ADVANCED ENGINEERING MATERIALS
www.aem-journal.com

Performance Evaluation of Upper-Level Ontologies in Developing Materials Science Ontologies and Knowledge Graphs

Hossein Beygi Nasrabadi, Ebrahim Norouzi, Harald Sack, and Birgit Skrotzki*

This study tackles a significant challenge in ontology development for materials science: selecting the most appropriate upper-level ontologies for creating application-level ontologies and knowledge graphs. Focusing on the use case of Brinell hardness testing, the research assesses the performance of various top-level ontologies (TLOs)—basic formal ontology (BFO), elementary multiperspective material ontology (EMMO), and provenance ontology (PROVO)—in developing Brinell testing ontologies (BTOs). Consequently, three versions of BTOs are created using combinations of these TLOs along with their integrated mid- and domain-level ontologies. The performance of these ontologies is evaluated based on ten parameters: semantic richness, domain coverage, extensibility, complexity, mapping efficiency, query efficiency, integration with other ontologies, adaptability to different data contexts, community acceptance, and documentation and maintainability. The results show that all candidate TLOs can effectively develop BTOs, each with its distinct advantages. BFO provides a well-structured, understandable hierarchy, and excellent query efficiency, making it suitable for integration across various ontologies and applications. PROVO demonstrates balanced performance with strong integration capabilities. Meanwhile, EMMO offers high semantic richness and domain coverage, though its complex structure impacts query efficiency and integration with other ontologies.

1. Introduction

The digital representation of materials science and engineering topics has recently attracted significant attention for enhancing data sharing, accessibility, and analysis.[1] Digitalizing research data and related scientific content enables the integration of large datasets, the use of advanced computational tools, and supports multidisciplinary collaboration, thereby accelerating scientific discoveries and innovations.[1a] Ontologies are crucial for this digitalization process, as they establish a standardized knowledge representation within each domain.[1b,2] In other words, ontologies provide a common vocabulary for the domain that facilitates the semantic organization of data, thus supporting consistent data interpretation, global conceptualization for materials information integration, and the publishing of linked materials data.[2,3]

To date, a wide range of ontologies have been introduced in the field of materials science and engineering. The first ontologies that have been developed in this domain (like Ashino, PREMaP, ONTORULE, SLACKS, and MatOWL ontologies),[3] mainly remained unused due to their narrow and specific focus fields, inappropriate level of abstraction, and insufficient representation of instances.[2,3] Moreover, these ontologies exhibited poor interoperability with other ontologies to enable seamless integration and reuse of other ontologies' contents and create a global framework of reference ontologies.[2,3] To improve the ontologies' interoperability, their design as extensions of standardized ontologies was evaluated in the next years.[2,4] **Figure 1** provides an overview of different types of ontologies in the domain of materials science and engineering. The name, description, and repositories of such ontologies are listed in **Table 1**. Based on the degree of abstraction and formal expressiveness, the ontologies are classified into the following four levels:[1b,2] 1) Top-level ontologies (TLOs) describe common general concepts across various 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.[2] basic formal ontology (BFO),[5] elementary multiperspective material ontology (EMMO),[6] provenance ontology (PROVO),[7] descriptive ontology for linguistic and cognitive engineering (DOLCE),[8] suggested upper merged ontology (SUMO),[9] and semanticscience integrated ontology (SIO)[10] are TLOs that were mostly reused for developing the ontologies

H. Beygi Nasrabadi, B. Skrotzki
Department 5.2 - Metallic High-Temperature Materials
Bundesanstalt für Materialforschung und -prüfung (BAM)
Unter den Eichen 87, 12205 Berlin, Germany
E-mail: hossein.beygi-nasrabadi@bam.de

E. Norouzi, H. Sack
Information Service Engineering
FIZ Karlsruhe - Leibniz Institute for Information Infrastructure
Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen, Germany

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/adem.202401534.

DOI: 10.1002/adem.202401534

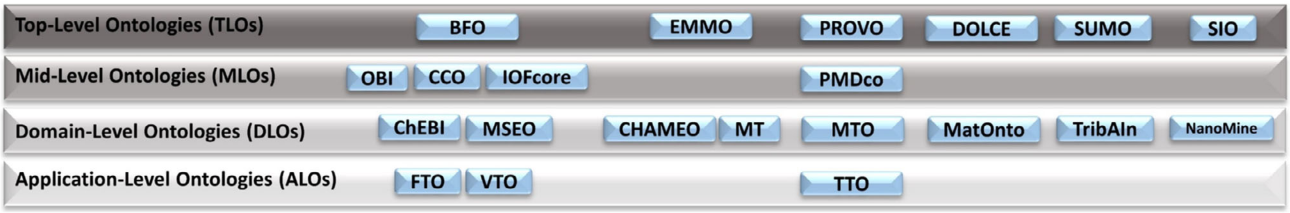

Figure 1. Overview of different level ontologies in materials science and engineering.

Table 1. List of frequently used TLOs, MLOs, and example DLOs in the domain of materials science and engineering.

Level	Ontology	Full name	Short description	Repository
TLO	BFO	Basic Formal Ontology	BFO is a small TLO that is designed for use in supporting information retrieval, analysis, and integration in scientific and other domains.	[5c]
	EMMO	Elementary Multiperspective Material Ontology	EMMO is a multidisciplinary effort to develop a standard representational ontology for materials sciences.	[6]
	PROVO	Provenance Ontology	PROVO provides a set of entities that can be used to represent and interchange provenance information generated in different systems and under different contexts.	[7]
	DOLCE	Descriptive Ontology for Linguistic and Cognitive Engineering	DOLCE is a foundational ontology developed and maintained by the ISTC-CNR Laboratory for Applied Ontology. It was originally developed within the WonderWeb project and was conceived as the first module of the WonderWeb Foundational Ontologies Library (WFOL).	[8b]
	SUMO	Suggested Upper Merged Ontology	SUMO is used for research and applications in search, linguistics, and reasoning. SUMO is the only formal ontology that has been mapped to the WordNet lexicon. SUMO is written in the SUO-KIF language.	[9]
	SIO	Semanticscience Integrated Ontology	SIO provides a simple, integrated ontology of types and relations for rich descriptions of objects, processes, and their attributes.	[10]
MLO	OBI	Ontology for Biomedical Investigations	An MLO based on BFO helps for clear communication about scientific investigations by defining more than 2500 terms for assays, devices, objectives, and more.	[11]
	CCO	Common Core Ontologies	CCO is a mid-level extension of BFO that comprises twelve ontologies designed to represent and integrate taxonomies of generic classes and relations across all domains of interest.	[12]
	IOF core	The Industrial Ontologies Foundry Core Ontology	IOF Core Ontology contains many intermediate-level terms that derive from BFO and are often domain-independent, meaning one can find them in other industries and fields.	[13]
	PMDco	Platform MaterialDigital Core Ontology	PMDco is a MLO based on PROV, for materials science and engineering.	[14a]
DLO	ChEBI	Chemical Entities of Biological Interest	ChEBI is a freely available dictionary of molecular entities focused on 'small' chemical compounds. ChEBI uses BFO as a TLO.	[57]
	MSEO	Materials Science and Engineering Ontology	MSEO utilizes the IOF Ontology stack giving materials scientists and engineers the ability to represent their experiments and resulting data.	[15]
	CHAMEO	Characterization Methodology Domain Ontology	A domain ontology based on EMMO for materials characterization based on the CHADA template.	[16b]
	MT	Mechanical Testing	A domain ontology for mechanical testing based on EMMO.	[17]
	MTO	Mechanical Testing Ontology	MTO was developed based on PMDco and by collecting the mechanical-testing vocabulary from several ISO mechanical-testing standards.	[18]

in the materials science domain. 2) 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 (DLOs).[2] For example, ontology of biomedical investigations (OBI),[11] common core ontologies (CCO),[12] and industrial ontologies foundry (IOF) core[13] are the MLOs that are established based on the BFO TPO. As another MLO in the materials science domain, PMDco has been developed recently based on PROVO.[14] 3) 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.[2] Until now, a variety of DLOs were introduced in the domain of materials science, which were designed based on different types of TLOs and MLOs.[4] For instance, MSEO is a DLO that reuses BFO and IOF core and represents extensive semantics in the domain of materials science and engineering.[15] Several DLOs were also developed based on the EMMO like CHAMEO (for the classification of materials, models, manufacturing processes, and software products related to materials characterization and modeling),[16] or MT (in the mechanical-testing domain).[17] With regards to the domain of materials mechanical testing, MTO has also been developed by reusing the PROVO and PMDco.[18] MatOnto (materials structure, properties, and processing domain),[19] tribAIn (tribology domain),[20] and NanoMine (polymer nanocomposite domain)[21] were also some DLO examples that reused DOLCE, SUMO, and SIO ontologies, respectively. 4) Application-level ontologies (ALOs) provide highly detailed semantics for specific use cases and support the development of knowledge graphs.[22] A knowledge graph is a structured representation of interconnected data, integrating information from diverse sources, and using graph structures to represent entities and their relationships, enabling advanced data querying, integration, and semantic understanding. For example, the knowledge graphs for the fatigue and Vickers hardness testing processes were designed along with the development of fatigue testing ontology (FTO)[23] and Vickers testing ontology (VTO)[24] that reused the concepts of BFO, IOF, and MSEO. Tensile test ontology (TTO) is another ALO example that has reused PROVO and PMDco for representing the tensile testing knowledge graph.[25]

The term upper-level ontologies (ULOs) refer to all ontologies whose levels are higher than that of intended ontology. For example, all the TLOs, MLOs, and DLOs, are essentially ULOs for designing the ALOs. A variety of ALOs were introduced in the materials science domain which have the advantage that the use of ULOs facilitates the extension of domain knowledge in an organized and sustainable way.[4] Although such ontologies represented various topics of materials science and engineering (like additive manufacturing, battery, crystallography, materials microstructure, materials characterization, or modeling),[4] many new ontologies and knowledge graphs are still needed in this domain. In this context, materials science experts aiming to create new ontologies and knowledge graphs encounter a fundamental challenge: determining which ULOs are most suitable for modeling specific DLOs or ALOs.

Along with this problem, the assessment of ULOs in the materials science domain remained pending. Ontology evaluation involves assessing an ontology against specific criteria to determine its quality and effectiveness.[26] Ontology and knowledge graph evaluation methods include gold standard (comparison to high-quality graphs), data-driven (keyword extraction), application/task-based (task performance), user-based (user perspective), structure-based (structural metrics), and data quality (accuracy and consistency) evaluations.[27] Various ontology evaluation criteria like richness, adaptability, clarity, accuracy, modularity, consistency, coverage, cohesion, completeness, and computational efficiency were reviewed in several reports.[28] Degbelo[29] classified the ontology evaluation criteria into two categories design evaluation parameters (like accuracy, adaptability, clarity, cognitive adequacy, completeness, consistency, expressiveness, and grounding), and implementation evaluation variables (such as computational efficiency, congruency, practical usefulness, precision, and recall). Sabou et al.[30] studied the ontology selection based on three parameters popularity, richness of knowledge, and topic coverage. Lourdusamy and John[31] also introduced various quantitive metrics for evaluating the basic, schema, knowledgebase, graph, and complexity parameters of the ontologies.

Based on the literature survey done, the research gap in this domain can be specified by two following questions: Which ULOs are most suitable for modeling the materials science DLOs or ALOs? And how to evaluate ontologies and knowledge graphs in the materials science domain. To deal with these questions, the current research aims to evaluate the strengths and weaknesses of various ULOs by using them for constructing the ALOs and knowledge graphs of the same use case in the materials science domain. Rather than focusing on a single global evaluation criterion, the article aims to give the materials domain experts an overview of various ontology evaluation parameters that they have to consider for choosing their desired ULOs. Addressing the published reports, the majority of the materials science-related DLOs and ALOs that have been developed in the last 10 years reused one of the BFO, EMMO, or PROVO TLOs. The performances of these TLOs are evaluated in this research for developing a single, well-defined use case in the materials science domain (Brinell hardness testing). This approach provides a consistent basis for evaluation and ensures performance differences are due to the ontologies themselves rather than variability from multiple use cases. Therefore, three different versions of Brinell testing ontologies (BTOs) were developed by reusing BFO, EMMO, or PROVO TLOs. Furthermore, specific MLOs and DLOs were also utilized along with such TLOs to ensure comprehensive coverage of materials science concepts and relationships. Therefore, the candidate ULOs combinations are 1) BFO + IOF core + MSEO, 2) PROVO + PMDco, and 3) EMMO + CHAMEO + MT. By overviewing the existing ontology evaluation methods, we also designed ten fundamental metrics and parameters which can analysis different aspects of the ontology and knowledge graphs development in the materials science domain. The result of such evaluations shows that the materials science-related ALOs (like the use case of BTO) can be efficiently developed by reusing all candidate ULOs. Such ULO combinations not only facilitate accurate modeling of complex domain-specific knowledge, but also enhance interoperability, consistency, and semantic richness. Furthermore, knowing the strengths and drawbacks of each ULO allows each domain expert to select different desired ULOs based on the aims, datasets/tool structures, or community preferences. This research is novel from different aspects, as it is the first to use a series of quantitative and qualitative metrics for evaluating ontologies in the materials science domain. It introduces new parameters like mapping efficiency, query efficiency, and data adapting for knowledge graph evaluation and assesses new variables of integration, community acceptance, documentation, and maintainability for selecting TLOs. Additionally, it offers a comprehensive overview of all ULOs in materials science, evaluating them from various aspects to guide domain experts in choosing the most suitable ULO combinations for modeling knowledge graphs and ontologies in the materials science domain.

2. Use Case: Development of Brinell Hardness Ontologies Based on the Testing Standard

The Brinell hardness is a typical materials mechanical-testing method that provides valuable data about the material's resistance to deformation. In this testing method, a hard and spherical indenter (usually made of steel or tungsten carbide) is pressed into the material's surface under a predetermined load for a specified time. After removing the load, the vertical and horizontal diameters of the indentation left on the material's surface are measured using a microscope or other precision measuring device. Subsequently, the Brinell Hardness is calculated using Equation (1)[32]

$$\mathrm{HBW} = 0.102 \times \frac{2F}{\pi D^2\left(1 - \sqrt{1 - d^2/D^2}\right)} \tag{1}$$

where P is the applied load (N), D is the diameter of the indenter (mm), and d is the average diameter of the indentation (mm). Therefore, the Brinell hardness unit would be N mm^{-2}, however, the values of Brinell hardness are reported by the notations which represent their specific testing configurations. For example, HBW 2.5/62.5 refers to the Brinell hardness values that are measured with a tungsten carbide indenter (W) with a diameter of 2.5 mm and at an applied load of 62.5 kgf.

A detailed description of the Brinell hardness testing procedure is given in the DIN EN ISO 6506-1:2015 standard.[32] This standard has been used as the main resource for preparing the ontology terminology in this research, not only because it provides accurate definitions from the standardization committee, but also due to its ability to classify process entities, determine the level of detail, and provide technical relations between entities. Accordingly, the standard-extracted terminology for developing the Brinell hardness ontology can be categorized into the following six groups: 1) Apparatus: testing machine, indenter material/diameter, measuring system, certified reference material (CRM), 2) Test piece: identification, preparation, thickness, 3) Procedure: apparatus calibration, Brinell hardness test method (control temperature, confirm verification, choose test forces, place test piece, apply test force, check indentation distances, optical measurement of the indentation, calculate the Brinell hardness), 4) Properties: Brinell hardness, test force, depth of indentation, surface area of indentation, indentation horizontal/vertical diameter, mean indentation diameter, force-diameter index, loading time, load maintaining time, test points distance, test point-edge distance, Brinell hardness equation, Brinell hardness symbol (HBW), 5) Uncertainty analysis: average/standard deviation Brinell hardness, Brinell hardness uncertainty, CRM uncertainty, testing machine uncertainty, permissible uncertainty, measurement resolution uncertainty, 6) Test report: reference to a standard, test date, test operator, and testing laboratory.

Considering all the aforementioned entities, highly detailed knowledge graphs were developed for the Brinell hardness testing.[33] To simplify the storyline of this research, the smaller versions of Brinell hardness knowledge graphs are developed in this section which mainly contains the semantic representation of the following entities: 1) testing provenance metadata (test standard, date, laboratory, operator), 2) test piece (identification and composition), 3) equipment (machine identification, indenter material, shape, and diameter), 4) test procedure (checking temperature, machine calibration, choosing test force, test piece positioning, loading, unloading and optical measurement of indentation), and 5) calculating the average diameter and Brinell hardness values. In addition to the similar content, the different versions of the Brinell testing knowledge graphs were also designed with almost identical knowledge representation patterns to allow for a more accurate assessment of the relevance of various ULOs in modeling BTOs.

The modeled BTOs are global ontologies that are developed based on international standards to ensure interoperability and consistency. Brinell testing terminologies were extracted from the content of DIN EN ISO 6506-1:2015 standard,[32] and then the ontologies were developed adhering to established frameworks of W3C's OWL. The knowledge graphs and ontologies were designed by reusing the entities of global upper-level ontologies and evaluated by various academic and industrial partners. It should be noted that BAM Institute is a participant of the standards committee and host of accredited materials testing laboratories that support the development of ontologies that align with test standards and laboratory data structures. Therefore, the article introduced the development of such standardized ontologies which can be used by all ontologists around the world.

2.1. BTO V.5.x.x by Reusing the BFO, IOF Core, and MSEO Ontologies

Along with the development of different version BTOs, the Brinell testing knowledge graphs were also constructed which enables the consistent schema definition, semantic accuracy, and efficient data integration, enhancing the overall quality and utility of the knowledge representation. The knowledge graphs were produced by collecting the Brinell testing terminology from the DIN EN ISO 6506-1:2015 standard[32] as well as the metadata of testing reports. It uses a graph structure that presents the Brinell testing knowledge as a network of entities (nodes), their relationships (edges), and other additional information (labels and properties), while the ontology defines the schema or structure of the knowledge graph, specifying the types of entities and relationships that exist, and the rules for how they can interrelate. The ontologies and knowledge graphs development were performed using a mid-range PC with an Intel Core i5 CPU, 8 GB of RAM, and Windows 64-bit operating system for running the ontology and query editors. The Brinell hardness knowledge graphs were designed via the Ontopanel graphical editing tool.[34] Here, the Ontopanel library provides different shapes for designing the semantic representations between the entities. The tool is also equipped with the "EntityManager" plugin, which allows searching and importing the entities of other different ontologies.[34] **Figure 2**, **3** and **4** illustrate the Ontopanel graphical view of the Brinell hardness knowledge graphs developed by reusing the different combinations of the ULOs. All these graphs were designed in distinct T- and A-boxes. The T-boxes (terminology boxes) represent the hierarchy of the ontology entities, define the concepts, properties, and constraints, and include axioms and rules that describe how concepts and properties interrelate, forming the backbone of the ontology's conceptual framework. In Figure 2–4, we just showed the classes' hierarchies, where the classes were defined into rectangles, and their hierarchy and sub-class relations are indicated by the black arrows. The reused classes from various ULOs are distinguished by different colored rectangles, while the new classes are indexed with the "bto" namespace and located in the

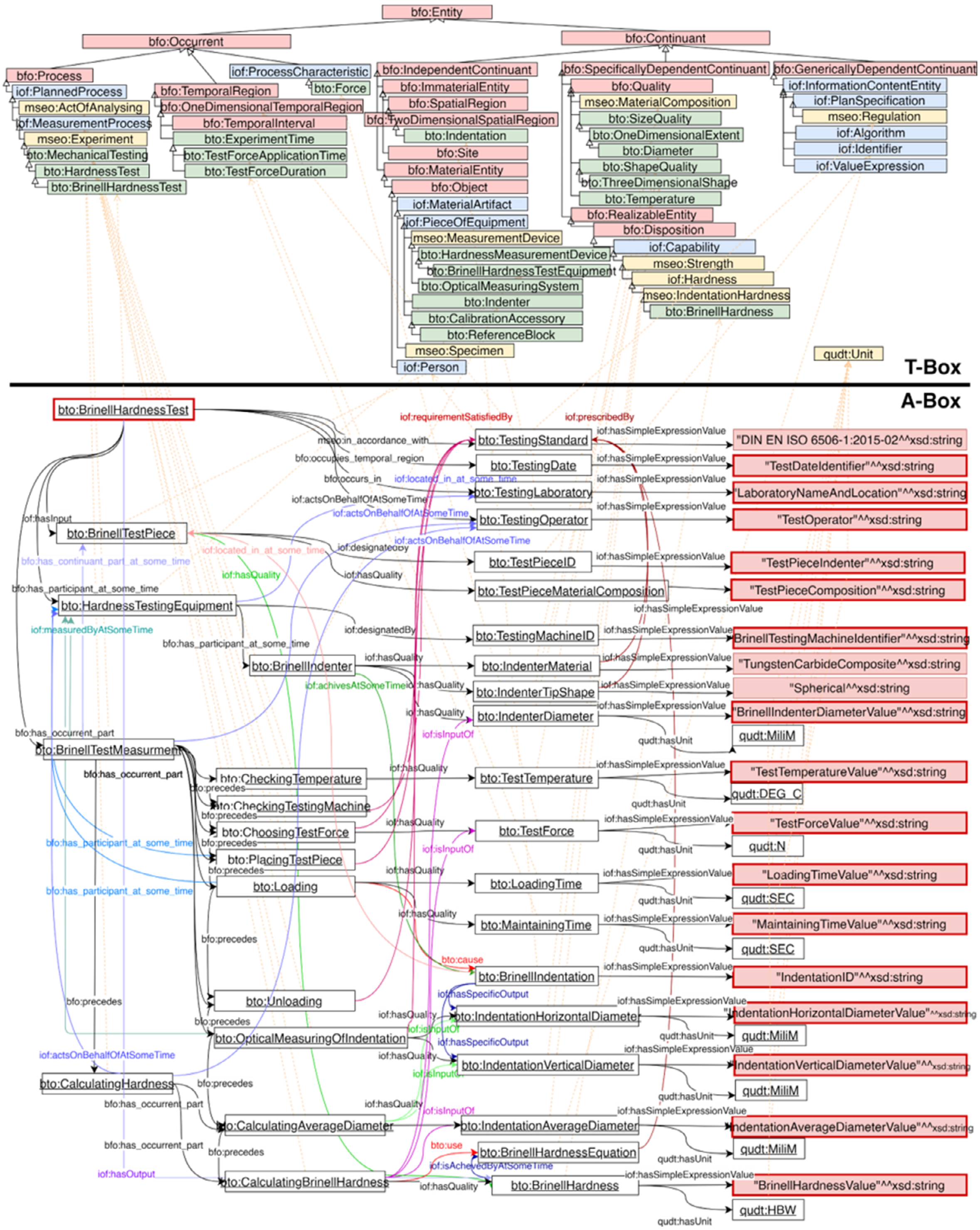

Figure 2. T- and A-boxes of the Brinell testing ontology (BTO V.5.0.3) developed by utilizing the BFO2020,[5c] IOF202401,[13] and MSEO2023.[15] ontologies. The high-resolution images and source files can be found in the project repository.[33]

green rectangles. On the other side, the A-boxes (assertion boxes) provide factual statements about specific instances within a knowledge graph. Here, the A-box entries populate the knowledge graph with data, asserting concrete information about entities and their attributes or relations, thereby enabling the practical application of the abstract schema defined in the T-box. The main entities of the A-boxes are the individuals (white rectangles), which are the types (brown dash lines) for some

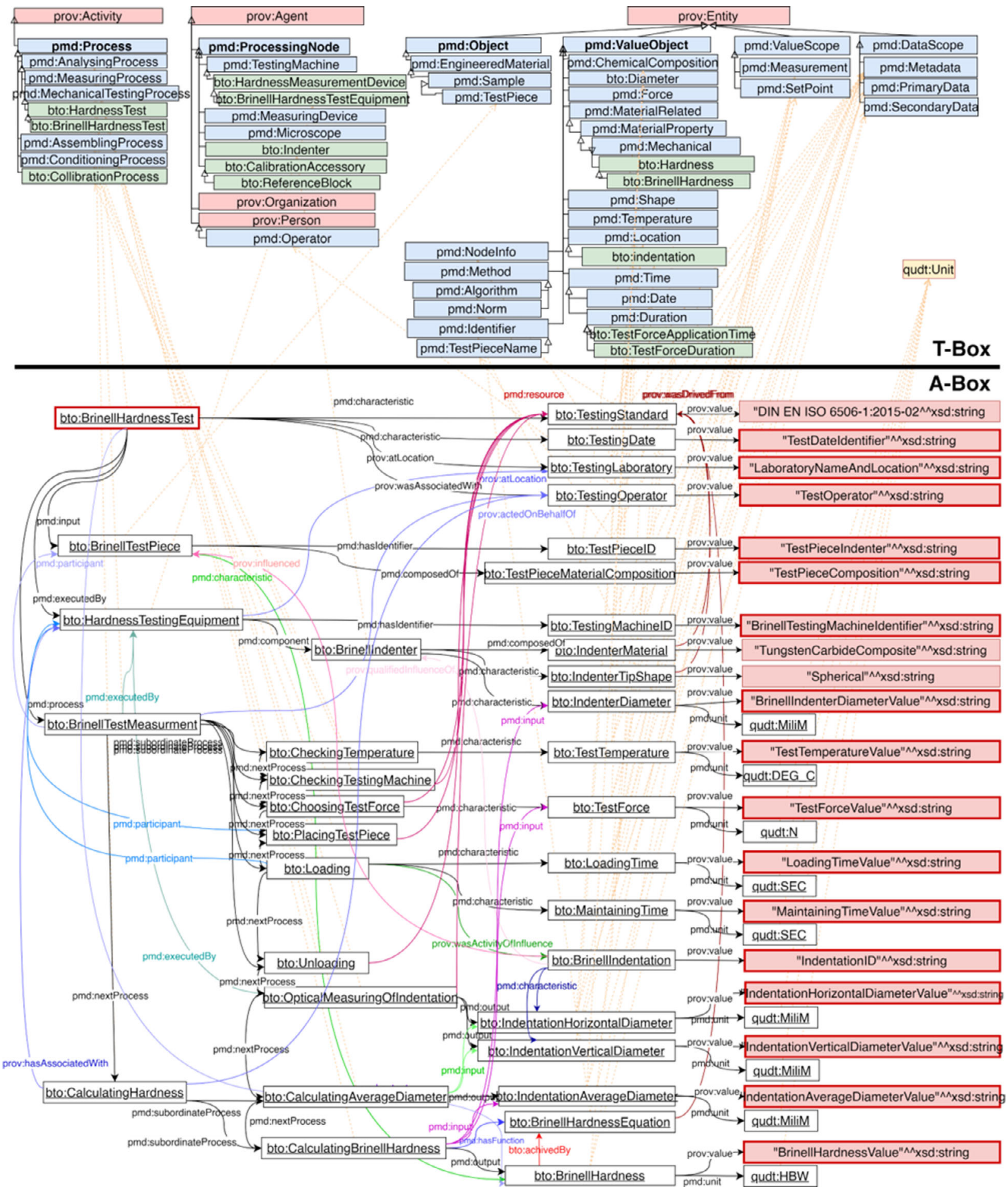

Figure 3. Graphical overview of the Brinell testing ontology (BTO V.4.0.3) developed based on the PROVO2013[7] and PMDco 2.0.7.[14a] ontologies. The high-resolution images and source files can be found in the project repository.[33]

T-box classes. Furthermore, any data related to the specific individuals is expressed by the data values (red rectangles). In this state, the object properties (arrows between white rectangles) express the semantic relationships between different individuals, and the data properties (arrows between white and red rectangles) connect the data values to their related individuals.

Figure 2 shows the Brinell hardness knowledge graph developed by reusing the combination of BFO,[5c] IOF,[13] and

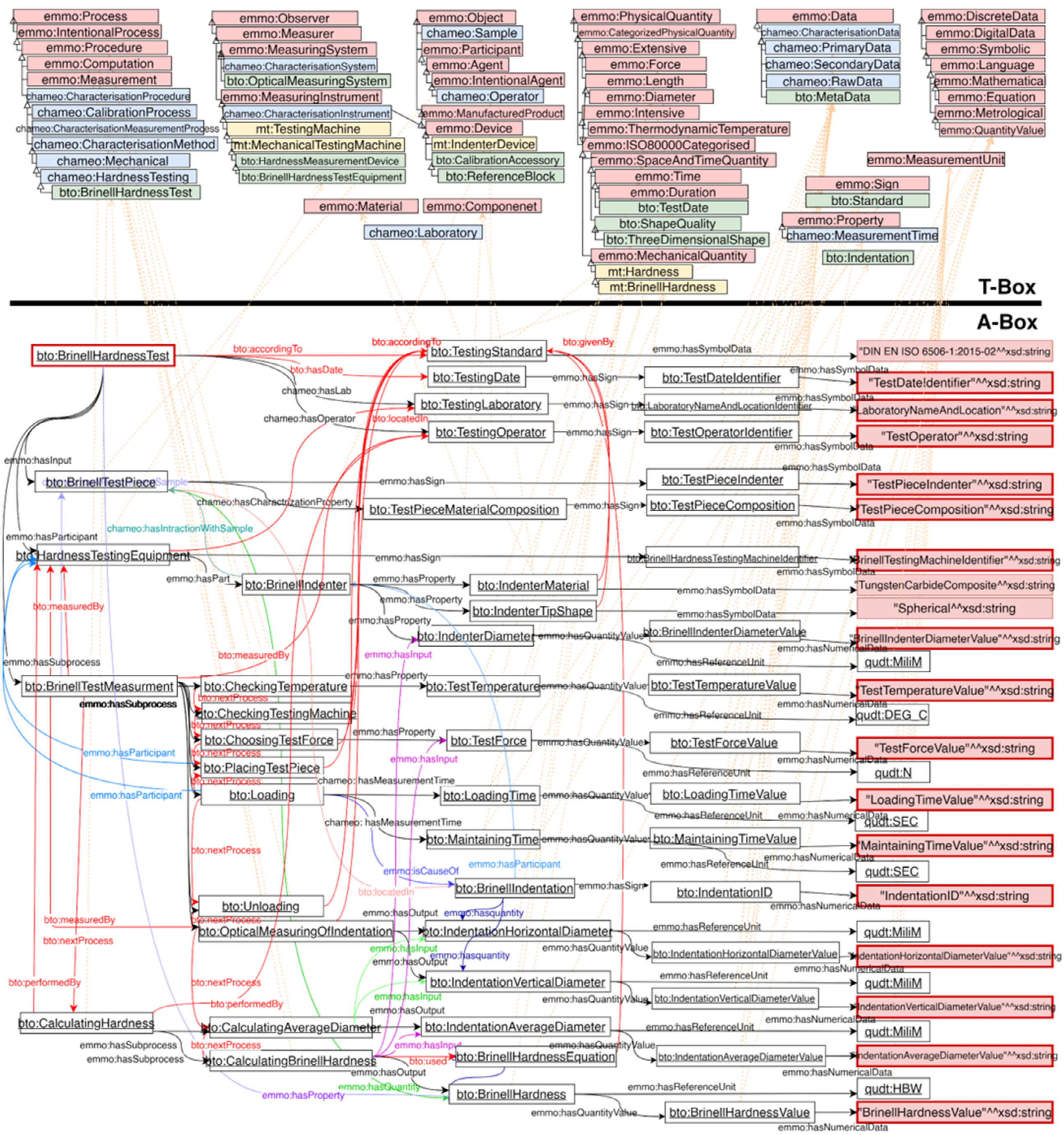

Figure 4. Brinell testing ontology (BTO V3.0.3) developed by reusing the EMMO 1.0.0-beta7,[6] CHAMEO2024,[16b] and MT 1.0.0.[17] ontologies. The high-resolution images and source files can be found in the project repository.[33]

MSEO[15] ontologies (red-, blue-, and yellow-colored classes, respectively). BFO is a TLO consisting of 36 classes, designed to support information integration, retrieval, and analysis across all domains of scientific investigation (no terms particular to material domains). It offers a systematic framework for categorizing entities based on their core characteristics and relationships, providing consistency in data classification and interoperability.[5] Accordingly, BFO was picked as the TLO in the ISO/IEC 21838-2:2021 standard because of its robust theoretical foundations, widespread usage in a variety of domains, and proven efficiency in facilitating data interoperability and standardization.[35] At the upper side of the BFO classes hierarchy, there are two Entity subclasses Continuant and Occurrent. Continuant contains the entities that persist through time including three subclasses SpecificallyDependentContinuant, GenerallyDependentContinuant, and IndependentContinuant.

On the other side, Occurrent contains the entities that unfold through time, and two of its main subclasses are Process and TemporalRegion. As an MLO, the IOF core ontology utilizes BFO and borrows some terms from various domain-independent or MLOs to provide basic entities in various industrial and manufacturing domains, and serves as a foundation for ensuring consistency and interoperability across various domain-specific reference ontologies.[13] MSEO[15] is a DLO that was designed based on the BFO and IOF core ontologies and was also reused for importing materials-related entities. Moreover, we partially utilized the quantities, units, dimensions, and types ontology (QUDT)[36] for expressing the required data units.

The combination of BFO, IOF, and MSEO provides an excellent framework for expressing the application-level entities in the best possible class classifications and hierarchies. For example, the BrinellHardnessTest class was created as the subclasses of bfo:Process > iof:PlannedProcess > mseo:ActOfAnalyzing > ... > bto:HardnessTest. In the same way, classes related to testing apparatus and test pieces were created as the subclasses of iof:MaterialArtifact. Furthermore, the classification logic of BFO classes allows for the optimal organization of different variables related to the Brinell hardness testing procedure. For example, entities related to time, such as ExperimentTime, are classified under the bfo:TemporalRegion subclasses. By a similar logic, the Force class was created as an iof:ProcessCharacteristic subclass, and the BrinellHardness class was developed under the bfo:Disposition > iof:Capability >...> mseo:IndentationHardness classes hierarchies. The other classes related to size, shape, and temperature entities also emerged by the bfo:Quality subclasses.

As shown in the A-box of Figure 2, numerous semantic relations were also employed for designing the Brinell hardness testing knowledge graph. For example, bto:BrinellHardness is an instance for the BrinellHardness class, which is an important output of the bto:BrinellHardnessTest process and a specific characteristic of the bto:BrinellTestPiece (representative object properties are iof:hasOutput and iof:hasQuality, respectively). The value of bto:BrinellHardness has been measured through different steps, the last one is bto:CalculatingBrinellHardness. Here, bto:IndenterDiameter, bto:TestForce, and bto:IndentationAverageDiameter are the input values and the Brinell hardness calculation process uses bfo:BrinellHardnessEquation that is prescribed by bto:TestingStandard. The calculated Brinell hardness value was provided by a data value ("BrinellHardnessValue"), which was linked to bto:BrinellHardness with iof:hasSimpleExpresionValue data property. This value also has the HBW unit, which is defined by the QUDT ontology. As can be seen in this A-box, the utilized BFO and IOF core ontologies provided almost all the object and data properties required for the design of the Brinell hardness knowledge graph, and only a few object properties such as bto:cause and bto:use have been developed in by the ontology.

2.2. BTO V.4.x.x by Reusing PROVO and PMDco Ontologies

Figure 3 graphically shows the T- and A-boxes of the BTO, which was designed by utilizing the PROVO and PMDco ontologies. PROVO is a lightweight ontology that can be adopted in a wide range of applications and serves as a reference model for creating domain-specific provenance ontologies thereby facilitating interoperable provenance modeling. The PROVO classes and properties are defined such that they can not only be used directly to represent provenance information but also can be specialized for modeling application-specific provenance details in a variety of domains.[7] For example, the basic layout of PMDco ontology is aligned with the PROVO framework, and this MLO provides extensive concepts in the domain of materials science and engineering for a detailed description of processes, experiments, and computational workflows.[14]

Three fundamental classes of Entity, Activity, and Agent provide the logic for the PROVO. An Entity is a physical, digital, conceptual, or other kind of thing with some fixed aspects that relate to the Activity and Agent classes by prov:wasGeneratedBy and prov:wasAttributedTo object properties, respectively. Activity is also something that occurs over a period of time and acts upon or with an Entity and associates with an Agent.[7] PMDco has also been designed by main four classes of Process (subclass of prov:Activity), ProcessingNode (subclass of prov:Agent), as well as Object and ValueObject (subclasses of prov:Entity). In other words, this model represents the Processes with some input/output Objects that are executed by some ProcessingNodes, and the characteristics of all these classes are expressed by the ValueObjects.[14b] In the case of the Brinell hardness test ontology for instance, the classes hierarchy of Figure 3 shows that the entities related to the process (e.g., bto:BrinellHardnessTest) are created as the subclasses of prov:Activity > pmd:Process, the entities related to the hardness testing equipment (like bto:Indenter) are the subclasses of prov:Agent > pmd:ProcessingNode classes and any instances related to the hardness test piece linked to the pmd:TestPiece (subclass of prov:Entity > pmd:Object). However, pmd:ValueObject lacks further detailed and logical classifications for its subclasses, and numerous entities of different natures (e.g., time, location, mechanical and physical properties, and even algorithm and identifier) were developed next to each other, and by similar hierarchies. The PMD community is currently working on upcoming versions of PMDco that not only offer more extensive terminology and classifications but also integrate with BFO and IOF ontologies.[14a]

On the other side, it can be observed in the A-box of Figure 3 that PROVO and PMDco ontologies provided almost all the object and data properties needed to model the semantic relationships of the Brinell testing knowledge graph. For example, object properties like wasAssociatedWith, actedOnBehalfOf, and influenced from PROVO and characteristic, input, output, executedBy, and process from PMDco ontologies suggest wide domains for expressing the semantic relationships. Accordingly, just a few object properties have been added for developing the ontology (such as bto:achievedBy). Furthermore, PROV's data properties (like prov:value) are sufficiently comprehensive and generic that satisfy all the requirements for developing the hardness test knowledge graph.

2.3. BTO V.3.x.x by Reusing EMMO, CHAMEO, and MT Ontologies

Figure 4 illustrates a graphical view from BTO V.3.x.x that was developed by reusing the EMMO, CHAMEO, and MT ontologies.

EMMO is a TLO that provides a standard representational ontology framework based on the fundamental concepts of physics, chemistry, and materials science. EMMO is designed to pave the road for semantic interoperability providing a generic common ground for describing materials, models, and data that can be adapted by all domains.[6] Unlike the other TLOs, EMMO is not an abstract ontology but provides very detailed semantics in the materials science domain. Instead of starting from general upper-level concepts, as done by other ontologies, EMMO has grown from the bottom (scientific application field) to the top (conceptualization), staying focused on the actual picture of the physical world and materials science, while at the same time maintaining an approach as general as possible.[6] Therefore, EMMO terminology is more detailed but less comprehensive and logical than BFO and PROVO. For example, the T-box of Figure 4 shows that the highest level of the EMMO classes' hierarchy consisted of many incoherent classes from various natures (like Process, Observer, Object, PhysicalQuantity, EncodedeData, Sign, and Property). Several DLOs like CHAMEO and MT were also built by extending this foundation of EMMO.[16b,17] We also utilized this foundation for developing the BTO V.3.x.x. For instance, the BrinellHardnessTest class was created as a subclass of emmo:Process <...< chameo:CharacterizationProcedure, and the instances related to the hardness test piece were introduced by the emmo:Material and emmo:Component classes. Since EMMO and CHAMEO are specifically built for the materials science and materials analysis domains, they also offer more diversified entities and classified hierarchies for modeling the Brinell testing knowledge graph. For instance, different classifications are available to introduce the equipment related to hardness testing; such as subclasses of emmo:MeasuringSystem < chameo:CharacterizationSystem (for creating bto:OpticalMeasuringSystem class), subclasses of emmo:MeasuringInstrument < chameo:CharacterizationInstrument < mt:TestingMachine (to establish bto:BrinellHardnessTestEquipment) or emmo:Device for developing subclasses like bto:CalibrationAccessory. Moreover, the emmo:PhysicalQuantity class provides a suitable range of material- and testing-related quantities with detailed logical classifications, therefore it fulfilled the majority of requirements for expressing the Brinell testing quantities. However, more emphasis on the material science domain was accompanied by challenges with the development of some generic or out-of-domain entities. Dealing with standards or regulations for example, and this issue can be solved by creating such concepts with a limited semantics relationship and at the top part of the entity's hierarchy.

The mentioned topics are also evident with the EMMO and CHAMEO object properties. The EMMO object properties originate from one of the roots of casual, mereotopological, and semiotical. The relation hierarchy extends more vertically (subrelations) facilitating the categorization and inferencing of individuals.[6] Therefore, EMMO offers a wide range of general object properties (e.g., hasInput, hasProperty, and hasPart) and specialized ones (like hasLab, hasOperator, and hasMeasurementTime), which greatly support the development of specialized knowledge graphs in the materials science domain. In some cases however, specified semantic relationships between individuals were expressed by the development of new object properties such as bto:accordingTo, bto:hasDate, bto:givenBy, bto:locatedIn, and bto:performedBy. Furthermore, due to the variety of the entities and challenges with object properties' domain/range limitations, the semantic relationships of the Brinell testing knowledge graph were modeled by utilizing relatively more semantic relationships than those developed by BFO + IOF core+MSEO or PROVO + PMDco.

3. Evaluation of ALOs and Knowledge Graphs Developed by Reusing Different ULOs

The small application-level ontologies with limited semantic relationships were developed in this research to evaluate the efficiency of upper-level ontologies in representing the Brinell testing entities. However, the developed ontologies not only provide the hierarchical classification systems that organize entities into nested categories (taxonomy) and map out the testing process and its sequences (workflow) but also furnish a more detailed and relational structure for the representation of Brinell testing domain knowledge and the relationships between the relating domain concepts (ontology). In other words, the developed ontologies provide the systematic arrangement for categorizing and retrieving information utilizing the ontological components of classes, properties, relationships, and rules.

Using the Ontopanel converter,[34] the graphically designed Brinell hardness knowledge graphs (XML-typed) were converted into the serialization formats for resource description framework (RDF) data (like RDF/XML or Turtle syntax) represented by the web ontology language (OWL). Subsequently, different versions of BTO were developed by editing such RDF files with Protégé ontology editor software.[37] In this case, the main ontology editing activities include importing the reused ULOs (for collecting the reused concepts and assuring their correct naming, URI, definition, relation, and hierarchy), checking the hierarchical order of new entities with regards to those of ULOs, adding explicit definitions, definition references and further restrictions to newly created entities, completing the basic ontology annotations (like name, version, and creators), reviewing, and testing the developed ontology.[2] The different versions of the developed Brinell hardness ontologies are publicly available in the project repository.[33]

In should be noted that, although the knowledge graphs and ontologies were built manually, they were reviewed several times by a group of scientists, checked with some tools like Ontopanel and Ontoflow,[2,34] and tested by different SPARQL queries. However, a low number of human errors like typos may accompany the developed knowledge graphs and ontologies. To make such human errors minimized and comparable in all cases, all three versions of BTOs were designed with the same taxonomy, same structure, same ontologists, and as much as possible fair and objective constructions. As a result, relatively few and equal quantities of human errors are expected in these three ontologies, ensuring that subsequent ontology evaluations are highly fair.

To investigate the role of utilized ULOs on the ALOs and knowledge graphs development, the developed ontologies and their building procedures were quantitatively and qualitatively evaluated in several manners. We performed the ontology and knowledge graph evaluation process by considering ten important parameters that logically cover different aspects of the ontology and knowledge graph development procedures and give the domain experts an overview for analyzing the quality and suitability of different ULOs for developing their application-level semantics. The planned ontology evaluation parameters are semantic richness, domain coverage, extensibility, complexity, integration with other ontologies, acceptance by the community, and documentation and maintainability. Furthermore, the mapping efficiency, query efficiency, and data adapting parameters evaluate the influences of reusing different ULOs on the performance of developed knowledge graphs. The evaluations of all these parameters are individually discussed in the following subsections.

3.1. Semantic Richness

Semantic richness refers to the number of an ontology's concepts and relationships, encompassing the diversity and depth of these connections, detailed descriptions, and the extent of meaning and associations captured within the structured knowledge representation. **Table 2** represents some semantic richness metrics of the utilized ULOs and developed ALOs evaluated using the OntoMetric tool.[38] Here, the semantic richness of the ontologies can be quantitatively analyzed by considering the number of different ontological entities like axioms, classes, individuals, and properties. Furthermore, the semantic richness scores (SR_{score}) were derived by averaging the number of object, datatype, and annotation properties.[39]

Evaluating the semantic richness parameters of BFO, PROVO, and EMMO TLOs, it can be seen that although BFO and PROVO have almost similar semantic richness metrics, EMMO exhibits significantly higher metrics across most categories (like SR_{score} and counts of axioms, classes, and object properties). Indeed, TLOs are expected to be abstract and provide a limited semantic richness, but EMMO is not abstract, and its structure has extended to the mid and domain levels of materials science and engineering, resulting in a higher amount of ontology entities and consequently higher semantic richness metrics. The IOF core and PMDco MLOs however provide higher semantic richness compared to BFO and PROVO TLOs. Comparing these two MLOs, IOF core provides more properties and SR_{score}, while PMDco contains higher classes. These differences arise because PMDco is tailored for more specialized entities within the materials science domain, while IOF encompasses a broader and more general range of manufacturing and industrial applications, leading to varying degrees of semantic richness. At the domain level also, CHAMEO provides more semantic richness than MT and MSEO.

The performed analyses show that although richer semantics are expected for lower-level ontologies, the semantic richness of the ontologies at the same level may be significantly different due to the domains that each ontology covers. As an overall comparison between the ULO combinations, EMMO + CHAMEO + MT provides higher semantic richness than BFO + IOF core+MSEO and PROVO + PMDco. However, all these ULO combinations provide a very high amount of semantic richness for developing our desired ALOs. The semantic richness analysis also confirms that the developed ALOs were designed in a comparable way that provided almost similar semantic richness. In this case, to better evaluate the semantic richness of the ALOs, the reported quantities excluded the semantics of imported ULOs. In other words, the quantities mentioned in Table 2 just report the number of ALO constructing semantics without importing the utilized ULOs. ALOs commonly generate low semantic as they focus on application-level taxonomy for a specific use case but provide most of their semantics from richer ULOs that deal with the broader, more general-purpose domains. Therefore, by importing the reused ULOs, all three proposals will show very high semantic richness. According to semantic quantities of Table 2, the three BTO versions were developed to represent the same domain content but with different semantic representations, so they provided almost similar semantic richness. However, a bit higher SR_{score} of BTO 3.0.3 originates from the more object properties that were used for representing the semantic relationships within the A-box of the Brinell testing process based on the EMMO and CHAMEO ontologies.

Table 2. Semantic richness metrics (ontology entities count and SR_{score}) of different level ontologies.

Ontologies				Counts						SR_{score}
TLOs	MLOs	DLOs	ALOs[a)]	Axiom	Class	Individual	Object property	Data property	Annotation property	
BFO 2020	–	–	–	602	36	0	40	0	34	24.67
	IOF 202 401	–	–	2458	114	0	128	3	32	54.33
	–	MSEO 2023	–	890	150	0	2	0	15	5.67
	–	–	BTO 5.0.3	539	62	50	23	1	9	11.00
PROVO 2013	–	–	–	971	31	1	44	6	7	19.00
	PMDco 2.0.7	–	–	1808	214	12	38	9	14	20.33
	–	–	BTO 4.0.3	544	59	50	23	1	9	11.00
EMMO 1beta7	–	–	–	19 898	2215	1	123	15	18	52.00
	–	CHAMEO 2024	–	1633	212	3	50	2	31	27.67
	–	MT 1.0.0	–	1764	410	0	12	5	13	10.00
	–	–	BTO 3.0.3	607	75	64	26	2	9	12.34

a)For better evaluating the semantic quantities of developed ALOs, entities of the utilized ULOs were not imported.

3.2. Domain Coverage

Although semantic richness is an important metric for selecting the appropriate ULOs, one more important metric would be the measure that shows how well such ULOs provide the basic concept entities required for developing the specific knowledge graphs and domain ontologies. In this regard, "domain coverage" is introduced as a metric that quantifies how effectively the candidate ontologies cover the terms extracted from the corpus.[39] Therefore, while a higher semantic richness reflects a richer ontology, it must be complemented by good domain coverage metrics to ensure that the ontology covers the necessary terms and concepts of the relevant domain knowledge.

To evaluate the domain coverage metric, ontology developers list a set of terms that they need from the ULOs to design their specific knowledge graphs and ontologies. For the use case of Brinell testing, these terms can be "measurement", "standard", "material", "equipment", "time", "force", "diameter", "shape", "has quantity", and "has value". Such entities or their synonyms from WordNet (a lexical database grouping words into synsets by concepts)[40] are then scanned from candidate ontologies, and the amounts of found matching items are used to calculate the domain coverage metrics. **Table 3** represents the quantitative outputs (Recall, Precision, and F_1 measure) of such analyses for the combination of BFO + IOF core + MSEO, PROVO + PMDco, and EMMO + CHAMEO + MT ontologies. The recall parameter is determined by dividing the number of matching terms between the extracted terms (including synonyms obtained using WordNet) and the ontology concepts by the total number of nominated terms, which is 10.[39] Similarly, the precision parameter is measured as the number of matching terms divided by the total number of extracted terms from candidate ULOs. Furthermore, the F_1 score is calculated by Equation (2) to represent the harmonic mean of Recall and Precision as a single metric and provide a comprehensive assessment of the ontologies coverages[39]

$$F_1(i) = 2 \times \frac{\text{Precision}(i) \times \text{Recall}(i)}{\text{Precision}(i) + \text{Recall}(i)} \tag{2}$$

The analysis of domain coverage metrics in Table 3 shows that the ULO combinations cover most of the nominated domain terms relevant to the Brinell testing use case. Covering 10, 9, and 8 nominated terms, PROVO + PMDco, BFO + IOF core + MSEO, and EMMO + CHAMEO + MT respectively provided the highest recall and precision metrics. Accordingly, PROVO + PMDco achieved the highest F_1 scores indicating that the combinations of such ULOs resulted in the highest domain coverage effectiveness. It should be noted that the outputs of the domain coverage analysis can vary with changing the size and content of the nominated terms list. In other words, the more accurate way to analysis the domain coverage is to evaluate the matching of all required ALO terms (Section 2) with the candidate ULOs. However, due to the time-consuming analysis process, a list with ten nominated terms was evaluated in this research.

Evaluating the Brinell testing graphs of Figure 2–4, it can be seen that the utilized ULOs provide most domain-related terminologies, but the main challenges come from those terms from multidisciplinary domains that link the use-case concept with the generalized terms of other domains (e.g., Standard). For highly multidisciplinary use cases, this issue can be addressed by reusing the integratable ontologies from different domains. However, gathering particular multidisciplinary terminology from some ULOs, such as EMMO, might be challenging due to their unique domain coverage and the difficulty of their integration with other desired ontologies.

Table 3. Domain coverage metrics (recall, precision, and F_1 score) for candidate ULO combinations (the nominated ten terms for testing the domain coverage of Brinell testing are: measurement, standard, material, equipment, time, force, diameter, shape, quantity, and has value).

Candidate ULO combinations	Number of matching terms	Recall [%]	Precision [%]	F_1 Score [%]
BFO + IOF core+MSEO	9	90	82	86
PROVO + PMDco	10	100	83	91
EMMO + CHAMEO + MT	8	80	80	80

3.3. Reuse and Extensibility

In addition to providing sufficient semantics and high coverage of the domain, the desired ULOs should allow experts to integrate their domain semantics into the ontologies and extend their desired new knowledge within the appropriate hierarchy and classifications. Confirming the results of Section 3.2, the quantitative analysis of developed BTOs in **Table 4** shows that most of the entities for developing the ALOs are driven from ULOs (BFO + IOF core + MSEO, PROVO + PMDco, and EMMO + CHAMEO + MT were respectively provided 73.3%, 81.9%, and 80.6% of the entities required for developing BTO). Therefore, the candidate ontology combinations provided good domain coverages, and only some specialized entities in the domain of Brinell testing were added to BTOs.

According to Table 4, ULOs allow for significant extension and growth, enabling the creation of new domain-specific terminology. The amounts of new classes and properties were more for BTO 5.0.3 (BFO + IOF core + MSEO) and BTO 3.0.3 (EMMO + CHAMEO + MT), respectively. In other words, extending BTO based on BFO + IOF core + MSEO ontologies has mostly been performed by adding new classes, while the properties offered by these ontologies were almost enough for representing the Brinell testing knowledge graph. On the other side, developing BTO based on EMMO + CHAMEO + MT ontologies was also performed by extending the ontology properties, and fewer new classes were required for BTO modeling. This suggests that while EMMO effectively includes detailed properties, it highlights the need for more properties to come from the ULOs to maintain consistency.

It should also be noted that the newly added entities are mostly domain-specific, but there are minor terms across all candidates that are not directly related to the domain (e.g., standard and force). Ideally, these terms and more importantly the required object and data properties should be provided by the TLOs to ensure uniformity. This issue has partially been observed with utilizing EMMO + CHAMEO + MT ontologies, as their focus on the materials science domain limited the presentation of

Table 4. Quantitative analysis of the BTOs entities reused from ULOs or newly added into the ontologies.

Ontologies	Percentage of entities reused from ULOs	Number of new entities			Assertion Axiom Count		
		Classes	Object Properties	Data Properties	Annotation	Data Property	Object Property
BTO 5.0.3 (BFO + IOF core+MSEO)	73.3	21	2	0	42	145	312
BTO 4.0.3 (PROVO + PMDco)	81.9	14	1	0	26	145	312
BTO 3.0.3 (EMMO + CHAMEO + MT)	80.6	12	8	2	24	145	327

more general terminology and even multidisciplinary concepts for developing specific ontologies.

Table 4 also provides additional metrics for the extension of different Brinell testing ontologies. The annotation assertion axiom count indicates the number of annotations within the ontology, which is proportional to the number of newly added classes. This means that ontologies with more new classes have correspondingly higher annotation counts. All three ontologies have similar data property assertion axiom counts, indicating comparable levels of detail, as all knowledge graphs are modeled similarly. The object property assertion axiom count, which indicates the relationships between instances, is slightly higher for BTO 3.0.3 (EMMO + CHAMEO + MT), highlighting more complex interconnections within its ontology. This concern comes from the fact that domain and range constraints of object properties in EMMO limit the expression of the semantic relationship, so additional assertions are required to fully capture these relationships.

3.4. Complexity

Complexity refers to how comprehensive and interrelated an ontology is, and it is vital to consider while picking TLOs since higher structural complexity might affect query processing efficiency and overall performance.[41] Complexity also affects the understandability of the ontology because higher complexity can make it more difficult to comprehend and navigate the relationships and structure within the ontology. The complexity degrees of ontologies are determined by their Description Logic (DL) levels. In other words, DL levels directly influence the complexity of an ontology by determining the expressiveness and computational properties of the ontology. More expressive DLs allow for more complex relationships and constraints but also increase the computational complexity of reasoning tasks. This can lead to longer reasoning times and greater difficulty in managing the ontology. But simpler DLs offer more straightforward reasoning processes but may limit the expressiveness and richness of the ontology.[42] In this research, all developed BTOs have the simplest DL level of Attributive Language (AL), meaning that basic properties and concepts were represented in the ontologies without complex constructs like negation, incorporates roles, role chaining, inverse roles, and qualified cardinality restrictions.[42] While all developed ontologies have simple and similar DL expressivity of AL, their complexity can be measured using structural complexity metrics. These metrics provide insight into the organization and intricacy of the ontology's knowledge base and graph structure and significantly affect the query efficiency and reasoning capability of the developed ontologies.

Table 5 presents a detailed analysis of various structural complexity metrics (absolute cardinality, depth, breadth, and tangledness) related to the knowledge base and graph structure of different developed BTOs. These metrics were evaluated using the OntoMetric tool.[38] Absolute cardinality metrics include the number of key structural elements (roots, leaves, and siblings) in the ontology. The root represents the starting point or the topmost node in the hierarchy, the leaves are the terminal nodes with no children, and the siblings are nodes that share the same parent. Depth measures the levels of hierarchy within the ontology, breadth assesses the number of nodes at each level, and tangledness indicates the degree of interconnection among elements.

BTO 3.0.3 (EMMO + CHAMEO + MT) shows the highest absolute root cardinality with 14 root classes, suggesting a less cohesive ontology structure compared to BTO 5.0.3 (BFO + IOF core + MSEO) and BTO 4.0.3 (PROVO + PMDco), which have 2 and 4 root classes respectively. Conversely, the absolute leaf cardinality of BTO 3.0.3 and BTO 4.0.3 is also high enough that result in the detailed and granulated ontologies, enhancing the specificity and relatedness among ontological entities. The absolute leaf cardinality is lower for BTO 5.0.3 indicating a potentially less detailed representation of concepts, while the absolute sibling cardinality is highest in BTO 3.0.3 denoting that a well-organized structure with many concepts grouped under common parent classes, which could enhance clarity if managed well but might also indicate overly broad categories.

The breadth, depth, and tangledness metrics further illustrate the hierarchical and horizontal structures of the ontologies, impacting their usability and performance efficiency. BTO

Table 5. Complexity metrics of the developed BTOs. Absolute cardinality, depth, breadth, and tangledness, respectively, introduce the number of key structural elements within the ontology, hierarchy levels, width level, and degree of interconnection among elements.

Ontologies	Absolute cardinality [Root/Leaf/Sibling]	Depth [Absolute/Average/Maximal]	Breath [Absolute/Average/Maximal]	Tangledness
BTO 5.0.3 (BFO + IOF core+MSEO)	2/24/62	343/5.5/10	62/1.5/4	0
BTO 4.0.3 (PROVO + PMDco)	4/32/59	195/3.3/7	59/5.1/11	0
BTO 3.0.3 (EMMO + CHAMEO)	14/34/ 75	302/3.6/8	82/1.7/14	0.02

5.0.3 (BFO + IOF core + MSEO) has the highest depth metrics pointing to a well-layered ontology with a deep and understandable hierarchical structure. On the contrary, BTO 3.0.3 (EMMO + CHAMEO + MT) has the highest breadth metrics, suggesting a complex ontology with limited usability and performance efficiency. Tangledness is low across all ontologies, with BTO 3.0.3 (EMMO + CHAMEO) showing a slight degree (0.02), indicating minimal complexity due to classes with multiple superclasses. BTO 5.0.3 (BFO + IOF core + MSEO) and BTO 4.0.3 (PROVO + PMDco) have no tangledness, suggesting straightforward hierarchical structures without complex interconnections. These metrics generally negatively impact usability and resulting performance efficiency due to increased complexity.

3.5. Mapping Efficiency

The knowledge graphs need to be modeled in such a way that they can efficiently map the various content of the datasets. Data mapping in a knowledge graph involves aligning and transforming data from various sources to a unified schema, ensuring consistency, semantic integration, and accurate representation of entities and relationships within the graph. In this regard, before selecting the ULOs, ontology developers need to know which methods they want to use for mapping the data into their knowledge graphs. Depending on using the scripts or different data management systems for performing the data mapping process, there are sometimes limitations to adapting the mapping method with developed ontologies. Furthermore, some scripts and data management systems are also developed based on specific ontologies. For example, ckan.kupferdigital[23] and Mat-O-Lab toolchain[2] successfully map the data to knowledge graphs that are developed based on the BFO, IOF, and CCO ontologies. PMD ontodocker[43] and data space management system (DSMS)[44] were also tested for mapping the data with PMDco- and EMMO-base knowledge graphs, respectively.

In this research, the Ontopanel tool[34] was used for graphically mapping the tabular Brinell datasets (see Section 5) into the Brinell testing knowledge graphs which were developed by reusing different ULOs. This mapping allowed for seamless integration between the ontology and the relational database, which subsequently can enable efficient querying. Following the graphical data mapping process, the data-mapped graphs were also converted to the RDF data using the Ontopanel converter plugin.[34]

3.6. Query Efficiency

Querying a knowledge graph involves retrieving specific information from the graph by using query languages such as SPARQL (SPARQL Protocol and RDF Query Language). The query efficiency parameter evaluates how efficiently the developed knowledge graphs can be processed, including reasoning time and query performance. In other words, an efficient knowledge graph can be successfully and easily processed by a reasoner, and this can be measured by query sentences, response time, and memory consumption during query answering and consistency checking.[45] The query efficiencies of the developed knowledge graphs were evaluated by checking their tendency to respond to the following competency questions: 1) Which alloy has the maximum Brinell hardness? What is its average Brinell hardness? And what is its unit? 2) What is the Brinell hardness of sample CuZn38As? 3) Which value of force was used for the Brinell hardness measurements? 4) Based on which standard was the Brinell hardness measured? 5) List the materials from the dataset that have measured Brinell hardness. 6) Which testing machine is used for Brinell hardness measurement?

Table S1–S6, Supporting Information, represent the SPARQL scripts that were used for querying the aforementioned competency questions from three versions of Brinell hardness knowledge graphs. The SPARQL queries were evaluated by storing the RDF files in Apache Jena Fuseki[46] triple stores (RDF database) and utilizing its SPARQL Endpoint. Furthermore, to assess query time and reasoning performance, RDF files were loaded into memory, and query scripts were tested with five replications in Python.

Concerning the first competency question, the experimental Brinell testing dataset of Section 5 displays that the desired answer to this question should be “alloy CuNi12Al3 with Brinell hardness of 201.23 HBW”. As can be seen in **Figure 5**, all the developed Brinell testing knowledge graphs successfully supported efficient data retrieval and correctly responded to this competency question. However, the length of query scripts which is a sign of ontology complexity differs for these three cases and increased from BTO 5.0.3 (BFO + IOF core + MSEO to BTO 4.0.3 (PROVO + PMDco) and BTO 3.0.3 (EMMO + CHAMEO + MT), respectively.

To have a more detailed assessment of the query efficiency of the developed knowledge graphs, **Table 6** summarizes the query lengths and times of six SPARQL scripts that match the given competency questions. The length of SPARQL scripts is less while querying all competence questions from BTO 5.0.3 (BFO + IOF core+MSEO). Furthermore, the SPARQL query lengths increased for BTO 4.0.3 (PROVO + PMDco) and were the longest among the produced knowledge graphs in BTO 3.0.3 (EMMO + CHAMEO + MT). Related, the average query times were also increased from BTO 5.0.3 to BTO 4.0.3 and BTO 3.0.3, respectively. In this regard, the lower query efficiency of BTO 3.0.3 (EMMO + CHAMEO + MT) can be related to its higher structural complexity metrics and more assertions that are utilized for the semantic representation of relationships within its A-box.

The reasoning performance was also measured as another ontology evaluation metric to determine how effectively and efficiently the developed BTOs can infer new information, check for consistency, and answer queries based on the defined concepts, relationships, and axioms. The reasoning efficiency was calculated using the owlready2 Python library[47] with the Pellet reasoner[48] (see Script S1, Supporting Information). The reasoning efficiency was tested by running each ontology three times, and the average reasoning times were reported in Table 6. In agreement with the query length and time results, BTO 5.0.3 (BFO + IOF core+MSEO), BTO 4.0.3 (PROVO + PMDco), and BTO 3.0.3 (EMMO + CHAMEO + MT) have respectively higher reasoning times. Because all ontologies were developed consistently and based on the less expressive language of AL, they provide rapid reasoning at the cost of expressiveness. These observations suggest that all developed ontologies can handle complex inferencing tasks effectively. However, the better

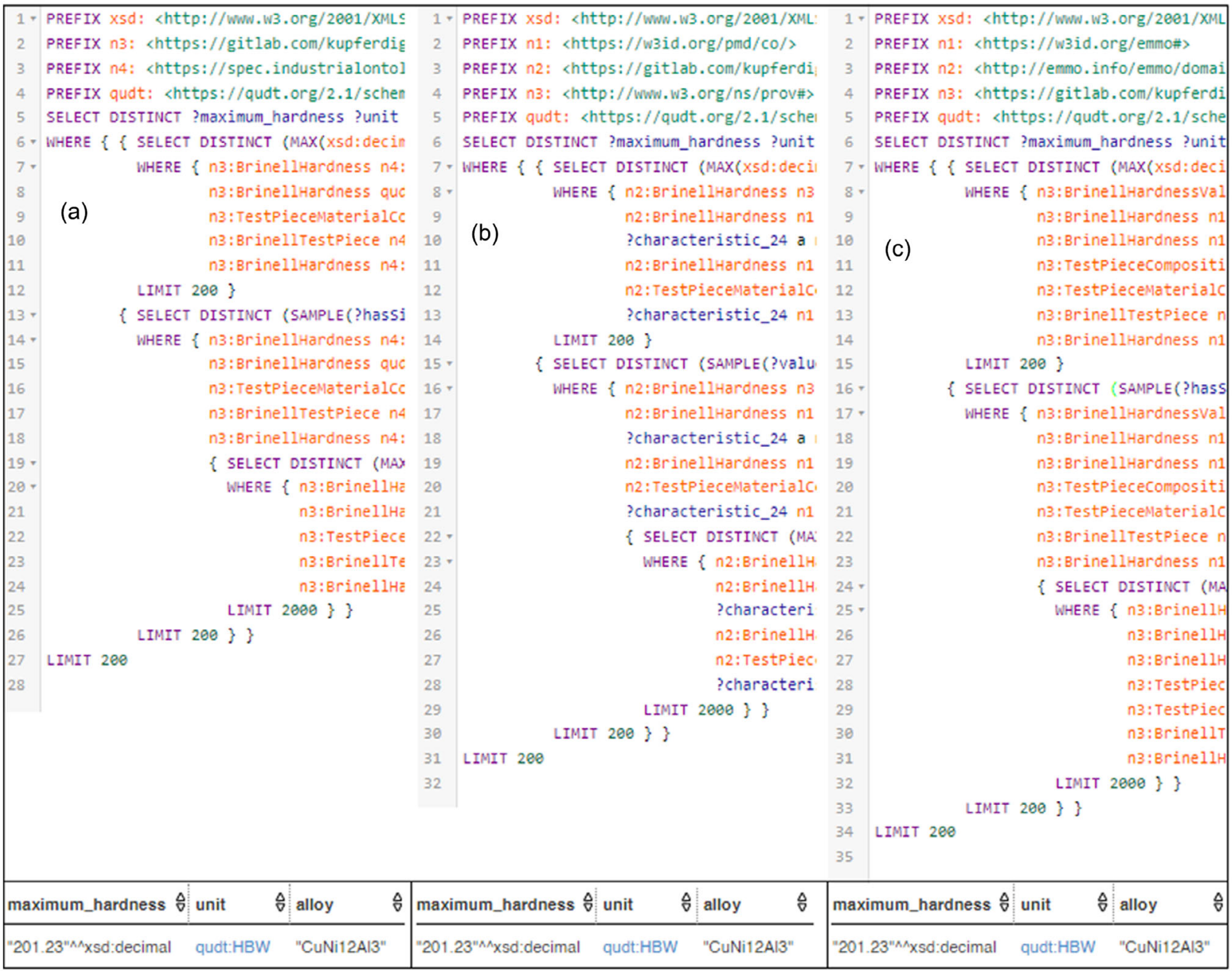

Figure 5. Apache Jena Fuseki-based SPARQL query for the alloy with maximum Brinell hardness (answer: CuNi12Al3 alloy with average Brinell hardness of 201.23 HBW) in different graphs: a- BTO 5.0.3 (BFO + IOF core+MSEO), b-BTO 4.0.3 (PROVO + PMDco), c-BTO 3.0.3 (EMMO + CHAMEO + MT).

Table 6. Query efficiency metrics of the developed BTOs. The query length and time were measured for SPARQL scripts that queried six competency questions.

Ontologies	Query length [number of lines]	Query time [s]	Reasoning Time [s]
BTO 5.0.3 (BFO + IOF core+MSEO)	27, 8, 20, 7, 10, 7 (average = 13)	0.59, 0.037, 0.044, 0.030, 0.034, 0.029 (average = 0.039)	2.311
BTO 4.0.3 (PROVO + PMDco)	31, 9, 19, 7, 10, 8 (average = 14)	0.75, 0.039, 0.049, 0.025, 0.040, 0.036 (average = 0.044)	2.211
BTO 3.0.3 (EMMO + CHAMEO)	34, 11, 21, 7, 11, 8 (average = 15)	0.091, 0.039, 0.052, 0.025, 0.042, 0.036 (average = 0.048)	1.457

reasoning performance of BTO 5.0.3 (BFO + IOF core + MSEO) can be explained by lower complexity, well-structured and modular design, minimized redundancy, and optimized hierarchy of the developed ontology. In this regard, higher reasoning efficiency not only results in faster and more accurate responses to complex queries, but also in the ability to detect inconsistencies quickly and accurately, classify new instances or concepts fast and correctly, and maintain efficiency as the ontology grows.

3.7. Integration with Other Ontologies

One of the most important considerations when selecting ULOs for making ALOs and knowledge graphs is the extent to which such ontologies may be integrated with other relevant ontologies. Integration with other ontologies ensures interoperability, data consistency, and collaboration across domains. In other words, this parameter promotes reuse, scalability, and comprehensive data integration, hence increasing the efficiency and usability of the developed ontology in enabling varied applications and interdisciplinary research. Integration of developed ontologies with existing knowledge can be evaluated by the mapping analysis data of ontology repositories like BioPortal[49] and MatPortal.[50] Furthermore, such analyses conclude that BFO, PROVO, and EMMO provided high mapping with existing ontologies. BFO is one of the TLOs that offers a high degree of integration with a wide range of ontologies. BFO is aligned with the open biomedical ontologies (OBO) foundry principles, which promote its interoperability and shared standards among participating ontologies. Furthermore, its modular, abstract, and high-level nature allows easy integration with other ontologies, providing a common framework that ensures interoperability across different fields.[5b] In addition, aligning IOF with BFO enhanced its integration with further ontologies in the manufacturing and industrial domains.[51] Because of the well-defined generic structure and naming, PROVO can also easily integrate with other ontologies without extensive restructuring, facilitating their interoperability in web technologies and data management.[52] Furthermore, due to the alignment of domain-independent PROVO with the semantics of the materials science domain, PMDco offers facilitating integration with ontologies in materials science and even further cross-domain connections.[14b] Eventually, due to its particular structure and focus on a domain of materials science, EMMO is specifically tailored to integrate with some ontologies in this domain, such as those related to physical properties, processes, and materials characterization. Although the multiperspective nature of EMMO also allows for integration in various scientific and technical disciplines, this integration requires more effort to adapt.[53]

3.8. Adapting to Different Data Contexts

One of the other important parameters that domain ontology developers should consider is to evaluate whether their developed knowledge graphs can be adopted with the data contexts. The knowledge graphs should be modeled in such a way that can integrate and unify data from various heterogeneous sources. Designing knowledge graphs is critical for data integration because it assures consistency in the schema, accurate semantic representation, and effective connecting of disparate data sources, resulting in a unified, interoperable, and informative data model. For example, different semantic representation approaches need to be utilized for designing knowledge graphs that address the data contexts of images, notes, or relational datasets. Furthermore, multiple heterogeneous data contexts should sometimes be modeled in the knowledge graphs. In this regard, the utilized ULOs should assist ontology developers with a wide range of data-related classes, data types, and data properties that can efficiently represent the semantics of planned data contexts. iof:InformationContentEntity, pmd:ValueScope/DataScope, and emmo:EncodedData/DiscreteData are the most important classes from IOF, PMDco, and EMMO that provide such basic classes for dealing with the different data contexts. According to Table 2 and 4, the selected ULOs provide enough data-related classes, data types, and data properties for designing different versions of BTO. However, it should be considered that the developed knowledge graphs of this research planned to address the entities of the simple relational databases, so the result of this section may be different for the other use cases of more complex datasets.

3.9. Acceptance by the Community

Communities, companies, or collaborated projects sometimes decide to use particular ULOs. Such ontologies are mostly recommended for reasons like community strategic plans, ease of collaboration and standardization, integration with their other ontologies, or better adapting to their data and processes. This can also enhance the domain ontologies' credibility, reliability, and potential for integration into other frameworks, hence enhancing their reusability and interoperability across diverse systems and domains.

3.10. Documentation and Maintainability

The following paragraphs provide brief descriptions of the maintainability characteristics (e.g., comprehensibility, documentation, versioning, updatability, and supporting) of the reused top- and mid-level ontologies in this study:

3.10.1. BFO and IOF Core

The BFO project was initiated in 2002, and its last version which was released in 2020[5c] forms the basis of the ISO/IEC 21838-2:2021 standard.[35] BFO GitHub repository[5c] provides detailed information about BFO versions, lists of changes, publications, supports, and developments. Furthermore, all BFO GitHub, ISO/IEC 21838-2:2021 standard, and BFO community publications provide very good documentation of BFO, with detailed specifications, tutorials, and examples. The IOF also released the newest version of core ontology in January 2023.[13] All IOF ontology sources are maintained on GitHub,[54] where the users can also report issues for the core ontology.

3.10.2. PROVO and PMDco

By April 2013, the latest published version of PROVO was released. PROVO is well-documented, with comprehensive guides, examples, and specifications provided by the W3C.[7] As an MLO based on PROVO, PMDco ontology advantages from attractive and well-structured documentation. It includes the PMDco's purpose, scope, design decisions, and limitations, as well as step-by-step visual guidelines for its usage and contribution, which serves as a valuable resource to facilitate its usability. The PMDco documentation is easily accessible via Platform MaterialDigital GitHub.[14a] Furthermore, this GitHub page supports the users' issues and controls the ontology versioning. The last version of PMDco (v. 2.0.7) was published in February 2024.[14a]

3.10.3. EMMO

The last version of EMMO (v. 1.0.0-beta7) was published in February 2024. Browsable documentation and preinferred versions of EMMO are available on its repository.[6] The repository provides a very clear overview of EMMO versioning and addresses perfect documentation of all EMMO versions. Furthermore, the users can access further EMMO products and create their ontology-related issues via the EMMO GitHub repository.[55]

The documentation and maintainability of other frequently used TLOs and MLOs in the materials science domain can be reviewed through the repositories listed in Table 1. Although there is no metric for evaluating the documentation quality of different ontologies, considering some parameters like comprehensibility, versioning, updatability, and supporting it has qualitatively concluded that BFO, EMMO, and PMDco have provided more detailed, comprehensive, and updated documentation than the other ontologies reported in Table 1. It should also be noted that all three developed ontologies of this research offer similar documentation and maintainability features, as all are documented in the same GitLab repository,[33] which provides the terminology documentation, supports the users' issues, and controls the ontology versioning and developments. Furthermore, different versions of the developed BTOs were documented in the ontology repositories like MatPortal.[50]

4. Conclusion

This research evaluated the performance of different TLOs for developing ALOs and knowledge graphs in the materials science and engineering domain. BFO, PROVO, and EMMO are the most frequently used TLOs in the materials science domain which offer well-structured documentation and maintainability. For the use case of Brinell testing, three different versions of Brinell test ontologies (BTOs) were developed by utilizing such TLOs and their integrated MLOs/DLOs. Although all candidate ontologies successfully modeled the Brinell testing knowledge graph, specific ULOs offer unique advantages based on criteria such as semantic richness, domain coverage, extensibility, complexity, query efficiency, integration with other ontologies, and documentation and maintainability. BFO + IOF core+MSEO ontologies yield the best overall performance. Although this ontologies combination has lower semantic richness compared to other candidates, they provide good domain coverage, and their low-complex and well-designed structures allow for simple extension of the domain entities across the hierarchy of ULOs. Furthermore, BFO offers a high degree of integration with a wide range of ontologies and efficient data mapping and query performances. PROVO + PMDco ontologies suggest an acceptable level of semantic richness, domain coverage, extensibility, complexity, and query efficiency. Furthermore, the well-defined generic structure of PROVO allows for efficient integration with other ontologies. Eventually, EMMO + CHAMEO + MT ontologies offer highly rich semantics in the materials science domain. Although these ontologies have high domain coverage and extensibility, their structure is a bit complex which can also negatively influence the mapping and query efficiency. In addition, integration of EMMO-based ontologies with ontologies from other domains appears to be a bit challenging due to their unique structure and focus on specific domains of materials science.

5. Experimental Section

The experimental Brinell hardness datasets were prepared by testing different grades of copper alloys according to the DIN EN ISO 6506-1:2015 standard.[32] **Table 7** represents the ID and composition of utilized test pieces as well as their measured Brinell hardness values. Here, the hardness measurements were repeated 5–6 times for each test piece, and the average, standard deviation and uncertainty values were statistically calculated. All the Brinell hardness tests were done using an Emco Test M4C 025 G3 machine, equipped with a 2.5 mm spherical tungsten carbide composite indenter. The tests were performed at room temperature by applying a test force of 612.9156 N and a loading time of 14 s. Further detailed metadata of the test pieces, Brinell hardness measurement, and uncertainty calculation along with the primary and secondary Brinell hardness testing datasets can be found in the public data repositories.[56]

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

The authors thank the German Federal Ministry of Education and Research (BMBF) for financial support of the project Innovation-Platform

Table 7. Dataset for Brinell hardness measurement of different copper alloys.

Test piece		Brinell Hardness [HBW 2.5/62.5]			
ID	Composition	Average	Standard deviation	Uncertainty	Average hardness ± Uncertainty
A	CuZn38As	111.48	6.71	5.5	111 ± 6
B	CuZn21Si3P	186.45	9.34	5.5	187 ± 6
D	CuSn6	82.60	10.22	5.5	83 ± 6
E	CuSn12	115.28	8.81	5.5	115 ± 6
F	CuNi12Al3	201.23	4.69	5.5	201 ± 6
G2	CuNi6Sn4	124.12	10.23	6.5	124 ± 7
G14	CuNi6Sn4	123.91	10.09	6.5	124 ± 7
H2	CuSn8Ni2	97.64	5.44	6.5	98 ± 7
H14	CuSn8Ni2	94.23	4.72	6.5	94 ± 7
I2	CuZn23Si2.5	141.01	3.42	6.5	141 ± 7
I13	CuZn23Si2.5	99.26	4.19	6.5	99 ± 7

MaterialDigital through project funding grant nos. 13XP5119C (BAM) and 13XP5094F (FIZ).

Open Access funding enabled and organized by Projekt DEAL.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Hossein Beygi Nasrabadi: Data curation: (lead); Formal analysis: (lead); Investigation: (lead); Methodology: (lead); Writing—original draft: (lead); Writing—review and editing: (equal). **Ebrahim Norouzi**: Conceptualization: (supporting); Formal analysis: (supporting); Writing—original draft: (supporting). **Harald Sack**: Funding acquisition: (lead); Supervision: (lead); Writing—review and editing: (equal). **Birgit Skrotzki**: Funding acquisition: (lead); Project administration: (lead); Supervision: (lead); Writing—review and editing: (equal).

Data Availability Statement

The data that support the findings of this study are openly available in Zenodo at https://doi.org/10.5281/zenodo.10820299, ref. [57].

Keywords

Brinell hardness, knowledge graph, materials science, ontology evaluation, top-level ontology

Received: June 28, 2024
Revised: September 13, 2024
Published online: October 4, 2024

[1] a) L. Himanen, A. Geurts, A. S. Foster, P. Rinke, *Adv. Sci.* **2019**, *6*, 1900808; b) B. Bayerlein, T. Hanke, T. Muth, J. Riedel, M. Schilling, C. Schweizer, B. Skrotzki, A. Todor, B. Moreno Torres, J. F. Unger, *Adv. Eng. Mater.* **2022**, *24*, 2101176.
[2] H. Beygi Nasrabadi, T. Hanke, M. Weber, M. Eisenbart, F. Bauer, R. Meissner, G. Dziwis, L. Tikana, Y. Chen, B. Skrotzki, *Comput. Ind.* **2023**, *153*, 104016.
[3] X. Zhang, C. Zhao, X. Wang, *Comput. Ind.* **2015**, *73*, 8.
[4] A. De Baas, P. Del Nostro, J. Friis, E. Ghedini, G. Goldbeck, I. M. Paponetti, A. Pozzi, A. Sarkar, L. Yang, F. A. Zaccarini, *IEEE Access* **2023**, *11*, 120372.
[5] a) R. Arp, B. Smith, A. D. Spear, *Building Ontologies with Basic Formal Ontology*, MIT Press, Cambridge, MA **2015**; b) J. N. Otte, J. Beverley, A. Ruttenberg, *Appl. Ontol.* **2022**, *17*, 17; c) https://github.com/bfo-ontology/BFO-2020 (accessed: August 2024).
[6] https://emmo-repo.github.io/ (accessed: August 2024)
[7] https://www.w3.org/TR/prov-o/ (accessed: August 2024)
[8] a) S. Borgo, R. Ferrario, A. Gangemi, N. Guarino, C. Masolo, D. Porello, E. M. Sanfilippo, L. Vieu, *Appl. Ontol.* **2022**, *17*, 45; b) https://www.loa.istc.cnr.it/index.php/dolce/ (accessed: August 2024).
[9] https://github.com/ontologyportal/sumo (accessed: August 2024)
[10] https://github.com/MaastrichtU-IDS/semanticscience (accessed: August 2024).
[11] https://obi-ontology.org/ (accessed: August 2024).
[12] https://github.com/CommonCoreOntology/CommonCoreOntologies (accessed: August 2024).
[13] https://github.com/iofoundry/ontology (accessed: August 2024).
[14] a) https://github.com/materialdigital/core-ontology/ (accessed: August 2024); b) B. Bayerlein, M. Schilling, H. Birkholz, M. Jung, J. Waitelonis, L. Mädler, H. Sack, *Mater. Des.* **2024**, *237*, 112603.
[15] https://github.com/Mat-O-Lab/MSEO (accessed: August 2024).
[16] a) P. Del Nostro, G. Goldbeck, D. Toti, *Appl. Ontol.* **2022**, *17*, 401; b) https://github.com/emmo-repo/domain-characterisation-methodology (accessed: August 2024).
[17] https://github.com/emmo-repo/domain-mechanical-testing (accessed: August 2024).
[18] https://gitlab.com/kupferdigital/ontologies/mechanical-testing-ontology (accessed: August 2024).
[19] K. Cheung, J. Drennan, J. Hunter, in *AAAI Spring Symp.: Semantic Scientific Knowledge Integration*, AAAI Publication, Washington, DC **2008**, p. 9.
[20] P. Kügler, M. Marian, B. Schleich, S. Tremmel, S. Wartzack, *Appl. Sci.* **2020**, *10*, 4421.
[21] https://github.com/tetherless-world/nanomine-ontology/blob/master/nanomine.ttl (accessed: August 2024).
[22] A. Haller, A. Polleres, *Semant. Web* **2020**, *11*, 87.
[23] H. Beygi Nasrabadi, T. Hanke, B. Skrotzki, *Adv. Eng. Mater.* **2024**, 2400675, https://doi.org/10.1002/adem.202400675.
[24] H. Beygi Nasrabadi, B. Skrotzki, *Adv. Sci. Technol.* **2024**, *149*, 33.
[25] M. Schilling, B. Bayerlein, P. von Hartrott, J. Waitelonis, H. Birkholz, P. Dolabella Portella, B. Skrotzki, *Adv. Eng. Mater.* **2024**, 2400138.
[26] A. Gómez-Pérez, in *Handbook on Ontologies*, Springer, New York, NY **2004**, pp. 251–273.
[27] J. Brank, M. Grobelnik, D. Mladenic, in *Proc. of the Conf. on Data Mining and Data Warehouses (SiKDD 2005)*, Citeseer, Princeton, NJ **2005**, pp. 166–170.
[28] a) C. Reyes-Pena, M. Tovar-Vidal, *Res. Comput. Sci.* **2019**, *148.3*, 257; b) H. Hlomani, D. Stacey, *Semant. Web J.* **2014**, *1*, 1.
[29] A. Degbelo, in *Proc. of the 13th Int. Conf. on Semantic Systems*, CEUR **2017**, pp. 1-8.
[30] M. Sabou, V. Lopez, E. Motta, V. Uren, in *15th Int. World Wide Web Conf., Edinburgh, Scotland*, Association for Computing Machinery, New York, NY **2006**.
[31] R. Lourdusamy, A. John, in *2018 2nd Int. Conf. on Inventive Systems and Control (ICISC)*, IEEE, Piscataway, NJ **2018**, pp. 1415–1421.
[32] ISO, ISO 6506-1:2014 Metallic Materials — Brinell Hardness Test — Part 1: Test Method, **2014**.
[33] https://gitlab.com/kupferdigital/process-graphs/brinell-hardness-test (accessed: August 2024).
[34] Y. Chen, M. Schilling, P. von Hartrott, H. Beygi Nasrabadi, B. Skrotzki, J. Olbricht, *Integrating Materials and Manufacturing Innovation* **2022**, p. 1–12, https://doi.org/10.1007/s40192-022-00279-y.
[35] ISO, *ISO/IEC 21838-2:2021 Information Technology — Top-Level Ontologies (TLO) — Part 2: Basic Formal Ontology (BFO)*, ISO, Geneva, Switzerland **2022**.
[36] D. Zherebetskyy, M. Scheele, Y. Zhang, N. Bronstein, C. Thompson, D. Britt, M. Salmeron, P. Alivisatos, L.-W. Wang, *Science* **2014**, *344*, 1380.
[37] https://protege.stanford.edu/ (accessed: August 2024).
[38] https://ontometrics.informatik.uni-rostock.de/ontologymetrics/ (accessed: August 2024).
[39] M. N. Demaidi, M. M. Gaber, in *Proc. of the ArabWIC 6th Annual Int. Conf. Research Track*, Association for Computing Machinery, New York, NY **2019**, pp. 1–10.
[40] https://wordnet.princeton.edu/ (accessed: August 2024).
[41] S. I. Wilson, J. S. Goonetillake, A. Ginige, A. I. Walisadeera, *IEEE Access* **2022**, *10*, 12889.
[42] F. Baader, I. Horrocks, C. Lutz, U. Sattler, in *Introduction to Description Logic*, Cambridge University Press **2017**.
[43] https://materialdigital.github.io/pmd-server/pages/services/onto-docker/ (accessed: August 2024).

[44] Y. Nahshon, T. Huschle, D. Helm, *An Intuitive and Expandable Dataspace Management System (DSMS) for Materials Science and Manufacturing Processes*, EMMC ASBL, The European Materials Modelling Council, Brussels, Belgium **2023**.
[45] A. Gangemi, C. Catenacci, M. Ciaramita, J. Lehmann, in *SWAP*, Vol. *166*, CEUR Workshop Proceedings, Aachen, Germany **2005**, p. 16.
[46] https://fuseki.kupferdigital.org/ (accessed: August 2024).
[47] https://owlready2.readthedocs.io/en/latest/index.html (accessed: August 2024).
[48] https://www.w3.org/2001/sw/wiki/Pellet (accessed: August 2024).
[49] https://bioportal.bioontology.org/ (accessed: August 2024).
[50] https://matportal.org/ (accessed: August 2024).
[51] C. Franciosi, A. Polenghi, M. Lezoche, A. Voisin, I. Roda, M. Macchi, in *16th IFAC/IFIP Int. Workshop on Enterprise Integration, Interoperability and Networking*, SCITEPRESS-Science and Technology Publications, Rome, Italy **2022**, pp. 218–229.
[52] S. Cox, in *22nd Int. Congress on Modelling and Simulation, Modeling and Simulation Society of Australia and New Zealand*, Tasmania, Australia **2017**, pp. 403–409.
[53] C. Feltus, D. Nicolas, C. Kavka, D. Khadraoui, S. Belouttar, N. Konchakova, H. A. Preisig, P. Klein, in *2023 18th Conf. on Computer Science and Intelligence Systems (FedCSIS)*, IEEE, Piscataway, NJ **2023**, pp. 471-479.
[54] https://github.com/iofoundry/ontology/tree/master/core (accessed: August 2024).
[55] https://github.com/emmo-repo (accessed: August 2024).
[56] a) H. Beygi Nasrabadi, F. Bauer, L. Tikana, P. Uhlemann, S. Thärig, B. Rehmer, B. Skrotzki, *Zenodo* **2023**, https://doi.org/10.5281/zenodo.10820299; b) H. Beygi Nasrabadi, F. Bauer, P. Uhlemann, S. Thärig, B. Rehmer, B. Skrotzki, *Data Brief* **2024**, *55*, 110687.
[57] https://obofoundry.org/ontology/chebi.html (accessed: August 2024).

