutions. Novel applications may include modelling that combines data from different domains, simulations of data in novel use case scenarios, or comparison applications which aid procurement. The ability to build hybrid industrial innovations as part of a hyperconnected cross-domain industry system is unprecedented. The resulting incremental or breakthrough innovations are fully trackable and traceable through the entire ecosystem. For example, data sets from two different domains may be combined in a simulation of a novel use case. The third data set which is thus generated indicates new market possibilities. The system notifies the proprietor of the intellectual property used in the simulation and allows them to make an informed decision on any further investment. Thus, the system enables continuous learning and decision-making control for organisations, and a sustainable integration of cross-domain innovation in decision-making processes.

5.5.5 Future work

During the second half of the OntoCommons project, several routes to innovation will be drawn from the analysis of results from the OntoCommons Demonstrators. Areas of focus include: (i) best practices for expansion across domains; (ii) potential novel business models; (iii) the role of interfaces in supporting work with ontologies; and (iv) positioning of the OCES within the cross-domain data-driven landscape.

5.5.6 References

Ref 5.5.1 Open Innovation 2.0 is an initiative by the European Commission Open Innovation Strategy and Policy Group (OISPG), whose legacy is continued by its founders as part of the Industry Commons Foundation Steering Board.

Ref 5.5.2 Tangible User Interfaces (Ishii and Ullmer 1997; Holmquist et al, 2019)

Ref 5.5.3 IP tracking tests have been successfully completed in decentralised systems in 2017 (Magas 2018).

Ref 5.5.4 Effective system design takes a holistic view of data marketplaces that drive productisation, and allows for constant adaptation to environmental conditions.

Ref 5.5.5 Vernadat provides a comprehensive review of Enterprise Modelling covering four decades of its evolution leading up to developments in the context of smart manufacturing and Industry 4.0 (Vernadat, 2020).

Ref 5.5.6 Bousdekis and Mentzas present the state of the art in Enterprise Integration and Interoperability in the frame of Industry 4.0 (Bousdekis and Mentzas, 2021).

6. Acknowledgements

The road-mapping effort has been supported by the OntoCommons CSA project which has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No. 958371.

The number of contributors to the OntoCommons Roadmap is too long to be listed here, but particular thanks go to the main contributors of the OntoCommons-CSA project: TU Wien (Nadja Adamovic, Andreas Rauber, Florina Piroi), ENIT (Hedi Karray, Arkopaul Sarkar), Goldbeck Consulting (Gerhard Goldbeck, Alex Simperler), UNIBO (Emanuele Ghedini), University of Galway (John Breslin, Lan Yang), Trust-IT (Silvana Muscella, Cristina Mancarella), ICF (Michela Magas), UiO (Dimitris Kiritsis), UIBK (Anna Fensel, Umutcan Simsek), Fraunhofer (Janne Haack, Ebrahim Norouzi, Joana Francisco Morgado), UPM (María Poveda-Villalón), ATB (Correia Ana Teresa), other members from the OntoCommons Consortium, members of the OntoCommons Exter nal Advisory Board and external experts who participated and provided valuable contributions within a number of OntoCommons workshops.

7. Conclusion

This Roadmap summarises the outcomes from a number of events organised by the project in the first 18 months of the project. It represents the first version of the OntoCommons Roadmap that will be updated and re-published by the end of the project (in M36).

The Roadmap considers Needs, State of the Art, Gaps, Definition of Success and Recommended Actions for a number of topics contributing to an Ontology Commons Ecosystem for ontology-based data documentation grouped into:

1. Ontology Foundations: Top Reference, Middle, Domain and Application Levels
2. Integrated Development Environment (Tools) and Infrastructures
3. Industrial Impact including marketplaces, standardisation, education and human resources

The collected feedback from OntoCommons stakeholders in upcoming project events will contribute to the updated version of this document.