Semantic Knowledge Management for Materials:
the benefits of a FAIR data and model-based approach in industrial research and development

Nadja Adamovic[1], Francesca L. Bleken[2], Alessandro Calvio[3], Vikki Cantrill[4], Vinicius Carrillo Beber[5], Seán Courtney[6], Anders Eklund[2], Gerhard Goldbeck[4]*, Fajar J. Ekaputra[7], Luca Foschini[3], Jesper Friis[2], Emanuele Ghedini[8], Thomas F. Hagelien[2], Guntur B. Herwanto[7], Natalia Konchakova[9], Daniel Höche[9], Welchy Leite Cavalcanti[5], Alessio Mora[3], Michael Noeske[5], Florina Piroi[7,11], Peter Schiffels[5], Alexandra Simperler[4] and Sean Stijven[10]

[1] *Institut f. Sensor- und Aktuatorsysteme, TU Wien, Austria*

[2] *SINTEF AS, Norway*

[3] *Department of Computer Science and Engineering, University of Bologna, Italy*

[4] *Goldbeck Consulting Ltd, St John's Innovation Centre, Cambridge, UK*

[5] *Fraunhofer Institute for Manufacturing Technology and Advanced Materials IFAM, Bremen, Germany*

[6] *Computational Modelling Cambridge Ltd, Cambridge UK*

[7] *Institute of Software Technology and Interactive Systems, TU Wien, Austria*

[8] *Department of Industrial Engineering, University of Bologna, Italy*

[9] *Institute of Surface Science, Helmholtz-Zentrum Hereon, Germany*

[10] *DataStories International nv, Belgium*

[11] *Center for Research Data Management, TU Wien, Austria*

ONTOTRANS

* Corresponding author: gerhard@goldbeck-consulting.com

Contents

1. Introduction

Semantic data and knowledge management lies at the heart of the web-based applications used by most people in their everyday lives. The semantic web (also known as Web 3.0), the brainchild of Tim Berners-Lee, enables machines to analyse items they find on the web based on content-related data and metadata, and moreover, the relations between them. These changes turned simple connections between documents — when only a web address was known — into a web of data with meaning. It is this semantic web that enables internet searches to come up with not only potential documents, but also with a range of data. For example, a search for a person that is an author will recognise them as being an author and display a lot of related information about them. Hence, any "data point" becomes part of a web of concepts and not just data. The technology that manifests this web is a knowledge graph. Knowledge graphs are a core technology for the web and for many applications that work with similar principles, like the operations of Google[1], Amazon[2], and Airbnb[3].

One of the primary advantages of applying semantic web technologies to manage data is the generation of an overarching, flexible data model that covers all the concepts. Such a system can contain and manage all available data sources; and, once implemented, costs less, saves time and improves data quality and maintenance. In addition, it is easy to incorporate new data and information into a semantic knowledge management (SKM) system relative to traditional technologies, such as relational databases. Consequently, industrial, systems-engineering solutions are expected to transition from traditional document-based practices to data-driven, model-based ones [1,2].

The formal conceptualisation of the data model underlying a knowledge graph is an ontology. In information science, for any given entity, an ontology can be based on the emerging metadata, concepts and relations that best describe a given set of data, and/or be based on foundational concepts, which provide a framework into which the specific entities and relations can be integrated.

Semantic data and knowledge management leads to a 'model of the knowledge area', which is much more than data management. It can be used to model operations, business processes, comprise rules and standards, and support complex queries about data.

These capabilities have led to the increasing use of semantic data management and knowledge graph technologies in many areas of business, for two major reasons: (a) the requirement to connect data from a range of different sources that cannot easily be integrated with traditional database techniques, (b) a need to capture and share semantic meaning, business knowledge and best practice, and deploy business operations. Benefits include efficiency, reliability, better decision making, less reliance on individual-person knowledge, and compliance.

SKM approaches have been used in a variety of situations to bring about benefits to industrial environments within manufacturing industries. Real-world examples of the benefits of adopting SKM include:

- The use of standardized ontologies to significantly simplify a data integration problem associated with microchips at the Bosch Salzgitter factory[4]
- The use of ontologies to support trade-off decision making, represent domain and process knowledge explicitly, and improve traceability of the decisions during product design and

[1] https://www.google.com/
[2] https://www.amazon.co.uk/
[3] https://www.airbnb.co.uk/
[4] https://ontocommons.eu/demonstrators/bosch-manifacturing

assembly processes. Such an approach helps to overcoming interoperability and data standardisation concerns at Airbus[5]

- The use of ontologies and reasoning to improve flexibility in manufacturing and automation engineering. Here, SKM systems aim to increase transparency and trust of AI systems by using methods for explaining AI-based decision-making at Siemens[6]
- Det Norske Veritas used ontology-based management technologies for requirements, designs, and physical assets, and semantic technologies for exchange and reasoning services during construction of the Johan Sverdrup Drilling Platform situated in the North Sea[7].

A report by Feng et al. [3] identified that increasing variety and complexity of product lifecycle applications as a driver for change in manufacturing industries. They noted that progress in smart manufacturing has been hampered because of a lack of mechanisms to integrate, share, and update domain-specific knowledge. However, as industries work to ensure they are competitive and innovative, SKM can now look to provide deeper, semantic data models to enable this. SKM combines diverse content into a single, connected data source. A single source is more transparent, so it is quicker for researchers to find and use relevant information.

On a strategic level, SKM makes it possible to uncover insights within the data to better inform business decisions and helps to retain and reflect enterprise-level knowledge. By moving to SKM, businesses move towards systematic innovation management, which enables oversight of knowledge management tasks and workflows, and provides additional opportunities for collaboration.

Although research environments require data integration and knowledge capture and can, in principle, benefit from SKM, they also bring further complexities that need to be overcome for SKM use to be successful. In materials research, these challenges include the wide range of material types, the multi-scale nature of materials where all structural levels may affect properties, the strong process(ing) dependence of materials structure and properties, which means that materials cannot simply be defined by their composition or atomic structure. In research, numerous methods are used to investigate and characterise materials, and in addition to physical characterisation and testing data, there are a range of materials models that play a role.

In engineering, due to challenges from light-weight construction, function integration and versatile joining of pre-manufactured polymeric materials are advancing along with constituents like fillers or additives. This challenge requires attention to detail with regards to materials composition and characterisation through careful control and documentation during the manufacturing process. Here, decision makers depend on well-defined requirements and performance criteria. Onward manufacturers and end-users also need to be informed and face the challenge of managing the material’s lifecycle.

The field of materials science covers many different research domains, each supported by their own community, that use highly specialised terminology. Sometimes, terms have a specific meaning that results in an incompatibility because the same term may be used in other domains with differing meanings. Research processes, procedures and types of data may change more frequently than they do in other parts of the business, so the cost of implementing and maintaining SKM may be seen as prohibitive. Also, despite some advances in the last decade, a widely agreed set of data models and ontologies for materials sciences that can be readily adapted by enterprises is lacking. Hence, each

[5] https://ontocommons.eu/demonstrators/airbus-design-and-manifacturing
[6] https://ontocommons.eu/demonstrators/siemens-digital-manufacturing
[7] https://sirius-labs.no/wp-content/uploads/2017/04/SubseaValley2017.pdf

organisation going down the SKM route was faced with a steep initial investment. In contrast, other fields, including oil and gas industries, bioinformatics, pharmaceutical analytical techniques and financial services have shared the development burden collaboratively.

In this paper, we will further discuss the benefits for materials industries to adopt SKM into its research and development (R&D) functions and beyond, and demonstrate how SKM goes hand-in-hand with the FAIR data principles and their recent extension to FAIR 2.0 [4]. In this context, we present the OntoTrans approach of combining SKM with a data- and model-based approach to addressing complex innovation challenges. The approach translates innovation cases into semantic models that, in turn, are connected to a multitude of data sources and models to arrive at a system in which data and knowledge about innovation cases can be queried by research scientists, engineers and managers.

2. Benefits of Semantic Knowledge Management in a Materials R&D Context

Materials R&D involves a wide range of knowledge, materials processing technologies, characterisation and modelling techniques. In the context of an increasing and urgent emphasis on industrial resilience, materials safety and sustainability, the range, sources, complexity of data and levels of understanding required to bring materials into products on the market is rising steeply. This means that more traditional ways of gathering data and information from various sources, including different databases and different experts and their interactions, are no longer adequate [5].

Integration of different data with their meaning, similar to that of the semantic web offers numerous opportunities. It can make inherent knowledge explicit, machine readable and actionable, converge on common vocabularies and concepts, and thereby integrate data from a wider range of sources. Also, the ability to capture and model procedures and organisational knowledge about materials and their processes is needed to keep track and optimise developments. Consequently, knowledge management within materials research, manufacture and engineering is a rapidly developing area.

Industries often have an array of data sources that they use from classic physical experiment data to computer simulations. The ability to intelligently work with this data and garner knowledge and understanding from it is key to a business remaining competitive. Yet, if a multitude of large data sources could be integrated, this could provide a more powerful way to inform business-critical decisions. For example, integrated data could extend product lifespans based on a combination of usage data and materials models or be related to the safety and sustainability of potential chemicals, materials and manufacturing choices. Materials science is an area that, in particular, stands to benefit from the integration of disparate data sources into a single query system.

By using semantics, it is possible to explore complex data more readily and uncover previously unknown relationships, which in turn supports multi-criteria decision making. For example, Wang et al. highlight that data models and knowledge management for digital twins and cyber-physical systems are central to intelligent decisions [6]. Another example comes from failure analysis as described in the materials design and production guideline VDI 3822 "Failure analysis – Fundamentals and performance of failure analysis".[8] The document outlines how failures can be documented and analysed. Failure analysis provides knowledge that can be applied in quality control to help avoid and

[8] https://www.vdi.de/en/home/vdi-standards/details/vdi-3822-failure-analysis-fundamentals-and-performance-of-failure-analysis

prevent failures and facilitate innovation. Ontologised knowledge in this instance is helpful because a thorough understanding of failures can assist in many areas of a business.

Ontologies also make data extraction and integration easier, for example literature sources or external data can be more easily integrated with internal, perhaps proprietary, data. Over time, by consistently mapping data to ontologies or shared semantic models, it becomes easier for industries to manage and curate information exchange between departments throughout the value chain. In turn, changes and concomitant harmonisation foster mutual understanding and enable new applications, products and innovations to develop faster and reach the market sooner. Furthermore, ontologies and knowledge graphs can be used to understand and express both scientific and organisational knowledge, including corporate memory.

Overall, SKM systems enable technically less-skilled users to quickly find, access and integrate data outside of their day-to-day line of work, for example from other departments or organisations. It also enables subject-matter experts to extend their remit, contribute to data modelling processes and work efficiently. As a result, use times are reduced significantly by 30% for data integration and deployment and data maintenance by 70%[7]. SKM also significantly reduces data search times in a way that is scalable and sustainable, such that all new projects that provide or require data can be easily incorporated.

Semantics provides information classification, interoperability without requiring expertise in syntactic standards, reasoning and retrieval.

Semantic data management works when…

- Research may generate new data facts/factors
- The data being managed is added to or changes
- Data evaluation criteria change regularly
- Cross-organizational collaboration takes place
- Diverse data sources need to be queried together
- There is a lack of industry data standards.

2.1. Addressing Semantic Challenges in Materials

Although semantic models and ontologies could bring about industrial benefits, there are significant challenges. Although numerous knowledge graphs and ontologies exist in different disciplines, one challenge is that the terminology used by researchers on this area lacks cohesive and explicit scientific definitions.[9] This means that researchers in different communities use a variety of terms and descriptors that are not clearly defined to others, which highlights why semantic models could be of significant benefit. Moreover, researchers, manufacturing engineers, managers and product users may intrinsically approach the same challenge and the involved materials and processes from different perspectives and with different intentions.

[9] https://eosc-portal.eu/eosc-interoperability-framework

Consequently, a review of domain-level ontologies within materials science to align descriptors across communities is needed to enable innovation. In 2018, several practitioners within materials science under the governance of the European Materials Modelling Council (EMMC)[10] expressed the need to develop a knowledge framework consistent with scientific principles and methodologies to complement the existing physical-mathematical approach. As a result, the Elementary Multiperspective Material Ontology (EMMO)[11] was developed. Such an overarching conceptualisation framework can aid the creation of well-developed, fit-for-purpose ontologies to integrate with diverse academic and industrial data sources. Here, the EMMO provides a solid ontology base that is open to all.

The development of EMMO in different projects is shown in Figure 1.

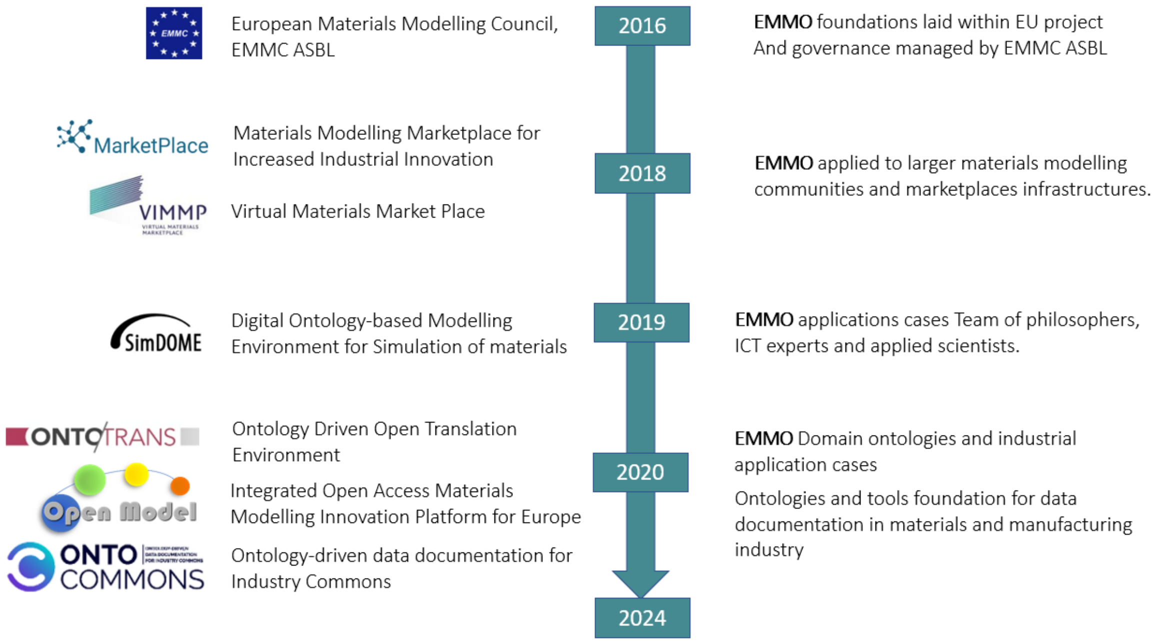

Figure 1.Timeline for the development of EMMO

To date, several European Union projects have incorporated EMMO; SimDOME[12] used and developed EMMO for application cases. Two materials modelling marketplaces, MarketPlace[13] and VIMMP[14] applied EMMO to infrastructures. In the open innovation platform project, OpenModel[15], EMMO-based ontologies both facilitate semantic interoperability for success stories and enable seamless integration with materials modelling marketplaces, whereas OntoCommons[16] uses EMMO as the foundation for data documentation in the materials and manufacturing industry. In OntoTrans, we created EMMO-compliant domain ontologies for industrial application cases. Furthermore, EMMO is

[10] https://emmc.eu/
[11] Formerly European Materials & Modelling Ontology, https://github.com/emmo-repo/EMMO
[12] https://simdome.eu/, https://cordis.europa.eu/project/id/814492
[13] https://cordis.europa.eu/project/id/760173, https://www.the-marketplace-project.eu/
[14] https://cordis.europa.eu/project/id/760907, https://www.vimmp.eu/
[15] https://cordis.europa.eu/project/id/953167, https://open-model.eu/
[16] https://cordis.europa.eu/project/id/958371, https://ontocommons.eu/

used as top-ontology for domain-ontologies in ongoing projects, such as OpenModel, VIPCOAT[17], MatCHMaker[18] and CoBRAIN[19]. The most advanced EMMO-based ontology that is being used and further developed in new projects is the BattINFO[8][20] ontology developed as part of the BigMAP[21] project, which comprises the domains "battery"[22] and "electrochemistry"[23]. If ontologised knowledge is available for a product, then onward innovation is facilitated.

3. The OntoTrans Approach

The OntoTrans[24] project started through funding from the H2020 call DT-NMBP-10-2019 – Adopting materials modelling to challenges in manufacturing processes,[25] which had a focus on development of novel materials and products that "*should allow reuse of materials modelling software, knowledge and expertise in different industrial domains, by use of the models, protocols and systems in other relevant areas or sectors*."

OntoTrans aims to advance SKM within materials science and in particular aims to:

- Facilitate the capture of curated knowledge in an EMMO-based ontology
- Provide reasoning, recommendations and knowledge graphs to support data exploration
- Present models that use 'live' simulation data to demonstrate the potential of SKM systems, in particular coupled to machine learning and artificial intelligence applications
- Demonstrate the benefits and flexibility of SKM through industrial use cases.

OntoTrans is a portmanteau of ontology and translation. Here, ontology refers to the formalisation of concepts, their meaning and interrelations, i.e., materials, processes, properties, models, and characterisation. For this purpose, OntoTrans contributed to the EMMO foundations and developed domain and application ontologies for industrial cases. **Error! Reference source not found.**Translation i n materials modelling, within the context of the EMMC, is the process of converting an innovation challenge into questions to be solved through the use of modelling and simulations tools [9]. Translation in OntoTrans involves setting up templates to capture the innovation challenge beyond materials modelling and include other aspects, such as industrial data, live quality-control information, and customer expectations and feedback.

OntoTrans provides a general-purpose ontology-based Open Translation Environment (OTE) that supports the development of dedicated applications that deliver smart guidance for materials producers and product manufacturers and enables their semantic knowledge frameworks. OntoTrans emphasises the use of data-based (ML/AI) modelling and demonstrates the wide range of data sources that may be required for decision making in R&D as well as during the product lifecycle. Different application cases (see Section 4) integrate datasets from materials characterisation, process monitoring, as well as product performance in the market.

[17] https://cordis.europa.eu/project/id/952903, https://ms.hereon.de/vipcoat/
[18] https://cordis.europa.eu/project/id/101091687, https://he-matchmaker.eu/
[19] https://cordis.europa.eu/project/id/101092211, https://www.cobrain-project.eu/
[20] https://github.com/BIG-MAP/BattINFO
[21] https://cordis.europa.eu/project/id/957189, https://www.big-map.eu/
[22] https://emmo-repo.github.io/domain-battery/
[23] https://emmo-repo.github.io/domain-electrochemistry/
[24] https://ontotrans.eu/, https://cordis.europa.eu/project/id/862136
[25] https://cordis.europa.eu/programme/id/H2020_DT-NMBP-10-2019

3.1. From Innovation Challenge to Semantic Data Models

OntoTrans provides a methodology for research organisations to capture their innovation challenges in the form of semantic data models that are connected to a range of data sources and models in such a way that the researchers and managers can more easily and quickly interrogate different scenarios. The process, based on the translation steps [9],is made up of several levels (Figure 2). At the outset, it is essential to detail the innovation challenge, to fully understand the problem and what the aims of the end-user client are. This task requires a team of translators skilled in materials modelling, data-based modelling and ontology-based data documentation and knowledge management, cf. the materials modelling translator role [9] and semantic competencies from the knowledge management translator role [10].

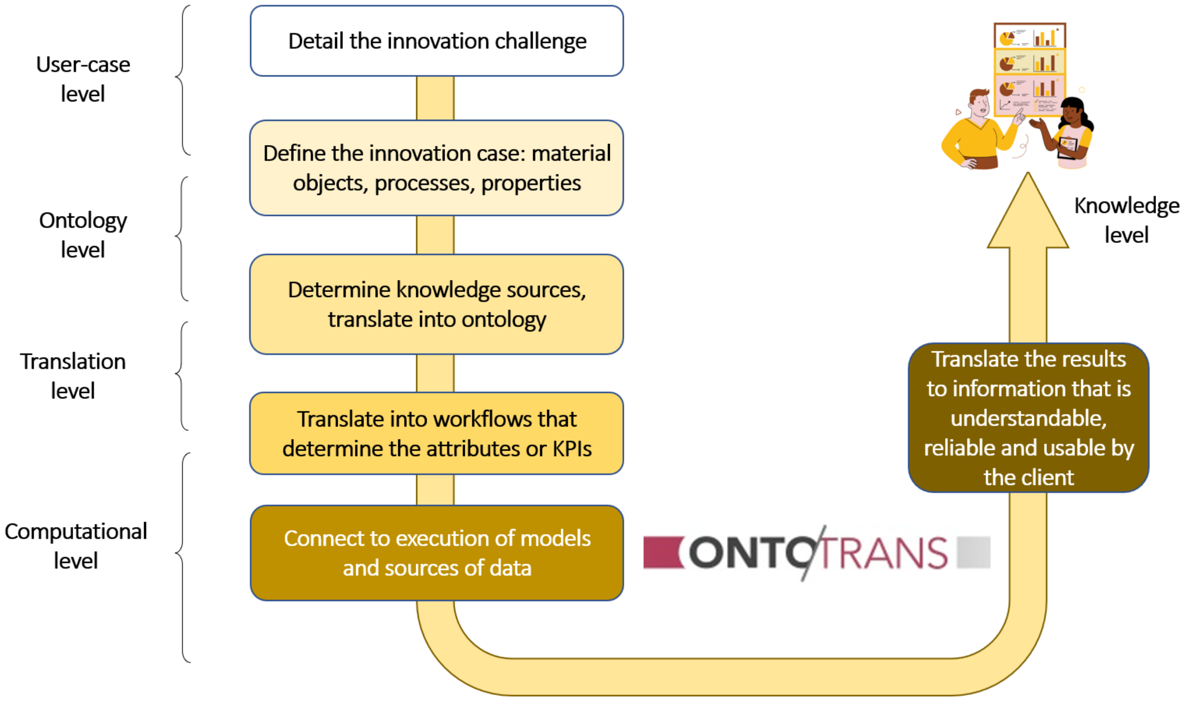

Figure 2. The steps from innovation challenge to knowledge.

Before moving from the *user-case level* to the *ontology level,* care must be taken when defining the innovation case to capture all the relevant conceptual details of the material objects, processes, their properties and relations among each other and with stakeholders involved. Should any of the details gathered not be covered within the existing ontologies, then these will need to be added, which requires time and additional expertise. OntoTrans offers tools to support that additions and data integrations are made in a consistent manner. OntoTrans is expected to be a lasting and evolving addition to materials or manufacturing organisations, rather than implemented for a single business need, therefore additional effort to conceptualise the innovation case and formalise related vocabularies and ontologies is repaid through its reusability.

Next, all internal and external knowledge sources are identified and incorporated. All added data must be addressed and handled in an ontology-compliant way, so some time and effort are needed to ensure that. Once all data has been collated, the project moves on to the *translation level* to define the workflows that determine the property-related attributes or key performance indicators of the innovation case. Here, workflows should be reported as a Modelling Data template (MODA)[26], which is a text and graphical workflow template.

Last is the *computational level,* in which models are assigned and sources of data are connected to the chosen workflow. At this point, the client can access new knowledge as reliable and easy-to-use information.

3.2. Innovation Case Ontologies in OntoTrans

A major step in our methodology is the development of the semantic model, i.e. the ontology, which captures the innovation case. It needs to be constructed in such as a way that it can address the potential queries users may have about the innovation case, so users can make better and more informed decisions.

Concerning the process of ontology development, a translator must be skilled in working with domain and ontology experts in transforming real world entities into an ontology-based representation that captures the requirements of the innovation case[27]. EMMO may be considered a language that is mastered by the experts involved in an ontology-based translation process to express their knowledge and their data understandable for machines ("machine readability"; Figure 3). In case they are guided by ontology-affine translators [10], it will be sufficient for industry experts to understand this language and read (rather than write or edit) statements expressed in the both human and computer-readable EMMO.

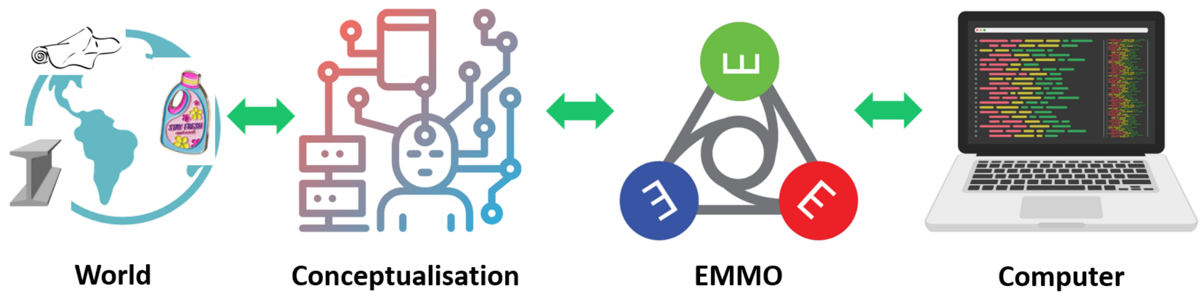

Figure 3. The path from real world to ontological representation on a computer.

However, people often generate several conceptualisations for the same entities. Hence, they need to agree on a common taxonomy. In OntoTrans, definitions are recorded based on categories provided by EMMO and EMMO-compliant domain-specific ontologies, an approach that supports consistency and its maintenance in further development of the knowledge system. EMMO also facilitates semantic representations based on semiotics – a communication-related approach that explicitly uses an

[26] https://emmc.eu/moda/
[27] https://ontotrans.eu/wp-content/uploads/2023/09/2.Emanuele_Ontology_OntoTrans_Workshop_2_Ghedini.pdf

interpreter to assure that real world occurrences are described correctly with an ontology. EMMO is designed in a way to be systematically expandable, so all real-world materials and manufacturing processes can be captured and documented in a persistent and FAIR way [5].

As described above, the OntoTrans project has been using and further developing the EMMO for this purpose. With EMMO, it has been possible to describe the multiple perspectives of the relevant entities, for example their object/process nature, or the roles they play, their material/physics-based aspects, the way that properties are determined, and the representation of data.

For the application of EMMO to industrial cases, a conceptualisation methodology has been developed, which is available through the EMMO wiki.

Key first steps include the identification of the objects and processes involved, and their relationships in terms of parthood and time/causality. Examples of parthood include objects that overlap other objects or processes, and examples of time/causality is a process following and being dependent on a previous process finishing, or a process producing an output.

A simple diagram of the object process relationships is shown in Figure 4.

Figure

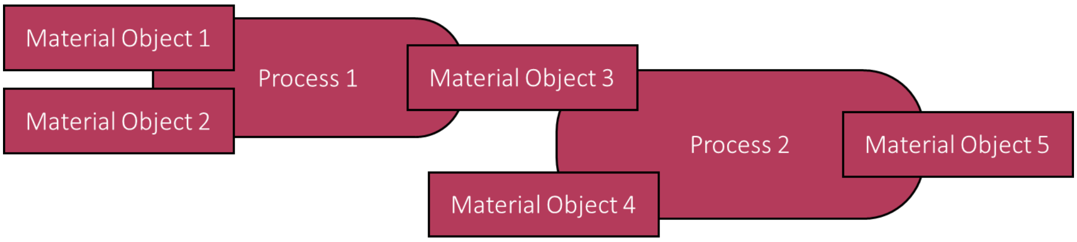

Figure 4: A simple workflow diagram of an innovation case that identifies all the object and process entities that play a role in the innovation case to be referenced by properties and models. Time/causality progresses from left to right.

Next, the properties of objects and processes and how they are determined come into focus. Here, the semiotic framework of EMMO is used, as shown in Figure 5. In particular, **emmo:Semiosis** can be used to describe the many ways in which properties are arrived at, be it by assignment, simulation, or measurement. It also makes the relationship between the object, its properties and the underlying declarer (e.g. person or instrument) of the properties and the actual data transparent.

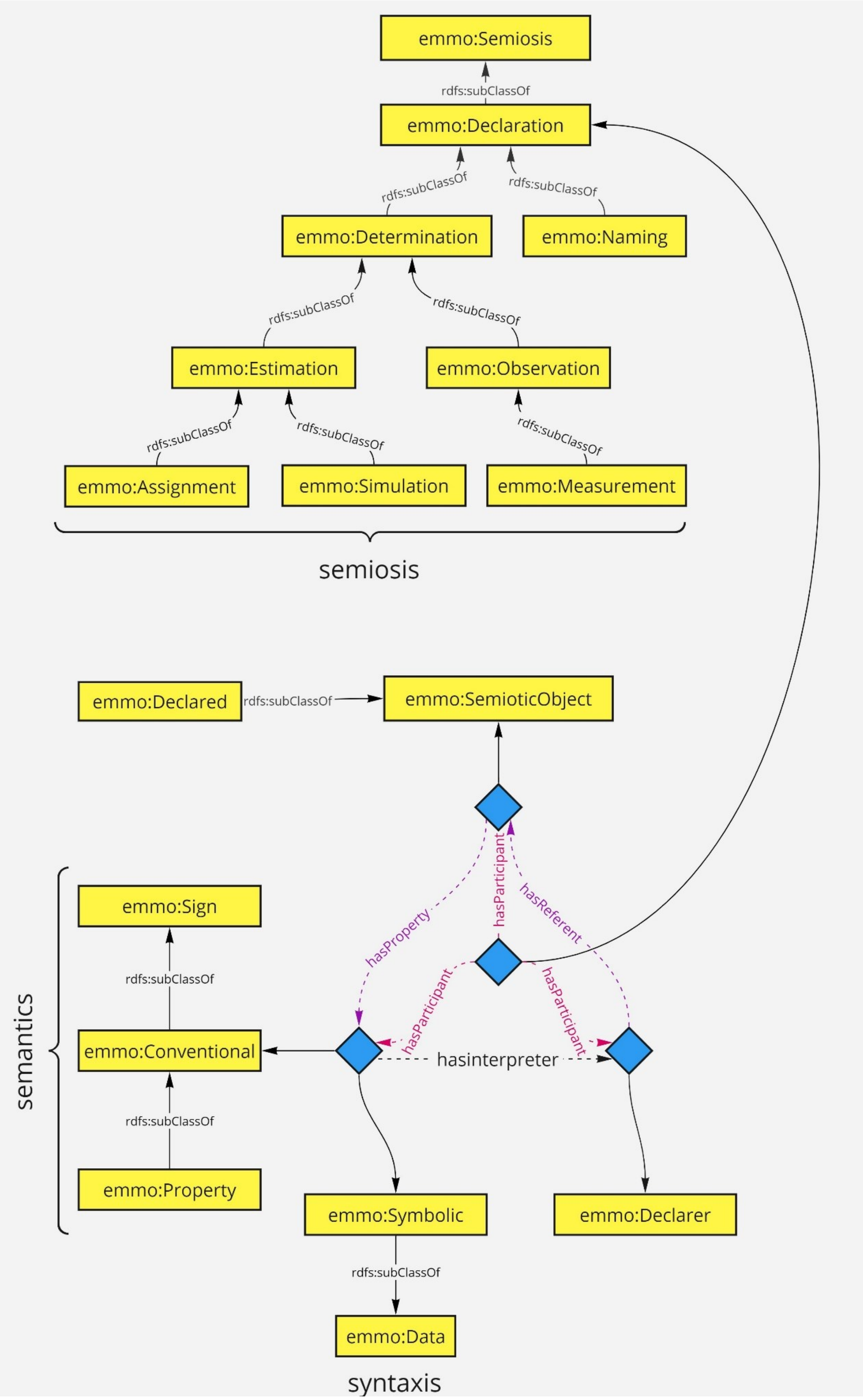

Figure 5: Semiotic representation of the property declaration for a generic object. Note that the yellow boxes depict classes in the ontology, the blue diamonds represent instances of the classes to which they are connected by a solid arrow.

3.3. How the OntoTrans OTE Works

A number of software applications are required for the OntoTrans OTE to work. Therefore, we will provide more details about the tools that were developed in OntoTrans[11].

The OTE can be conceptually divided into several parts that handle data provided by a user and aggregate knowledge such that understandable and useable results are available to the user (Figure 6).

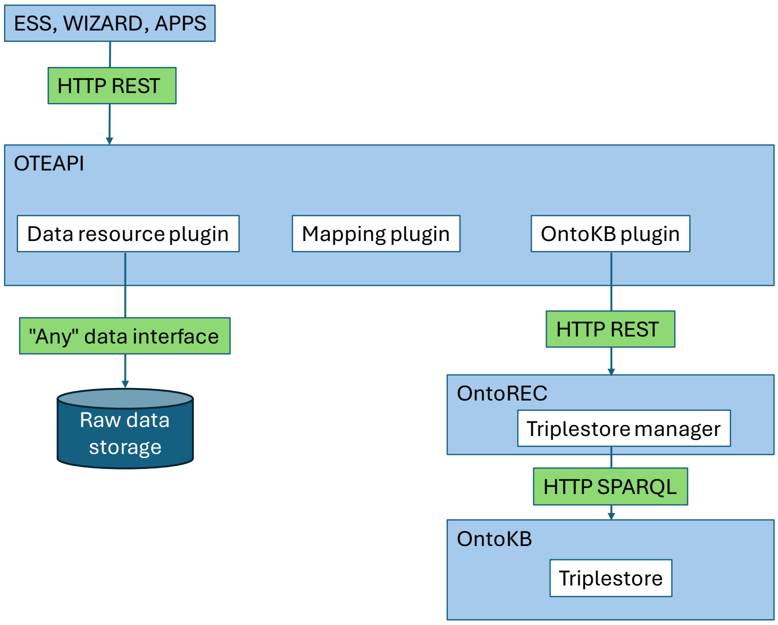

Figure 6. Structure of the Open Translation Environment in OntoTrans.

The user may be working with non-ontologised, and often low-level organised information, such as disconnected lists of materials properties and business data. Data generated in research environments is often at a low level of semantic interoperability; perhaps non-structured or structured according to non-standard classifications. Here, good cooperation between data owners and the translator is crucial because the OntoTrans OTE system requires data that is ontology-ready. OntoTrans provides interfaces that permit the transfer of data into the knowledge base system, OntoRec and OntoKB.

OntoKB and OntoRec are at the technological heart of the OTE. These are two (logical) components, whereby OntoRec works as an application programming interface (API), and OntoKB handles the knowledge base. Thus, the "single source of knowledge" is the OntoKB, which hence supports the interoperability (see also Figure 9). OntoRec ensures consistency of the incoming knowledge entities and will infer relations between new and existing entities through reasoning. To keep the overall system running, the interfaces must effectively populate OntoRec and OntoKB with all relevant, available data. A relatively simple case has been constructed by OntoTrans to demonstrate how the OTE works[28]. More technical detail is described in the following section.

[28] https://github.com/EMMC-ASBL/OTE-demonstration

3.4. OntoRec and OntoKB – The Heart of OntoTrans

The OntoTrans systems comprises two, (logical) components, OntoRec and OntoKB. OntoRec, which is the main way to interact with the system, facilitates, standardises, and controls user-driven access to the knowledge base through a set of APIs (Figure 7). Similarly, a SPARQL endpoint may be used to access OntoKB. OntoRec Proxy is the software part of the recommendation system that implements the OntoRec APIs. It may also embed historical data and data that reflects corporate memory to facilitate interaction with the OntoKB or to not query it.

OntoKB leverages a resource description framework (RDF) triplestore database (through semantic triples composed of subject, object and predicate) and its native reasoning engine. In addition, an external tailored reasoner is used and integrated inside the OntoKB component to enable customised levels of expressivity in conjunction with the triplestore and the semantically presented datasets. This means that the system can support the required level of reasoning without any detriment in all-round performance.

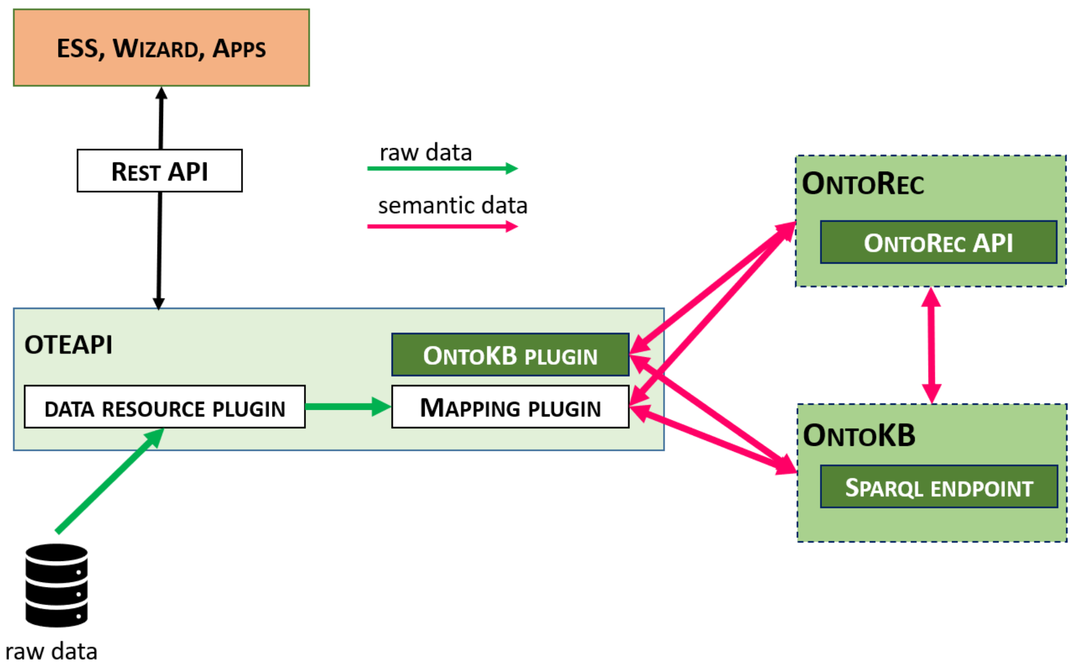

Figure 7. Structure of the OTE. The Wizard guides the user through an application case. The ESS provides different perspectives on the knowledge graph stored and managed in the core components of the OTE (OntoKB, OntoRec). The red arrow connected to the SPARQL endpoint indicates limited and controlled access to the OntoKB module.

OntoRec is a form of AI[29] (in this case an intelligent system that uses machine learning and semantic reasoning) similar to that which is widely used in web-based applications to provide personalised content to users by e-commerce companies. Establishing similarities as defined by the user, content-based filtering is used by companies such as Amazon and Google to analyse the behaviour of users to find similarities (collaborative filtering). In contrast, OntoRec focuses on content-based analysis by using the EMMO Web Ontology Language (OWL)-DL ontology as the knowledge source to provide

[29] https://medium.com/oxford-semantic-technologies/ask-the-oxford-professors-the-interplay-between-machine-learning-and-semantic-reasoning-f61f14f47be

information about connections. Inferences will be obtained by OntoRec by means of reasoning tools that are already widely used in ontology-based systems and ontology development.

Practical implementation of an OWL-based recommendation system relies on existing tools, such as OWL API (Java), Owlcpp (C++) or Owlready2 (Python), that will parse existing OWL ontologies and provide API for querying and reasoning.

OntoRec is devoted to user-case definition by taking into account ontological constraints, such as "Si is not part of a water molecule" and "viscosity is not a property of a solid body". User-case definitions can also be used as profiles to model end-user preferences and log interactions between end users and OntoRec. OntoRec infers relations between the declared, ontologised user case and existing knowledge stored in the OntoKB (including new user cases as part of the knowledge base), to provide suggestions to users about potential solutions f a given manufacturing challenge, e.g. by suggesting models suited for a particular material, by providing a list of software that can implement such models, by suggesting property databases to feed models or by suggesting similar user cases that the current user could learn from. OntoRec primarily uses forward-chaining inference, i.e., to perform inferences based on known facts provided by a user (explicit statements). OntoRec can also provide broader inferences, in the form of implicit statements whenever new explicit statements are added to the OntoKB.

OntoKB uses an RDF triplestore, to store the semantic data model (ontology) and metadata. An RDF triplestore is a semantic graph database that stores semantic-based data as a network of objects with links, so that data are always interconnected. The links improve the flexibility of these types of databases with respect to traditional relational databases. Such technologies overcome the typical limitations of data reuse and integration based on proprietary APIs by using consumer-friendly open data to ease integration between various sources of data of interest to the consumer. The RDF triplestore architecture will support powerful semantic queries and inferences to build new knowledge and relations from existing information. Alongside the RDF, OntoKB also defines an interrogation language for the data called SPARQL Protocol and RDF Query Language (SPARQL).

There are a wide range of knowledge graph systems available for the RDF triplestore. We utilised Stardog[30] in this instance because it has characteristics well suited to the backbone for OntoKB. In particular:

- It can store ontological data, perform reasoning on it and provide SPARQL endpoint and querying functionalities
- It natively handles RDF and ontologies, and provides native SPARQL support
- It supports all OWL profiles, including owl-dl.
- It provides convenience features for developers and general users alike, including a user-friendly web interface to administer, check and work on managed repositories
- Additional software components (e.g., OntoRec Proxy) can be developed in Java with any RDF-handling library
- It provides support for semantic web rule language
- It comes with commercial licensing and support.

[30] https://www.stardog.com/

3.5. Connecting to Data from Different Sources

OntoTrans works with a wide range of data sources and connects to models that can generate data on demand. Typically, input data and the data elaborated by OntoTrans are NOT stored in the OntoKB triplestore. Instead, the OntoKB contains the metadata necessary to identify and use raw data from various data sources.

For the purposes of connecting different data sources with data consumers, OntoTrans includes the OTEAPI package[31], which can be described as an actionable data documentation system for semantic interoperability. OntoTrans also functions as an extendable library and framework for execution of software plugins for the operations of data access, processing and storage with a range of data formats and communication protocols. The OTEAPI plugin-based architecture facilitates the implementation of various data interfaces in one place, rather than them being spread out and duplicated across all of the OntoTrans software components. The plugin-based system also facilitates the connection of multiple plugin actions in data processing chains, referred to as data pipelines. The pipeline configurations are serialised into triples for storage in the OntoKB.

The creation (i.e., configuration) of OTEAPI pipelines provide a simple interface for data providers to document their data according to the first three levels of the four possible data documentation levels (Figure 8). These levels include:

- **Cataloguing** is a level of data documentation that enables basic findability and accessibility. Cataloguing in OTEAPI requires documenting how the data can be accessed by using terms from the dataset catalog vocabulary (DCAT)[32], like downloadUrl, mediaType, accessUrl and accessService. Basic searches of the data are also possible by allowing standard terms from the Dublin Core vocabulary[33], like title, abstract, subject, creator, publisher, and licence.
- **Structural documentation** describes how the data is structured and represented numerically. In OTEAPI, this is documented by referring to a formal data model. By using the DLite interoperability framework[34] developed by SINTEF, this data model can be instantiated to provide a complete representation of the dataset. Structural documentation is essential for interoperability at a numerical level.
- **Semantic documentation** enhances the structural documentation by adding shared meaning to the data models. In OTEAPI, this is done by relating the properties in the data models to ontological concepts through so-called mappings. Understanding the meaning of data is essential for reusability and semantic interoperability. Semantic data documentation also enables data exploration through ontologies.
- **Contextual documentation** relates the data with other resources by using semantically well-defined relations from ontologies. This step is key for linked data and enables contextual search and improves reusability. Contextual documentation is not covered by OTEAPI because it does not contribute to interoperability. Instead, contextual documentation is provided by the OntoTrans wizard and can be entered directly into the OntoKB.

[31] https://emmc-asbl.github.io/oteapi-core/latest/
[32] https://github.com/w3c/dxwg/
[33] https://www.dublincore.org/specifications/dublin-core/dcmi-type-vocabulary/
[34] https://sintef.github.io/dlite/index.html

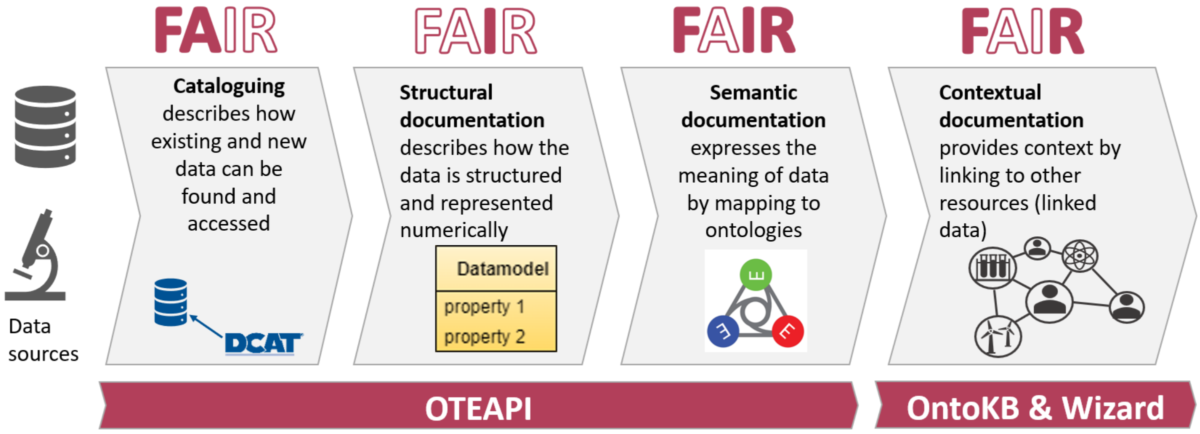

Figure 8. The four levels of data documentation. The aspects of FAIRness that each level contribute to is also indicated.

In OTEAPI, both the data provider and data consumer are expected to document the data source (where the data comes from) and data sink (who the consumers of the data are), respectively. By documenting a data sink, the form and representation of the desired data is specified.

OTEAPI uses the pipes and filters software design pattern, in which each of the three levels of data documentation are represented by dedicated filters. Figure 9 shows a pipeline that consists of two partial pipelines — the first documents experimental grain orientations (data source) and the second documents how the data consumer wants to receive it (data sink). Each of the partial pipelines are built from three filters and each filter corresponds to one level of data documentation. Partial pipelines can be mix and matched to provide flexibility between data providers and consumers. Another advantage of OTEAPI is that the data consumer does not need to know anything about how the data provider represents data.

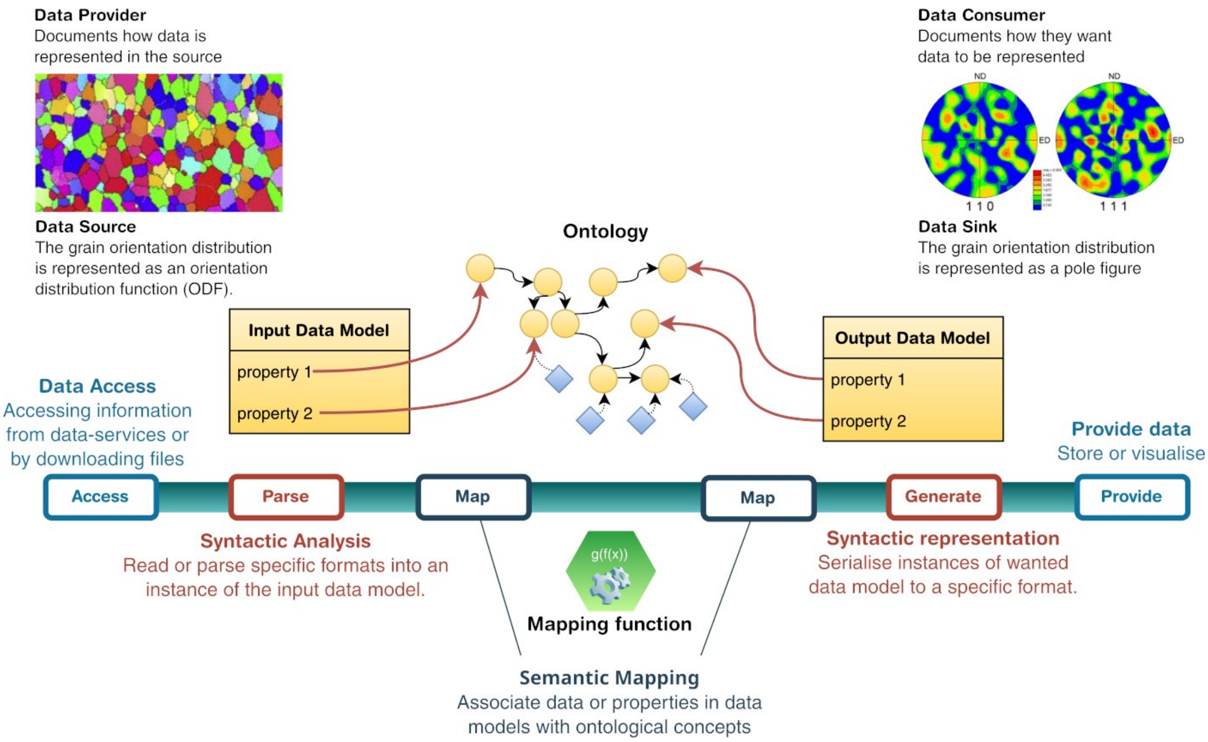

Figure 9. The different steps of data documentation in OTEAPI represented as a pipeline. The data provider documents the data source and the data consumer documents the sink, i.e. how the consumer wants the data to be represented.

OTEAPI is, by design, agnostic with regards to the choice of an underlying interoperability framework, although a choice must be made and implemented to achieve full interoperability. DLite is the interoperability framework used in OntoTrans — a light-weight semantic interoperability framework designed to be data-centric. This means that the low-level data documentation should be as close to the original data as possible to simplify parsing of the data source.

In DLite, data are described by *Instances* and *Instance* is described by *Metadata*, which in turn is described by a *MetadataSchema* based on a *BasicMetadataSchema*. The latter is self-explanatory and thus, no further meta levels are needed. This construction ensures a finite but flexible metadata structure. For the end users, who are responsible for and need to document their data, the *Metadata* and *Instance* concepts are crucial. This construction is best described by the example in Table 1, in which the left side shows an Instance of actual data with values, and the right shows the corresponding DataModel (i.e., its Metadata).

Instance (actual data)	DataModel (Metadata)
meta: http://onto-ns.com/meta/0.1/Molecule **dimensions:** • **natoms:** 2 • **ncoords:** 3 **properties:** • **name:** carbon monoxide • **positions:** [[0,0,0],[0,1,0]] • **symbols:** [C, O]	**uri:** http://onto-ns.com/meta/0.1/Molecule **description**: A minimal description of a molecule **dimensions:** • **natoms:** Number of atoms • **ncoords:** Number of spatial dimensions. Always 3 **properties:** • **name:** **type**: string **description**: Name of the molecule. • **positions:** **type**: float64

	shape: [natoms, ncoords] **unit**: "Å" **description**: Atomic positions in Cartesian coordinates. • **symbols:** **type**: string **shape**: [natoms] **description**: Chemical symbols.

Table 1. An instance of actual data with values and the corresponding DataModel

The DataModel is a reusable description of the data. It has a unique *uri* together with a description that is intended for human understanding. Furthermore, the dimensionalities or features connected to the properties are defined. The dimensions should be orthogonal. For instance, a molecule can be described either by the number of atoms or by the number of 3D spatial arrangements. Both are descriptions of the same molecule but each description provides discrete information. Finally, all properties are described, each has a name, type and description. If the dimension is not given, it is 1D by default. Also, optionally, the unit can be defined, e.g. Å or nm.

An important aspect of this framework is that that data can be described locally and independent of an ontology. Mapping the metadata to an ontology is, however, a requirement to reach full interoperability to tie OTEAPI and DLite together.

3.6. End-user applications: Wizard, Data Analytics and Exploratory Search System

The end-user applications and translator tools form the digital tools the client needs to tap into the structured knowledge.

In practical terms, OntoTrans, facilitates interactions between the user and the digital environment by means of a step-by-step wizard that guides the user through the workflows and to the data sources resulting from an ontologised translation process (Figure 10).

To set up the wizard, the profile, requirements and intentions of the user were identified based on an overarching translation approach that considered the views of experts in translation, ontologies, the respective (material science) domain or (industrial) application. Intuitive interactive use of the OTE is facilitated by means of a correspondingly designed graphical user interface. From a diverse set of innovation challenges detailed in the OntoTrans project, the user can select an illustrative scenario. In this case, as shown in the bottom part of Figure 10, a process optimisation has been selected – the manufacture of composite prepregs[35,36]. The user clicks on the Composite Prepreg tile and will select to register data sources that can then be used to query properties, such as porosity or mechanical properties.

[35] https://ontotrans.eu/wp-content/uploads/2022/03/01-PRO-S3.1-Beber-PUB-min.pdf
[36] https://ontotrans.eu/wp-content/uploads/2023/09/7.APP3_OntoTrans-Workshop.pdf

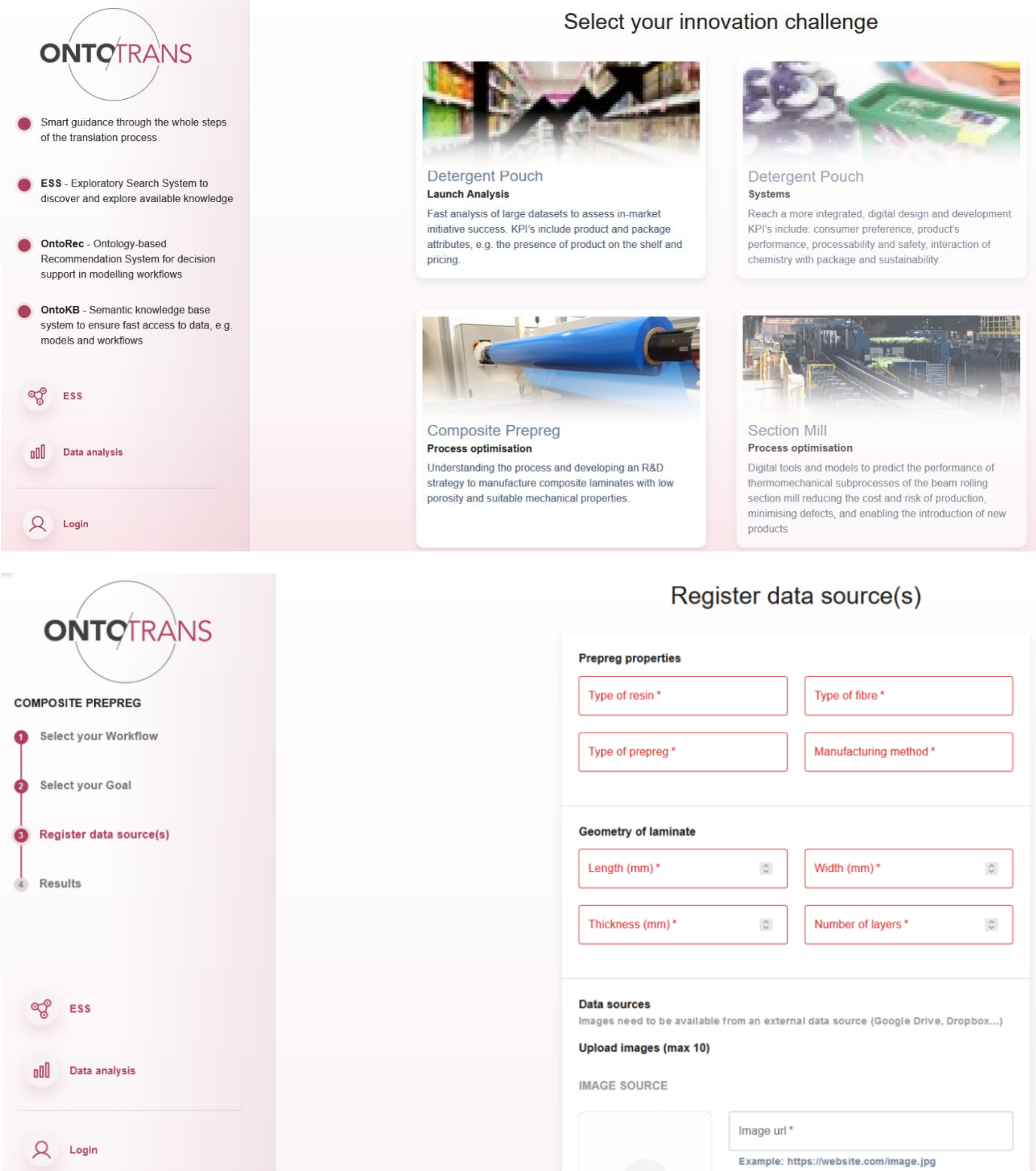

Figure 10. A prototype OntoTrans wizard illustrating the prepreg innovation case.

The toolbox accessible from the Wizard includes Data analytics tools to delve into existing data and generate knowledge in real time. In OntoTrans, these include the Augmented Analytics Software from DataStories[37] and the Model Development Suite (MoDS) from CMCL Innovations[38].

The creative side of the innovation process can lead researchers to learn about or investigate a topic and discover new information in serendipitous way. Traditional search systems rely on query-response

[37] https://datastories.com/products
[38] https://cmcl.io/software-suite/mods/

tasks that work well for information needs for which the researcher is able to formulate a precise question.

Exploratory Search Systems (ESSs) enable information seeking tasks, such as learning and investigating, through increased interaction between user and search system. Here the user is more actively engaged in the search process. Consequently, ESS user interfaces are more advanced and capitalise on new technological capabilities. The OntoTrans ESS[39], applied to the material science domain, integrates analytical services for knowledge graphs within that domain.

An ESS for materials and manufacturing [12] should permit users to choose between various perspectives on a materials and manufacturing knowledge graph, and thus permit multi-perspective explorations. It must permit users to identify the original data source for certain properties (data provenance) and enable users to navigate through the deep hierarchies present in materials and manufacturing knowledge graphs.

Figure 11 describes how the user arrives in OntoTrans at an ESS Instantiation. For an innovation case, (experimental) data must be collected and prepared to be fit for serialisation into an RDF format, which is then used and processed by an ESS instance to be ready for exploration.

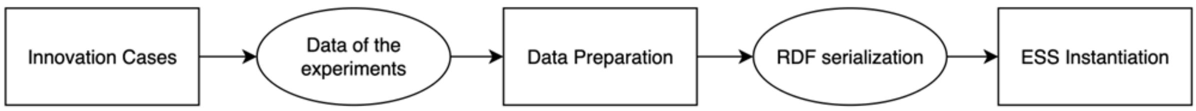

Figure 11. Workflow from innovation case to ESS.

[39] http://data.ifs.tuwien.ac.at/essdoc/

The landing page for the ESS (Figure 12) was designed for users that are working with pre-pregs[35].

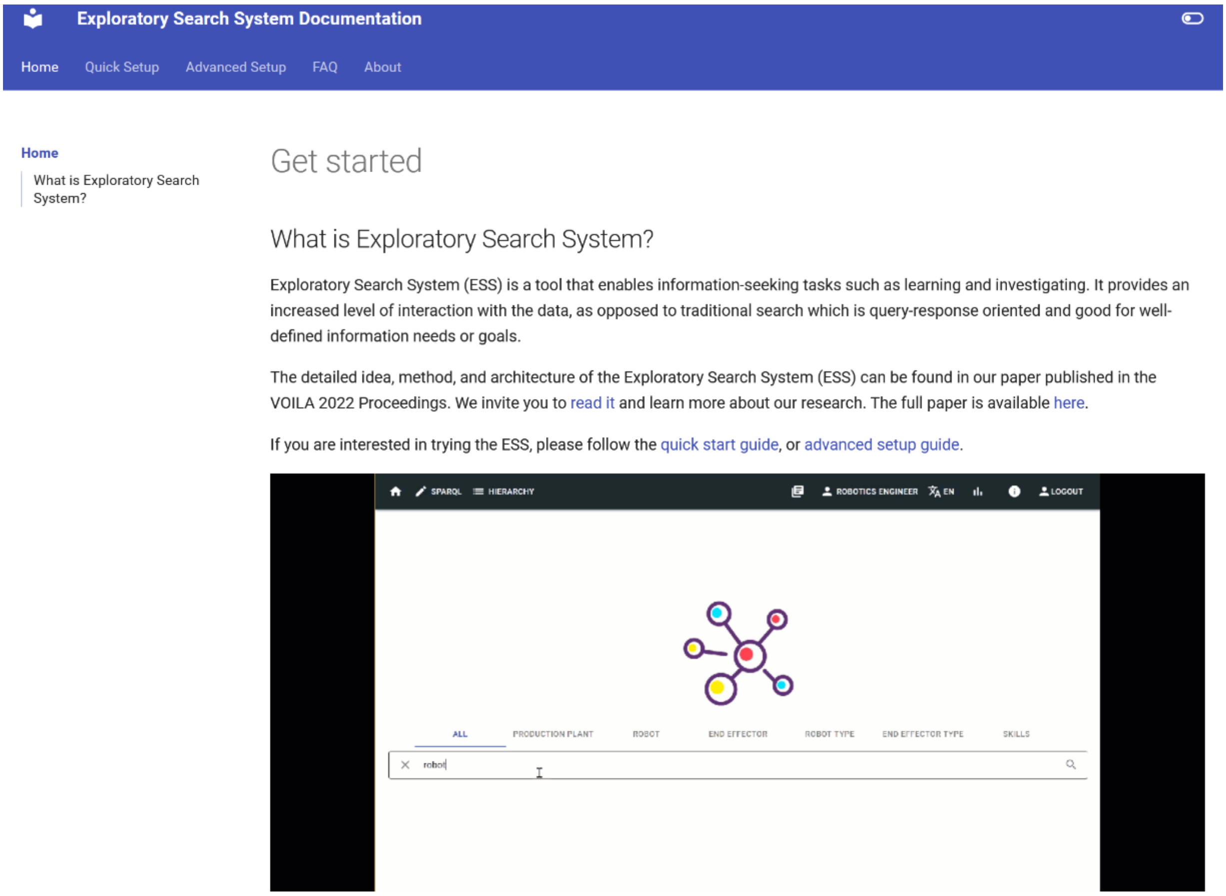

Figure 12. The ESS landing page.

The landing page is shown with a person icon and next to it a role can be defined. In this case, it is a Translator. Depending on their role, a person sees relevant topics to search.

Figure 13 shows an example search with Tensile Testing[35] (with the identifier ID: 1) in the input field and the output produced by the ESS.

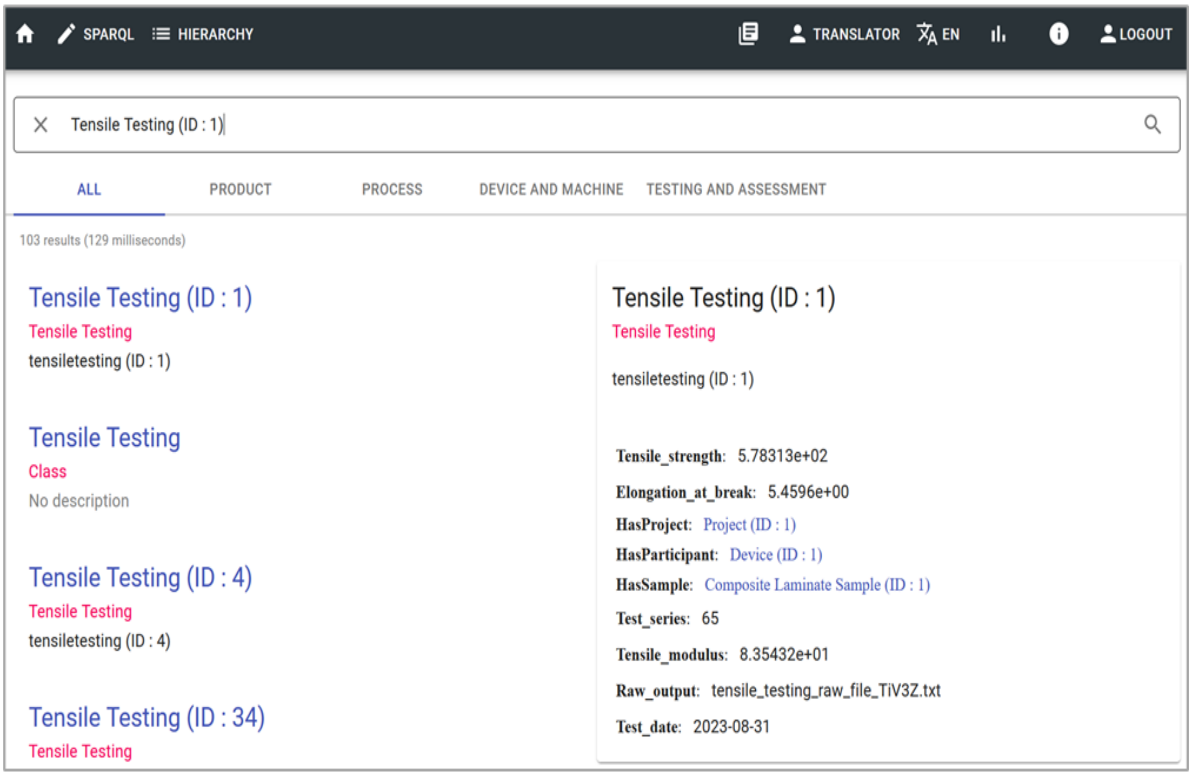

Figure 13. Text search with keyword "Tensile Testing (ID: 1)"

The ESS provides to the user the numeric results of this test, i.e., tensile strength and elongation at break, and also information about the whole experimental scenario. The user may be interested in the whole project and can find out which one it was. For example, the user may be interested in the device that was used for a particular experiment. Hence, the ESS permits a user to search with a keyword they remember and uncover a plethora of knowledge related to it.

4. Industrial Applications for OntoTrans

Four industrial applications were selected to demonstrate the benefits that the OntoTrans approach can bring organisations. Each application was based around an industrial innovation challenge that could benefit from transitioning to a fully digitised, ontological problem-solving approach. The project industrial partners are Procter & Gamble (P&G)[40], Composites Evolution[41], and Arcelor Mittal[42].

4.1. Post-Launch Analysis of Pouch Detergents

Pods (unit dose detergents) are an innovative and fast-growing laundry detergent sector. Understanding the in-market dynamics of consumer goods is critical for consumer-goods organisations, like P&G, to assess product success. Relevant factors include point of sale country and retailer, time period, pricing and distribution, and product performance profile in relation to brand equity. By understanding what drives sales of a product, an organisation can make sure any new product has the best chance of market success and reduce the risk of product failure. The key objective

[40] https://us.pg.com/
[41] https://compositesevolution.com/
[42] https://corporate.arcelormittal.com/

of this project was to develop machine learning models to evaluate sales as a function of in-market factors and product characteristics through the power of a unified ontology framework. Such a system would enable rapid adaptation to in-market changes to offer consumers a good range of products that suit their needs, lifestyles and budgets.

4.2. Detergent Pouch Systems

This case study with P&G leverages an end-to-end consumer centric framework, which uses consumer and market data/models in the translation challenge and includes chemistry and processability. Consumer preference and freshness/fragrance models are continuously refined as a result of tailored consumer research studies that reflect consumer behaviour in several European countries. Machine learning models are trained on synthetic data by P&G before using a company-specific installation of OntoTrans to populate the OntoKB with business-critical data.

P&G aim to achieve a five-fold reduction in the number of iterations taken to design new products and the product development cycle time. This would enable them to extend their range of new products in terms of performance and sustainability. OntoTrans is expected to better inform product design and innovation strategy through understanding of fundamental relationships between product formula and consumer needs. P&G estimate that introducing OntoTrans, or a similar system, could lead to a two-fold reduction in costs by reducing the amount of laboratory and consumer testing, prototyping, and performance evaluation.

4.3. Composite Prepregs

The third use-case deals with manufacturing composite prepregs based on polyfurfuryl alcohol (PFA) resins produced by Composites Evolution. In this case, the process of translation was carried out in cooperation with Fraunhofer IFAM and substantial outcomes as well as data were recently published [5]. The innovation challenge is based on successful materials design for the manufacturing of composite prepregs using autoclave-based processes. The challenge was to determine process parameters to avoid porosity of laminates when using vacuum bagging as an alternative and sustainable out-of-autoclave method. Furthermore, this case demonstrates how an industrial manufacturer and an organisation that offers translation services can mutually profit from OntoTrans to increase a small and medium-sized enterprise’s responsiveness in view of expected market needs.

During prepreg curing — a polycondensation reaction — of PFA-bioresins, exceeding a specified, low level of porosity of cured composite laminates is not desirable but can be safely avoided by applying both pressure and heat during autoclave-based manufacturing. Undesired porosity levels during manufacture can affect the process output such that it might not fulfil the product requirements. So, on the one hand PFA is a fire-retardant eco-friendly alternative to phenolic resins, which is suitable for aircraft and train interiors. On the other hand, the curing process of PFA generates water, which can cause some undesired porosity in vacuum-bagging-based manufacturing processes. Hence, Composites Evolution wanted to benefit from contributions from OntoTrans by achieving:

- Improved understanding and control of processes
- Improved quality of end products
- Reduced development time and costs
- Creating a model-based approach for future product developments.

In Figure 14, we show how different concepts, such as simulation, manufacturing and testing can be semantically connected via an ontology. The manufacturing process under a given set of process parameters generates a composite laminate as an output. For testing, a small part of the composite laminate is cut to obtain a specimen, which is then used for characterisation to determine its materials properties. Porosity is an essential property that can be measured by using light microscopy of cross-sections of laminates or obtained from density measurements. Material performance is characterised by mechanical testing, yielding the interlaminar shear strength. The interlaminar shear strength provides a numerical value about mechanical performance of the composite laminate. Later on, the mechanical testing can be modelled and simulated, which validates and enables a model-based approach for the assessment of new composite laminates.

Further details of the translation approach in general and its application to this innovation case are discussed in Reference [5].

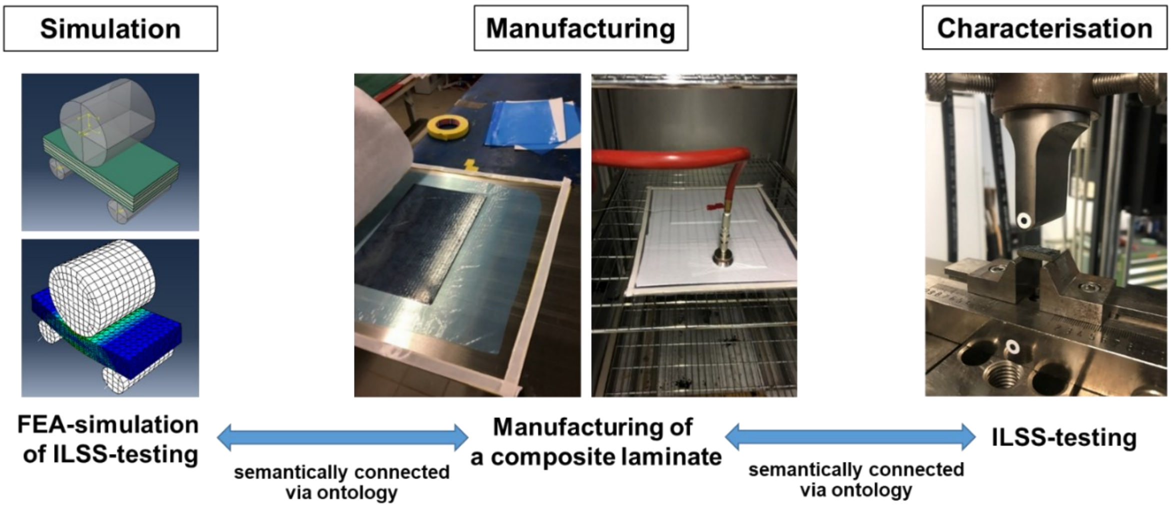

Figure 14. Composite prepreg: simulation, manufacturing and characterisation semantically connected through an ontology. ILSS = interlaminar shear strength.

4.4. Section Mill

Innovations in steel are essential for energy efficient, lightweight, and sustainable products, which results in stringent requirements with regards to mechanical properties. For steel producer Arcelor Mittal, the challenge is to master the relationship between process parameters and properties of hot-rolled H-beams[43] at a section mill. Currently, process and properties are physically tested, which is time and resource intensive and generates large data sets. A coherent, seamless, and unified system can address manufacturing challenges more efficiently in a single framework that covers the entire product development process. The objective is to predict the final mechanical properties of a beam given a set of process parameters (the forward problem) and to determine the process parameters given the required mechanical properties of a target product (the backward problem).

At Arcelor Mittal, time to market for new products can be up to 10 years. The current approach relies on expert knowledge to identify similar products, and potential chemistry and process changes to target desired properties. Any process changes are evaluated through detailed physical tests that take several months. Hence, a shorter cycle will improve the company's ability to respond promptly to new demands. OntoTrans allowed to eliminate several rounds of analysis of candidate solutions. And, over

[43] steel beams made of rolled steel in the shape of the letter "H"

time, data and feedback from new tests will further refine the accuracy of the models, which will result in additional time and cost savings and may lead to a 5-fold reduction in the time to market for new products.

5. Reaping the Benefits of Semantic Knowledge Management

Implementing a SKM approach requires an organisation to undertake a change to working processes. Peter Drucker[44] stated that "*Doing the right thing is more important than doing the thing right*" and "If you want something new, you have to stop doing something old." Historically, the companies involved with this project have been doing "things right", and through OntoTrans we help to provide the "right thing" to do next.

Within an organisation, OntoTrans provides a set of enablers that facilitate change in how people, tools, processes and data work (Figure 15). The enabling changes are entities that can be changed in some way, such as material products, processes, organisations or people's skills. These changes allow for onward innovation[13].

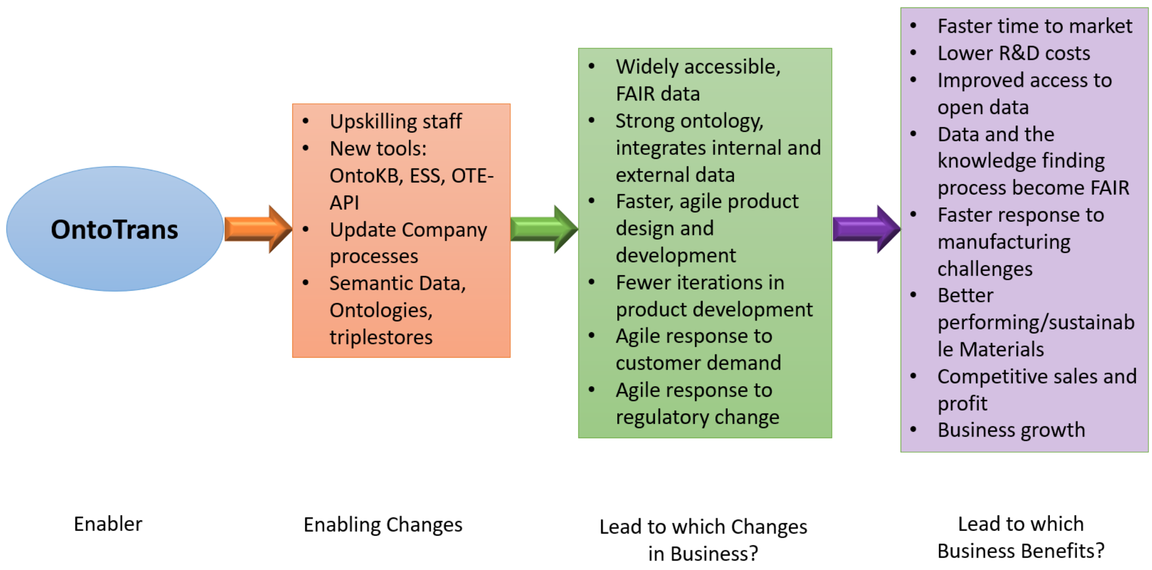

Figure 15. An example benefits dependency network, adapted from Ward & Daniel[14]

Enabling changes are identified along with any barriers and risks that may occur during the enabling phase. *Enabling changes* must be supported by sufficient training. It is important to note that an *enabler* works alongside the *enabling changes* to affect business changes and realise potential benefits. Business changes can positively influence the work life of individuals within the organisation. Once the *business benefits* are realised, the value of the *enabler* (in this case the OntoTrans approach) becomes evident to management and stakeholders. It is in this way that a thorough understanding of the *business benefits* can act as incentive to undergo change. Also, tangible benefits help organisations to justify why old ways of working have to change or stop.

[44] Peter Ferdinand Drucker (1909-2005) was an Austrian management consultant, educator, and author.

Within large organisations, OntoTrans requires staff with appropriate skills to fulfil the roles of Translators and Data Scientist during adoption of the system. If these skills do not exist within the company, then staff need to be upskilled or hired. For the OntoTrans OTE, as a software suite, to be successful, it must be implemented and fully integrated into the software landscape of an organisation. This requires close cooperation between the developers of the OntoTrans components and internal IT staff.

All data need to be curated to an extent that an ontology can help to extract knowledge from them. Once these *enablers* are in place, businesses experience exciting changes. R&D staff will work with a common knowledge base (OntoKB) taking out and feeding in knowledge. Data should be democratized i.e., made accessible to all employees and stakeholders. Once in place, materials modellers can provide workflows and log the knowledge about how they were developed. Workflows should be standardised so that they can be used by any skilled co-worker. Once implemented, the benefits of the OntoTrans approach for the employee and business are:

- Reduction of iterations in the design of new products
- Elimination of several rounds of characterisation/analysis
- A choice of better performing and more sustainable materials for new products
- Faster responses to new customer demands
- Faster time to market
- Effective and efficient achievement of consistent and predictable results.

A small enterprise will unlikely be able to afford new staff, so existing personnel may need to outsource work and find external translator and data science services, such as MarketPlace[45] and DOME4.0[46]. Also, if there is need for modelling, they may want to look into provision of modelling and simulation as a service such that model-based data sources are integrated into the IT landscape. However, small enterprises must curate data such that knowledge can be extracted. These enablers used will lead to improved processes specific to the market segment they are in. This is a business change that can reap a significant market advantage. The overall business benefits of using OntoTrans include a reduction in iterative testing and increased production speed and faster responses to new customer demands.

6. Conclusions and Outlook

The OntoTrans project developed a prototype for the application of Semantic Knowledge Management (SKM) to materials innovation challenges.

OntoTrans supports making data FAIR; the ESS makes knowledge findable, and the user interfaces ensure accessibility. The OntoKB knowledge base works in conjunction with the EMMO to provide mapping tools for data sources (and sinks), supporting interoperability and re-usability. Data sources include data from materials modelling, which enables plugging data gaps, playing through what-if scenarios, optimisation of products and processes. It aids an organisation to react faster to changes in markets and regulatory demands.

The industrial partners involved in this project anticipate a 2–3-fold improvement in data reuse through improved access to a wider pool of curated data. OntoTrans will be of interest to companies

[45] https://materials-marketplace.eu/
[46] https://dome40.eu/

involved in digitalisation and FAIR data (involving metadata, semantics, knowledge base, data analytics and translators) because it will strengthen their competitiveness and growth opportunities in global markets.

Stakeholder engagement by OntoTrans[15] has confirmed the strong interest by industry in SKM. From agreed vocabularies across functions and value chains to interoperable systems such as demonstrated by OntoTrans, materials-based industries could benefit from a "Semantic Materials" approach to better capture their knowledge, interlink data from processing, characterisation and modelling as well as product performance, and hence make smarter and faster decisions.

There is a clear request and needs from the end-user community to be prepared to use OntoTrans and other digital tools & platforms. That is why a multidisciplinary training program especially focused on digitalization aspects including SKM for both the young generation and experienced researches and industrial users should be further developed.

Going forward, OntoTrans partners plan to set up Semantic Materials Consortia to collaboratively develop ontologies for SKM in materials industries and adopt a tried and tested approach from other industries.

7. Acknowledgements

The authors acknowledge the valuable contributions of the manufacturing industry users in the OntoTrans project, in particular Zhiwu Liang, Oreste Todini and colleagues at P&G, Julian Dizy and colleagues at ArcelorMittal, as well as Tasmin Alliott, Lucy Bull, Brendon Weager and colleagues at Composites Evolution.

This work is part of the OntoCommons and OntoTrans project that have received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements no. 958371 and 862136, respectively.

8. Disclaimer

All information contained in this study and any opinions expressed in it are intended to share the insights the authors have gained on a survey on social media usage during their work on the H2020 Project OntoTrans. All statements of fact, opinion, or analysis expressed in the report are those of the authors. The information used and statements of fact made are not guarantees, warranties or representations as to their completeness or accuracy. The authors assume no liability for any short-term or long-term professional decisions made by any reader based on analysis included in this report.

9. Abbreviations

API – Application programming interface

DLite – A light-weight semantic interoperability framework designed to be data-centric and used as part of the OntoTrans project

EMMC – European Materials Modelling Council

EMMO – The Elementary Multiperspective Material Ontology

ESS – Exploratory Search System

FAIR – Findable, accessible, interoperable, reusable

H2020 – Horizon 2020 is the financial instrument implementing the Innovation Union, a Europe 2020 flagship initiative aimed at securing Europe's global competitiveness. (2014–2020).

OTE – Open translation environment

OTEAPI – An actionable data documentation system for semantic interoperability

OWL – Web Ontology Language

P&G – Procter & Gamble

PFA – polyfurfuryl alcohol

R&D – Research and development

RDF – Resource Description Framework

SKM – Semantic knowledge management

SPARQL – SPARQL Protocol and RDF Query Language

10. References

1. Kenett, R. S., Zonnenshain, A. & Swarz, R. S. Systems Engineering as a Data-Driven and Evidence-Based Discipline. *INCOSE Int. Symp.* **30**, 757–770 (2020).
2. Henderson, K. & Salado, A. Value and benefits of model-based systems engineering (MBSE): Evidence from the literature. *Syst. Eng.* **24**, 51–66 (2021).
3. Feng, S. C., Bernstein, W. Z., Hedberg, T. & Barnard Feeney, A. Toward Knowledge Management for Smart Manufacturing. *J. Comput. Inf. Sci. Eng.* **17**, 031016 (2017).
4. Vogt, L. *et al.* FAIR 2.0: Extending the FAIR Guiding Principles to Address Semantic Interoperability. Preprint at https://doi.org/10.48550/ARXIV.2405.03345 (2024).
5. Noeske, P.-L. M. *et al.* Towards Advancing Translators' Guidance for Organisations Tackling Innovation Challenges in Manufacturing within an Industry 5.0 Context. *Sustainability* **16**, 3486 (2024).

6. Wang, H. & Peng, G. The Merging of Knowledge Management and New Information Technologies. in *Collaborative Knowledge Management Through Product Lifecycle* 229–283 (Springer Nature Singapore, Singapore, 2023). doi:10.1007/978-981-19-9626-9_6.
7. Panetta, K. *Gartner Top 10 Data and Analytics Trends for 2021*. https://www.gartner.com/smarterwithgartner/gartner-top-10-data-and-analytics-trends-for-2021 (2021).
8. Clark, S. *et al.* Toward a Unified Description of Battery Data. *Adv. Energy Mater.* **12**, 2102702 (2022).
9. Klein, P. *et al. Translation in Materials Modelling – Process and Progress*. https://zenodo.org/record/4729918 (2021) doi:10.5281/ZENODO.4729918.
10. Goldbeck, G. *et al. The Translator in Knowledge Management for Innovation – towards Industry Commons*. https://zenodo.org/record/7041697 (2022) doi:10.5281/ZENODO.7041697.
11. Ekaputra, F. J. *et al.* Towards an Open Translation Environment for Supporting Translators in the Materials Domain. in *CEUR Workshop Proceedings* vol. 3240 Poster 3 (CEUR, Tarbes, France).
12. Haller, K., Ekaputra, F. J., Sabou, M. & Piroi, F. Enabling Exploratory Search on Manufacturing Knowledge Graphs. in *CEUR Workshop Proceedings* vol. 3253 Paper 2 (Online).
13. *ISO 9000:2015 Quality Management Systems Fundamentals and Vocabulary*. https://www.iso.org/standard/45481.html.
14. Ward, J. & Daniel, E. *Benefits Management: Delivering Value from IS & IT Investments.* (2006).
15. Cantrill, V., de Baas, A., Roscioni, O. & Goldbeck, G. *Developing Ontologies in Materials Science: Semantic Materials Workshop Report - Semantic Data and Knowledge Management for Chemicals and Materials Industries*. https://zenodo.org/doi/10.5281/zenodo.11127892 (2024) doi:10.5281/ZENODO.11127892.