Computers in Industry 153 (2023) 104016

Contents lists available at ScienceDirect

Computers in Industry

journal homepage: www.sciencedirect.com/journal/computers-in-industry

Toward a digital materials mechanical testing lab

Hossein Beygi Nasrabadi [a,*], Thomas Hanke [b], Matthias Weber [b], Miriam Eisenbart [c], Felix Bauer [c], Roy Meissner [d], Gordian Dziwis [d], Ladji Tikana [e], Yue Chen [a], Birgit Skrotzki [a]

[a] Bundesanstalt für Materialforschung und -prüfung (BAM), Berlin, Germany
[b] Fraunhofer-Institut für Werkstoffmechanik (IWM), Freiburg, Germany
[c] Forschungsinstitut Edelmetalle + Metallchemie (fem), Schwäbisch Gmünd, Germany
[d] Institut für Angewandte Informatik (InfAI), Leipzig, Germany
[e] Deutsches Kupferinstitut (DKI), Düsseldorf, Germany

ARTICLE INFO

Keywords:
Industry 4.0
Digitalization
Ontology
Mechanical testing
FAIR data

ABSTRACT

To accelerate the growth of Industry 4.0 technologies, the digitalization of mechanical testing laboratories as one of the main data-driven units of materials processing industries is introduced in this paper. The digital lab infrastructure consists of highly detailed and standard-compliant materials testing knowledge graphs for a wide range of mechanical testing processes, as well as some tools that enable the efficient ontology development and conversion of heterogeneous materials' mechanical testing data to the machine-readable data of uniform and standardized structures. As a basis for designing such a digital lab, the mechanical testing ontology (MTO) was developed based on the ISO 23718 and ISO/IEC 21838-2 standards for the semantic representation of the mechanical testing experiments, quantities, artifacts, and report data. The trial digitalization of materials mechanical testing lab was successfully performed by utilizing the developed tools and knowledge graph of processes for converting the various experimental test data of heterogeneous structures, languages, and formats to standardized Resource Description Framework (RDF) data formats. The concepts of data storage and data sharing in data spaces were also introduced and SPARQL queries were utilized to evaluate how the introduced approach can result in the data retrieval and response to the competency questions. The proposed digital materials mechanical testing lab approach allows the industries to access lots of trustworthy and traceable mechanical testing data of other academic and industrial organizations, and subsequently organize various data-driven research for their faster and cheaper product development leading to a higher performance of products in engineering and ecological aspects.

1. Introduction

Digitalization of industrial activities and expansion of data-driven technologies are the main essentials of the fourth industrial revolution (Industry 4.0) with the aim of sustainable economic, environmental, and social development (Ghobakhloo, 2020). Within manufacturing-related industries especially, digitalizing the production processes is expected to facilitate increased productivity and production efficiency, reduced manufacturing costs, energy and resource sustainability, and carbon/-harmful gas emission reduction (Ghobakhloo, 2020; Xu et al., 2018). The rapidly and continuously growing computing power in the industry also offers sustainable product development and the designing of new functional products. In the traditional manufacturing concept, product design and development are achieved by producing and testing numerous test pieces. For example, the development of mechanically stressed components for applications in the automotive or aerospace industries requires deep knowledge about or an extensive investigation of the materials' mechanical properties. Materials mechanical testing laboratories are the places for evaluating the materials' mechanical properties (like tensile, compression, stress relaxation, creep, and fatigue) and providing manufacturers with the required information related to hardness, strength, ductility, fatigue limit, and other important mechanical properties of the produced materials.

Plenty of mechanical testing data are generated daily in thousands of materials test labs around the world, and most of these data cannot be utilized by external parties for technical purposes, data-driven research, or product development. For example, Fig. 1 demonstrates that the valuable tensile test data that is frequently generated in the materials

* Correspondence to: Bundesanstalt für Materialforschung und -prüfung (BAM), Unter den Eichen 87, 12205 Berlin, Germany.
E-mail address: hossein.beygi-nasrabadi@bam.de (H.B. Nasrabadi).

https://doi.org/10.1016/j.compind.2023.104016
Received 5 April 2023; Received in revised form 1 September 2023; Accepted 5 September 2023
Available online 9 September 2023

mechanical testing labs may not often find research and computing possibilities for a variety of reasons. The first reason can be raised in the case that such data are not findable because the output results are in the form of printed papers or files stored in unstructured and non-publicly available repositories, or maybe the files cannot be found by search engines due to the lack of appropriate identification and documentation. Even access to the findable data may be regulated for authorized users or user groups with the necessary permissions or credentials while considering specific legal, ethical, or privacy constraints. Another obstacle is that such data may be findable and accessible, but their unreadable or image- or media-like data formats, distinct data languages, or a lack of the appropriate data structure and test metadata make them difficult to interoperate. In the case of those interoperable data sets, there are also sometimes concerns about their reusability. Some of these reusability concerns include inadequate documentation of the processing and testing parameters, inaccurate or non-standardized testing procedures, limited consent or usage agreements, and a lack of data quality control or long-term preservation strategies (Bayerlein et al., 2022; Zhao and Qian, 2017). As a result, a high amount of data generated daily in worldwide materials testing labs does not meet the criteria of being Findable, Accessible, Interoperable, and Reusable (FAIR) (Scheffler et al., 2022), and therefore cannot be (re-)used for data-driven product development purposes. Related, it has been estimated that the non-existence of FAIR data invokes costs to the European economy of at least €10.2 billion per year (European Commission, 2018).

The main motivation of this research is to develop a lab digitalization approach for the generation and storage of FAIR data in mechanical testing labs (Fig. 1). The desired outputs of such digital labs would be testing results that are represented digitally and expressively using the standardized structures of common vocabulary and include the measured materials data along with the process metadata and materials testing standards they are based on. Hence, testing reports are machine-processable semantic descriptions of data and their generation processes, while efficient data management systems are utilized to store them as structured datasets in repositories and enrich and set up comprehensive, documented, and connected data spaces. As a result of such an approach, the advanced semantic search will enable materials engineers to easily and fully access all FAIR testing data made available in open-access data repositories. As shown in Fig. 1, the generation and storage of such FAIR data in the digital materials mechanical testing lab can accelerate data-driven materials research such as machine learning and deep learning approaches (Himanen et al., 2019) for the development and discovery of new materials.

A prerequisite for implementing the FAIR principles and building a comprehensive database infrastructure is formulating the formal standardized knowledge representation through the ontologies (Bayerlein et al., 2022). Studer et al. (1998). defined ontology as a formal, explicit specification of a shared conceptualization, where "formal" means machine-readable abstraction of a domain, "explicit" refers to well-defined constraints, relationships, and dependencies between the concepts, "shared" constitutes the building and agreement by a group of different people, and "conceptualization" is an abstract model describing a particular field of knowledge (Fensel, 2001). In other words, ontologies form a sustainable representation of expert knowledge by using a formal, standardized language to express a vocabulary as an interconnected collection of concepts, attributes, relations, constraints, individuals, and axioms (Ekaputra et al., 2017). Ontologies semantically structure and annotate materials data, represent knowledge sustainably, facilitate storing the data and information in semantically organized ways, and support data handling (Bayerlein et al., 2022). As a result of these functions, ontologies enable knowledge organization through hierarchical structure and relationships between concepts, global conceptualization for materials information integration, linked materials data publishing, semantic query supports, interoperability between related domains, and automatic inference support for discovering new materials knowledge (Studer et al., 1998; Zhang et al., 2015).

2. Related works

Although relatively mature domain ontologies have been established recently in biology, medicine, environment, and art domains, semantic web technology's growth in materials science and engineering was accompanied by the development of relatively few ontologies (Zhang et al., 2015). MASON manufacturing's semantics ontology (Lemaignan et al., 2006), MatOWL ontology transformed from MatML schema (Zhang et al., 2009), materials ontology (Ashino, 2010), ONTORULE ontology for the steel domain (Sainte Marie et al., 2011), SLACKS ontology for laminated composites (Premkumar et al., 2014), STSM steel semantic model (Zhang et al., 2014), PREMaP ontology (Bhat et al., 2013), and MMOY metallic materials ontology (Zhang et al., 2016) are some of the first ontologies that were established in the materials science domain. The mentioned ontologies often suffer from limited usability for various reasons. One significant limitation of these ontologies is their narrow and specific focus fields. By targeting specific applications or domains within materials science, these ontologies may not have been designed with broad applicability in mind. This restrictiveness arises from the fact that the ontology's concepts, relationships, and axioms are tailored to a specific domain and may not generalize well to other domains or real-world applications. Another factor contributing to the limited usability of the mentioned ontologies is the inappropriate level of abstraction. Those abstract ontologies mostly missed the necessary level of detail required for practical application, while some of the granular and specific ontologies failed to capture the broader context and relationships between entities. Furthermore, the lack of identifiable

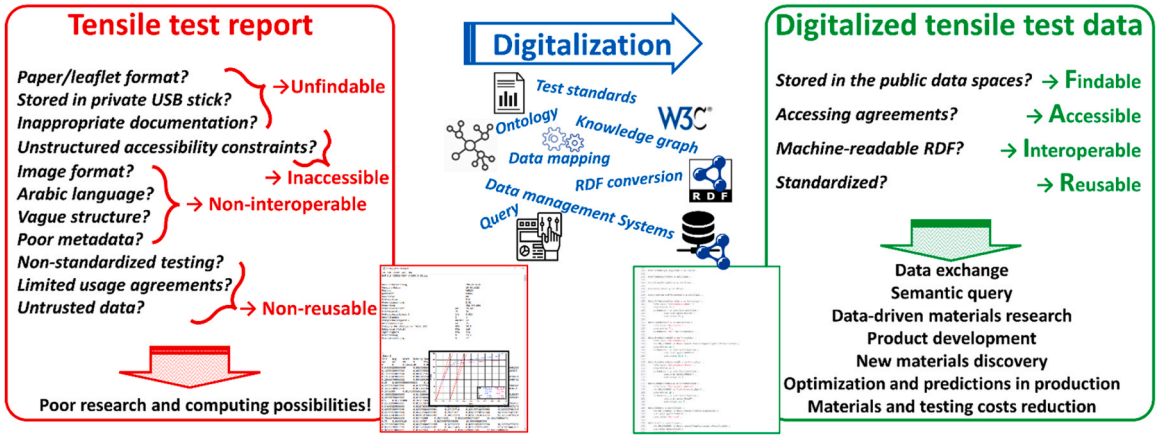

Fig. 1. Lab digitalization: a motivation to benefit the generation and storage of FAIR data in mechanical testing labs.

and concrete definitions within some of these ontologies prevented their accurate interpretation and made it challenging to establish a common understanding and consistent usage of concepts and relationships within the ontology. Insufficient knowledge of instances in some of the ontologies mentioned also contributed to the serious problems in validating, testing, and applying such ontologies in practical settings. Additionally, some of the ontologies mentioned lacked comprehensive documentation, detailed guidelines, or standardized formats, making them less accessible and difficult to integrate into existing systems or applications. Finally, practical ontologies should be able to seamlessly integrate with other ontologies, leveraging and reusing their content to create a global framework of reference ontologies. However, the mentioned ontologies may not adequately incorporate such interoperability, preventing the advancement of data interoperability across different contexts and organizations in their respective fields (Zhang et al., 2015; Ali et al., 2019; Mutz et al., 2022; Fraga et al., 2020).

Extending the interoperability between the materials science domain ontologies can provide a possibility for linking the data from different domain ontologies, enabling more semantic context for domain experts, and exploring higher knowledge about the process-structure-property relationships of materials. One idea for building highly interoperable ontologies is to develop them as extensions of standardized upper-level ontologies. Generally speaking, different kinds of ontologies can be divided into three levels based on the degree of abstraction and formal expressiveness (Bayerlein et al., 2022):

i) Top-Level Ontologies (TLOs) describe common general terms across many domains at the highest possible level of abstraction. TLOs establish semantic standards and incorporate universal and fundamental concepts to ensure the connection and interoperability of a wide range of conceivable domain ontologies. Basic Formal Ontology (BFO) (Arp et al., 2015; Otte et al., 2022), Elementary Multiperspective Material Ontology (EMMO) (European Materials Modelling Counsil, 2020), and Descriptive Ontology for Linguistic and Cognitive Engineering (DOLCE) (Borgo et al., 2022) are some of the most applicable TLOs in the materials science domain.
ii) Mid-Level Ontologies (MLOs) add finer granular entities to the TLOs and make them more modular to enable interconnecting of the complex and expressive domain-level ontologies. For example, Common Core Ontology (CCO) (Rudnicki, 2019) is an MLO that has been established based on the BFO top-level ontology.
iii) Domain-Level Ontologies (DLOs) contain highly expressive and explicit expert knowledge and represent concepts, definitions, facts, statements, axioms, rules, and relations that belong to specific domains. Most of the DLOs (not basically all of them) are developed by extending the mentioned concepts of top-level ontologies to facilitate the extension of domain knowledge in an organized and sustainable way.

A few materials science domain ontologies were introduced in recent years that offer the construction of highly interoperable, formal, and specific knowledge based on their connection and integration with other existing top-level ontologies. For example, the Li-ion battery ontology (Mutz et al., 2022) and the Materials And Molecules Basic Ontology (MAMBO) (Piane et al., 2021) were structured based on the EMMO. The Materials Design Ontology (MDO) that represents the materials' structures, properties, calculations, and relationships has also used some entities of PROV-O (PROV Ontology), QUDT (Quantities, Units, Dimensions, and data Types ontologies), and Dublin Core Metadata Initiative (DCMI) (Lambrix et al., 2023; Li, 2022). The materials characterization methodologies ontology (CHAMEO) has also been aligned with some EMMO-based domain ontologies for the classification of materials, models, manufacturing processes, and software products related to materials modeling (Del Nostro et al., 2022). The NanoMine ontology for polymer nanocomposites (McCusker et al., 2020) and the Additive Manufacturing Processing Ontology (AMPO) (Li et al., 2017) were developed by using top-level ontologies like PROV-O, QUDT, and semantic Science Integrated Ontology (SIO). Sanfilippo et al. (2019). utilized DOLCE as the TLO for ontology-based knowledge representation of additive manufacturing. Cheung et al. (2008, 2009). also developed the MatOnto ontology based on the DOLCE top-level ontology, for structured knowledge representation of general materials science and particular fuel cell domains. However, they revised the next MatOnto versions based on the BFO top-level ontology to improve its logical consistency and commonality (Zhang et al., 2015). Recently, BFO has been employed for building interoperable domain ontologies within disparate engineering fields (Hagedorn et al., 2019). Furini et al. (2016). developed a manufacturing ontology for functionally graded materials by using the BFO and the Ontology for Biomedical Investigation (OBI) as the TLO and MLO, respectively. Hagedorn et al. (2018). also established the Innovative Capabilities of Additive Manufacturing (ICAM) ontology according to the BFO and some BFO-compliant MLOs like CCO and the Information Artifact Ontology (IAO). In the same way, and reusing the BFO and CCO, the product life cycle ontology was represented by Otte et al. (2019). Furthermore, the BFO was used as the TLO for developing the NMR relaxometry ontology (Moreno Torres et al., 2021), and ontology-based input, thermal, microstructure, and mechanical properties metamodels of metal additive manufacturing (Roh et al., 2021).

To enhance the reuse and adoption of modern formal ontologies in industrial applications, the Industrial Ontologies Foundry (IOF) core ontology was launched recently (The IOF Core Ontology et al., 2023). IOF is an initiative that aims to promote the reuse and interoperability of formal ontologies in industrial domains and provide an open-source platform for developing, validating, aligning, sharing, and curating industrial ontologies (Ali et al., 2019; Kulvatunyou et al., 2018). The IOF core is a mid-level ontology, which was extended based on the BFO (The IOF Core Ontology et al., 2023). In this regard, a few of the very recently developed ontologies organized based on the BFO upper-level ontologies provided high compatibility with the IOF core ontology. The Additive Manufacturing Ontology (AMO) (Ali et al., 2019), the Functionally Graded Materials (FGM) ontology (Mohd Ali et al., 2021), 4D printing knowledge in design (Dimassi et al., 2021), the Zero Defect Manufacturing ontology (ZDM) (Psarommatis et al., 2023), and the Materials Science and Engineering Ontology (MSEO) (Materials Open Lab, 2021) are to date the only introduced ontologies in the materials science domain that provide industrial interoperability based on the IOF core ontology.

Concerning mechanical testing, as one of the important domains of materials manufacturing and development, a small number of ontologies were presented. The Creep ontology (Ashino and Fujita, 2006) was the first trial example for describing the creep properties and creep data analysis process. In 2010, the European Committee for Standardization provided a guide for developing the materials tensile testing ontology based on existing ISO standards (Austin et al., 2010). Although a well-defined ontology was provided, this approach remained unused due to its low interoperability with other existing ontologies and the impossibility of linking the process-structure-property knowledge of different repositories. To address this limitation, Morgado et al. (2020). suggested building a tensile testing ontology based on the EMMO. Currently, the Materials Open Lab (Mat-O-Lab) initiative is also working on developing the tensile and Brinell hardness ontologies based on the BFO, CCO, and MSEO upper-level ontologies (Chen et al., 2022; Materials Open Lab, 2022).

Based on the literature survey done, the ontology development methodologies in the domain of materials and processes are transitioning to the introduction of higher interoperable and industrial applicable ontologies. However, few ontologies, especially in the domain of materials testing processes, have been presented until now that provide these desirable capabilities. Moreover, those few existing ontologies on the materials testing processes are not applicable for building the digital

lab concept due to limitations like non-compliance with testing standards, low level of details and relations for the description of the test process, lack of instances and process knowledge bases, low interoperability with BFO and other industrial applicable ontologies, and incompatibility with real test reports of mechanical testing labs. On the other side, the lack of both suitable data spaces for data exchange and appropriate tools and infrastructures for linking the testing data and ontologies acted as barriers to the efficient usage of ontology-based approaches for the digitalization of industrial processes until now.

To address the development and industrial application of ontologies in the materials and processes domains, the Platform MaterialDigital[1] (PMD) was initiated in 2019 by the German Federal Ministry of Education and Research (BMBF). Within the PMD projects, big communities of materials and data scientists are working in a collaborative environment to define highly interoperable and industry-compatible ontologies and establish a virtual materials data space, databases, and toolchains for sustainably digital representation of the materials and their production. As one of the projects funded in the framework of MaterialDigital, Kupfer*Digital*[2] consists of different partners from science and industry to develop the data ecosystem for ontology-based digitalization of the entire life cycle of copper materials, with a focus on industrial applications.

Concerning materials testing as one of the main chains of such a life cycle, this research aims to introduce innovative infrastructures for establishing the digital materials mechanical testing lab. To reach this aim, a series of materials testing standards and well-defined upper-level ontologies are utilized for developing a mechanical testing ontology (MTO), and highly detailed and standard-compliant materials testing knowledge graphs. Here, for a better presentation of the ontology and knowledge graphs in the process of data retrieval and connecting it with the identified problem, some typical competency questions were planned as below:

- What is the ultimate tensile strength (or Brinell/Vickers hardness, fatigue strength, creep life, etc.) of the material "x"?
- Which processing parameter (or material, datasets, etc.) fulfills some specific mechanical properties (e.g., Vickers hardness more than "y")?
- Which load (or temperature, testing machine, standard, etc.) was used for experiment "z"?

To evaluate how the developed ontology and knowledge graphs can address such competency questions, this research also introduces some tools for mapping different types and formats of mechanical testing data and their conversion to RDF ones.

The findings of this research are highly novel and interesting from different aspects. Here, the innovative idea of the digital mechanical testing lab was organized based on highly detailed and standard-compliant materials testing process graphs and ontologies. The process graphs and ontologies were developed with the collaboration of academic and industrial partners and based on testing standards and scientific best practices. A highly interoperable and industry-applicable ontology was established based on the BFO and IOF core ontology, which is compatible with all types of test reports and can link data from other domains and facilitate the prediction of process-structure-property dependencies of materials. In addition, we introduce a well-designed methodology and several open-access user-friendly toolchains that simplify the development of process graphs and ontologies by domain experts. Furthermore, the introduced data ecosystems concepts for data transformation, mapping, and storage were utilized for the first time to digitalize the different formats/types of lab-generated testing reports.

The proposed infrastructure of this study is generic enough to cover all types of materials testing labs at academic and industrial scales and enable their digital representation. The developed methodology and toolchains can also be extended into other chains of the material life cycles to accelerate digitalization and contribute to further developing the Industry 4.0 approach. The comprehensive machine-readable FAIR data that are generated in the digital lab, not only facilitate the structured storage, retrieval, and exchange of the test data, but also enable complex semantic searches, as well as reusability, integration, and analysis of data from different repositories. Eventually, the generated FAIR data of such digital labs allows industrial partners to better understand the properties and behavior of materials via data-driven materials research, optimize and develop their products, make predictions in production, reduce the materials preparation and testing costs, and develop new advanced materials. The biggest obstacles to incorporating these advantages into most industrial parties' current activities are a lack of user-friendly infrastructure for data management and the reluctance of most industries to share their data. On the other hand, several industries are still lacking an overall comprehension of the principles of digitization and data management, as well as the current methods and technologies. The presented approach and examples of this research not only provide an explanation of the principle and advantages of data digitalization to both industry and academic communities but also expose them to a simplified approach and user-friendly toolchains to profit from making their data FAIR.

3. Digital lab architecture (methodology and toolchains)

The materials mechanical testing lab is an important unit of materials production lines for quality assurance and more specifically evaluating the mechanical properties of produced materials based on standardized mechanical testing methods. Within its traditional format, different kinds of test samples are prepared and tested in mechanical testing labs, and then the test data of heterogeneous structures are generated and stored in private, local, or maybe exclusive resources and delivered to materials engineers for further evaluation of the products. To address the generation and storage of FAIR data in such units, the schematic architecture of a digital materials mechanical testing lab is presented in Fig. 2. In the proposed digital lab concept, all the infrastructures needed for process digitalization are prepared in a digitalization workflow as the basis of the lab (first level of Fig. 2), so the lab technicians will follow the regular standardized testing processes and will perform just one more step to digitalize their testing data by using user-friendly and open access software and store them in the special data repositories (second level of Fig. 2). The generated FAIR data in such digital labs then can be used by the materials engineers (third level of Fig. 2) for purposes like data exchange, data discovery from other sources, machine learning, data-driven materials research, and product development.

The main idea of designing the digital lab infrastructure (first level of Fig. 2) is to provide an environment for mapping the materials testing data to the modeled test process graphs (more accurate: knowledge graph of processes) and storing the resulting digitalized data in suitable data spaces. Here, the process graphs are the graphical and semantic representation of the experimental procedure which are developed based on materials testing standards and enriched by valuable metadata about the used materials, equipment, specifications, and objects relevant to producing the results. The process graphs for different materials' mechanical testing procedures were prepared with the cooperation of a group of materials and data scientists, and are available in Kupfer*Digital* GitLab (see data availability). Regarding, the minimum level of detail required for designing such process graphs, essential concepts, and their definitions, as well as the relationship between different testing concepts, were extracted from the materials testing standards (glossary

[1] https://www.materialdigital.de/.
[2] https://www.materialdigital.de/project/1.

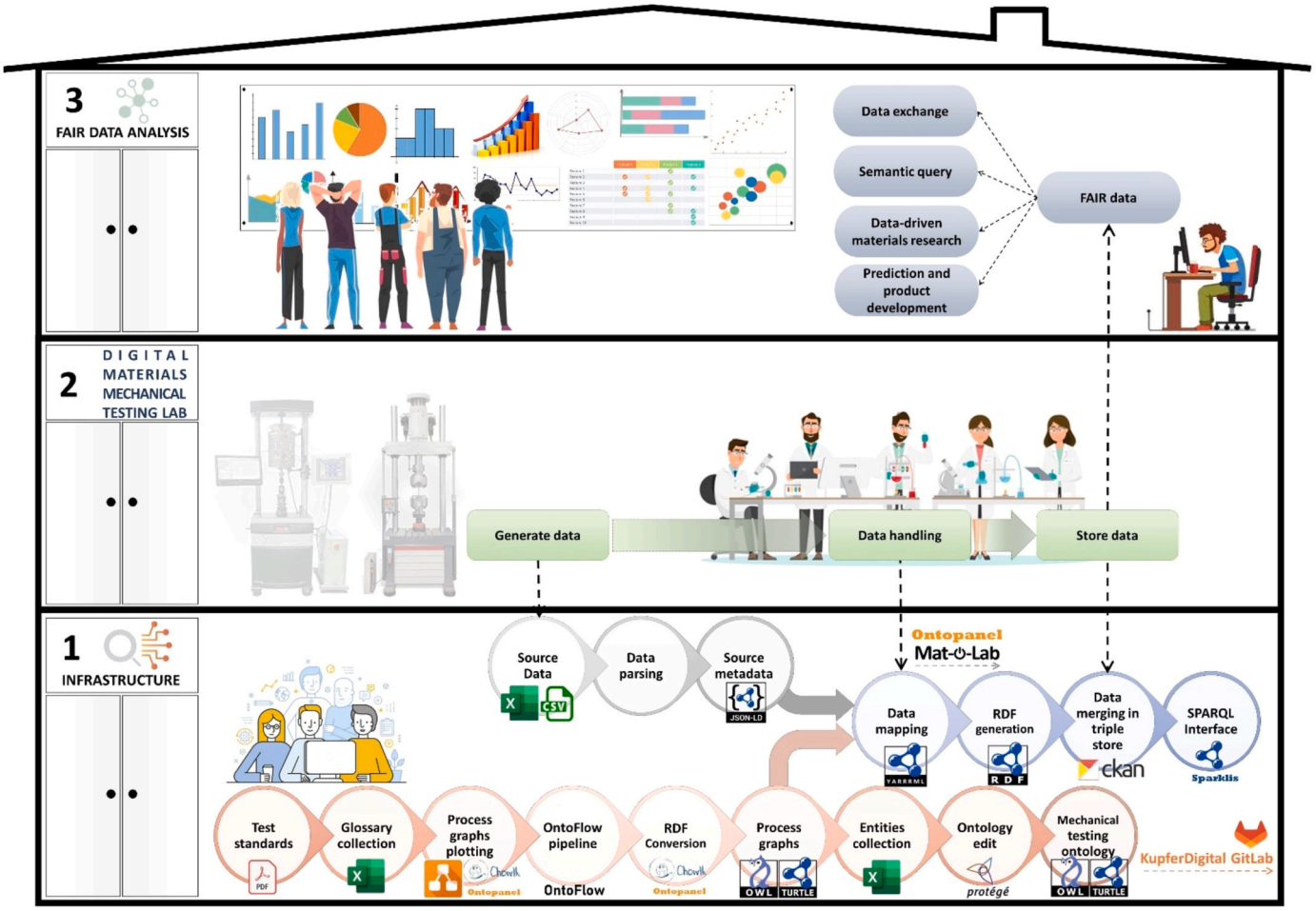

Fig. 2. Digital materials mechanical testing lab: infrastructures, FAIR data creation, and data usage. The main proposed approach of this research, digital lab infrastructure for FAIR data generation and application, is basically established based on the following steps and toolchains: i) collaboratively designing and validating the knowledge graphs of the mechanical testing processes according to the test standards via GitLab-based tools like Ontopanel and OntoFlow; ii) Developing the standard-compliant mechanical testing ontology in Protégé; iii) Metadata generation, data mapping, and RDF conversion with Mat-O-Lab and Ontopanel tools; and iv) Data management and SPARQL query using Sparklis. (Reproduced with permission: Draw.io and Chowlk logo under Apache License 2.0; Protégé logo licensed by The Board of Trustees of the Leland Stanford Junior University; GitLab logo under the terms of the MIT license, CKAN logo under GNU Affero General Public License, AGPL).

collection). The process graphs were subsequently designed in the open-access graphical editing tool of diagrams.net (Draw.io).[3] In this case, the shapes required for presenting the necessary semantic representations can be provided by the standard libraries of Chowlk[4] or Ontopanel,[5] and then the resulting process graphs of XML formats are converted into the Ontology Web Language (OWL) formats like Resource Description Framework (RDF) or turtle syntax (TTL), using the Chowlk/Ontopanel converter. Furthermore, OntoFlow[6] automatically validated the process graphs and ontologies against SHACL shapes[7] and documented them as HTML with pyLODE.[8] Finally, MTO was prepared in the software Protégé,[9] by collecting all the concepts used in the process graphs of various mechanical tests, explicit definition of entities, merging the planned upper-level ontologies, and construction of a hierarchical order of new concepts and classes. The design and preparation of the process graphs and MTO are discussed in detail in Section 4.

To evaluate the generation of FAIR data in a digital lab, extensive mechanical tests such as tensile, Brinell, and Vickers hardness tests were performed on different cast copper samples (Beygi Nasrabadi et al., 2023) and the resulting testing data of various structures and formats (like CSV and XLSX) were mapped in the process graphs. It should be noted that trial digitalization of materials testing data was performed in an accredited materials testing laboratory of the German Federal Institute for Materials Research and Testing (BAM), which assured the generation of test results according to the testing standards, and attached with maximum possible metadata about the sample history, equipment, calibration, etc.

While the current research mostly focuses on the standard-based ontology development and evaluation of the approach by addressing the competency questions, the data mapping tools are required to map the testing data of different types and formats in the developed graphs and generation of RDF data. Section 5 evaluates two methods of Mat-O-Lab and Ontopanel for mapping the test data in the process graphs. By the Mat-O-Lab approach, CSVToCSVW[10] tool was used to convert the

[3] https://app.diagrams.net.
[4] https://chowlk.linkeddata.es.
[5] https://github.com/yuechenbam/yuechenbam.github.io.
[6] https://gitlab.com/infai/ontoflow.
[7] https://github.com/MaastrichtU-IDS/shacl-shapes.
[8] https://github.com/RDFLib/pyLODE.
[9] https://protege.stanford.edu/.
[10] http://csvtocsvw.kupferdigital.org/.

raw test data to machine-readable formats like JSON-LD. Following, the MapToMethod[11] tool allows the user to link the different data of the test report to the related items of the process graphs, and the YARRRML files resulting from mapping are converted to the RDF dataset by the RDFConverter[12] tool. In the second method, Ontopanel (Chen et al., 2022) was used for graphically mapping the secondary test datasets in the process graphs and converting the resulting data to RDF.

The RDF models of the test data resulting from the digital mechanical testing lab were placed into the CKAN-data management repository (see data availability), where it can be easily processed by a knowledge base triple store. Section 6 represents how materials engineers and data scientists can now easily use, exchange, reproduce, and analyze such generated FAIR data. Furthermore, this section evaluates how the introduced approach can support the data retrieval, connecting it with the stated problem, and response to competency questions. In this case, Sparklis[13] (Ferré, 2017) has been utilized for the SPARQL (a semantic query language, SQL) query of data stored in the RDF format.

4. Mechanical testing knowledge graphs and ontology

The first step in developing the mechanical testing knowledge graphs is to collect all the important terms related to the mechanical testing processes. There are various resources for collecting ontology terminologies, such as test standards, test reports, manuals, scientific literature, and interviews with experts. Meanwhile, the materials testing standards are the most important resource for preparing the ontology vocabulary, since they provide the specific names and definitions for the different concepts of the processes, and orient the structuring and categorization of test process entities based on the information that has already been agreed by an expert standardization committee (Bayerlein et al., 2022). In addition, developing the ontologies based on the materials testing standards ensures their acceptance and applicability in the industry. A variety of materials testing standards have been used in the construction of mechanical testing process graphs and ontology, like ISO 6892–1 (tensile testing) (ISO, 2019), ISO 6506-1 (Brinell hardness test) (ISO, 2014), ISO 6507–1 (Vickers hardness test) (ISO, 2018), DIN EN 10319-1 (relaxation test) (DIN, 2003), and ISO 12106 (fatigue test) (ISO, 2017). Furthermore, ISO 23718 (mechanical testing vocabulary) (ISO, 2007) was used as the main resource for collecting the mechanical testing vocabulary. Accordingly, the minimum relevant terminology of the mechanical testing processes was gathered and well defined from the materials testing standards and other mentioned resources, and the resulting collection was stored in a Kupfer*Digital* GitLab repository (see data availability). For example, Fig. S1 shows a part of such terminology collection for the tensile test process.

In the next step, the knowledge graphs of different mechanical testing methods were designed based on the previously prepared testing terminologies and the semantic relationships between them. Designing the process graphs was graphically performed in the Draw.io tool with the help of the Ontopanel plugin (Chen et al., 2022). In this regard, the Ontopanel plugin facilitates the process graph development by using the shapes provided in its library, reusing the entities of the upper-level ontologies in EntityManager, and converting the graphical schema into OWL codes in its convertor (Fig. S2).

Fig. 3 shows an example of the process graph that has been designed for the tensile testing method. As seen in its yellow namespace box (Fig. 3A), the upper-level ontologies BFO, CCO, IOF, QUDT, and MSEO were utilized for designing the tensile test process graph. We selected BFO as the top-level ontology not only because it is very well structured for learning and usage, but also for its wide applications and capabilities of merging with a variety of ontologies, and therefore its potential to make the developed ontologies highly interoperable. It is worth noting that according to the recent ISO/IEC 21838-2 standard, BFO is conformant to the requirements specified for top-level ontologies (ISO, 2022). In the mid-levels also, CCO and IOF ontologies enable high compatibility with BFO and provide lots of manufacturing and industrialization concepts for designing process graphs. The QUDT ontology has also been used particularly for its valuable units of measurement. Eventually, the application of MSEO enables the process graphs to access a wide entity in the field of materials and their mechanical parameters. Furthermore, the namespaces of MTO and KD (from Kupfer*Digital*) were utilized for indexing the newly created class/properties and individuals, respectively. As can be seen in Fig. 3, the description of the complex tensile test process can be mostly performed by reusing the entities of mentioned higher-level ontologies, and only specific domain entities need to be added. This use case shows why reusing as many as possible higher-level ontologies have been recommended for developing domain-level ontologies (Arp et al., 2015). Reusing these higher-level ontologies not only facilitates the development of process graphs, but also establishes the usage of well-defined entities and standardized ontology design with a universal hierarchy, and eventually higher interoperability of designed process graphs with other ontologies.

For the sake of enabling the knowledge graphs to answer the planned competency questions, the ontology and knowledge graphs should be designed in such a way that provide a general description of the test method according to the level of detail given in the materials testing standards. In other words, they should describe the concept of a testing process in terms of its relationships to other materials concepts like materials characteristics, process history, test equipment, and data analyzing methods. On the other side, the process graphs need to be compatible with the different types of test data, be able to fully map the reported metadata of the test reports, and further use universal terms to enable the users to easily manage the data mapping process. To cover all the requirements raised by the competency questions, materials testing standards, and useful metadata of test reports, the tensile test process graph has been designed in several boxes:

- The provenance box (Fig. 3B) contains the basic semantics of the test standard, test date, and test data file.
- The test pieces box (Fig. 3C) defines the specimen metadata like its name, composition, manufacturing information, processing history, shape, and dimensions.
- The equipment box (Fig. 3D) with entities around the name and type of the testing equipment.
- The test process box (Fig. 3E) can be divided into two parts of “the act of measuring” and “the act of analyzing”. The first part provided the measurement-related concepts of the testing process, such as test temperature, gauge length, stress/strain rate, test speed, and measured time-force-displacement dataset. In the second part, the parameters like stress, strain, stress-strain curve, yield point, elastic modulus, tensile strength, elongation, reduction of area, and test uncertainty are calculated using the provided semantics and by using the analysis methods given in the algorithm box.

The expression of ontological statements in Draw.io is performed using the different shapes that are assigned to classes, individuals, properties, axioms, and restraints. These shapes and their embedded texts are subsequently converted to OWL or TTL codes using the Ontopanel converter. For example, the upper first row of the provenance box (extracted in Fig. S3a) was converted to TTL codes of the machine-processable subject–predicate–object triples format in Fig. 3F. The OWL codes of the same context were also represented in Fig. S3b. Here, *Regulation* and *NonNameIdentifier* are respectively classes from MSEO and CCO, while *TensileTest* is a new class that will be added to MTO. In addition, *designated_by* and *in_accordance_with* are object properties from CCO and MSEO, respectively. Moreover, *TensileTestExperiment*, *TestStandard*, and *TestStandardIdentifier* are individuals from *mto:TensileTest*,

[11] http://maptomethod.kupferdigital.org/.
[12] http://rdfconverter.kupferdigital.org/.
[13] http://www.irisa.fr/LIS/ferre/sparklis/.

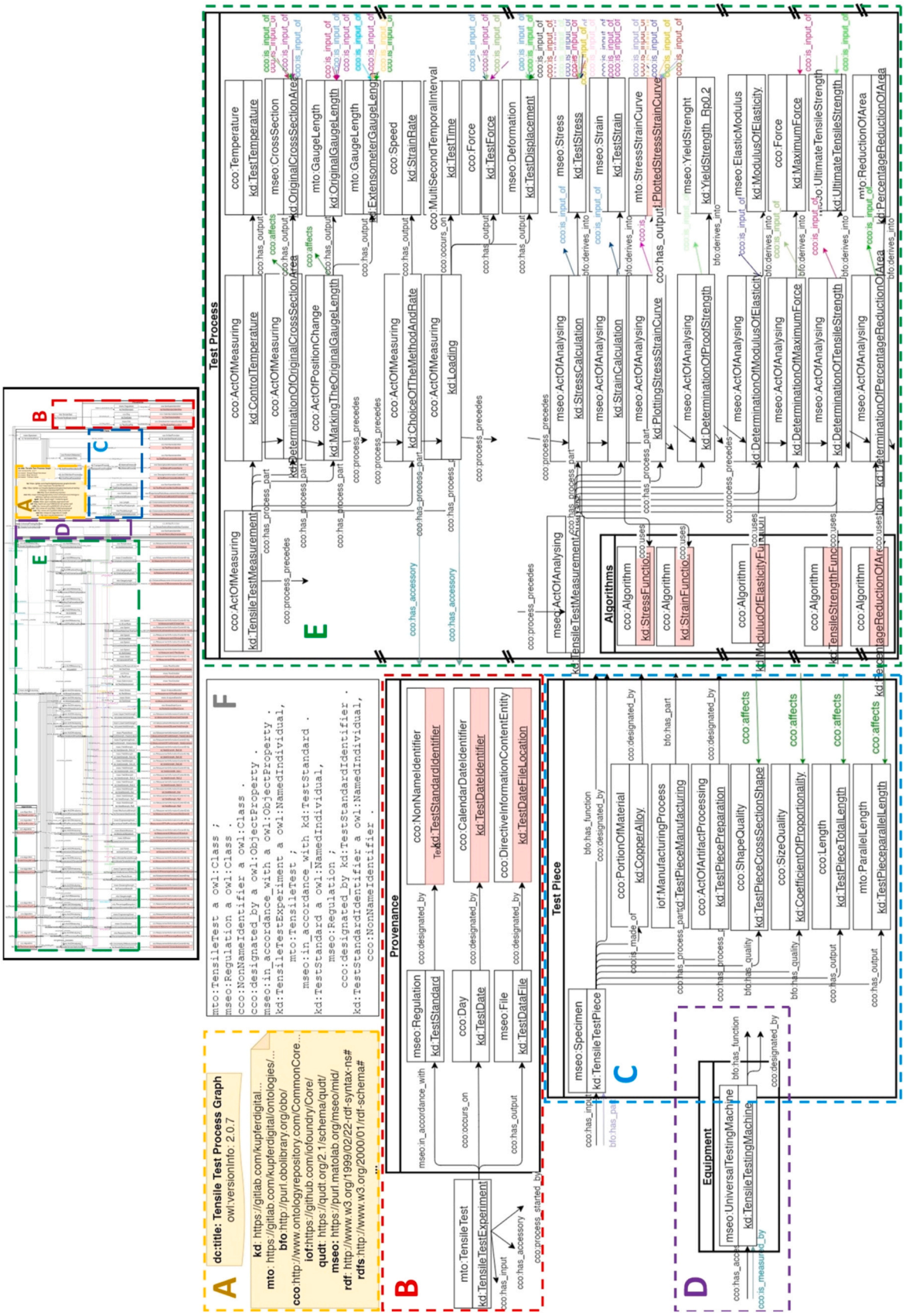

Fig. 3. Knowledge graph of the tensile testing method and some of its magnified parts: A) namespaces and metadata, B) provenance, C) test piece, D) equipment, E) test process, F) TTL codes representing the first row of the knowledge graph. The full version of the developed knowledge graph can be found in the Kupfer*Digital* GitLab repository (see data availability).

mseo:Regulation, and *cco:NonNameIdentifier* classes. In this case, those individuals in red boxes are planned for mapping the testing data. For example, we will map the *TestStandardIdentifier* individual to the metadata like "DIN EN ISO 6892–1″ from the test data report. For this purpose, we will show later that all the planned individuals for mapping the metadata should be instances of one of the *cco:InformationContentEntity* classes.

In Fig. 3, the colored arrows also model the complex relationships between the mentioned concepts of the process graph boxes. For example, calculating the ultimate tensile strength is part of the *mseo: ActOfAnalysing* whose process graph and TTL codes are represented in Fig. S4. Here, the maximum force, test piece cross-section, and stress are the input concepts for calculating the ultimate tensile strength, and it uses a *cco:Algorithm* for the calculation method that is described in the test standard. The *mseo:UltimateTensileStrength* result of this calculation is a value that can be mapped in the *kd:MeasurementOfUltimateTensileStrength*. Later and during the data mapping process, the users just see the final item of *kd:MeasurementOfUltimateTensileStrength*, and they can easily map the tensile strength data of the test report in it.

It should be noted that although more concepts can be added to different boxes of the process graph, the presented graphs modeled the level of detail that fulfills the requirement of the competency questions and test standards. In other words, the creation and review of process graphs by the collaboration of BAM (as a participant of the standards committee and host of accredited materials testing laboratory) and well-known industrial/academic partners leads to the development of workflows that meet both the needs of test standards and data structures of industrial labs.

Moreover, the process graph was validated against SHACL shapes employing the OntoFlow tool (Fig. S5). Finally, Ontopanel is utilized for converting the whole graphical process graph to OWL and TTL files. Similarly, some other process graphs were developed for different mechanical testing methods like Brinell and Vickers hardness, and the corresponding modeling files were stored in the Kupfer*Digital* GitLab repository (see data availability).

MTO was built by collecting the entities of developed mechanical testing process graphs and based on ISO 23718 standard. Fig. 4 represents part of the MTO new classes hierarchy. According to ISO/IEC 21838-2 standard, BFO was used as the top-level ontology for the development of MTO. As seen in Fig. 4, BFO is constructed by an upper class entitled *bfo:Entity*, which has two subclasses of *bfo:Continuant* and *bfo:Occurrent*. A *bfo:Continuant* is an entity that persists, endures, or continues to exist through time while maintaining its identity, while the *bfo:Occurrent* is an entity that unfolds itself in time. The *bfo:Continuant* has three subclasses: *bfo:IndependentContinuant* with two main subclasses of *bfo: ImmaterialEntity* (like *bfo:SpatialRegion*) and *bfo:MaterialEntity* (*especially bfo:Object* and *bfo:ObjectAggregative*), *bfo: GenerallyDependentContinuant*, and *bfo:SpecificallyDependentContinuant* that provides important subclasses of *bfo:Quality* (*bfo:RelationalQuality*) and *bfo:RealizableEntity* (like *bfo:Disposition* and *bfo:Function*) for developing the ontology. On the other side, *bfo:Occurrent* subclasses of *bfo: Processes* and *bfo:TemporalRegion* highly contribute to the construction of testing process concepts of MTO.

As a mid-level ontology, CCO added a set of conservative extensions to the different parts of BFO. Particularly, four parts of CCO were widely

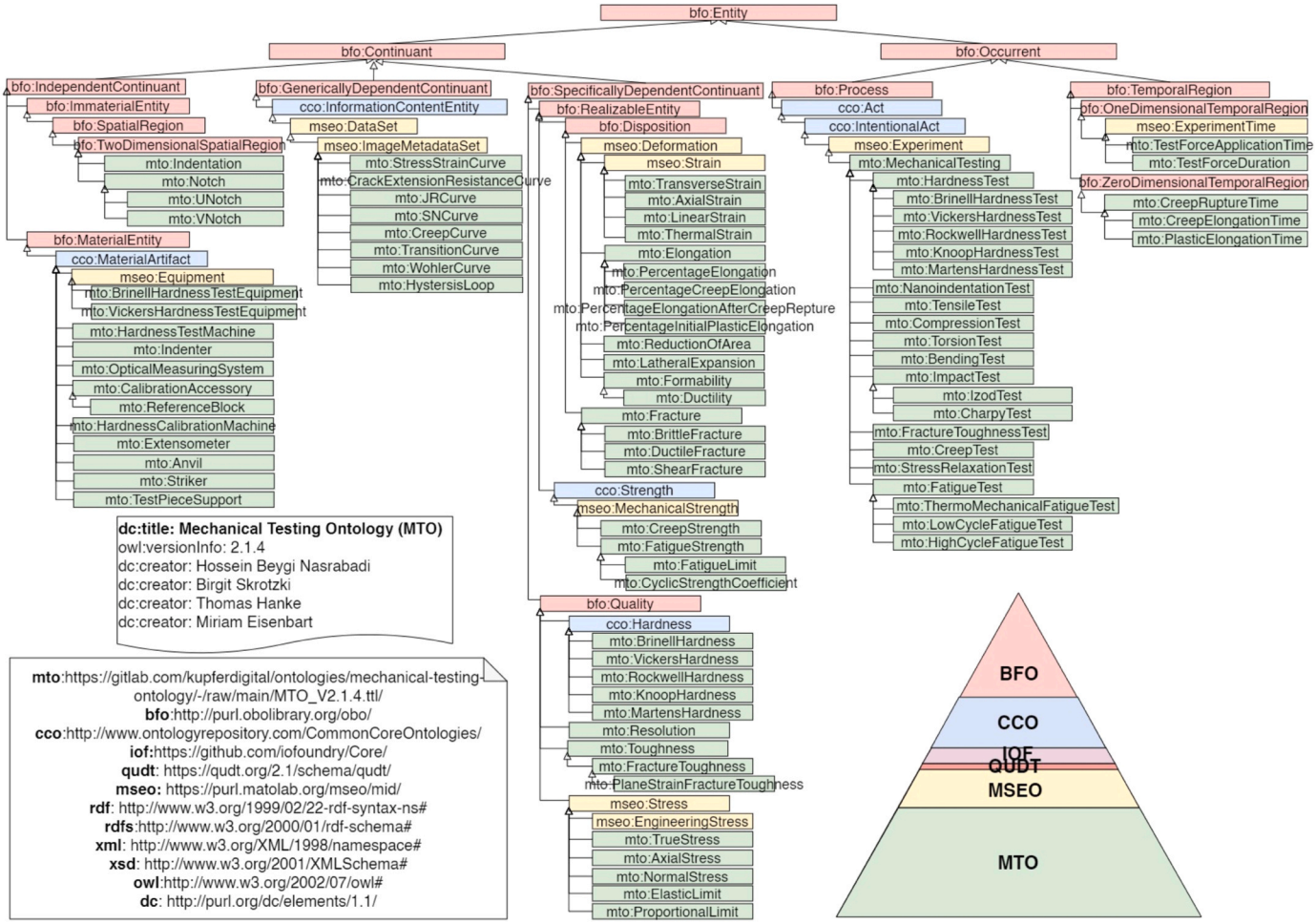

Fig. 4. Part of the MTO new classes hierarchy.

used for the development of MTO: artifact ontology with concepts like *cco:PortionOfMaterial*, information entity ontology that represents *cco: InformationContentEntity*, a quality ontology for representing qualitative concepts (such as size, shape, temperature, amount, strength, hardness), and event ontology which deals about the acts in processes (like *cco: ActOfMeasuring*, or *cco:ActOfManufacturing*). As already mentioned, *cco: InformationContentEntity* is one of the important entities of the MTO since it represents the generic types of information like reports, sentences, data values, algorithms, and specifications that can map a wide variety of metadata in the process graphs. As another high- and middle-level ontology, IOF added concepts like *iof:ManufacturingProcess* and *iof: MaterialResource* to MTO and made it more applicable for industrial purposes. QUDT was also utilized for mapping a wide variety of measurement units. In the case of new units which were not available there (like HBW and HV, the Brinell and Vickers hardness units) we also added them to the *qudt:DerivedUnit*. At a lower level, MSEO provided many entities about materials and their mechanical parameters, such as specimen, equipment, experiment, deformation, strain, stress, mechanical strength, ultimate tensile strength, yield strength, and Elastic modulus.

It can be found from Fig. 4 that appropriate concepts of such mentioned upper-level ontologies covered around 50 % of the MTO class entities. Reusing these upper-level ontologies not only provides a global hierarchy and well-defined concepts for MTO but also makes it more interoperable with other ontologies in the fields of materials science and engineering and desirable for industrial applications. As a domain-level ontology, MTO was developed in four main parts:

- Mechanical testing experiments: Entities for mechanical testing methods like tensile, hardness, creep, and fatigue tests were created as the subclasses of *mseo:Experiment*.
- Mechanical testing quantities: Entities related to the parameters of mechanical testing were placed in the appropriate hierarchies of *bfo: Disposition* and *bfo:Quality* classes. As examples of such entities *mto: Toughness*, *mto:Elongation*, and *mto:FatigueStrength* can be named.
- Mechanical testing artifacts: Entities of mechanical testing devices and their components and accessories (like *mto:HardnessTestMachine*, or *mto:Indenter*), according to the level of requirements details mentioned in the test standards. These entities were classified as subclasses of *cco:Artifact*.
- Mechanical testing data: As subclasses of *cco:InformationContentEntity,* mechanical testing data like stress-strain, S-N, or creep curves provide the basic data representation of mechanical testing processes.

Table S1 lists parts of the 110 newly created classes of MTO, along with their definitions and correct hierarchies in the ontology. It should also be noted that the level of defined properties in the upper-level ontologies was enough for developing MTO, in such a way that no new property entity was defined for MTO.

As shown in Fig. S6, MTO was built in the Protégé ontology editor software by importing the reused concepts of higher-level ontologies (for assuring correct naming, URI, definition, relation, and hierarchy of reused entities), adding newly created entities of Table S1 along with their definitions and restrictions, and importing the designed process graph of various test methods (for providing the individuals and expressing the relationships between different concepts). Fig. 5 shows the schemas of the MTO classes and object properties structures in the Protégé environment. MTO is publicly available via the Kupfer*Digital* GitLab repository (see data availability).

It should be noted that the MTO development based on the test standards and reusing the well-developed upper-level ontologies can also highly resolve the semantic heterogeneities in the data sources.

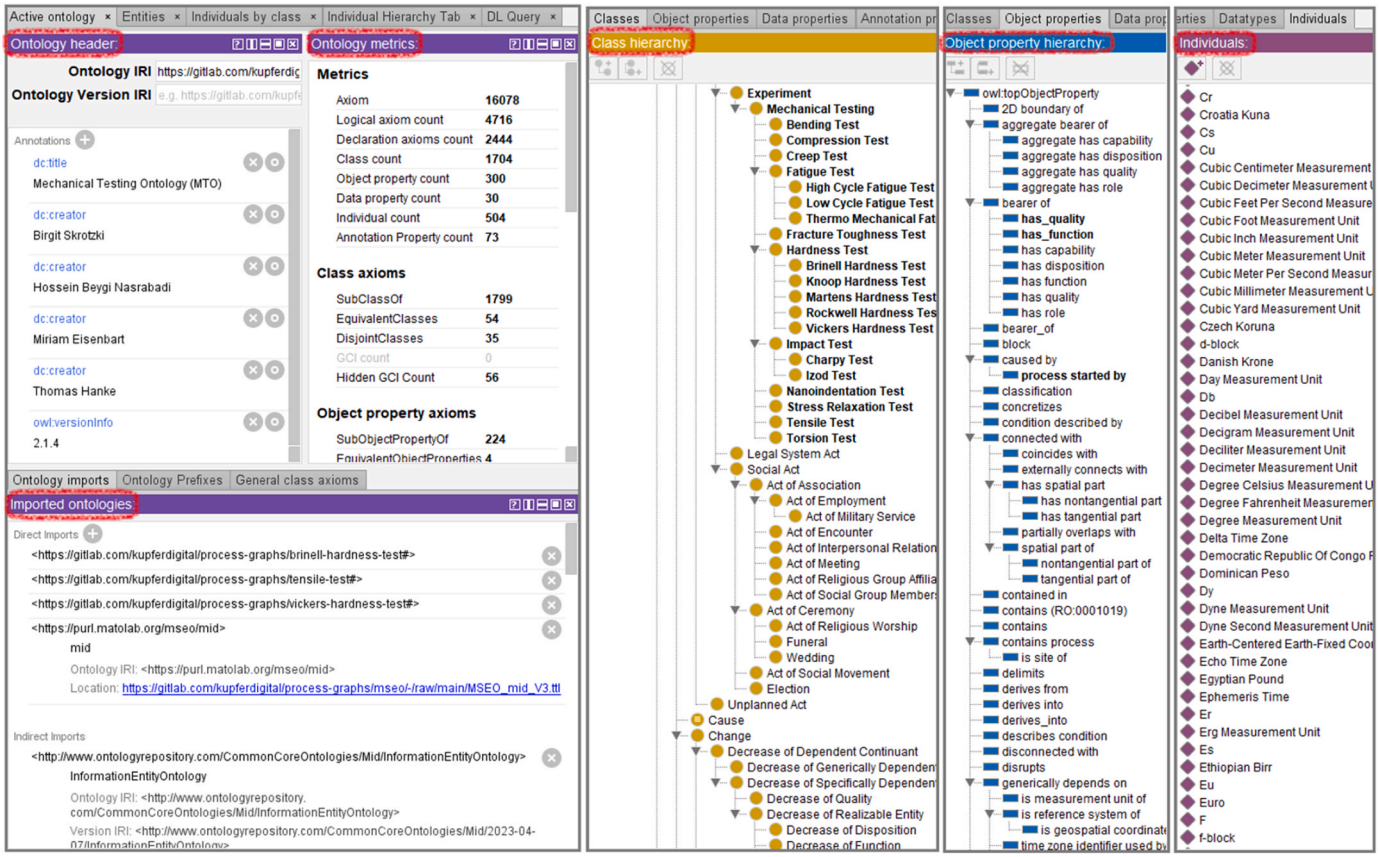

Fig. 5. Images of MTO in Protégé showing imported process graphs, ontology metric, and the hierarchy of classes, object properties, and individuals.

Semantic heterogeneity in data refers to the variations in the meaning and interpretation of datasets that are provided by different sources. Differences in terminology, concept classification, and data models are the main reasons for the semantic data heterogeneity. Here, the key concept for resolving the semantic heterogeneities is that both the ontology and datasets are prepared based on the test standards. It means that almost the same terminology, data processing equations, and units are used for preparing the test data from different sources, which are identical to the concept of ontology derived from the standards. For example, different data sources may use varying terminologies to refer to "ultimate tensile strength", like maximum tensile strength, peak tensile strength, ultimate strength, tensile limit, etc. Here, the standard-derived ontology provides a universal and community-accepted terminology, description, equation, and unit for this concept. Therefore, any data with such structures can be mapped to this common concept and resolve this heterogeneity. Furthermore, designing the MTO based on the BFO upper-level ontology ensures its unified and standard classification and resolves semantic data heterogeneity from concept classification. Finally, employing the tools developed in the next sections, three examples of different heterogeneous test data are mapped into their knowledge graphs. It will be shown that mapping the metadata-contained XML to RDF using the MTO ensures interoperability and semantic consistency across different XML-based data sources.

5. FAIR data generation in the digital materials mechanical testing lab

Generation of FAIR data in the digital lab can be achieved by data mapping of each test result in its corresponding process graph and the construction of machine-readable RDF test data. In the trial phase of materials testing data digitalization, different cast copper samples were tested by extensive mechanical testing methods (like Brinell and Vickers

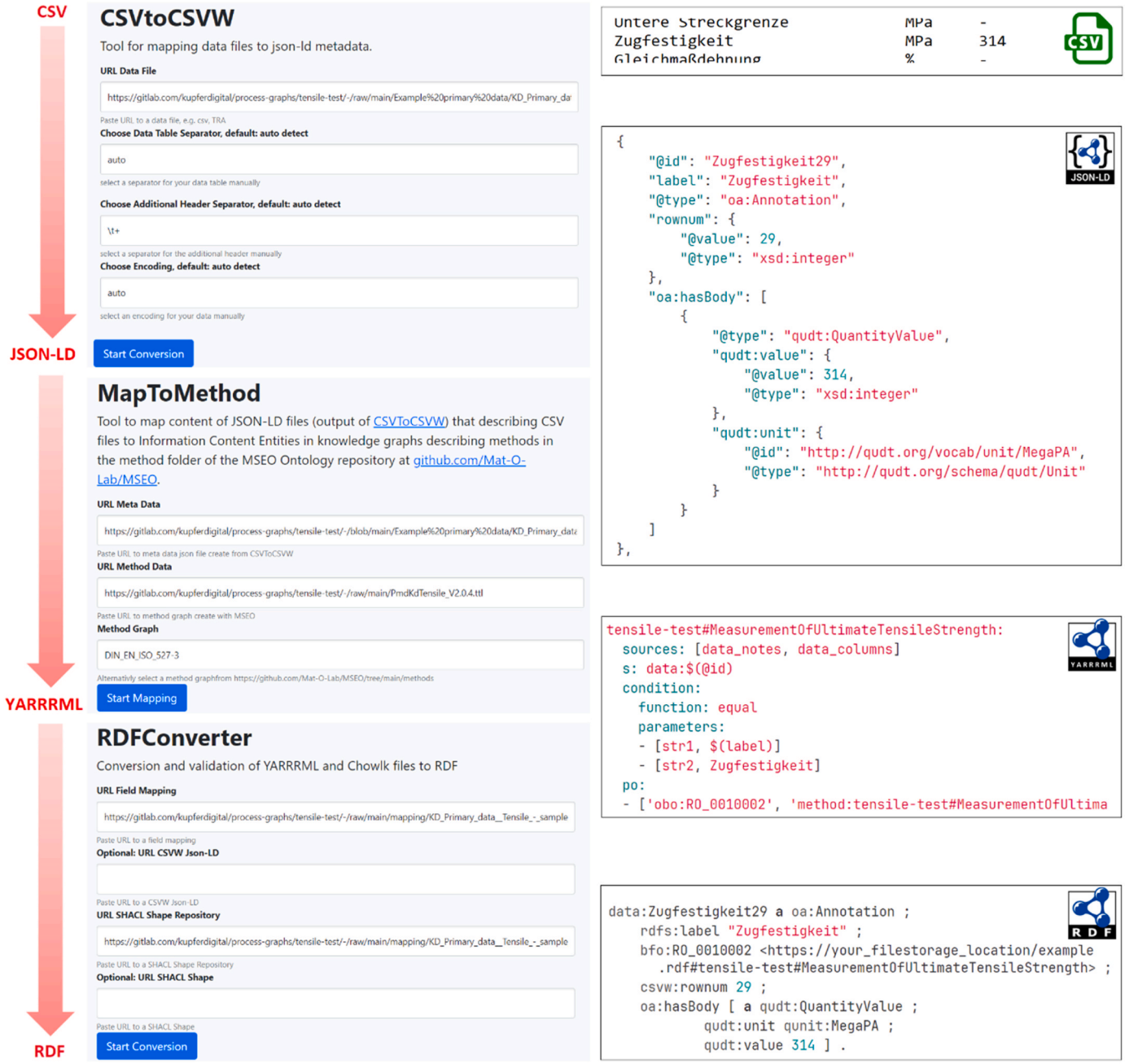

Fig. 6. Mat-O-Lab approach for mapping an example tensile test data in its process graph.

hardness tests, and tensile tests), and according to related ISO standards. The heterogeneous test data of various structures, languages, and formats (like CSV, XLSX, and complex Excel sheets) resulted from different testing processes (Beygi Nasrabadi et al., 2023). Within this project, different ontology-based approaches like Platform MaterialDigital (PMD), Data Space Management System (DSMS), Mat-O-Lab, and Ontopanel are utilized for the data parsing, mapping, and RDF conversion. Although the current paper essentially doesn't aim to rely on RDF data conversion tools and their comparison, we address this interesting topic in our next papers. There, various data mapping and RDF conversion tools are evaluated by considering the requirements of the data sharing, different test data types and formats, compatibility with varied ontologies and knowledge graphs, and their user-friendly and simple-applicability features. However, to evaluate the capabilities of developed knowledge graphs in mapping a wide range of test data, just two approaches of Mat-O-Lab and Ontopanel are utilized in the following sections for data mapping the datasets of different structures and formats. The mentioned tools also enable direct RDF data generation. Here, data generating in RDF format offers several benefits, like providing a meaningful and detailed data representation by explicitly capturing the ontology-based data and semantic relationships between resources, being highly qualified for web applications and highly interoperable with most other linked data on the web, enabling powerful querying capabilities using standardized query languages (e.g., SPARQL), supporting the representation of provenance information and metadata, and serialization into other different formats (e.g., RDF/XML, Turtle, JSON-LD).

5.1. Mat-O-Lab approach

Fig. 6 represents the serialization of CSV tensile test data into other different formats and the generation of machine-readable RDF data utilizing the Mat-O-Lab toolchain. As seen in Fig. S7, the example CSV file of tensile test data contains a variety of metadata about the sample and testing along with the dataset of the measurements in a human-created structure and German language. Therefore, CSVToCSVW tool was utilized in the first step to convert such test data into standardized and machine-readable structures. For example, it is shown in the right corner of Fig. 6 that the content of one line of the CSV file (about tensile strength, "Zugfestigkeit" in German language, data of the test with a value of 314 MPa) has been converted into the machine-readable codes of the JSON-LD format. The CSVToCSVW tool has also the ability to convert all other data of the CSV test report into the standardized structure of a uniform language, unit, and numerical system. In the next step and using the MapToMethod tool, the user can import the JSON-LD metadata file and select the corresponding process graph from the provided options (in this case tensile test process graph) to start the mapping procedure. By running the mapping procedure, another window (Fig. S8) appears that lists all the *cco:InformationContentEntity* individuals of the tensile test process graph. Clicking on each item, the user can select the correct option from the *cco:InformationArtifactBearing* individuals of the test data file. For example, "Zugfestigkeit" should be mapped in the "MeasurementOfUltimateTensileStrength" item (Fig. S8). In the same way, all the other items of test data can be mapped into their corresponding individuals on the process graph. The result of such a data mapping process is a YARRRML file, and its codes for mapping the ultimate tensile strength data are shown in Fig. 6. In the last step, the YARRRML file obtained from the data mapping is converted into the machine-readable RDF data by the means of RDFConverter tool. Here not only all the items of the tensile test report were successfully mapped in their corresponding process graph concepts (like mapping the "Zugfestigkeit" metadata in "MeasurementOfUltimateTensileStrength"), but also the data values and their units were represented in uniform structures.

Using the RDF Grapher[14] tool, Fig. S9 graphically shows how the test data were mapped in the tensile test process graph and what kind of complex relationships and semantics exist between different testing concepts and their data.

Furthermore, the XLSX Brinell hardness test data of a copper sample was also successfully converted to the machine-readable JASON-LD, YARRRML, and RDF files by utilizing the Mat-O-Lab toolchain (Fig. S10), indicating the high capability of this approach for converting the different and heterogeneous types and formats of the test data.

5.2. Ontopanel approach

Ontopanel is a Draw.io plugin that not only helps with the graphical development of the process graphs and ontologies but also enables graphical data mapping of big datasets and converting the mapped graphs to RDF files. An example process for mapping the Vickers hardness dataset to its process graph is represented in Fig. 7. Here, the Vickers hardness dataset was collected for different copper alloys in a complex and human-created XLSX sheet and contains a variety of columns dealing with the materials and testing metadata, along with their measurement and analysis values (see complete dataset in Fig. S11). Such a complex dataset can be easily and graphically mapped into its corresponding process graph via the Ontopanel tool. The orange part of Fig. 7 shows part of a Draw.io page which contains the Vickers hardness process graph and Ontopanel-Converter plugin. Here, the mentioned complex Excel datasets can be imported into the plugin, and the data mapping process can easily be performed by choosing a *cco:InformationContentEntity* individual and linking it (clip sign) to the related data column (Fig. 7). After mapping all the columns of the dataset in the process graph, the dataset mapped process graph is converted in RDF or TTL via the Ontopanel-Convertor plugin (right side of Fig. 7).

6. Utilization of FAIR mechanical testing data

The RDF mechanical testing data of standardized content and structure are stored in functional data spaces like Kupfer*Digital* CKAN (see Data availability) to ensure data sovereignty and the availability of distributed data (Fig. S12). The data spaces offer the users to store their real metadata which cannot be edited by other parties, prevent any data leakage, and allow large, decentralized amounts of heterogeneous data to be implemented by providing generic data formats and structures. By the storage of such RDF files in publicly available data spaces, the mechanical testing data satisfies all the requirements of FAIR data. The following paragraphs describe some potential examples of utilizing the generated FAIR mechanical testing data in the industrial domains.

6.1. Data sharing

The published datasets in the data spaces are referenceable with a uniform resource identifier (URI) or more specifically a uniform resource locator (URL). Here, the data providers ensure full data sovereignty and can manage their triple stores, establish the applicable publication policy, and even identify and authorize their data users by a central identity provider. Accordingly, industries can protect their private data in the data spaces and share publicly available data with other industries and academic centers. On the other side, the companies obtain the opportunity to access lots of trustable and traceable mechanical testing data of other organizations that are available in the data spaces.

[14] https://www.ldf.fi/service/rdf-grapher.

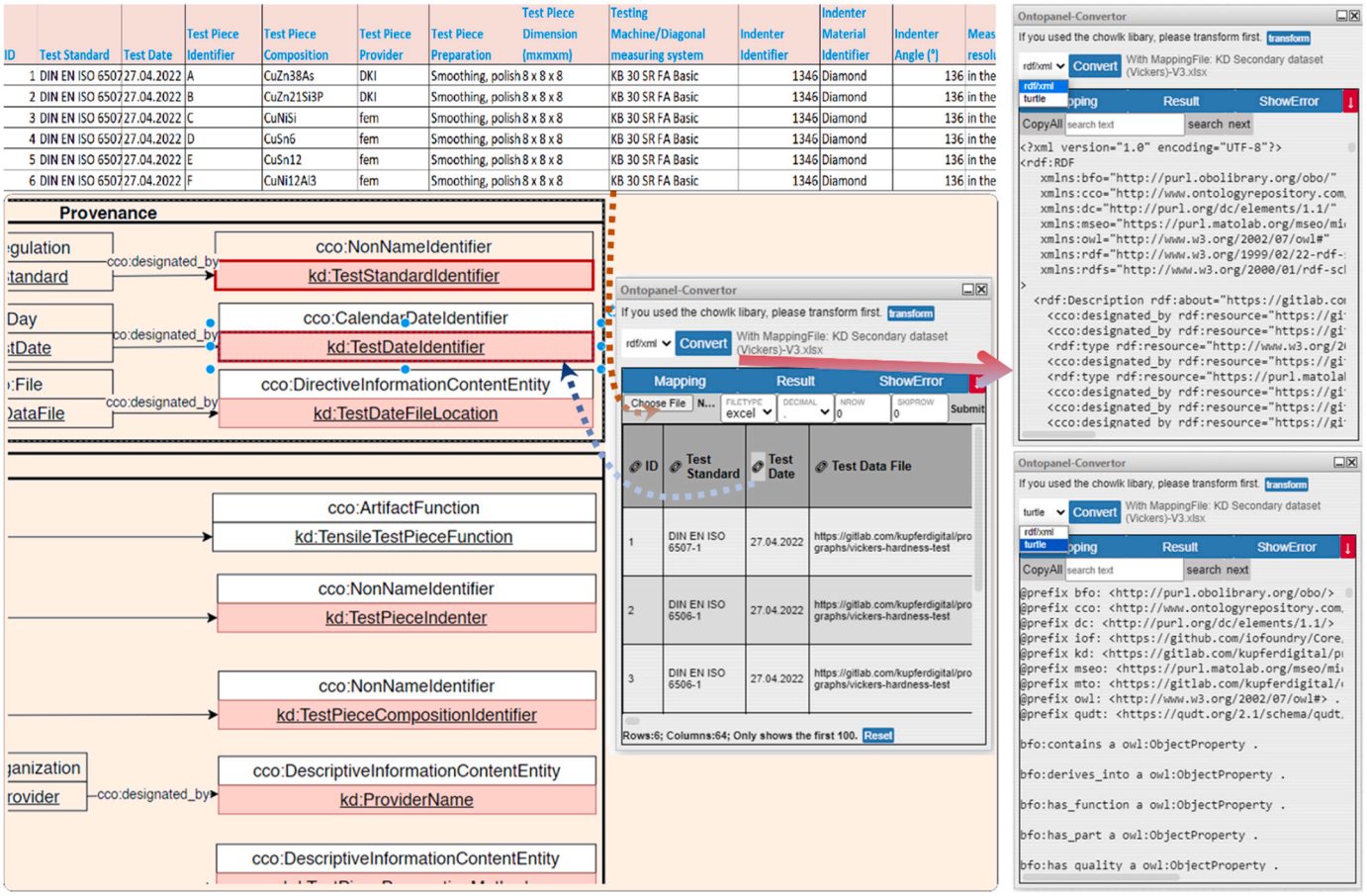

Fig. 7. Mapping of a complex Vickers hardness dataset into its process graph and conversion of data-mapped graphs to RDF/TTL files by the Ontopanel approach.

6.2. Data reuse

One big advantage of sharing FAIR mechanical testing data is that the industries and academic environments can reduce duplication of costly compound preparation and testing, and reuse, compare, and reproduce the already existing experimental databases, which paves the way for the creation of automated data pipelines. Other potential users of mechanical testing data are the data scientists and model developers who need data for the calibration and validation of their models. Furthermore, data management platforms provide a logical and digital architecture that enables the retrieval, integration, visualization, and analysis of data between different networks.

6.3. Data-driven materials research, prediction, and product development

With the digitalization of materials testing data, extensive and comprehensive machine-readable data will be available for data-driven materials research, such as machine learning and deep learning approaches. This will lead to a better understanding of the properties and behavior of materials, optimizing the process, and predicting the most efficient production processes. The solutions developed give industrial partners also the opportunity to strengthen their international competition, and proactively evaluate new materials of better properties for their value chains.

6.4. Semantic query

The FAIR mechanical testing data can be dynamically queried by complex semantic searches. For instance, the below examples show SPARQL queries of some competence questions from the prepared RDF files of Section 5, which can eventually utilized for evaluating the introduced ontology development and data mapping approaches:

- Example 1 – What is the ultimate tensile strength value of an alloy with a composition of CuSn6? This query aims to evaluate whether the developed tensile test knowledge graph and mapping the CSV tensile test data in it (utilizing the Mat-O-Lab approach) can result in successful data retrieval. As seen in Fig. S7, the ultimate tensile strength value of this alloy is 314 MPa, while the correct answer was also successfully obtained by the SPARQL query (Fig. 8a).
- Example 2 – What is the final Brinell hardness and uncertainty value for the alloy with a composition of CuSn6? According to Fig. S10, the answer to this question is 83 ± 6 HBW 2.5/62.5, which was also correctly answered by the SPARQL query (Fig. 8b). This finding suggests that the XLSX Brinell hardness test data were successfully mapped in the Brinell hardness knowledge graph utilizing the Mat-O-Lab approach.
- Example 3 - What is the maximum Vickers hardness among the measured values in the dataset? In this case, the complex Excel sheet from the various Vickers hardness measurements was mapped into the Vickers hardness knowledge graph using the Ontopanel tool. According to this dataset (Fig. S11), the alloy CuNi12Al3 has the maximum Vickers hardness value of 273.6 HV5. As shown in Fig. 8c, this value was also correctly recognized through the natural language query of the mentioned question within the Sparklis interface.

More competency questions that have been considered for the evaluation of the introduced approach can be found in Kupfer*Digital* GitLab repository (see data availability). As a result of successful data retrieval by the SPARQL queries, it can be concluded that different types

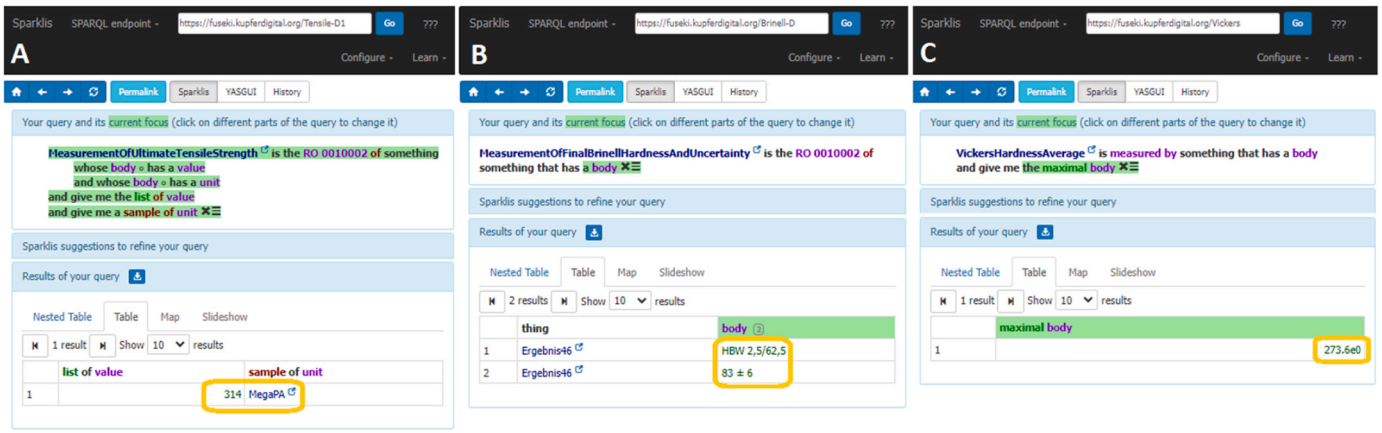

Fig. 8. Example SPARQL queries using the Sparklis natural language interface from generated RDF files: a) What is the ultimate tensile strength value of an alloy with a composition of CuSn6? b) What is the final Brinell hardness and uncertainty value for the alloy with the same composition? c) What is the maximum Vickers hardness among the measured values in the dataset?

and formats of the test data were successfully mapped in the developed knowledge graphs. On the other side, considering both the test standards and competency questions' requirements led to the development of the mechanical testing ontology and knowledge graphs that represent the highly detailed materials testing process with standardized and well-defined entities. Accordingly, such important findings can clarify how the introduced approach connects with the stated research problem; i- The way that the ontology and knowledge graph are developed makes them compatible with all types of test reports, industrial data, and other kinds of ontologies in the materials science domain. ii- The introduced methodology and toolchains for ontology development, data transformation, mapping, and storage enable the digital representation of different types of lab-generated testing reports. iii- The introduced approach can covert the experimental test data of heterogeneous structures, languages, and formats to the comprehensive machine-readable FAIR data.

7. Conclusion

A schema for designing a digital mechanical testing lab was introduced in this research. As a prerequisite for designing such labs, the highly detailed process graphs of a variety of mechanical testing methods (such as the tensile test, Brinell, and Vickers hardness) were developed based on the requirements of test standards and the test metadata of industrial labs. Regarding, the Kupfer*Digital* GitLab and its developed workflows like OntoFlow have been used for the collaborative design of process graphs and ontologies, and their automatic validation. MTO was prepared by collecting all the designed entities of different process graphs and gathering the mechanical testing vocabulary from ISO 23718 standard. Based on ISO/IEC 21838-2 standard, MTO was developed by utilizing the BFO, CCO, IOF, QUDT, and MSEO as upper-level ontologies. Reusing these upper-level ontologies and materials testing standards not only makes MTO highly interoperable with other ontologies but also ensures its acceptance and applicability in the industry. MTO represents the mechanical testing methods in the 230 classes via four main parts of mechanical testing experiments, quantities, artifacts, and report data. The infrastructure of the mechanical testing lab provides an environment for mapping the materials testing data in the modeled process graphs and storing the resulting data in suitable data spaces. In this regard, two approaches were evaluated for mapping the test data in their corresponding process graphs. While the Mat-O-Lab approach provided a professional and powerful environment for the parsing, mapping, and RDF conversion of test data, the Ontopanel approach can graphically map the complex test datasheets in the process graphs and convert them to RDF files. Although the current research basically doesn't rely on the tools' evaluation, our next paper will compare the functionality of Mat-O-Lab, Ontopanel, PMD, and DSMS tools for the mapping, conversion, and storage of the created RDF datasets. However, both Mat-O-Lab and Ontopanel approaches resulted in perfect data mapping performances and ensured that the different test data reports of heterogeneous structures, languages, and formats could successfully be converted to machine-readable RDF data of uniform and standardized structures. Efficient data management systems were utilized to store such test data within the structured, comprehensive, and documented data repositories. The RDF data stored in such data spaces can allow industries in the future to access lots of trustable and traceable mechanical testing data of other academic and industrial organizations. Such FAIR data were also successfully queried by the Sparklis natural language interface, indicating that the introduced approach can support data retrieval, response to competency questions, and eventually address the research problem. The generation and storage of such FAIR data in the digital materials mechanical testing lab can eventually accelerate data-driven materials research such as machine learning and deep learning approaches.

CRediT authorship contribution statement

The article was conceived and structured by all the authors. **Hossein Beygi Nasrabadi:** Methodology, Software, Validation, Investigation, Data curation, Writing – original draft. **Thomas Hanke:** Methodology, Software. **Matthias Weber:** Conceptualization. **Miriam Eisenbart:** Conceptualization, Supervision. **Felix Bauer:** Resources. **Ladji Tikana:** Resources. **Roy Meissner:** Software. **Gordian Dziwis:** Software. **Yue Chen:** Software. **Birgit Skrotzki:** Conceptualization, Data curation, Supervision. All the authors reviewed the manuscript, process graphs, and ontology.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

The related data of this research are publicly available in below repositories:

Process graphs of mechanical tests: https://gitlab.com/kupferdigital/process-graphs.

Mechanical testing ontology (MTO): https://gitlab.com/

kupferdigital/ontologies/mechanical-testing-ontology, http://matportal.org/ontologies/MTO, and http://industryportal.enit.fr/ontologies/MTO.

Raw materials mechanical testing data (CSV, XLSX): https://doi.org/10.5281/zenodo.7670583.

Digitalized materials mechanical testing data (RDF): http://ckan.kupferdigital.org/.

Acknowledgments

This work was funded by the German Federal Ministry of Education and Research (BMBF) through the project 'KupferDigital-Datenökosystem für die digitale Materialentwicklung auf Basis Ontologie-basierter digitaler Repräsentationen von Kupfer und Kupferlegierungen' (grant number 13XP5119). Funding provided by Bundesanstalt für Materialforschung und -prüfung (BAM) in the framework of the project 'Materials-open-Lab (Mat-o-Lab)' enabling the participation of Yue Chen is gratefully acknowledged. The authors would like to acknowledge Mr. Patrick Uhlemann and Mr. Steffen Thärig (BAM) for performing mechanical testing, Mr. Patrick Wais (fem) for assisting in figures drawing, as well as Dr.-Ing. Jürgen Olbricht, Dr.-Ing. Markus Schilling (BAM), and Mr. Leon Steinmeier (Helmholtz-Zentrum Dresden-Rossendorf), for reviewing the manuscript. The authors also thank colleagues in the Kupfer*Digital* project, Mat-O-Lab, and PMD initiatives, as well as industrial project partners (Aurubis, C.Hafner, Indutherm, wbk, Schaeffler, Wieland, and KME) for sharing their valuable perspectives related to the digital transformation of industrial data.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.compind.2023.104016.

References

Ali, M.M., et al., 2019. A product life cycle ontology for additive manufacturing. Comput. Ind. 105, 191–203. https://doi.org/10.1016/j.compind.2018.12.007.
Arp, R., Smith, B., Spear, A.D., 2015. Building Ontologies with Basic Formal Ontology. Mit Press. https://doi.org/10.7551/mitpress/9780262527811.001.0001.
Ashino, T., 2010. Materials ontology: an infrastructure for exchanging materials information and knowledge. Data Sci. J. 9, 54–61. https://doi.org/10.2481/dsj.008-041.
Ashino, T., Fujita, M., 2006. Definition of a web ontology for design-oriented material selection. Data Sci. J. 5, 52–63. https://doi.org/10.2481/dsj.5.52.
Austin, T., et al., 2010. A Guide to the Development and Use of Standards-Compliant Data Formats for Engineering Materials Test Data.
Bayerlein, B., et al., 2022. A perspective on digital knowledge representation in materials science and engineering. Adv. Eng. Mater. 2101176. https://doi.org/10.1002/adem.202101176.
Beygi Nasrabadi, H., et al., 2023. KupferDigital mechanical testing datasets. DOI: https://doi.org/10.5281/zenodo.7670583.
Bhat, M., et al., 2013. Premλp: knowledge driven design of materials and engineering process. In: ICoRD'13. Springer, pp. 1315–1329. https://doi.org/10.1007/978-81-322-1050-4_105.
Borgo, S., et al., 2022. DOLCE: a descriptive ontology for linguistic and cognitive engineering. Appl. Ontol. (Preprint), 1–25. https://doi.org/10.3233/AO-210259.
Chen, Y., et al., 2022. Ontopanel: a tool for domain experts facilitating visual ontology development and mapping for FAIR data sharing in materials testing. Integr. Mater. Manuf. Innov. 1–12. https://doi.org/10.1007/s40192-022-00279-y.
Cheung, K., J. Drennan, J. Hunter.Cheung, K., Drennan, J., Hunter, J., 2008. Towards an Ontology for Data-driven Discovery of New Materials. In: AAAI Spring Symposium: Semantic Scientific Knowledge Integration.
Cheung, K., Hunter, J., Drennan, J., 2009. MatSeek: an ontology-based federated search interface for materials scientists. IEEE Intell. Syst. 24 (1), 47–56. https://doi.org/10.1109/MIS.2009.13.
Del Nostro, P., Goldbeck, G., Toti, D., 2022. CHAMEO: an ontology for the harmonisation of materials characterisation methodologies. Appl. Ontol. (Preprint), 1–21.
Dimassi, S., et al., 2021. An ontology-based framework to formalize and represent 4D printing knowledge in design. Comput. Ind. 126, 103374 https://doi.org/10.1016/j.compind.2020.103374.
DIN, 2003. DIN EN 10319–1:2003 Metallic materials - Tensile stress relaxation testing - Part 1: Procedure for testing machines. DOI: ⟨https://doi.org/10.31030/9271616⟩.
Ekaputra, F., et al., 2017. Ontology-based data integration in multi-disciplinary engineering environments: a review. Open J. Inf. Syst. 4 (1), 1–26.
European Commission, 2018. Cost of not having FAIR research data: Cost-Benefit analysis for FAIR research data.
European Materials Modelling Counsil, 2020. Elementary Multiperspective Material Ontology (EMMO). Available from: ⟨https://github.com/emmo-repo/EMMO⟩.
Fensel, D., 2001. Ontologies. In: Ontologies. Springer, pp. 11–18. https://doi.org/10.1007/978-3-662-04396-7_2.
Ferré, S., 2017. Sparklis: an expressive query builder for SPARQL endpoints with guidance in natural language. Semant. Web 8 (3), 405–418. https://doi.org/10.3233/SW-150208.
Fraga, A.L., Vegetti, M., Leone, H.P., 2020. Ontology-based solutions for interoperability among product lifecycle management systems: a systematic literature review. J. Ind. Inf. Integr. 20, 100176 https://doi.org/10.1016/j.jii.2020.100176.
Furini, F., et al., 2016. Development of a manufacturing ontology for functionally graded materials. in International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. American Society of Mechanical Engineers.DOI: ⟨https://doi.org/10.1115/DETC2016-59964⟩.
Ghobakhloo, M., 2020. Industry 4.0, digitization, and opportunities for sustainability. J. Clean. Prod. 252, 119869 https://doi.org/10.1016/j.jclepro.2019.119869.
Hagedorn, T.J., et al., 2019. Interoperability of disparate engineering domain ontologies using basic formal ontology. J. Eng. Des. 30 (10–12), 625–654. https://doi.org/10.1080/09544828.2019.1630805.
Hagedorn, T.J., Krishnamurty, S., Grosse, I.R., 2018. A knowledge-based method for innovative design for additive manufacturing supported by modular ontologies. J. Comput. Inf. Sci. Eng. 18 (2) https://doi.org/10.1115/1.4039455.
Himanen, L., et al., 2019. Data-driven materials science: status, challenges, and perspectives. Adv. Sci. 6 (21), 1900808. https://doi.org/10.1002/advs.201900808.
ISO, 2014. ISO 6506–1:2014 Metallic materials — Brinell hardness test — Part 1: Test method.
ISO, 2007. ISO 23718:2007 Metallic materials — Mechanical testing — Vocabulary.
ISO, 2022. ISO/IEC 21838–2:2021 Information technology — Top-level ontologies (TLO) — Part 2: Basic Formal Ontology (BFO).
ISO, 2018. ISO 6507–1:2018 Metallic materials — Vickers hardness test — Part 1: Test method.
ISO, 2019. ISO 6892–1:2019 Metallic materials — Tensile testing — Part 1: Method of test at room temperature.
ISO, 2017. ISO 12106:2017 Metallic materials — Fatigue testing — Axial-strain-controlled method.
Kulvatunyou, B.S., et al., 2018. The industrial ontologies foundry proof-of-concept project. In: IFIP International Conference on Advances in Production Management Systems. Springer. https://doi.org/10.1007/978-3-319-99707-0_50.
Lambrix, P., et al., 2023. The materials design ontology. Semant. Web (Preprint), 1–35.
Lemaignan, S., et al., 2006. MASON: A proposal for an ontology of manufacturing domain. in IEEE Workshop on Distributed Intelligent Systems: Collective Intelligence and Its Applications (DIS'06). IEEE,. https://doi.org/10.1109/DIS.2006.48.
Li, C., et al. Developing ontologies and web-based data management system for additive manufacturing processes. in Proceedings of the 2017 ACM on Web Science Conference. 2017. ⟨https://doi.org/10.1145/3091478.3098863⟩.
Li, H., 2022. Ontology-Driven Data Access and Data Integration with an Application in the Materials Design Domain. Linköping University Electronic Press.
Materials Open Lab, 2021. Materials Science and Engineering Ontology (MSEO). Available from: ⟨https://github.com/Mat-O-Lab/MSEO⟩.
Materials Open Lab, 2022. MatOLab Tensile Test Ontology (MOL_TENSILE). Available from: ⟨https://matportal.org/ontologies/MOL_TENSILE⟩.
McCusker, J.P., et al., 2020. NanoMine: a knowledge graph for nanocomposite materials science. In: International Semantic Web Conference. Springer. https://doi.org/10.1007/978-3-030-62466-8_10.
Mohd Ali, M., et al., 2021. Enriching the functionally graded materials (FGM) ontology for digital manufacturing. Int. J. Prod. Res. 59 (18), 5540–5557. https://doi.org/10.1080/00207543.2020.1787534.
Moreno Torres, B., et al., 2021. An ontology-based approach to enable data-driven research in the field of NDT in civil engineering. Remote Sens. 13 (12), 2426. https://doi.org/10.3390/rs13122426.
Morgado, J.F., et al., 2020. Mechanical Testing Ontology for Digital-Twins: a Roadmap Based on EMMO. in SeDiT@ ESWC.
Mutz, M., et al., 2022. Toward a Li-ion battery ontology covering production and material structure. Energy Technol. 2200681. https://doi.org/10.1002/ente.202200681.
Otte, J.N., et al., 2019. An ontological approach to representing the product life cycle. Appl. Ontol. 14 (2), 179–197. https://doi.org/10.3233/AO-190210.
Otte, J.N., Beverley, J., Ruttenberg, A., 2022. BFO: basic formal ontology. Appl. Ontol. (Preprint), 1–27. https://doi.org/10.3233/AO-220262.
Piane, F.L., et al., 2021. Introducing MAMBO: materials and molecules basic ontology. arXiv Prepr. arXiv. https://doi.org/10.48550/arXiv.2111.02482.
Premkumar, V., et al., 2014. A semantic knowledge management system for laminated composites. Adv. Eng. Inform. 28 (1), 91–101. https://doi.org/10.1016/j.aei.2013.12.004.
Psarommatis, F., Fraile, F., Ameri, F., 2023. Zero defect manufacturing ontology: a preliminary version based on standardized terms. Comput. Ind. 145, 103832 https://doi.org/10.1016/j.compind.2022.103832.
Roh, B.-M., et al., 2021. Ontology-based process map for metal additive manufacturing. J. Mater. Eng. Perform. 30 (12), 8784–8797. https://doi.org/10.1007/s11665-021-06274-2.
Rudnicki, R., 2019. An Overview of the Common Core Ontologies. CUBRC, Buffalo, NY.
Sainte Marie, Cd, Iglesias Escudero, M., Rosina, P., 2011. The ONTORULE project: where ontology meets business rules. In: International Conference on Web Reasoning and Rule Systems. Springer. https://doi.org/10.1007/978-3-642-23580-1_3.

Sanfilippo, E.M., Belkadi, F., Bernard, A., 2019. Ontology-based knowledge representation for additive manufacturing. Comput. Ind. 109, 182–194. https://doi.org/10.1016/j.compind.2019.03.006.

Scheffler, M., et al., 2022. FAIR data enabling new horizons for materials research. Nature 604 (7907), 635–642. https://doi.org/10.1038/s41586-022-04501-x.

Studer, R., Benjamins, V.R., Fensel, D., 1998. Knowledge engineering: principles and methods. Data Knowl. Eng. 25 (1), 161–197. https://doi.org/10.1016/S0169-023X(97)00056-6.

The IOF Core Ontology, 2023; Available from: ⟨https://github.com/iofoundry/ontology/tree/202301/core⟩.

Xu, L.D., Xu, E.L., Li, L., 2018. Industry 4.0: state of the art and future trends. Int. J. Prod. Res. 56 (8), 2941–2962. https://doi.org/10.1080/00207543.2018.1444806.

Zhang, X., et al., 2016. MMOY: towards deriving a metallic materials ontology from Yago. Adv. Eng. Inform. 30 (4), 687–702. https://doi.org/10.1016/j.aei.2016.09.002.

Zhang, X., Hu, C., Li, H., 2009. Semantic query on materials data based on mapping MATML to an OWL ontology. Data Sci. J. 8, 1–17. https://doi.org/10.2481/dsj.8.1.

Zhang, X., Lv, P., Wang, J., 2014. STSM: an infrastructure for unifying steel knowledge and discovering new knowledge. Int. J. Database Theory Appl. 7 (6), 175–190. https://doi.org/10.14257/ijdta.2014.7.6.16.

Zhang, X., Zhao, C., Wang, X., 2015. A survey on knowledge representation in materials science and engineering: an ontological perspective. Comput. Ind. 73, 8–22. https://doi.org/10.1016/j.compind.2015.07.005.

Zhao, S., Qian, Q., 2017. Ontology based heterogeneous materials database integration and semantic query. AIP Adv. 7 (10), 105325 https://doi.org/10.1063/1.4999209.