

RESEARCH ARTICLE

Review and Alignment of Domain-Level Ontologies for Materials Science

ANNE DE BAAS¹, PIERLUIGI DEL NOSTRO¹, JESPER FRIIS², EMANUELE GHEDINI³,
GERHARD GOLDBECK¹, ILARIA MARIA PAPONETTI³, ANDREA POZZI⁴, (Member, IEEE),
ARKOPAL SARKAR⁵, LAN YANG⁶, FRANCESCO ANTONIO ZACCARINI³,
AND DANIELE TOTI^{1,4,7}, (Member, IEEE)

¹Goldbeck Consulting Ltd., St. John's Innovation Centre, CB4 0WS Cambridge, U.K.

²SINTEF AS, 7034 Trondheim, Norway

³Alma Mater Studiorum, University of Bologna, 40126 Bologna, Italy

⁴Faculty of Mathematical, Physical and Natural Sciences, Catholic University of the Sacred Heart, 25133 Brescia, Italy

⁵Ecole Nationale d'Ingénieurs de Tarbes, Institute for Corrosion and Multiphase Technology, 65000 Tarbes, France

⁶Insight SFI Research Centre for Data Analytics, Data Science Institute, University of Galway, Galway, H91 TK33 Ireland

⁷Department of Sciences, Roma Tre University, 00146 Rome, Italy

Corresponding author: Daniele Toti (daniele.toti@unicatt.it)

This work was supported by the European Union's Horizon 2020 Research and Innovation Programme, in part by OntoCommons under Grant Agreement n. 958371, and in part by OntoTrans under Grant Agreement n. 862136.

ABSTRACT The growing complexity and interdisciplinary nature of Materials Science research demand efficient data management and exchange through structured knowledge representation. Domain-Level Ontologies (DLOs) for Materials Science have emerged as a valuable tool for describing materials properties, processes, and structures, enabling effective data integration, interoperability, and knowledge discovery. However, the harmonization of DLOs, and, more generally, the establishment of fully interoperable multi-level ecosystems, remains a challenge due to various factors, including the diverse landscape of existing ontologies. This work provides, for the first time in literature, a comprehensive overview of the state-of-the-art of DLOs for Materials Science, reviewing more than 40 DLOs and highlighting their main features and purposes. Furthermore, an alignment methodology including both manual and automated steps, making use of Top-Level Ontologies' (TLO) capability of promoting interoperability, and revolving around the engineering of FAIR standalone entities acting as minimal data pipelines ("bridge concepts"), is presented. A proof of concept is also provided. The primary aspiration of this undertaking is to make a meaningful contribution towards the establishment of a unified ontology framework for Materials Science, facilitating more effective data integration and fostering interoperability across Materials Science subdomains.

INDEX TERMS Ontology, domain-level ontology, top-level ontology, alignment, data, harmonization, standardization, bridge concepts, materials science, materials modeling, materials characterization.

I. INTRODUCTION

Materials Science is an interdisciplinary domain that encompasses the study of the properties, processing, and applications of various materials, including metals, ceramics, polymers, and composites. The rapid advancement of Materials Science research, coupled with the escalating requirement for streamlined data management and exchange, has resulted

in an increasing demand for structured knowledge representation in the form of ontologies.

In the context of information science and computer science, an ontology is a formal, explicit representation of shared knowledge, taking the form of a set of classes, relationships and axiomatic constraints. It serves as a tool for structuring data, enabling efficient information retrieval, data integration, interoperability, and knowledge discovery.¹

The associate editor coordinating the review of this manuscript and approving it for publication was Ali Kashif Bashir¹.

¹Compare with [1] and [2] for an introduction to ontologies and applied ontology as a discipline. See also [3] and [4].

In general, it is common practice in the literature to classify ontologies hierarchically depending on the generality of the concept they include, *i.e.* their domain of application. The “levels” so individuated have vague boundaries, yet the presence of borderline cases does not undermine the practical utility of the criterion.²

By definition [8], a Domain-Level Ontology (DLO henceforth), or simply a domain ontology, is an ontology that focuses on concepts, properties, and relationships relevant to a particular area of knowledge or field of study. A DLO can be either a specialized module of an upper-level ontology, or a standalone ontology targeting a specific domain (*e.g.* additive manufacturing, composite materials). While they are intended to retain a measure of neutrality with respect to specific use cases, these ontologies are developed to cater to the informational needs of the domain they revolve around; as such, they are hinged on more fine-grained concepts, and they are not useful beyond their respective domains.

Ontologies at this level are often lightweight, inasmuch as their classes are not thoroughly conceptualized formally, and often lack proper informal characterizations (*i.e.* documentation) as well. This is chiefly due to the fact that these ontologies are engineered to fit the practical needs of a specific community that takes a certain jargon for granted – hence, interoperability is made to rest on common ground rather than semantic transparency. In fact, lower-level ontologies might not necessarily take the form of fully-fledged axiomatic theories; instead, they tend to offer only a basic taxonomical organization of the relevant domain (focusing on hierarchical relations), and minimal terminological services.

Conversely, a Top-Level Ontology (TLO henceforth) provides very general concepts that can be used across various domains. TLOs usually offer systematic and inference-supporting conceptual schemata resting on few principles; thus, they can act as a framework to ensure semantic consistency and are frequently employed to harmonize, or ground, Domain-Level Ontologies (see [8] again).

DLOs for Materials Science provide a structured and standardized vocabulary for describing materials properties, processes, and structures, enabling effective data integration, interoperability, and knowledge discovery. However, DLOs frequently grapple with the absence of generally accepted definitions/elucidations for terms, and often support diverse interpretations, also due to their lack of generality. Furthermore, even ontologies whose scope covers the same, or overlapping, domains, can be widely different: pluralism is a direct result of stakeholders' heterogeneous desiderata, resulting in a diverse landscape. This poses well-known challenges for their harmonization,

²This classification often goes hand in hand with the more rigorous distinction between Foundational and Light-Weight Ontologies [5], [6], [7]. Given the aims of the present work, the analysis is based on the hierarchical criterion.

compromising data exchanges among ontologies as well as their assimilation within the frameworks offered by Top Level Ontologies.

In the context of ontologies, harmonization refers to the process of resolving inconsistencies to allow for efficient collaboration and interoperability. At the semantic and semiotic level [9], this involves ensuring the comparability of the concepts and relations employed by different ontologies as well as the removal of potential sources of ambiguity and friction, up to the achievement of uniformity with respect to formal and informal characterizations – thereby fostering coherence and unity in the knowledge representation, as well as satisfying a precondition for the consistent usage of ontologies in practical scenarios. Ontology harmonization is often understood in terms of alignments, *i.e.* formal connections among entities (like classes, relations, and instances) in different ontologies, up to correspondence/equivalence. These links can be used to transfer knowledge from one ontology to another or to merge ontologies together. Ontology alignment can be a core step in ontology integration [10], which involves merging multiple ontologies into a single, coherent framework, possibly repairing inconsistencies, resolving semantic conflicts, and filling in eventual gaps. The “new” ontology resulting from the integration processes contains the knowledge of all the original ontologies, and can be used in their place, making it easier to manage and navigate the associated information. However, notably, the ontologies resulting from integration processes might not be strictly compatible with the ones they are based on – which might be the ones *de facto* employed in the field. Consequently, alignment can often be more effective from a practical point of view than stricter forms of harmonization, involving ontology integration.

The purpose of this work is to review DLOs for Materials Science and propose (and showcase) a methodology for their alignment with one another and with salient TLOs, focusing on semantic and semiotic aspects. The first contribution of this paper consists of an overview of the existing DLOs for Materials Science, including details concerning their main features and purposes, such as the language in which their latest version is published, the state of their update, their actual inference capacity of reasoning, their use in Materials Science, whether they are based on/aligned with TLOs *etc.* This should increase said ontologies' findability and reusability, among other things, going ways towards the establishment of a defined ontology library for Materials Science.

The focus then turns to mediated alignments, taking inspiration from the fact that TLOs are often successfully employed to facilitate the establishment of connections among DLOs, and taking into account the results of the coverage analysis. A harmonization approach specifically tailored to the establishment of ecosystems of interoperable ontologies is thus presented: a key aspect in this approach concerns the individuation and informal definition of concepts capable of supporting salient, informative, formal

connections across the network: “bridge concepts”. The analysis of Materials Science’s ontologies, and of the TLOs they are based on, can be considered a prerequisite for both the individuation of bridge concept candidates and bridge concepts engineering.

Expounding the philosophical principles underlying bridge concepts is beyond the scope of this paper; nonetheless, given the aims of this work, a brief overview will be offered, focusing on practical aspects. Bridge concepts can be understood as degenerate content ontology design patterns [11], standalone entities engineered independently of any particular ontological framework and then employed as data pipelines. Formal connections are established between bridge concepts and ontology entities belonging to the knowledge representation artifacts to be harmonized; thus, the seamless exchange of information across varied ontological structures is mediated by the proposed tools, which operate as hubs in hub-and-spoke structures. Bridge concepts’ inherent modularity and ontology-neutrality have been deemed a precondition for a sustainable approach to the harmonization of a plurality of diverse ontologies, possibly operating at different levels of abstraction and expressing different worldviews. It is also worth pointing out that bridge concepts are FAIR by design: they are jointly engineered by teams of domain experts and ontologists to ensure accessibility; their informal characterization in natural language (elucidation) is tailored for ontology usage, and the rationale underlying each engineering and mapping choice is properly documented.

In this regard, this work proposes a template for defining bridge concepts, and proposes a method to extract *prima facie* effective candidates via the statistical analysis of the terms appearing in a set of relevant DLOs, which can support a subsequent step of semantic analysis to engineer bridge concepts with comprehensible elucidations both linked to standards and supporting informative links with concepts from said DLOs. Salient bridge concepts resulting from the aforementioned process are also presented as a proof of concept.

By providing a comprehensive analysis of the state-of-the-art of DLOs in Materials Science and offering a methodology for their harmonization based on bridge concepts, this work aims to contribute to the establishment of a unified yet pluralistic ontology framework for Materials Science; such a framework is expected to facilitate more effective data integration and interoperability across Materials Science subdomains, promoting collaborative research, innovation, and the accelerated discovery of new materials and their applications.

This work is carried out within the context of the Onto-Commons H2020 project [12], a European initiative aiming to establish a common foundation for the development, harmonization, and application of ontologies in materials and manufacturing research.

This paper is structured as follows. In Section II, an overview of the state of the art concerning Materials

Science ontologies is provided. A review of the relevant DLOs is offered; given TLOs’ role in facilitating interoperability, the ontologies are grouped according to the TLOs they are aligned with, if any. The section also covers the most salient TLOs and ontology hubs related to Materials, as well as a discussion of relevant tools, approaches and initiatives found in the literature. Section III presents the methodology for analyzing existing ontologies’ coverage for Materials Science. Additionally, it provides a brief outline of a harmonization approach based on bridge concepts, which is implemented in the following two sections, respectively. Section IV offers a partition of Materials Science into subdomains, and implements the coverage analysis of Materials Science DLOs, situating the latter in the landscape based on a statistical method resting on terminological grounds. Conclusions in the form of limitations and opportunities for DLO harmonization are presented. Section V implements the approach for harmonizing DLOs based on bridge concepts, offering more details in the process. The section includes an automated statistical approach for the individuation of effective candidate bridge concepts, a template for defining bridge concepts, and the description of a process of semantic analysis carried out by experts, which can be used to come up with salient bridge concepts from a set of relevant DLOs and standards. The described methodology is applied to the reviewed DLOs, and a sample bridge concept is presented as the output (Section V-C); in light of that, the methodology is reassessed, focusing on matters pertaining to its role in ontology harmonization, practices to ensure efficient reuse of the results and potential future improvements. Finally, in Section VI conclusions are drawn.

II. STATE OF THE ART: DOMAIN OVERVIEW, RELATED TLOS AND HARMONIZATION ATTEMPTS

The sole library/repository targeting specifically Materials Science ontologies is MatPortal <https://matportal.org/>, which covers a total of 28 ontologies. IndustryPortal <http://industryportal.enit.fr/>, also part of the OntoPortal <https://ontportal.org/> initiative, includes Materials Science ontologies as well; however, it does not revolve specifically around that domain. Likewise, to the best of the authors’ knowledge, the domain hasn’t been previously subjected to a systematic coverage analysis. Various TLOs and ontology hubs have been put forward as frameworks to facilitate the harmonization of ontologies pertaining to Materials Science and to support the creation of new ontologies already harmonized with a given ecosystem.

However, perhaps surprisingly, there seem to have been no systematic attempts at harmonizing the various subdomains; in fact, there is also a significant scarcity of alignments between overlapping ontologies (even pairwise), despite the potential benefits. This might be in part due to the lack of a systematic analysis improving ontologies’ findability and reusability.

A. EXISTING TLOs AND HUBS RELATED TO MATERIALS SCIENCE

TLOs are often employed as tools to create new consistent and well-organized ontologies. DLOs based on (or otherwise aligned to) TLOs are more easily harmonized, and already minimally interoperable; as such, TLOs are often employed as the core of hubs of ontologies in the attempt to create harmonized networks of ontologies.

Harmonization strategies resting on semantic alignments with a specific TLO can be particularly effective when all the involved ontologies are based on the same TLO; however, the usual challenges immediately arise if the alignments have to be established. That said, it is possible to establish connections among TLOs to facilitate alignments among ontologies based on different ontologies, a strategy which was also employed in the context of the OntoCommons project [12]. Nevertheless, harmonization strategies based solely on TLOs have limitations when it comes to establishing alignments among DLOs: while they can entail minimal mediated connections, TLOs' role is chiefly that of facilitators and safety checks; in order to establish connections supporting the exchange of meaningful data at the domain level, the standard issues have to be addressed. That said, TLOs' potential should not be underestimated, especially given the possibility of employing mixed harmonization strategies.

A number of TLOs have been deemed especially salient for Materials Science, as a result of a bottom-up coverage analysis, namely: BFO [13], EMMO [14], SIO [15] and SUMO [16]. A detailed description of each of them can be found in the following paragraphs.

1) BASIC FORMAL ONTOLOGY (BFO)

BFO [13] is a top-level ontology that provides a foundational structure consisting of thirty-six classes, including no terms particular to Materials Science. BFO has been employed as a reference by a number of lower-level ontologies – especially in the biomedical domain, but also including relevant hubs e.g. the Industry Ontology Foundry (IOF).

BFO rests on a distinction between universals and particulars; in line with that, it adopts three basic relation types: universal-universal, universal-particular, and particular-particular (where the first relates subtypes to parent types and cannot be time indexed; and the only universal-particular relation is that of instantiation).

BFO's taxonomy is hinged on criteria borrowed from analytic philosophy, concerning whether entities have temporal parts (or, alternatively, whether entities have their properties temporally, or atemporally) and relations of ontological dependence, among other things. BFO endorses a (i) realist, (ii) fallibilist, and (iii) adequatist stance: (i) it represents the world rather than language and concepts; (ii) it concedes that our understanding of the world (and specifically of universals) can change in light of new discoveries – and, consequently, it is committed to tracking said changes over

time; (ii) and it assumes a form of ontological and semantic anti-reductionism on the basis of the parity of (scientific) disciplines.

2) ELEMENTARY MULTIPERSPECTIVE MATERIAL ONTOLOGY (EMMO)

EMMO [14] is an interdisciplinary ontological framework that serves as a platform for applied sciences, specifically designed to cater to Materials Science. Its ultimate purpose is to standardize the representation of materials modeling information, improving interoperability among diverse materials models, data, and software tools.

EMMO combines a sciences-friendly core, comprising of formal distinction supporting the qualitative representation of space-time and offering well-defined criteria of identity, with a plurality of “perspectives” – mutually compatible conceptual schemata augmenting the ontology's expressiveness.

The following are some of EMMO's main perspectives.

- Holistic perspective: this perspective focuses on the dependence relations between an entity, seen as an integrated system (a *whole*), and its parts, playing functional *roles* within the system.
- Semiotics perspective: this perspective takes inspiration from Charles S. Peirce's semiotic theory [17]; it is employed, among other things, to ground a nominalistic, metrology-friendly, approach to the representation of properties, by conveying meaning in the establishment of relationships among entities playing the role of *sign* and *semiotic object* for an *interpreter*.
- Physicalistic perspective: this perspective classifies entities by referring to concepts taken from Physics, and, specifically, from the Standard Model of Particle Physics and Materials Science.
- Reductionistic perspective: this perspective supports the analysis of systems at different levels of granularity, facilitating mechanistic approaches and fine-grained descriptions of systems' evolution in time.

3) SEMANTISCIENCE INTEGRATED ONTOLOGY (SIO)

SIO [15] is a comprehensive and versatile ontology framework that supports collaborative, integrative, and translational research in diverse scientific disciplines. SIO is intended to provide a semantic foundation for the representation of complex scientific knowledge, bridging the gap between the heterogeneity of data representation and the need for integrated knowledge in various research areas.

SIO shares many similarities with the already cited BFO; comparatively, it adopts a more commonsense-friendly approach when it comes to architecture and labeling, and it is more permissive with respect to the representation of fictional and virtual entities.

4) SUGGESTED UPPER MERGED ONTOLOGY (SUMO)

SUMO [16] is an IEEE-sanctioned, Top-Level Ontology created by merging a number of publicly available ontologies in a single cohesive structure. Taking into account its domain

ontologies, it is one of the largest, free formal ontological resources. It is public, and aligned with WordNet. Like other TLOs, SUMO provides a general framework for the representation of knowledge across various domains, facilitating cross-domain integration and interoperability, and the ontology has been used to support the development and integration of domain-specific ontologies in Materials Science.

B. EXISTING MATERIALS SCIENCE DLOS

This subsection offers an overview of the DLOs developed for Materials Science, as well as those used by its community of stakeholders and practitioners.

The authors surveyed the landscape of DLOs related to Materials Science, enlisting them in the study, taking the available resources listed above as a starting point. Domain experts and community practitioners individually made the initial selections; these decisions were then jointly validated to reduce the risks of false positives and false negatives, as well as to ensure that consistent criteria for inclusion were applied. In the process, information regarding the ontologies was also collected for subsequent analysis.

Table 1 includes the name of the ontologies, their acronym (when present), their URLs, as well as the TLO they are aligned with (if any). The alignment of the ontologies with a TLO was inferred from the presence of subsumption relationships involving classes of the DLOs and classes of a TLO, along with bibliographic research. The latest public version of the ontologies was taken as a reference.

In what follows, each ontology is briefly described, paying particular attention to characteristics pertaining to the number of classes and relations, as well as their intended domain of application. When available, data related to the annotations of ontology entities has been included, along with a FAIR score coefficient endorsed by the OntoPortal initiative. The FAIR score depends, among other things, on the quantity and quality of annotations and documentation pertaining to an ontology. Relevant projects associated with the development of the ontologies are also addressed, with the aim of providing contextual information.

1) DLOS IN MATERIALS SCIENCE NOT ALIGNED WITH A TLO

a: AMONTOLOGY: ADDITIVE MANUFACTURING ONTOLOGY [18], [19], [20]

The Additive Manufacturing (AM) Ontology is a modular ontology that employs Basic Formal Ontology as its top level. AMO serves as an advanced framework to represent knowledge about additive manufacturing processes, particularly crucial for metal-based AM, where intricate interconnections among process parameters are yet to be fully understood. It comprises three main constituents: AMProcessOntology, ModelOntology, and AMOntology. AMProcessOntology encompasses a collection of entities capturing the specifics of additive manufacturing processes, aiming to untangle the web of interconnected parameters and

establish a foundation for reliable process control models. ModelOntology, on the other hand, constitutes modeling concepts encapsulating potentially multi-physics, multi-scale processes. Its role is to standardize terminologies and modeling protocols, thus providing a more intuitive graphical view of these complex relationships. In essence, AMOntology weaves together the entities from AMProcessOntology and ModelOntology, creating a robust knowledge base delineating the characteristics of computational models for AM processes. By unifying these diverse aspects, AMOntology presents an integrated framework for understanding and modeling the AM processes. This ontology primarily consists of undefined, horizontally organized independent classes. It comprises 85 classes, of which 83 lack annotations, and includes 5 properties. The ontology features a relatively flat taxonomic structure with 55 child classes that do not further branch out. Only some of these classes are connected by restriction properties. With such a structure, the ontology readily responds to logical inference and deduction processes.

b: BUILDMAT: BUILDING MATERIAL ONTOLOGIES [21]

The Building Material Ontology defines the main concepts of materials, the elements as layers that identify materials design, and general properties. Most of the classes lack elucidations, and there is no nested taxonomy; only first-degree subclass relationships are present. It consists of 26 classes, 23 of which do not have an elucidation, and 63 properties. This ontology's top-level classes are not connected to any TLO, making it an independent low-level ontology. It provides a semantic framework for the representation and sharing of building material data, catering to various phases of engineering and construction projects. This ontology addresses the central role of building material information in design decisions and different simulation processes such as energy, acoustics, and lighting. Despite the availability of different metadata schemas like the Industry Foundation Classes (IFC) [22], efficient data sharing among stakeholders is limited due to inherent constraints. To address this issue, the authors propose an ontology-based approach that utilizes semantic web concepts to enhance interoperability and Building Information Modeling (BIM) [23] data sharing in collaborative workflows. The outcome is the BUILDMAT ontology, which offers improved management of building material information in BIM-oriented collaboration process.

c: DEB: DEVICES, EXPERIMENTAL SCAFFOLDS AND BIOMATERIALS ONTOLOGY [24]

The Devices, Experimental scaffolds, and Biomaterials Ontology (DEB) is an open-source tool designed to structure information related to biomaterials, covering their design, manufacturing, and biological evaluation. It comprises 601 classes, of which 597 lack elucidations, and includes 121 properties. It presents a nested structure, where classes are organized in a hierarchy with different levels of subsumption relationships. It was crafted using text analysis of

TABLE 1. DLOs in materials science.

N.	Acronym	Name	TLO	URL
1	AMONTOLOGY	Additive Manufacturing Ontology		https://matportal.org/ontologies/AMONTOLOGY
2	Atomistic	Atomistic	EMMO	https://github.com/emmo-repo/domain-atomistic
3	BattINFO	Battery INterFace Ontology	EMMO	https://github.com/BIG-MAP/BattINFO
4	BUILDMAT	Building Material Ontologies		https://matportal.org/ontologies/BUILDMAT
5	CAO	Chemical Analysis Ontology	BFO	https://champ.stuchalk.domains.unf.edu/images/ontology/cao.owl
6	CDO	Crystallography Domain Ontology	EMMO	https://github.com/emmo-repo/domain-crystallography
7	CHAMEO	Characterization Methodologies Ontology	EMMO	http://emmo.info/emmo/domain/chameo/chameo
8	CHEBI	Chemical Entities of Biological Interest	BFO	https://www.ebi.ac.uk/ols/ontologies/chebi
9	CHEMINF	Chemical Information Ontology	BFO	https://www.ebi.ac.uk/ols/ontologies/cheminf
10	CHMO	Chemical Methods Ontology	BFO	https://github.com/rsc-ontologies/rsc-cmo
11	CIFO	CIF Ontology	EMMO	https://github.com/emmo-repo/CIF-ontology
12	DEB	Devices, Experimental scaffolds and Biomaterials Ontology		https://github.com/ProjectDebbie/Ontology_DEB
13	EMMO Disciplines: Chemistry	EMMO Disciplines: Chemistry	EMMO	https://github.com/emmo-repo/EMMO/blob/1.0.0-beta4/disciplines/chemistry.ttl
14	EMMO Disciplines: Manufacturing	EMMO Disciplines: Manufacturing	EMMO	https://github.com/emmo-repo/EMMO/blob/1.0.0-beta4/disciplines/manufacturing.ttl
15	EMMO Disciplines: Materials	EMMO Disciplines: Materials	EMMO	https://github.com/emmo-repo/EMMO/blob/1.0.0-beta4/disciplines/materials.ttl
16	EMMO Disciplines: Metrology	EMMO Disciplines: Metrology	EMMO	https://github.com/emmo-repo/EMMO/blob/1.0.0-beta4/disciplines/metrology.ttl
17	EMMO Disciplines: Models	EMMO Disciplines: Models	EMMO	https://github.com/emmo-repo/EMMO/blob/1.0.0-beta4/disciplines/models.ttl
18	ENM	eNanoMapper	BFO	https://bioportal.bioontology.org/ontologies/ENM
19	LPBFO	Laser Powder Bed Fusion Ontology	BFO	https://matportal.org/ontologies/LPBFO
20	MatOnto	MatOnto		https://matportal.org/ontologies/MATONTO
21	MDO-FULL	Material Design Ontology		https://github.com/LiUSemWeb/Materials-Design-Ontology
22	Metal-Alloy	Metal-Alloy	EMMO	https://github.com/arsarkar/ontorepo/tree/master/metal-alloy
23	Microstructure	Microstructure Domain Ontology	EMMO	https://github.com/emmo-repo/domain-microstructure
24	MM	MaterialsMine	SIO	https://matportal.org/ontologies/MM
25	MOCO	Mat-O-Lab Container Ontology		https://matportal.org/ontologies/MOCO/
26	MOL_BRINELL	Brinell Test Ontology		https://matportal.org/ontologies/MOL_BRINELL
27	MOL_TENSILE	Mat-O-Lab Tensile Test Ontology	BFO	https://matportal.org/ontologies/MOL_TENSILE
28	MPO	Material Properties Ontology		https://bimerr.iot.linkeddata.es/def/material-properties
29	MSEO	Materials Science and Engineering Ontology	BFO	https://github.com/Mat-O-Lab/MSEO
30	MTO	Mechanical Testing Ontology	EMMO	https://github.com/emmo-repo/domain-mechanical-testing
31	NanoMine	NanoMine	SIO	https://raw.githubusercontent.com/tetherless-world/nanomine-ontology/master/nanomine.ttl
32	NPO	NanoParticle Ontology	BFO	https://bioportal.bioontology.org/ontologies/NPO
33	OIE Ontologies: Characterization Methods	OIE Ontologies: Characterization Methods	EMMO	http://emmo.info/emmo/domain/oie/characterisation-methods
34	OIE Ontologies: Manufacturing	OIE Ontologies: Manufacturing	EMMO	http://emmo.info/emmo/domain/oie/manufacturing
35	OIE Ontologies: Materials	OIE Ontologies: Materials	EMMO	http://emmo.info/emmo/domain/oie/materials
36	OIE Ontologies: Models	OIE Ontologies: Models	EMMO	http://emmo.info/emmo/domain/oie/models
37	OIE Ontologies: Software	OIE Ontologies: Software	EMMO	http://emmo.info/emmo/domain/oie/software
38	PDO	Photovoltaics Domain Ontology	EMMO	https://github.com/emmo-repo/domain-photovoltaics
39	PMDco	PMD Core Ontology		https://github.com/materialdigital/core-ontology
40	Precipitation Model	Precipitation Model	EMMO	https://github.com/arsarkar/ontorepo/blob/master/metal-alloy/precipitation-model.owl
41	RXNO	Reaction Ontology	BFO	https://github.com/rsc-ontologies/rxno
42	TribAln	TribAln Ontology	SUMO	https://github.com/snow0815/tribAln
43	VIMMP Ontology	Virtual Material MarketPlace Ontology	EMMO	https://gitlab.com/vimmp-semantic/vimmp-ontologies

a carefully curated biomaterials corpus to accurately capture the sector's terminology. Subject matter experts from the biomaterials research field validated the coverage of topics within the ontology. Stored in .owl format, DEB is versatile,

with utilities that extend to term searching, machine learning annotation, standardizing metadata indexing, and facilitating cross-disciplinary data use, thereby promoting improved interoperability and data exploitation.

d: MATONTO [25]

MatOnto is an ontology created by the eResearch Lab at the University of Queensland, aimed at facilitating materials science. It seeks to integrate materials databases, model provenance data, and support knowledge extraction. Initially using DOLCE as its upper ontology, it weaves in pre-existing ontologies to describe diverse elements like units, time, and scientific experiments. Key concepts include Material, Property, Family, Process, Structure, and Measurement. It is composed of 848 classes, 351 of which do not have an elucidation, and 96 properties. It presents a nested structure, where classes are organized in a hierarchy with different levels of subsumption relationships. MatOnto also provides a unique way to describe relational database relationships, allowing dynamic SQL statement creation. Its primary application is in the fuel cell domain, and it's employed to discover new compounds for fuel cell electrolytes. MatOnto also serves as the backbone for MatSeek, a web search tool. Currently, MatOnto is being updated to version 2.0, planning to switch its upper ontology to BFO and enhance its logical consistency and commonality.

e: MDO-FULL: MATERIAL DESIGN ONTOLOGY [26]

The Materials Design Ontology (MDO) aims to address the challenges associated with data accessibility and interoperability across disparate materials databases, particularly prevalent in the materials design domain. It is composed of 37 classes, 8 of which do not have an elucidation, and 64 properties. The majority of the classes are superclasses without children. It sports a FAIRness score of 242. Given that materials databases typically have different data models, users often struggle to locate and integrate data from various sources. The MDO leverages ontologies and ontology-based techniques to improve data availability and interoperability by formalizing domain knowledge representation. Crafted using domain knowledge from materials science, particularly solid-state physics, and guided by data sourced from several materials design databases, MDO is a comprehensive ontology encapsulating knowledge within the materials design field. The application of MDO has been demonstrated on data extracted from established materials databases, showcasing its practical utility in enhancing data interoperability.

f: MOCO: MAT-O-LAB CONTAINER ONTOLOGY [27], [28]

This ontology offers a lightweight solution for detailing the structure of tabular data (series data) stored in hdf5 containers. It has been effectively applied to describe time-force-displacement data from tensile tests. It consists of an ongoing initiative to model an ontology prototype for representing Materials Science experiments, as a common standard formalization for materials knowledge that has not been realized yet. A use case was presented where such an ontology was used to facilitate the curation and comprehension of experiments, outlining the expectations

from the ontology. It is composed of 8 classes, with only one without an elucidation, and 21 properties.

g: MOL_BRINELL: BRINELL TEST ONTOLOGY [29]

This modeling is based on ISO standards. The Mol Brinell ontology is a knowledge representation system that focuses on the domain of material characterization, measurement processes, and associated quantities of interest. This ontology has been developed by distributing a significant number of individuals across a limited number of classes. It is composed of 37 classes, all of which do not have an elucidation, and 21 object properties. It hosts, indeed, more than 2,000 instances. For this reason, it can be regarded as an ontology used primarily at the application level, with a large proportion of its classes being connected to the classes of the MSEO Ontology.

h: MPO: MATERIAL PROPERTIES ONTOLOGY [30]

The Material Properties Ontology is made up of about 150 classes and about 13 object properties, in order to provide the vocabulary to describe the building components, materials, and their corresponding properties, relevant within the construction industry. More specifically, the building elements and properties covered in this ontology support applications focused on the design of building renovation projects. The Building Ontology at the same time reused the SAREF4 Building Ontology to construct this classification. SAREF4 Building Ontology focuses on the concept of *device*, which is defined as a tangible object designed to accomplish a particular task in households, common public buildings, or offices.

i: PMDco: PMD CORE ONTOLOGY [31]

The Platform Material Digital (PMD) project aims to store FAIR data in alignment with a standard-compliant ontological representation (application ontology) of a tensile test of metals at room temperature (ISO 6892-1:2019-11). The process includes developing an ontology per the relevant standard, converting standard test data into the interoperable RDF format, and establishing a connection between the ontology and data. This semantic association promotes interoperability and boosts querying capabilities. To enhance data and knowledge reusability, the PMD core ontology (PMDco), a DLO in Materials Science, has been developed. The interconnection of the tensile test application ontology with the PMDco is highlighted by the ontology's authors. Additionally, a tool designed to aid domain experts in visual ontology development and mapping for FAIR data sharing in materials science and engineering has been reported as developed (Ontopanel).

2) DLOs IN MATERIALS SCIENCE ALIGNED WITH BFO

a: CAO: CHEMICAL ANALYSIS ONTOLOGY [32]

The Chemical Analysis Ontology (CAO) is structured around the Basic Formal Ontology (BFO) framework. It primarily

comprises entries under various categories such as concepts, material entities, information content entities (data items), roles, and processes. At this stage, the ontology mostly resembles a vocabulary, as the development of predicates (ontology properties) linking subjects to objects has not been a central focus. This ontology is associated with ChEBI, CHEMINF and CHMO.

b: CHEBI: CHEMICAL ENTITIES OF BIOLOGICAL INTEREST [32]

The Chemical Entities of Biological Interest (ChEBI) ontology serves as a comprehensive dictionary for small molecular entities, providing a freely accessible resource for researchers across various disciplines. These molecular entities include any atom, molecule, ion, ion pair, radical, radical ion, complex, conformer, and more, which can be distinguished as a separate entity. The entities covered by ChEBI are broad: in fact, it is composed of about 183,000 classes, 130,000 of which without any elucidation, and 10 object properties. Ranging from products found naturally in biological organisms to synthetic products that have been designed to interact with living systems. This could encompass substances that are intentionally introduced into an organism, such as pharmaceutical drugs, or unintentional environmental chemicals that may have biological effects. ChEBI, however, specifically excludes entities directly encoded by the genome, which means that nucleic acids, proteins, and peptides that are derived from proteins by cleavage are generally not incorporated within this ontology. In addition to these individual entities, ChEBI also includes classes of molecular entities and part-molecular entities. These are forms of substituent groups or atoms, providing a more granular level of detail in the ontology's coverage.

c: CHEMINF: CHEMICAL INFORMATION ONTOLOGY [33]

The Chemical Information Ontology (CHEMINF) is an ontology developed to represent and describe chemical information. It is composed of 850 classes, 509 of which not elucidated, and 118 object properties. Cheminformatics, which involves the application of informatics techniques to solve chemical problems in silico, necessitates accurate data exchange, increasingly being accomplished through the use of ontologies. CHEMINF is particularly focused on data-driven research and the integration of calculated properties (descriptors) of chemical entities within a semantic web context. It distinguishes algorithmic, or procedural information, from declarative, or factual information, making the annotation of provenance to calculated data particularly important. This ontology primarily aims to standardize the representation of chemical information. Its primary goals include producing an ontology to represent chemical structure and richly describe chemical properties, whether intrinsic or computed.

d: CHMO: CHEMICAL METHODS ONTOLOGY [32]

The Chemical Methods Ontology (CHMO) consists of over 3,000 classes, 1,000 of which without any elucidation, and less than 30 object properties. It outlines methods employed in chemical experiments. It encompasses methods for data collection (like mass spectrometry and electron microscopy), methods for material preparation and separation (like sample ionization, chromatography, and electrophoresis), and methods for material synthesis (like epitaxy and continuous vapor deposition). Additionally, it describes the instruments employed in these experiments, such as mass spectrometers and chromatography columns, and their outputs. The ontology is represented in both OBO and OWL formats, and both can be edited in Protégé, an ontology editor as well as a knowledge management system.

e: ENM: eNanoMapper [34]

The eNanoMapper project provides a comprehensive computational infrastructure for the management of toxicological data associated with engineered nanomaterials (ENMs). It is made up of 26,000 classes, 3,000 of which not elucidated, and sports 55 object properties. Utilizing ontologies, open standards, and interoperable designs, eNanoMapper promotes a unified approach to European nanotechnology research. Its database supports diverse data and provides an infrastructure for data sharing, analysis, and modeling of ENMs, accessible through an API. A key feature is a configurable spreadsheet parser that simplifies data preparation and upload. Additionally, a web application allows users to retrieve and analyze experimental data using machine learning algorithms. Importantly, the database supports the import and online publication of ENM data from various sources, facilitating the development of reproducible quantitative structure-activity relationships for nanomaterials (NanoQSAR).

f: LPBFO: LASER POWDER BED FUSION ONTOLOGY [35]

LPBFO, the Linked Product Basic Formal Ontology, was developed and effectively employed within the AluTrace project to establish digital connections between disparate data and knowledge silos that often emerge throughout the course of industrial product development and manufacturing cycles. The primary objective of the AluTrace initiative was to create seamless linkages between these silos, enabling a unified flow of linked data and knowledge. By successfully achieving this integration, the linked data was harnessed to address specific use cases, including empowering design engineers to optimize components with a focus on lightweight design for additive manufacturing. This streamlined approach facilitated the realization of enhanced efficiency and precision in the product development process. It has 509 classes, almost all with an elucidation, and 40 object properties.

g: MSEO: MATERIALS SCIENCE AND ENGINEERING ONTOLOGY [36]

MSEO utilizes the IOF Ontology stack giving materials scientists and engineers the ability to represent their experiments and resulting data. The goal is to create machine and human-readable semantic data that can be easily digested by other science domains. It is a product of the joint venture Materials Open Lab Project between the Bundesanstalt für Materialforschung und prüfung (BAM) and the Fraunhofer Group MATERIALS and uses the BWMD ontology created by Fraunhofer IWM as a starting point. It is considered in [37] an “Upper domain level Ontology”. As such, it classifies equipment in a particular domain. It constitutes a base to which ontologies on specific technologies are added. It has 231 classes, 139 of which without any elucidation, and 118 properties.

h: MOL_TENSILE: MAT-O-LAB TENSILE TEST ONTOLOGY [38]

This is an ontology for describing the tensile test process, made in the Materials Open Lab Project. It is a domain-specific Ontology [37] and can be considered aligned to BFO Ontology. Most of its classes are aligned with the MSEO Ontology. It has 327 classes, 50 of these without any elucidation, and 97 properties. It is a detailed digital representation of the classical mechanical tensile test. This widely used and well-standardized test method, with which comparable material parameters are generated.

i: NPO: NANOPARTICLE ONTOLOGY [39]

The NanoParticle Ontology (NPO) serves as a comprehensive repository of essential knowledge pertaining to the physical, chemical, and functional characteristics of nanotechnology in the context of cancer diagnosis and therapy. It is an ontology composed of about 2,000 classes with elucidations and about 80 object properties. Developed within the Basic Formal Ontology (BFO) framework and implemented using the Ontology Web Language (OWL) with well-defined design principles, NPO focuses on representing crucial information related to the preparation, chemical composition, and characterization of nanomaterials used in cancer research. To ensure accessibility and widespread usage, the NPO is made available to the public through the BioPortal website, which is maintained by the National Center for Biomedical Ontology. This facilitates seamless access to the ontology for interested parties. Furthermore, mechanisms for editorial and governance processes are under development to support the ongoing maintenance, review, and expansion of the NPO, ensuring its relevance and accuracy as the field of nanotechnology and cancer research progresses.

j: RXNO: REACTION ONTOLOGY [40], [41]

The RXNO Ontology, or more formally the Reaction Ontology, is a resource designed to provide a standardized classification and description of chemical reactions. This

ontology is designed to annotate and integrate reaction data across various databases and information sources, thereby promoting interoperability. It has about 1,000 classes and 40 object properties. The core of RXNO focuses on reaction types, providing a high-level classification based on reaction mechanisms, reactant or product classes, and roles. These descriptions include a broad array of reactions, from simple transformations like oxidation, reduction or cyclizations, to more complex processes like the Diels-Alder reaction. Among the presented ontologies, RXNO stands out as the most comprehensive and thorough. It boasts a vast array of over 500 name reactions, thoughtfully organized into a multi-layered ontology. Initially, reactions are categorized by their general types, such as oxidations or cyclizations. This primary layer serves as a broad grouping. Subsequently, a secondary layer further subdivides the reactions based on specific characteristics, such as the dedicated reactants involved. For example, oxidations are subcategorized into reactions that synthesize alcohols or alkenes. This hierarchical tree structure empowers users to select reactions explicitly tailored to their desired outcomes. However, despite the wealth of reactions and their organization within the system, the ontology lacks additional information beyond the reaction names and their parent relationships. Moreover, it fails to provide information about the chemicals required for each reaction and vice versa. Additionally, there is no mention of any licensing details associated with the ontology.

3) DLOs IN MATERIALS SCIENCE ALIGNED WITH EMMO

a: ATOMISTIC [42]

An EMMO-based domain ontology for atomistic and electronic modeling. It is a domain ontology aligned to the EMMO multiperspective alpha 2 version. It includes all the concepts related to the fundamental atomistic theory. Most of its concepts are under EMMO’s Reductionistic perspective, and use elucidations from IUPAC Goldbook, IEC standards, and Wikipedia. The ontology, including its imports (the TLO it is based on), encompasses 533 classes and 46 object properties.

b: BattINFO: BATTERY INTERFACE ONTOLOGY [43]

The Battery Interface Ontology (BattINFO) is a free, open-source domain ontology designed to address the current challenges in the battery data landscape characterized by heterogeneous sources and inconsistent metadata. BattINFO aims to enhance the interoperability, reusability, and machine-readability of battery data. Developed within the scope of the EU H2020 project BIG-MAP, it fosters data interoperability across over 30 battery research institutes and companies in Europe. BattINFO employs an ontology - a data model that encapsulates domain knowledge as a network of concepts and their relationships. This model permits data to be mapped to a common vocabulary, enabling the connection of two data pieces mapped to the same term. Expert knowledge can be expressed as machine-readable graphs

that can be queried to identify data linkages and facilitate knowledge sharing. The ontology, including its imports (*i.e.* the version of EMMO it is based on), encompasses 3284 classes and 116 object properties.

c: CDO: CRYSTALLOGRAPHY DOMAIN ONTOLOGY [42]

A crystallography domain ontology that imports EMMO and the CIF core dictionary. It is implemented in a formal language with the objective of archiving and distributing crystallographic information. It is an extremely extensive ontology, in which the fundamental categories of crystallographic models and CIF files are integrated. Both imported ontologies contain the fundamental concepts, while CIF Core, the more detailed one, takes a large percentage of the space. The ontology, including its imports, encompasses 1806 classes and 47 object properties.

d: CHAMEO: CHARACTERIZATION METHODOLOGIES ONTOLOGY [44], [45]

The CHAMEO ontology, a domain ontology, seeks to bridge the divergence in terminologies and data management approaches that exist in the wide-ranging field of materials characterization. Stemming from a recent CEN Workshop Agreement (CWA 17815), CHAMEO is grounded in a standardized terminology and the characterization Data (CHADA) documentation scheme. The overarching goal of CHAMEO is to serve as a unifying framework, harmonizing method-specific ontologies by reusing and specializing its generic constructs. CHAMEO is part of a broader European Materials Modelling Council (EMMC) initiative to develop interconnected materials modeling ontologies based on the Elementary Multiperspective Material Ontology (EMMO) root. Created within the scope of the NanoMECommons European project, CHAMEO's objective is to harmonize characterization protocols. Furthermore, it aligns with several newly developed EMMO-based domain ontologies for classifying materials, models, manufacturing processes, and software products relevant to materials modeling. The ontology's axiomatization can be accessed in a GitHub repository and is published online. It is composed of 479 classes and sports a FAIR score of 195.

e: CIFO: CIF ONTOLOGY [46]

The CIF Ontology is an ontologization of the CIF vocabulary. The ontology, not including its imports, is made up of 68 classes and 5 object properties.

f: EMMO DISCIPLINES [47]

EMMO Disciplines ontologies are positioned between top-level and domain-specific ontologies, representing a unique and versatile category. Some prominent examples include Conformity Assessment, Chemistry, Manufacturing, and Life Cycle Assessment ontologies. These ontologies span across various domains, effectively capturing concepts within a multidisciplinary context. Due to their broad scope and comprehensive nature, the concepts within these

ontologies can be applied in diverse and unrelated applications. By bridging the gap between middle-level and domain-specific ontologies, EMMO Disciplines ontologies provide a flexible framework for knowledge representation and exchange. Their rich conceptual structure allows for seamless integration of data from different domains, promoting interoperability and facilitating knowledge sharing across various fields. For this reason, these are ontologies frequently imported into domain-specific ontologies.

g: METAL-ALLOY [48]

This ontology hosts many main concepts from materials science and metallurgy, that characterize a metal alloy structurally. There are many references to structural types of cells and properties. The ontology, not including its imports, is comprised of 11 classes and 18 object properties.

h: MICROSTRUCTURE DOMAIN ONTOLOGY [49]

The Microstructure Domain Ontology is a specialized ontology designed to encapsulate all vital aspects of metallic microstructures within the field of physical metallurgy. This includes components such as composition, particles (both stable and metastable), grains and subgrains, grain boundaries, particle-free zones, texture, dislocations, and alloy systems. The ontology's goal is to support both microstructure modeling and characterization. In terms of modeling, the ontology differentiates between evolution models, which alter the microstructure via a sequence of states, and property models, which associate a microstructure state with a specific property. Furthermore, it supports the description of the same concept at varying spatial resolutions, such as mean-field, 1D, 2D, and 3D. The ontology can encapsulate mean field descriptions and offers the ability to detail quantities at different statistical levels like mean size, normalized size distribution, and full-size distribution. It also incorporates a common method to link a state to external conditions such as temperature, volume/shape, and pressure, which are vital for describing a process. The ontology, including its imports (*i.e.* the TLO it is based on), encompasses 1305 classes and 107 object properties.

i: MTO: MECHANICAL TESTING ONTOLOGY [50]

This ontology comprises an extensive collection of mechanical properties and testing procedures, which are essential for determining and characterizing the mechanical behavior of materials. It also includes terminology of Metallurgy treatments referring to a series of processes and techniques applied to metals and alloys to modify their physical, chemical, and mechanical properties. It is composed of 819 classes and features a FAIR score of 194.

j: OPEN INNOVATION ENVIRONMENT (OIE) ONTOLOGIES [51]

Five EMMO-compliant, domain-level ontologies developed for the purpose of the Open Innovation Environment platform (developed within the context of the EU-funded

OYSTER [52] and NanoMECommons [53] projects) tackling the areas of characterization methods, manufacturing processes, materials, models, and software products, respectively. The ontologies include 42 classes and no object property (Characterization Methods), 221 classes and 3 object properties (Manufacturing), 117 classes and no object property (Materials), 107 classes and 1 object property (Models), and 161 classes and no object properties (Software). A number of EMMO-based ontologies are aligned with them, including CHAMEO [44], [45] and MAEO [54], the latter an application ontology modeling experts, expertise and knowledge providers within EMMO's ecosystem.

k: PDO: PHOTOVOLTAICS DOMAIN ONTOLOGY [55]

This project aims to establish a comprehensive taxonomy of terminologies related to photovoltaic plants, starting with the fundamental concepts and gradually expanding to cover the detailed description of photovoltaic module composition. By systematically organizing and categorizing the terminology, this initiative seeks to enhance understanding, communication, and research within the field of photovoltaic technology. It imports another module called *Photovoltaics* that contains the classes related to the composition of the Photovoltaic cells layers and inherent measurements.

l: PRECIPITATION MODEL [56]

This ontology includes many physical, statistical and simulation models for the calculation for studying how dislocations move and their position. It is focused on linear defects in material structures. As well as the Mechanical Testing Ontology earlier described, it composes the metal alloy module in EMMO, covering both the experimental and numerical observations.

m: VIRTUAL MATERIAL MARKETPLACE (VIMMP) ONTOLOGY [57]

The Virtual Materials Marketplace project is an initiative that focuses on creating an open platform for offering and utilizing services related to materials modeling. A significant part of the project deals with the development of ontologies and advancing data technology. VIMMP has established a series of marketplace-level ontologies designed to describe services, models, and user interactions, all within the framework of the European Materials and Modelling Ontology, which serves as a top-level ontology. These ontologies play a vital role in data annotation, facilitating the storage of this information within the ZONTAL Space component of VIMMP. Additionally, these ontologies are instrumental in managing data and metadata intake and retrieval at the VIMMP marketplace's front-end.

4) DLOs IN MATERIALS SCIENCE ALIGNED WITH SIO

a: MM: MATERIALSMINE [58]

A materials ontology to support data publication involving nanomaterials and metamaterials. The MaterialsMine Team brings together expertise across five research institutions

in the fields of mechanics, materials, design, manufacturing, data science, and computer science to build and develop an open-source, user-friendly materials data resource guided by FAIR principles, with current modules geared toward research communities in the domains of polymer nanocomposites (NanoMine) and mechanical metamaterials (MetaMine). MaterialsMine is an extension of NanoMine and incorporates curated data from research articles in the field. It is composed of 2052 classes, 1711 of which do not have an elucidation, and 325 properties. The majority of the classes have at least a subclass. The platform utilizes a knowledge graph structure, which is built upon linked data conforming to semantic web ontologies and vocabularies. The knowledge graph in MaterialsMine consists of curated data including information about materials, processing techniques, characterization methods, and bibliographic details.

b: NANOMINE [59]

NanoMine is based on an XML-based data schema designed for nanocomposite materials data representation and distribution. The schema aligns with a higher-level polymer data core, ensuring its consistency with other centralized materials data efforts. Alongside the schema, an ontology and a knowledge graph framework are implemented to provide a more comprehensive representation of nanopolymer systems. The schema, ontology, and knowledge graph ensure ease of accessibility and compatibility with concurrent material standards, thus establishing a robust platform for data storage and search, tailored visualization, and machine learning tool integration for material discovery and design. This integration supports a more systematic approach toward material data handling and application within the materials science domain. The ontology, not including its imports, is comprised of 172 classes and 1 object property.

5) DLOs IN MATERIALS SCIENCE ALIGNED WITH SUMO

a: TRIBAIN ONTOLOGY [60]

The TribAIn ontology was developed to organize, standardize, and enhance accessibility to experimental data in the field of tribology, which focuses on understanding and controlling friction, lubrication, and wear in interacting surfaces. Due to inconsistencies in test procedures and terminology, as well as the practice of publishing results in natural language, accessing and reusing knowledge from tribological experiments can be challenging and time-consuming. Comparisons between different tribological systems or test conditions can be difficult to make. The TribAIn ontology seeks to address these issues by providing a formal, explicit specification of knowledge in the field of tribology. This allows for semantic annotation and searching of experimental setups and results, facilitating the selection of potential tribological pairings based on specific requirements and enabling comparative evaluations. To ensure generalizability, TribAIn is linked to the intermediate-level ontology EXPO, which focuses on the ontology of scientific experiments.

Additional subject-specific concepts are incorporated to address the unique needs of the tribology domain. The ontology's formalization is expressed using the OWL DL standard from the World Wide Web Consortium (W3C). In a use case demonstration, TribAIn effectively covers tribological knowledge from experiments, incorporating data from a range of sources, including natural language texts and tabular data. This illustrates the ontology's potential to improve the management, accessibility, and utility of experimental data in tribology. The ontology, including its imports (*i.e.* the version of SUMO it is based on), encompasses 238 classes and 59 object properties.

C. ONTOLOGY ALIGNMENT AND INTEGRATION APPROACHES

An analysis of the existing literature revealed no relevant deviations or innovations with respect to standard alignment and integration approaches or techniques [61] for the Materials Science area. Notably, automatic techniques used in many contexts (*e.g.* [62], [63], [64]) are strongly privileged, possibly due to the large number of concepts appearing in the relevant ontologies – one of the peculiarities of the domain.

On the other hand, in Materials Science the focus has been on Ontology-based data access, extraction and integration [65]. To cite two examples, the authors of [66] proposed a hybrid Natural Language Understanding-driven technique to extract data from journal articles, whereas the authors of [67] created a database (the Crystallography Open Database) and employed it as a basis to unify and harmonize a plurality of resources.

Notably, the authors of [66] take inspiration from the GENE ontology [68], *i.e.* a hub for biomedical ontologies, and their approach could, in principle, be used to ground alignments between ontologies covering Materials Synthesis. Likewise, an ontology version of the Crystallography Open Database could be used as a core to integrate ontologies, besides databases.

III. METHODOLOGY FOR ANALYZING AND HARMONIZING DOMAIN-LEVEL ONTOLOGIES IN MATERIALS SCIENCE

This section describes a methodology for analyzing and harmonizing DLOs in Materials Science, and evaluating the required steps towards the establishment of a unified ontology framework for Materials Science. This process consists of two distinct phases, which will be detailed in the following subsections.

A. COVERAGE ANALYSIS

The first phase consists of a coverage analysis apt to evaluate the existing Materials Science DLOs, and the effective coverage of the domain. This phase includes the following steps:

- 1) Sub-domains in Materials Science are defined; a glossary of characterizing terms for each sub-domain is compiled, taking inspiration from golden standards.

Domain experts and community practitioners are called to identify the standards and validate the assignments; the validation is carried out jointly by a group of experts in order to reduce the risks and to ensure that coherent criteria are employed.

- 2) An automatic check is executed on the Materials Science DLOs with respect to the characterizing terms, in order to assess each ontology's coverage of a certain sub-domain in Materials Science and situate them within the landscape.
- 3) The overall coverage of the domain is evaluated; the ontologies are analyzed singularly and comparatively to draw a better picture of the state of affairs.

B. HARMONIZATION OF DLOs VIA BRIDGE CONCEPTS

The second phase covers an approach to harmonize the ontologies individuated in the first phase, as well as the ontologies they are based on. The inclusion of upper-level ontologies, and specifically TLOs, serves both to act as a check for the consistency of the proposed alignments, as well as a facilitator. The choice is supported by standard strategies in the literature, and also goes ways towards the establishment of a more comprehensive and reasoning-friendly framework, possibly extending beyond Materials Science *per se* to cover other aspects of manufacturing value chains *etc.* This phase includes the following steps:

- 1) a statistical analysis of relevant terms' frequency is carried out on the individuated ontologies for the domain, in order to identify potential candidates for bridge concepts capable of supporting connections among a large number of ontologies.
- 2) A template for bridge concepts is defined, both as a guide for bridge concepts engineering and to guarantee a FAIR documentation of entities and alignments.
- 3) Bridge concepts are engineered by working groups including both domain experts and ontologists, following the results of the statistical analysis.

More details will be provided *infra*.

IV. MATERIALS SCIENCE: ANALYSIS OF THE DOMAIN AND OF THE DLOs

The following three subsections implement steps 1 to 3 of the coverage analysis in the methodology described above, respectively.

A. CLASSIFICATION OF SUB-DOMAINS IN MATERIALS SCIENCE

Materials Science is a very wide field encompassing a number of disciplines and sub-domains. In order to attain a clearer picture of DLOs' domain coverage, it was decided to assess the ontologies depending on the terms appearing in their documentation. Though indicative and susceptible to errors, this approach was deemed appropriate (and conducive to reasonably accurate results) given the sheer volume of data to be analyzed.

The well-respected, *de facto* standard textbook for Materials Science, *i.e.* Callister's Materials Science and Engineering (now at its 10th edition [69]) with its glossary of some 700 terms, was chosen as a point of reference. To ensure comprehensiveness, supplementary terms were sourced from the materials modeling CWA [70] (based on the Review on Materials Modeling (RoMM) [71]) and the materials characterization CWA. Key chemistry terms (atom, molecule, *etc.*) were also included in the list.

This collection of terms was then manually allocated by experts in Materials Science from the OntoCommons project in one of six subdomains partitioning Materials Science's logical space (listed below, alongside the number of relevant terms), depending on their usage. The operation was carried out by groups of experts subsequently subjected to cross-validation, to ensure that uniform criteria were adopted, and that borderline cases were subjected to thorough discussion.

- 1) Materials classes – 86 terms
- 2) Materials structure – 143 terms
- 3) Materials properties – 99 terms
- 4) Materials behavior – 60 terms
- 5) Materials technologies – 258 terms
- 6) Materials theories – 91 terms

The complete list of terms for each sub-domain is reported in Appendix A. It's worth noting that a term could be classified in more than one sub-domain. However, a principle of salience was adopted to ensure more insightful results. Additionally, it was decided to adopt a permissive approach with respect to polysemy and semantic indetermination. To provide a concrete example, a single term might be used both for the (functional description of a group of) materials and their characteristic property, *e.g. extrinsic semiconductors*; similarly, a term can stand both for a device and a material: *e.g. optical fiber*. In such cases, both uses of the terms were entered. Furthermore, alleged synonyms were identified (*e.g. environment* and *medium*). Naturally, within a domain ontology that describes a particular materials technology, such as characterization, it is expected that specific tools will be connected to the materials amenable to investigation by the latter. Likewise, the structurally measured attributes of the materials are expected to be presented alongside the recorded properties. Nevertheless, in the present analysis, interconnections among the six subject domains are intentionally omitted. This deliberate choice was made to provide a clear-cut assessment of each DLO's specific focus. The terms found in an ontology's documentation are strongly indicative of its contents and domain of application: for instance, it becomes evident whether an ontology describes particular materials, and also which materials structures it documents.

The identified list of terms is by any means not exhaustive (relevant glossaries from textbooks on physics and chemistry could be added to the list), and it does not aim to be so: however, if a domain ontology claims to concern Materials Science, it is arguably expected to include some of the identified terms. Most of the materials among the types

described in Callister's textbook actually appear as materials classes in this analysis.

B. CLASSIFICATION OF THE DLOs AND (SUB-)DOMAIN COVERAGE

The glossary and the assignment of terms to sub-domains formed the basis for situating each ontology in the landscape and analyzing the overall coverage of the Materials Science domain. While domain experts initially reviewed DLOs against the glossary and the terms identified for each sub-domain, this curated process was also supported by a machine-processed computational analysis, given the large number of ontologies involved.

The computational analysis aimed to determine whether these terms were present within each considered ontology, in order to evaluate each ontology's relevance with respect to identified sub-domains. From a technical standpoint, the terms were organized within an Excel file by sub-domain, and served as input for a procedure designed to read, parse, and scan each ontology for their presence. It's essential to note that the search was conducted broadly, including verbatim terms and slight syntactical variations (*e.g.* their corresponding adjectival form), against all the elements of each ontology including classes, properties and comments/annotations. The results of this search are presented in Table 2, where the numerical values provided for each ontology and domain correspond to the percentage of terms associated with a given sub-domain that have been detected within a given ontology. In line with what has been said, these figures provide insights into the extent to which an ontology covers a particular sub-domain.

The main findings of this analysis are as follows. In general, the coverage of terms from the list is relatively low. The highest percentage coverage was observed in the Materials Properties sub-domain, with a maximum coverage of approximately 35% of terms, with 9 ontologies covering more than 10%. The lowest coverage was found in the Materials Classes sub-domain, with a maximum value close to 19%. This result might appear *prima facie* surprising, given the importance of materials classes, though it may be attributed to the extensive diversity of materials types: Materials Science DLOs typically focus on just some types of materials (*e.g.* metals and alloys vs polymers), and research and application areas are typically separated by materials types.

These results show that the overall coverage of each sub-domain by each ontology is limited, with the highest percentages being no higher than about 36%, and with average values being much lower. Moreover, they stress the need for harmonization, as no single ontology can comprehensively cover any sub-domain within Materials Science, not to talk about the domain itself.

C. DLOs ANALYSIS: LIMITATIONS AND OPPORTUNITIES FOR HARMONIZATION

As an uptake of the previous analysis, several limitations and opportunities for harmonization can be discerned:

TABLE 2. Results of the automatic analysis of DLOs. The numerical values provided for each ontology and domain correspond to the percentage of terms associated with a given domain that have been detected within a given ontology.

N.	Acronym	Name	Materials classes [%]	Materials structure [%]	Materials properties [%]	Materials behavior [%]	Materials technologies [%]	Theories for Materials [%]
1	AMONTOLOGY	Additive Manufacturing Ontology	2.198	1.579	2.020	0.000	2.264	0.000
2	Atomistic	Atomistic	0.000	0.526	1.010	0.000	1.887	0.000
3	BattINFO	Battery INterFace Ontology	2.198	1.579	2.020	0.000	3.396	0.000
4	BUILDMAT	Building Material Ontology	2.198	3.158	8.081	8.197	3.019	3.297
5	CAO	Chemical Analysis Ontology	6.593	7.368	4.040	3.279	9.811	3.297
6	CDO	Crystallography Domain Ontology	0.000	4.211	1.010	0.000	1.887	0.000
7	CHAMEO	Characterization Methodologies Ontology	1.099	4.737	5.051	4.918	20.000	3.297
8	CHEBI	Chemical Entities of Biological Interest	3.297	4.211	0.000	0.000	2.264	0.000
8	CHEMINF	Chemical Information Ontology	4.396	6.842	5.051	3.279	4.528	7.692
10	CHMO	Chemical Methods Ontology	12.088	20.526	35.354	26.230	29.434	31.868
11	CIFO	CIF Ontology	4.396	11.053	8.081	4.918	12.075	5.495
12	DEB	Devices, Experimental scaffolds and Biomaterials Ontology	6.593	10.000	14.141	8.197	9.811	3.297
13	EMMO Disciplines: Chemistry	EMMO Disciplines: Chemistry	0.000	4.211	2.020	0.000	3.019	0.000
14	EMMO Disciplines: Manufacturing	EMMO Disciplines: Manufacturing	0.000	1.579	1.010	0.000	2.642	0.000
15	EMMO Disciplines: Materials	EMMO Disciplines: Materials	2.198	5.263	3.030	1.639	3.774	3.297
16	EMMO Disciplines: Metrology	EMMO Disciplines: Metrology	0.000	3.158	6.061	3.279	6.792	3.297
17	EMMO Disciplines: Models	EMMO Disciplines: Models	0.000	1.579	2.020	0.000	7.547	0.000
18	ENM	eNanoMapper	18.681	23.158	31.313	24.590	26.792	27.473
19	LPBFO	Laser Powder Bed Fusion Ontology	0.000	4.737	1.010	4.918	6.415	0.000
20	MatOnto	MatOnto	8.791	6.842	12.121	14.754	3.774	4.396
21	MDO-FULL	Materials Design Ontology	0.000	0.000	1.010	0.000	0.377	0.000
22	Metal-Alloy	Metal-Alloy	3.297	3.158	1.010	0.000	1.887	0.000
23	Microstructure	Microstructure Domain Ontology	3.297	10.526	1.010	1.639	4.906	1.099
24	MM	MaterialsMine	9.890	15.789	23.232	24.590	13.962	10.989
25	MOCO	Mat-O-Lab Container Ontology	0.000	0.526	1.010	0.000	1.132	0.000
26	MOL_BRINELL	Brinell Test Ontology	0.000	0.526	2.020	0.000	3.774	0.000
27	MOL_TENSILE	Mat-O-Lab Tensile Test Ontology	1.099	1.053	4.040	6.557	2.264	0.000
28	MPO	Material Properties Ontology	3.297	2.632	3.030	0.000	3.019	0.000
29	MSEO	Materials Science and Engineering Ontology	6.593	10.000	10.101	19.672	9.811	2.198
30	MTO	Mechanical Testing Ontology	12.088	10.000	13.131	8.197	14.340	2.198
31	NanoMine	NanoMine	6.593	13.158	21.212	21.311	10.566	9.890
32	NPO	NanoParticle Ontology	9.890	16.842	22.222	13.115	16.604	24.176
33	OIE Ontology: Characterization Methods	OIE Ontology: Characterization Methods	0.000	0.526	3.030	0.000	2.264	1.099
34	OIE Ontology: Manufacturing	OIE Ontology: Manufacturing	12.088	5.789	6.061	3.279	8.679	1.099
35	OIE Ontology: Materials	OIE Ontology: Materials	6.593	6.842	3.030	3.279	4.528	3.297
36	OIE Ontology: Models	OIE Ontology: Models	2.198	3.684	3.030	3.279	4.151	1.099
37	OIE Ontology: Software	OIE Ontology: Software	0.000	1.579	0.000	0.000	3.019	0.000
38	PDO	Photovoltaics Domain Ontology	0.000	2.632	2.020	1.639	1.887	2.198
39	PMDco	PMD Core Ontology	0.000	5.263	1.010	1.639	5.660	0.000
40	Precipitation Model	Precipitation Model	1.099	1.579	2.020	1.639	1.509	0.000
41	RXNO	Reaction Ontology	2.198	7.368	4.040	1.639	6.792	2.198
42	TribAIIn	TribAIIn Ontology	2.198	5.789	3.030	3.279	8.679	1.099
43	VIMMP Ontology	Virtual Material MarketPlace Ontology	6.593	7.895	6.061	8.197	18.491	6.593

• *Lack of a comprehensive and unified ontology framework.* Existing Materials Science ontologies often focus

on specific subdomains, resulting in fragmented knowledge representation. The establishment of a unified

ontology framework could facilitate more efficient data integration and interoperability across subdomains.

- *Limited coverage of Materials Science subdomains.* Some subdomains, such as biomaterials and energy materials, are not well-represented in existing ontologies. Harmonization approaches including a built-in way to individuate and fill conceptual gaps, might offer potential solutions to this issue, or contribute to that end.
- *Inconsistencies in terminology and semantics.* Different ontologies may use distinct terminologies and semantics for similar concepts; alternatively, they might use the same labels for different things: both phenomena lead to challenges for data integration and interoperability. Harmonization efforts should concentrate on aligning and standardizing terminologies and semantics across ontologies.
- *Opaque ontological commitments & semantics.* It is often unclear what a certain ontology entity stands for, given alternative equally reasonable interpretations. This is partly due to the fact that DLOs' architecture is often organized pragmatically. Opacity poses challenges for alignment, which require ways to check for consistency (e.g. links with TLOs).
- *Incompatible (and often unclear) ontological commitments.* As in other domains, Materials Science DLOs are partly task-oriented, and, thus, endorse the worldview which they find more suitable given specific aims. This poses another challenge when it comes to alignment, which can be partly addressed by making the commitments more transparent by establishing links with TLOs.
- *Scalability and maintainability.* With the continuous growth of Materials Science, it is crucial to develop a scalable, open, and maintainable ontology framework capable of adapting to the evolving demands of the entire domain.

V. HARMONIZATION OF DLOs THROUGH BRIDGE CONCEPTS

The points above can be seen as guidelines and desiderata for a harmonization methodology. Among other things, it appears pivotal to focus on specific ontology entities that can support informative data exchanges across the network, improve the FAIRness and transparency of the ontologies to be included in the ecosystem and their ontological entities, and provide a tool that also supports links with TLOs, to ensure consistency, validate the proposed alignments and deal with some of the issues outlined above.

In this section, a methodology based on bridge concepts is briefly outlined, following the approach sketched in Section III. Specifically, the following three subsections implement steps 1 to 3 of the harmonization of DLOs via bridge concepts, respectively, while a fourth subsection serves to discuss potential improvements and limitations of the outlined approach.

In short, a statistical analysis of the terms employed in the involved DLOs serves to individuate core nodes in the network of ontologies. Standalone entities are engineered taking the individuated terms as a starting point; they are informally characterized referring to golden standards in a FAIR and extensive fashion; semantic connections are established with ontology entities belonging to Materials Science's DLOs and relevant TLOs. Bridge concepts operate as minimal data pipelines, enabling data exchange at core junctures.

A. IDENTIFICATION OF THE MOST FREQUENT TERMS IN DLOs

In the previous section, the full list of 700 terms has been used to check which of those terms appear in which ontology; however, it can also be used with the purpose of determining the most frequent terms overall. While focusing on terms (labels) is fairly problematic, once again it should provide some hints on which concepts are most relevant within the context of Materials Science DLOs. The results of this automatic statistical analysis are shown in Table 3: the numerical values indicate the number of ontologies (and the percentage) in which each term appears.

Considering the terms' frequency, Table 3 shows the terms with high occurrence, down to the term "molecule" at 47%.

B. BRIDGE CONCEPT TEMPLATE

The following template is proposed for documenting bridge concepts. The template, acting as a human interface, can be divided into three main parts:

- 1) **General Information.** It contains five subparts:
 - **Concept Name.** The label, preferred label, or IRI³ title to be used for identifying the bridge concept.
 - **IRI.** IRI proposed for the bridge concept.
 - **OWL Type.** A value between *Class* | *ObjectProperty* | *Individual*. In this work, the focus will be on the former for the sake of simplicity.
 - **Concept Elucidation.** Natural language informal definition of the concept (elucidation), meant to be easily understandable by domain experts. Elucidations ought to be in line with common sense and knowledge domain resources; they must not refer to other ontology entities, and, ideally, they should be ontologically neutral, *i.e.* they should avoid taking a stance on matters which do not pertain to a specific domain the concepts belongs to. Diverse examples of usage are included whenever possible, and relevant ambiguities that could potentially compromise ontology usage are explicitly addressed. A final desideratum for elucidations is brevity.
 - **Labels.** Labels used to address the concept, ordered as: (i) preferred (one) – the label primarily

³Internationalized Resource Identifier: a unique string that identifies an ontology or an ontology entity in RDF spaces.

TABLE 3. Most frequent terms derived from the automatic analysis of DLOs.

Terms	Occurrences	Percentage (across ontologies) [%]
material	42	89
atom	33	70
structure	33	70
component	32	68
property	29	62
phase	26	55
quantity	25	53
composition	24	51
measurement	23	49
particle	23	49
experiment	22	47
molecule	22	47

TABLE 4. Part 1 of the Bridge Concept template filled for “Material Component”: General information.

Concept Name	Material Component
IRI	ontocommons:MaterialComponent
OWL Type	Class
Concept Elucidation	Constituent part of a material (the whole). A component is distinguishable from the whole for some characteristics (chemically, magnetically, by orientation, <i>etc.</i>), and takes part in the whole, adding its own characteristics and potentially modifying those of the whole. The component can be a chemical compound, a molecular entity, or a phase. A composite material can be made of filler and fiber components. A microstructure consists of components called grains. Magnetic materials also consist of grains with different magnetizations.
Labels	(i) preferred: Material Component; (ii) alternative labels: Material Part, Material Constituent.

used to refer to the concept in brief. It is meant to be both intuitive and informative; (ii) alternative (multiple) – labels that are commonly used to address the concept in practice, even if they are used with narrower or wider sense; (iii) deprecated (multiple) – labels that are misleading with respect to the concept, in that they are ambiguous or encourage misuse). Hidden labels can be included to support queries.

2) **Knowledge Domain Resources.** It contains two subparts:

- **Related Domain Resources.** Existing domain resources (*e.g.* standards, books, articles, dictionaries) that were taken into account for the engineering of the bridge concepts. The template includes static references to the resources and quotes of the relevant informational content. More than one resource can be reported; widely shared resources that are likely to have influenced the users are also covered. Not only do these resources act as a guiding light in the engineering phase, but they also help domain experts situate the bridge concepts, acting as points of reference. As such, a second hub-and-spoke structure pertaining to informal characterizations is established, improving conceptual clarity overall.
- **Comments.** They explain the motivations behind the concept definition with reference to the domain resources, underlying similarities and differences.

3) **Alignment to Existing Ontologies.** It contains six subparts:

- **Target Ontology.** IRI of the ontologies that encompass ontology entities supporting the establishment of semantic connections with the bridge concept.
- **Related Ontology Entities.** List of the identifiers of the specific ontology entities to which the bridge concept is semantically connected, included for the sake of FAIR-ness.
- **Mapping Elucidation.** Natural language discussion of the mapping choice and the underlying rationale, including possible alternative mappings considered and evidence gathered, facilitating the evaluation and validation of the proposed connection by third parties and contributing to the clarification of the bridge concept itself.
- **Semantic Relationship Level.** The strength of the semantic relationship that can be established with a certain ontology entity: (a) **Equivalence** (strong mapping), *e.g.* owl:equivalentClass, owl:equivalentProperty; (b) **Strong Hierarchical**, *e.g.* rdfs:subClassOf, rdfs:subPropertyOf; (c) **Weak Hierarchical**, *e.g.* skos:narrower, skos:broader; (d) **Similarity**, *e.g.* skos:related. Ideally, for each ontology, a bridge concept should be connected with a set of broader concepts (usually one), and a set of narrower concepts (at least one). The strength of the relations and the informativeness of the mappings (inversely proportional to the distance between said sets of concepts) determine the quality of the overall alignment, *i.e.* how much data

TABLE 5. Part 2 of the Bridge Concept template filled for “Component”: Knowledge domain resources.

Knowledge Domain Resources	<p>Component</p> <ol style="list-style-type: none"> [Wikipedia]: Component (thermodynamics): a chemically independent constituent of a phase of a system. [Merriam-Webster dictionary]: The constituent part, ingredient. [Callister, Materials Science and Engineering, Glossary; ISBN 978-1119721772:]: A chemical constituent (element or compound) of an alloy that may be used to specify its composition (the relative content of a particular element or constituent within an alloy, usually expressed in weight percentage or atom percent). [Wikidata]: Smaller, self-contained part of technical entity consisting of components. [IUPAC Goldbook]: Constituent of a mixture the amount or concentration of which can be varied independently. The number of components in a given system is the minimum number of independent species necessary to define the composition in all the phases of a system. It may vary with external conditions since additional chemical equilibria reduce the number of components. The term component is also often used in the more general sense as defined under constituent. [Britannica dictionary]: Mineral associations and phase equilibrium from the article mineral — components are the minimum number of independent chemical species that are necessary to describe the compositions of all the phases present in a system. [ISO Standards 14532:2014, 2.5.2.2.1]: Main component, major component: component whose content influences physical properties. <p>Compound</p> <ol style="list-style-type: none"> [Wikipedia]: Chemical compound — a combination of two or more elements; [Merriam-Webster]: Something formed by a union of elements or parts especially a distinct substance formed by the chemical union of two or more ingredients in definite proportion by weight. [Callister]: Not appearing. [Wikidata]: Chemical compound (Q11173): pure chemical substance consisting of two or more different chemical elements intermetallic compound: solid-state compound exhibiting metallic bonding, defined stoichiometry, and ordered crystal structure. [IUPAC Goldbook]: No entry but the term is used in composite terms (like carbon compound) without explanation of “compound”. <p>Part</p> <ol style="list-style-type: none"> [Cambridge dictionary]: A separate piece of something, or a piece that combines with other pieces to form the whole of something. [Wordreference]: A separate or distinct portion of a whole. [Oxford]: An amount or section which, when combined with others, makes up the whole of something.
Comments	<p>Component: 2. [Merriam-Webster] <i>Adopted as part of the elucidation of the bridge term.</i></p> <p>Component: 3. [Callister] <i>Component does not need to be restricted to alloy. For microstructures, a component could be a phase.</i></p> <p>Component: 5. [Wikidata] <i>This is a circular definition. There is no need to be self-contained. There is no need to be a technical entity; this seems more appropriate for parts of products.</i></p>

TABLE 6. Part 3.1 of the Bridge Concept template filled for “Material Component”: Alignment to existing ontologies – ENM.

Target Ontology	eNanoMapper (ENM)
Related Ontology Entities	<p>Chemical component: http://purl.bioontology.org/ontology/npo#NPO_1497</p> <p>Fiat material part: http://purl.bioontology.org/ontology/npo#NPO_1597</p> <p>Environmental material: http://purl.obolibrary.org/obo/ENVO_00010483</p> <p>Material entity: http://purl.obolibrary.org/obo/BFO_0000040</p> <p>Molecular entity: http://purl.obolibrary.org/obo/CHEBI_23367</p>
Mapping Elucidation	<p>The class chemical “component” is not described in ENM. It is defined as a subclass of “fiat material part”. In “fiat material part”, which is not elucidated as well, there are parts of material, from the macroscopic to the microscopic dimension. “Fiat material part” is a subclass of material entity, which is defined as “<i>An independent continuant that is spatially extended whose identity is independent of that of other entities and can be maintained through time.</i>” In material entity, there are also the environmental materials that are defined as “<i>A portion of environmental material is a fiat object part which forms the medium or part of the medium of an environmental system.</i>”</p> <p>Hierarchy of related entities:</p> <ul style="list-style-type: none"> • Chemical component (no elucidation) is a subclass of entity • Fiat material part is a subclass of material entity • Environmental material is a subclass of material entity
Semantic Relationship Level	ontocommons:MaterialComponent is in between chemical component and Fiat material part from ENM: strong hierarchical semantic relationship.
Mapping Axioms	<pre>ontocommons:MaterialComponent skos:broader enm:ChemicalComponent ontocommons:MaterialComponent skos:narrower enm:FiatMaterialPart enm:MolecularEntity rdfs:subClassOf ontocommons:MaterialComponent</pre>

and reasoning can be potentially shared with other ontologies.

- **Mapping Axioms.** Proposed mapping axiom (or axioms) between the Concept entity and the Target Ontology entities in an OWL2 compliant syntax (e.g. Turtle, Manchester, RDF/XML,

Functional-Style, OWL/XML). It should be noted that bridge concepts support complex alignments that could involve, for instance, properties.

The template developed by the OntoCommons Consortium serves as a guide for bridge concept engineering and alignment. Notably, it has been designed with ontology

TABLE 7. Part 3.2 of the Bridge Concept template filled for “Component”: Alignment to existing ontologies – MM.

Target Ontology	MaterialsMine (MM)
Related Ontology Entities	Component: http://materialsmine.org/ns/Component Constituent: http://materialsmine.org/ns/Constituent MatrixComponent: http://materialsmine.org/ns/MatrixComponent FillerComponent: http://materialsmine.org/ns/FillerComponent MaterialEntity: http://semanticscience.org/resource/MaterialEntity hasProperPart: http://semanticscience.org/resource/hasProperPart ChemicalSubstance: http://semanticscience.org/resource/ChemicalSubstance Ingredient: http://semanticscience.org/resource/Ingredient Composition: http://semanticscience.org/resource/Composition — no elucidation, subclass of entity
Mapping Elucidation	Neither Component nor its subclasses are elucidated in MM. Constituent is a subclass of MaterialEntity. MaterialEntity is defined as “ <i>something that has got a mass as an attribute, that has a proper part only material entity or spatial region</i> ”. Chemical substance is defined as “ <i>a chemical substance is a chemical entity composed of two or more weakly (non-covalently) interacting chemical entities</i> ”. Ingredient is defined as follows: “ <i>an ingredient is a chemical substance that forms part of a mixture.</i> ” Hierarchy of related entities: <ul style="list-style-type: none"> • Composition (no elucidation) is a subclass of Entity • Component is a subclass of Thing • Matrix and Filler are both subclasses of Component • Constituent is a subclass of MaterialEntity • ChemicalSubstance is a subclass of ChemicalEntity • hasProperPart is a subproperty of hasPart
Semantic Relationship Level	ontocommons:MaterialComponent is equivalent to mm:Component: strong hierarchical semantic relation
Mapping Axioms	ontocommons:MaterialComponent owl:equivalentClass mm:Component

implementation in mind and it is FAIR by design. It is worth noting that, as a result, bridge concepts can contribute to addressing issues related to ontologies’ lack of documentation, once mappings are established. Finally, the template is organized into three parts: the first part contains essential information relevant to all users, while the second and third parts cater to domain experts and ontologists, respectively: thus its structure is arguably user-friendly.

C. BRIDGE CONCEPT ENGINEERING

Based on all of the results from the automatic and manual analysis, as well as from further expert discussion, it would be very impactful for domain-level interoperability in chemistry and materials if bridge concepts were available for all of the terms in Table 3, in relation to their use in Materials Science. In addition, experts analysis proposed that the concept of “Materials Process” should also be included. While it does not occur with the highest frequency, the ontologies include a wide range of materials processes in a variety of ways, thus a general superclass would be helpful for bridging.

In the following discussion, some general observations regarding the representation of the concepts ‘material’ and ‘materials property’ in Materials Science’s DLOs are provided. These highlight some of the issues encountered throughout.

(a) Material: The concept of material is of course crucial to all of Materials Science studies. However, while the more generic ‘material entity’ (meaning anything of a material nature rather than abstract nature) is widely used, and also the concept of a chemical substance is reasonably well established (see *e.g.* <https://schema.org/ChemicalSubstance>), material in the sense of its use in Materials Science tends to be represented in different and sometimes contradictory ways.

In particular, the questions are how the concept ‘material’ relates to ‘matter’, and also how it relates to the field of chemistry, where the term ‘chemical substance’ has been defined in the IUPAC Goldbook [72] as a “Matter of constant composition best characterized by the entities (molecules, formula units, atoms) it is composed of”. Furthermore, the question is whether individual atoms and molecules (known as ‘molecular entities’ in the IUPAC Goldbook) are subclasses of materials or more broadly ‘matter’. In the bridge concepts developed in OntoCommons, the elucidation of a material states that it is “an amount of matter at the super-molecular level”, hence differentiating it from molecular entities, in line with IUPAC Goldbeck. It includes chemical substances or mixtures of substances in different states of matter or phases (‘continuum matter’), as well as nanomaterials such as nanoparticles (‘mesoscale matter’). Material is hence broader than or superclass of a chemical substance, since the latter requires a defined composition, whereas the composition of a material may not even be known. This is in contrast to the usage of materials classes in a number of ontologies.

(b) Materials Property is a particular characteristic associated with a material. It is often referred to as a quality, which is inherent in a material, but can also be regarded as the outcome of an observation using a certain method and involving an interpretation of that observation as the property. Materials Property is not necessarily quantitative and not necessarily related to measurement. Nevertheless, some ontologies do not clearly differentiate between materials properties and physical quantities, which require a standardized definition capable of supporting quantification.

So far, the bridge concepts for Material, Materials Property, Material Component and Experiment have been elaborated. As an example, the Material Component bridge concept is

TABLE 8. Part 3.3 of the Bridge Concept template filled for “Component”: Alignment to existing ontologies – EMMO.

Target Ontology	EMMO
Related Ontology Entities	Component: http://emmo.info/emmo#EMMO_f76884f7_964e_488e_9bb7_1b2453e9e817 Constituent: http://emmo.info/emmo#EMMO_f76884f7_964e_488e_9bb7_1b2453e9e817 Object: http://emmo.info/emmo#EMMO_90ae56e4_d197_49b6_be1a_0049e4756606 Material: http://emmo.info/emmo#EMMO_4207e895_8b83_4318_996a_72cfb32acd94 hasComponent: http://emmo.info/emmo#EMMO_3c7f239f_e833_4a2b_98a1_c88831770c1b Substance: http://emmo.info/emmo#EMMO_df96cbb6_b5ee_4222_8eab_b3675df24bea ChemicalEntity: http://emmo.info/emmo#EMMO_47338839_6cca_4a8e_b565_3c4d5517e2c0
Mapping Elucidation	From EMMO’s elucidations and comments, a Component is a “ <i>A constituent of a system</i> ”. A Component is a subclass of Constituent that is “ <i>an object which is a holistic spatial part of an object.</i> ” Object is “ <i>A whole that is identified according to criteria based on its spatial configuration that is satisfied throughout its time extension.</i> ” Material is “ <i>a matter individual that stands for a real-world object representing an amount of a physical substance (or a mixture of substances) in different states of matter or phases.</i> ” Substance is a “ <i>Matter of constant composition best characterized by the entities (molecules, formula units, atoms) it is composed of...</i> ” ChemicalEntity’s comments state that “ <i>A chemical entity comprises the two different ways to represent matter: as a single recognizable particle entity (molecular entity) and as a composition of particle entities (substance). This distinction is not well assessed in actual chemical nomenclature, in which an element name refers to both the pure elemental substance and the atom. In the EMMO the adoption of a stricter categorization based on mereotopology is enforced. The class Material hosts the subclasses for which a substance can be identified without necessarily considering its nature of molecule/atom or substance (e.g. hydrocarbon is the class of both hydrocarbon molecules or gases).</i> ” Hierarchy of related entities: <ul style="list-style-type: none"> • Component is a subclass of Constituent • Constituent is a subclass of Object • Material is a subclass of Matter • hasComponent is a subproperty of hasProperPart • Substance is a subclass of ChemicalEntity • ChemicalEntity is a subclass of Matter
Semantic Relationship Level	ontocommons:MaterialComponent is a subclass of both EMMO’s Component and EMMO’s Matter, or, in other words, is an overlap of both these two concepts: strong hierarchical semantic relation.
Mapping Axioms	ontocommons:MaterialComponent rdfs:subClassOf emmo:EMMO_f76884f7_964e_488e_9bb7_1b2453e9e817 (EMMO’s Component) ontocommons:MaterialComponent rdfs:subClassOf emmo: (EMMO’s Matter)

TABLE 9. Part 3.4 of the Bridge Concept template filled for “Component”: Alignment to existing ontologies – CHEBI.

Target Ontology	CHEBI
Related Ontology Entities	Chemical substance: http://purl.obolibrary.org/obo/CHEBI_59999 Chemical entity: http://purl.obolibrary.org/obo/CHEBI_24431 Food component: http://purl.obolibrary.org/obo/CHEBI_78295 Role: http://purl.obolibrary.org/obo/CHEBI_50906 has part: http://purl.obolibrary.org/obo/BFO_0000051
Mapping Elucidation	From CHEBI’s elucidations and comments, “ <i>A chemical substance is a portion of matter of constant composition, composed of molecular entities of the same type or of different types.</i> ” “ <i>A chemical entity is a physical entity of interest in chemistry including molecular entities, parts thereof, and chemical substances.</i> ” The elucidation of “has part” is missing. Regarding Role, “ <i>biochemical role</i> ” is a subclass of BFO:role. The prefLabel of “ <i>holistic part</i> ” is role; identification of components is based on their role. Food component is “ <i>A physiological role played by any substance that is distributed in foodstuffs. It includes materials derived from plants or animals, such as vitamins or minerals, as well as environmental contaminants.</i> ” Hierarchy of related entities: <ul style="list-style-type: none"> • Chemical substance is a subclass of Chemical entity • Chemical entity is a subclass of Thing • FoodComponent is a subclass of Role As a result, no direct mapping between CHEBI and ontocommons:MaterialComponent could be established, because CHEBI’s Chemical substance could function as a component, but it can also be simply the whole.
Semantic Relationship Level	None
Mapping Axioms	N/A

provided in Table 4, 5, 6, 7, 8, 9. Materials Component plays an important role in one of the largest ontology efforts in Materials Science, namely in NanoMine/MaterialsMine, the latter addressing composite materials. The study of (Material) Component also may enable further elaboration related to

other connected terms, such as ‘Constituent’ and ‘Part’. Furthermore, by taking a closer look at the ontologies that have a high degree of domain coverage in Materials Science, the following priority ontologies (or groups of ontologies) were identified:

- eNanoMapper, which already has some alignment with the Nanoparticle ontology
- CHEBI
- EMMO-based ontologies
- MaterialsMine and NanoMine

Also, MSEO as a BFO-based DLO in Materials Science was deemed important, but it currently does not cover the term ‘Component’. While CHEBI also does not include ‘Component’, CHEBI is the most widely used chemistry ontology, and hence it was important to discuss its mapping regarding such a relevant concept. Further bridge concepts are available and will be further developed on the following GitHub repository: <https://github.com/OntoCommons/OntologyFramework>.

D. POTENTIAL IMPROVEMENTS AND LIMITATIONS

The harmonization approach outlined in this section has several limitations. For instance, the focus on terms (labels) in the automatic analysis might lead to the selection of sub-optimal candidates, even more so considering the issues related to polysemy, semantic indeterminateness, and intra-domain variance considered in the course of the discussion. While the semantic connections are re-evaluated case by case, the selection of a sub-optimal candidate can lead to a waste of resources and to less informative connections.

In general, approaches revolving around pinpointed links will always be less informative (*i.e.* support the transmission of less information/data) than full mappings.

Another problem concerns the difficulty of establishing connections between the bridge concepts and ontology entities belonging to the ontologies part of the ecosystem: while having a properly characterized point of reference can help, this process is not per se any simpler than the establishment of a standard semantic connection between two ontology entities whatsoever. While the connections with TLOs might offer a consistency check, thus defining affordances and constraints and partially validating the links, extreme caution should be exercised for each and every alignment.

Lastly, inconsistencies might emerge due to mistakes in the ontologies considered, or due to incompatible ontology commitments. Given the focus on harmonization rather than integration, only a cooperative effort with the ontology developers can help resolve similar issues.

Overall, the methodology might appear as a fallback solution given the intrinsic difficulties involved in the establishment of a harmonized networked ecosystem; however, there could be ways to improve the approach without sacrificing scalability. For instance, automatic alignment tools could be employed to improve the selection of candidate terms, moving beyond terms (labels); the statistical analysis could also be improved by weighting the candidates making use of graph theory and network sciences; finally, it might be possible to partially automatize the engineering process, going ways towards a fully bottom-up approach starting from

a given set of ontologies and selected documentation. These points might be explored in future works.

VI. CONCLUDING REMARKS

This work has provided a comprehensive examination of the current state-of-the-art of DLOs for Materials Science, highlighting their main characteristics and their connections with existing TLOs

The harmonization of DLOs with TLOs has been identified as a crucial challenge in developing a unified ontology framework for Materials Science. The role of TLOs in supporting the harmonization of DLOs has been examined, mentioning the main TLOs currently available and/or under development for the area, which may be exploited for facilitating consistency and cross-domain integration, as well as to ensure interoperability.

Harmonization via TLOs is arguably not enough to establish strong DLO-DLO pipelines, therefore, this work has proposed an approach for the establishment of multi-level, harmonized ontology networks based on the identification of “bridge concepts”. A template for bridge concepts has been presented, and a methodology made up of both a manual and automatic semantic analysis has been proposed and applied to the identified DLOs in order to bring about a number of potential bridge concept candidates. Eventually, an example of a bridge concept has been proposed as a result of this process, showcasing actual semantic alignments mediated by the latter.

Ideally, this work is a step toward the creation of a more coherent and unified ontology framework for Materials Science, being the latter one of the core objectives of the OntoCommons project from which it stems. Such a framework is expected to facilitate more effective data integration and interoperability across Materials Science subdomains, ultimately promoting collaborative research, innovation, and the accelerated discovery of new materials and their applications.

APPENDIX A SUB-DOMAINS IN MATERIALS SCIENCE WITH THEIR CORRESPONDING LIST OF TERMS

A. MATERIALS CLASSES

- 1) alloy
- 2) alternating copolymer
- 3) atom
- 4) austenite
- 5) bainite
- 6) block-copolymer
- 7) brass
- 8) bronze
- 9) carbon-carbon composite
- 10) cement
- 11) cementite
- 12) ceramic-matrix composite
- 13) ceramics
- 14) cermet

- 15) chemical entity
- 16) chemical substance
- 17) coarse pearlite
- 18) colorant
- 19) compacted graphite iron
- 20) concrete
- 21) copolymer
- 22) crosslinked polymer
- 23) crystallite
- 24) dispersion strengthened materials
- 25) ductile iron
- 26) elastomer
- 27) electrolyte
- 28) extrinsic semiconductor
- 29) ferrite
- 30) ferrous alloy
- 31) fine pearlite
- 32) flame retardant
- 33) glass-ceramic
- 34) graft copolymer
- 35) gray cast iron
- 36) green ceramic body
- 37) hard magnetic material
- 38) high polymer
- 39) high-strength low-alloy steel
- 40) homopolymer
- 41) hsla steel
- 42) hybrid composite
- 43) hypereutectic alloy
- 44) insulator
- 45) malleable cast iron
- 46) martensite
- 47) material category
- 48) metal-matrix composite
- 49) nodular iron
- 50) non-ferrous alloy
- 51) optical fiber
- 52) particle
- 53) particle-reinforced composite
- 54) pearlite
- 55) plain carbon steel
- 56) plastic
- 57) plasticizer
- 58) prepreg
- 59) prestressed concrete
- 60) primary phase
- 61) processing
- 62) proeutectoid cementite
- 63) proeutectoid ferrite
- 64) refractory
- 65) reinforced concrete
- 66) repeat unit
- 67) semiconductor
- 68) soft magnetic material
- 69) solute
- 70) solvent
- 71) stabiliser
- 72) stainless steel
- 73) structural clay products
- 74) structural composite
- 75) structure
- 76) substitutional solid solution
- 77) superconductor
- 78) tempered martensite
- 79) thermoplastic elastomer
- 80) thermoplastic polymer
- 81) thermoset polymer
- 82) ultra-high-molecular-weight polyethylene
- 83) whisker
- 84) white cast iron
- 85) whiteware
- 86) wrought alloy

B. MATERIALS STRUCTURE

- 1) allotropy
- 2) alternating copolymer
- 3) amorphous
- 4) anion
- 5) anisotropic
- 6) annealing point
- 7) antiferromagnetic
- 8) atactic
- 9) atom percent
- 10) atomic packing factor
- 11) atomic weight
- 12) bifunctional
- 13) block-copolymer
- 14) body-centered cubic
- 15) bonding energy
- 16) branched polymer
- 17) brittle
- 18) burgers vector
- 19) carbon nanotubes
- 20) carbonizing
- 21) case hardening
- 22) chain folded model
- 23) cis
- 24) component
- 25) composition
- 26) concentration
- 27) concentration profile
- 28) concentratrion gradient
- 29) congruent transformation
- 30) continuouse cooling transformation
- 31) coordination number
- 32) copolymer
- 33) covalent bond
- 34) crosslinked polymer
- 35) crystal structure
- 36) crystal system
- 37) crystallinity
- 38) crystallisation

- 39) crystallite
 - 40) defect structure
 - 41) dislocation
 - 42) dislocation density
 - 43) dislocation line
 - 44) dispersed phase
 - 45) dispersion strengthened materials
 - 46) domain
 - 47) electric dipole
 - 48) electron configuration
 - 49) electroneutrality
 - 50) eutectic phase
 - 51) eutectic reaction
 - 52) eutectic structure
 - 53) eutectoid reaction
 - 54) face-centered cubic
 - 55) fiber
 - 56) fiber reinforcement
 - 57) fiber-reinforced composite
 - 58) filler
 - 59) fine pearlite
 - 60) foam
 - 61) frenkel defect
 - 62) glass-ceramic
 - 63) graft copolymer
 - 64) grain
 - 65) grain boundary
 - 66) grain size
 - 67) growth
 - 68) hexagonal closed packed
 - 69) high polymer
 - 70) hydrogen bond
 - 71) imperfection
 - 72) intergranular fracture
 - 73) interstitial solid solution
 - 74) isomerism
 - 75) isomorphous
 - 76) isotopes
 - 77) laminar
 - 78) lattice
 - 79) lattice parameters
 - 80) lattice strain
 - 81) linear polymer
 - 82) liquid crystal polymer
 - 83) liquidus line
 - 84) longitudinal direction
 - 85) macromolecule
 - 86) matrix phase
 - 87) metastable
 - 88) microconstituent
 - 89) miller index
 - 90) mixed dislocation
 - 91) molarity
 - 92) molecular structure
 - 93) molecular weight
 - 94) molecule
 - 95) monomer
 - 96) nanocarbons
 - 97) nanocomposite
 - 98) network polymer
 - 99) non-crystalline
 - 100) octahedral position
 - 101) particle growth
 - 102) particle-reinforced composite
 - 103) pearlite
 - 104) phase
 - 105) phase transformation
 - 106) pilling-bedworth ratio
 - 107) plane strain fracture toughness
 - 108) point defect
 - 109) polycrystalline
 - 110) polymer-matrix composite
 - 111) polymerization degree
 - 112) polymorphism
 - 113) primary phase
 - 114) quantity
 - 115) random copolymer
 - 116) recrystallisation
 - 117) relaxation modulus
 - 118) repeat unit
 - 119) sandwich panel
 - 120) schottky defect
 - 121) screw dislocation
 - 122) self-diffusion
 - 123) self-interstitial
 - 124) single crystal
 - 125) slip system
 - 126) solid solution
 - 127) solute
 - 128) solvent
 - 129) specific heat
 - 130) spheroidite
 - 131) spherulite
 - 132) stereoisomerism
 - 133) stress raiser
 - 134) substitutional solid solution
 - 135) syndiotactic
 - 136) tempered martensite
 - 137) tetrahedral position
 - 138) transgranular fracture
 - 139) transverse direction
 - 140) true strain
 - 141) unit cell
 - 142) vacancy
 - 143) whisker
- C. MATERIALS PROPERTIES**
- 1) abrasive
 - 2) adhesive
 - 3) allotrope
 - 4) anelastic
 - 5) anisotropic

- 6) annealing point
- 7) antiferromagnetic
- 8) atactic
- 9) bifunctional
- 10) capacitance
- 11) coercivity
- 12) color
- 13) conductivity electrical
- 14) critical resolved shear stress
- 15) crystallinity
- 16) design stress
- 17) diamagnetic
- 18) dielectric
- 19) dielectric breakdown strength
- 20) dielectric constant
- 21) diffusion coefficient
- 22) dispersion strengthened
- 23) doping
- 24) ductility
- 25) elastic
- 26) electric dipole
- 27) electrical conductivity
- 28) electron mobility
- 29) electronegative
- 30) electropositive
- 31) endurance limit
- 32) extrinsic semiconductor
- 33) fatigue limit
- 34) fatigue strength
- 35) ferrimagnetic
- 36) ferroelectric
- 37) ferromagnetic
- 38) flexural strength
- 39) fluorescent
- 40) fracture toughness
- 41) glass transition temperature
- 42) hard magnetic
- 43) hardenability
- 44) hardness
- 45) heat capacity
- 46) hole mobility
- 47) insulator
- 48) isotropic
- 49) lower critical temperature
- 50) magnetic susceptibility
- 51) magnetization
- 52) melting point
- 53) melting temperature
- 54) metastable
- 55) molarity
- 56) molecular structure
- 57) molecular weight
- 58) opaque
- 59) p-type semiconductor
- 60) paramagnetic
- 61) permeability
- 62) permittivity
- 63) phosphorescent
- 64) piezoelectric
- 65) polar
- 66) polarisation
- 67) property
- 68) proportional limit
- 69) recrystallization temperature
- 70) reflectivity
- 71) refraction index
- 72) refractory
- 73) relative magnetic permeability
- 74) relaxation frequency
- 75) residual stress
- 76) resilience
- 77) resistivity
- 78) resolved shear stress
- 79) safe stress
- 80) saturation flux density
- 81) saturation magnetization
- 82) softening point
- 83) solubility limit
- 84) specific modulus
- 85) specific strength
- 86) strain point
- 87) superconductivity
- 88) tensile strength
- 89) thermal conductivity
- 90) thermal expansion coefficient
- 91) toughness
- 92) translucency
- 93) transparency
- 94) upper critical temperature
- 95) viscoelasticity
- 96) viscosity
- 97) working point
- 98) yield strength
- 99) young's modulus

D. MATERIALS BEHAVIOR

- 1) design stress
- 2) diffusion
- 3) diffusion coefficient
- 4) dislocation
- 5) dislocation density
- 6) dislocation line
- 7) ductile-to-brittle transition
- 8) ductility
- 9) elastic deformation
- 10) elastic recovery
- 11) elastomer
- 12) electroluminescence
- 13) engineering stress
- 14) fracture toughness
- 15) interdiffusion
- 16) interstitial diffusion

- 17) lattice strain
- 18) magnetic field strength
- 19) magnetic flux density
- 20) magnetic susceptibility
- 21) magnetization
- 22) non-steady-state diffusion
- 23) paramagnetic
- 24) paramagnetism
- 25) permeability
- 26) permittivity
- 27) plane strain
- 28) plastic deformation
- 29) poisson's ratio
- 30) polarisation
- 31) proportional limit
- 32) recovery
- 33) relative magnetic permeability
- 34) remanence
- 35) residual stress
- 36) resilience
- 37) resolved shear stress
- 38) saturation flux density
- 39) saturation magnetization
- 40) shear strain
- 41) shear stress
- 42) slip
- 43) specific modulus
- 44) specific strength
- 45) specific heat
- 46) steady-state diffusion
- 47) strain
- 48) thermal conductivity
- 49) thermal expansion coefficient
- 50) thermal stress
- 51) thermally activated transformation
- 52) translucency
- 53) transparency
- 54) true strain
- 55) true stress
- 56) upper critical temperature
- 57) viscoelasticity
- 58) viscosity
- 59) vitrification
- 60) young's modulus
- 11) artificial aging
- 12) athermal transformation
- 13) atomic emission spectroscopy
- 14) atomistic model
- 15) austenitizing
- 16) bifunctional
- 17) bonding energy
- 18) branched polymer
- 19) brazing
- 20) calcination
- 21) calibration
- 22) calibration workflow
- 23) carbonizing
- 24) case hardening
- 25) cation
- 26) ccd cooling temperature
- 27) chada
- 28) chain of models
- 29) characterization environment
- 30) characterization experiment
- 31) characterization method
- 32) characterization validation
- 33) charpy test
- 34) cold working
- 35) component
- 36) composition
- 37) computational boundary conditions
- 38) computational representation
- 39) computational translation
- 40) concentration
- 41) concentration polarisation
- 42) concentration profile
- 43) concentratrion gradient
- 44) condensation
- 45) continuum model
- 46) data acquisition rate
- 47) data analysis
- 48) data filtering processes
- 49) data levels
- 50) data management
- 51) data normalization
- 52) data post-processing
- 53) data processing methodologies
- 54) data quality check
- 55) dea
- 56) design stress
- 57) deta
- 58) detector
- 59) dielectric thermal analysis
- 60) differential scanning calorimetry
- 61) differential thermal analysis
- 62) diffraction
- 63) diffraction-based methods
- 64) diffraction grating
- 65) diffraction slit width
- 66) dislocation

E. MATERIALS TECHNOLOGIES

- 1) access conditions
- 2) access conditions data
- 3) acquisition parameters
- 4) acquisition routine
- 5) acquisition time
- 6) action (characterization)
- 7) addition polymerization
- 8) ambient conditions
- 9) annealing
- 10) annealing point

- 67) dislocation density
- 68) dislocation line
- 69) dispersion strengthening
- 70) domain
- 71) doping
- 72) drawing
- 73) dsc
- 74) dta
- 75) dynamic light scattering
- 76) elaborated data
- 77) electrical characterization methods
- 78) electron configuration
- 79) electron diffraction
- 80) electron microscopy
- 81) electronic model
- 82) energy dispersive spectroscopy
- 83) environment
- 84) equations
- 85) equipment setup
- 86) eutectic reaction
- 87) eutectoid reaction
- 88) excitation wavelength
- 89) experiment
- 90) experimental plan
- 91) extrusion
- 92) fatigue life
- 93) fatigue limit
- 94) fatigue strength
- 95) fiber reinforcement
- 96) figure of merit
- 97) filler
- 98) flexural strength
- 99) focused ion beam microscopy
- 100) forging
- 101) galvanic corrosion
- 102) geometry
- 103) governing equations
- 104) grain growth
- 105) grain size
- 106) growth
- 107) hardness
- 108) hazard
- 109) hot working
- 110) hydroplastic forming
- 111) impact energy
- 112) in-line material characterization
- 113) in-situ micro/nano-mechanical testing
- 114) infrared spectroscopy
- 115) inservice conditions
- 116) interaction physics
- 117) interaction volume
- 118) isothermal transformation diagram
- 119) iterative coupling
- 120) izod test
- 121) jominy end-quench test
- 122) laser diffraction
- 123) laser power
- 124) level of expertise
- 125) lever rule
- 126) light scattering techniques
- 127) linking
- 128) lot
- 129) lower critical temperature
- 130) machine automated data reprocessing
- 131) macroscopic scale mechanical testing
- 132) magnetic characterization methods
- 133) manufacturing process
- 134) margin of error
- 135) mass spectroscopy
- 136) material
- 137) material characterization
- 138) materials characterization data
- 139) materials modelling data
- 140) materials relation
- 141) matthiessen's rule
- 142) measurement
- 143) measurement parameters
- 144) measurement time
- 145) mechanical testing method
- 146) medium
- 147) mesoscopic model
- 148) micro-indentation
- 149) microscope objective
- 150) microscopy
- 151) microscopy-based method
- 152) moda
- 153) model entity
- 154) model type
- 155) modelling
- 156) molecule
- 157) multi-equation modelling
- 158) multi-modal characterization
- 159) multi-scale modelling
- 160) multi-technique characterization
- 161) nanoindentation
- 162) nanomaterial
- 163) nanoparticle tracking analysis
- 164) near-infrared spectroscopy
- 165) neutron diffraction
- 166) noise
- 167) nuclear magnetic resonance
- 168) numerical solver
- 169) off-line characterization
- 170) operand characterization
- 171) operation
- 172) optical microscopy
- 173) overaging
- 174) particle growth
- 175) photomicrograph
- 176) physical system state
- 177) physics equation
- 178) physics-based model

- 179) physics-based model types
 180) polarized laser excitation
 181) post-processing
 182) pre-processing
 183) precipitation hardening
 184) precipitation heat treatment
 185) primary data
 186) principle of combined action
 187) probe
 188) process annealing
 189) processed output
 190) processing reproducibility
 191) production batch
 192) quantity
 193) raman spectroscopy
 194) raw data
 195) raw output
 196) real-time material characterization
 197) recrystallisation
 198) representative volume element
 199) rve
 200) safe stress
 201) sample
 202) sample material properties
 203) sample preparation
 204) sample preprocessing
 205) sampling process
 206) scanning electron microscope
 207) scanning probe microscope
 208) secondary data
 209) signal
 210) simulated input
 211) simulated output
 212) simulation
 213) slip casting
 214) software tool
 215) soldering
 216) solution heat treatment
 217) solver
 218) solver parameters
 219) specimen
 220) spectra accumulation
 221) spectroscopic ellipsometry
 222) spectroscopy-based methods
 223) spheroidising
 224) stage
 225) stand-alone model
 226) static light scattering
 227) strain hardening
 228) supercooling
 229) superheating
 230) survey sampling
 231) system manufacturer
 232) system owner
 233) tempering
 234) test plan
 235) tga
 236) thermal analysis methods
 237) thermal tempering
 238) thermogravimetric analysis
 239) thermosetting
 240) tight coupling
 241) time lapse
 242) time step
 243) transmission electroscop
 244) tribological characterization
 245) ultrasound spectroscopy
 246) units
 247) user
 248) user case
 249) user experiment
 250) uv spectroscopy
 251) validation protocol
 252) visible spectroscopy
 253) weld decay
 254) welding
 255) workflow
 256) x-ray diffraction
 257) x-ray spectroscopy
 258) yielding

F. MATERIALS THEORIES

- 1) absorption
- 2) acceptor state
- 3) activation energy
- 4) anion
- 5) antiferromagnetic
- 6) band gap energy
- 7) bohr atomic model
- 8) bohr magneton
- 9) boltzmann constant
- 10) bonding energy
- 11) coercivity
- 12) color
- 13) concentration
- 14) condensation
- 15) conduction band
- 16) conductivity electrical
- 17) congruent transformation
- 18) continuous cooling transformation
- 19) coulomb force
- 20) curie temperature
- 21) diamagnetic
- 22) dielectric
- 23) dielectric breakdown strength
- 24) dielectric constant
- 25) dielectric displacement
- 26) donor state
- 27) doping
- 28) driving force
- 29) ductile-to-brittle transition
- 30) electric field

- 31) electroluminescence
- 32) electron configuration
- 33) electron energy band
- 34) electron hole
- 35) electron mobility
- 36) electron state
- 37) energy band gap
- 38) equilibrium phase
- 39) eutectic phase
- 40) eutectic reaction
- 41) eutectic structure
- 42) eutectoid reaction
- 43) excited state
- 44) fermi energy
- 45) ferrimagnetic
- 46) ferroelectric
- 47) ferromagnetic
- 48) fluorescent
- 49) free electron
- 50) free energy
- 51) gas constant
- 52) gibbs phase rule
- 53) glass transition temperature
- 54) ground state
- 55) hall effect
- 56) heat capacity
- 57) hole mobility
- 58) hydrogen bond
- 59) isothermal transformation diagram
- 60) liquidus line
- 61) luminescence
- 62) magnetic field strength
- 63) magnetic flux density
- 64) magnetic susceptibility
- 65) magnetization
- 66) matthiessen's rule
- 67) melting point
- 68) melting temperature
- 69) metallic bond
- 70) metastable
- 71) molarity
- 72) pauli exclusion principle
- 73) permeability
- 74) phase diagram
- 75) phase transformation
- 76) phonon
- 77) phosphorescent
- 78) photoconductivity
- 79) photon
- 80) planck's constant
- 81) quantum numbers
- 82) reflectivity
- 83) refraction
- 84) solidus line
- 85) solvus line
- 86) supercooling

- 87) superheating
- 88) terminal solid solution
- 89) tie line
- 90) time-temperature-transformation diagram
- 91) working point

REFERENCES

- [1] N. Guarino, "Formal ontology in information systems," in *Proc. FOIS*, N. Guarino, Ed. Trento, Italy: IOS Press, 1998, pp. 399–407.
- [2] N. Guarino, D. Oberle, and S. Staab, *What is an Ontology*, S. Staab and R. Studer, Eds. Berlin, Germany: Springer-Verlag, 2009.
- [3] J. F. Sowa, *Knowledge Representation: Logical, Philosophical, and Computational Foundations*. Pacific Grove, CA, USA: Brooks/Cole Publishing, 2000.
- [4] K. Munn and B. Smith, *Applied Ontology: An Introduction*. Frankfurt, Germany: Ontos, 2008.
- [5] C. Masolo, S. Borgo, A. Gangemi, N. Guarino, and A. Oltramari, "WonderWeb deliverable D18 ontology library (final)," WonderWeb, Ontology Infrastructure for the Semantic Web, Tech. Rep. IST Project 2001-33052, 2003.
- [6] S. Borgo and L. Lesmo, "The attractiveness of foundation ontologies in industries," *Frontiers Artif. Intell. Appl.*, vol. 174, pp. 1–9, Jan. 2015.
- [7] C. Trojahn, R. Vieira, D. Schmidt, A. Pease, and G. Guizzardi, "Foundational ontologies meet ontology matching: A survey," *Semantic Web*, vol. 13, no. 4, pp. 685–704, May 2022.
- [8] *Information Technology—Top-Level Ontologies (TLO)—Part 1: Requirements*, Standard ISO/IEC 21838-1:2021, 2021. [Online]. Available: <https://www.iso.org/obp/ui/en/#iso:std:iso-iec:21838-1:ed-1:v1:en>
- [9] J. Euzenat, "Towards a principled approach to semantic interoperability," in *Proc. Workshop Ontology Inf. Sharing*, 2001, pp. 19–25.
- [10] I. Osman, S. B. Yahia, and G. Diallo, "Ontology integration: Approaches and challenging issues," *Inf. Fusion*, vol. 71, pp. 38–63, Jul. 2021.
- [11] V. Presutti and A. Gangemi, "Content ontology design patterns as practical building blocks for web ontologies," in *Conceptual Modeling—ER 2008*, Q. Li, S. Spaccapietra, E. Yu, and A. Olivé, Eds. Berlin, Germany: Springer, 2008, pp. 128–141.
- [12] OntoCommons. (2023). *Ontology-Driven Data Documentation for Industry Commons, Grant Agreement N. 958371*. [Online]. Available: <https://cordis.europa.eu/project/id/958371>
- [13] R. Arp, B. Smith, and A. D. Spear, *Building Ontologies With Basic Formal Ontology*. Cambridge, MA, USA: MIT Press, 2015.
- [14] T. EMMC-Consortium. (2022). *The European Materials Modeling Ontology*. [Online]. Available: <https://emmc.info/emmo-info/>
- [15] M. Dumontier, C. J. Baker, J. Baran, A. Callahan, L. Chepelev, J. Cruz-Toledo, N. R. Del Rio, G. Duck, L. I. Furlong, N. Keath, D. Klassen, J. P. McCusker, N. Queralt-Rosinach, M. Samwald, N. Villanueva-Rosales, M. D. Wilkinson, and R. Hoehndorf, "The semantic science integrated ontology (SIO) for biomedical research and knowledge discovery," *J. Biomed. Semantics*, vol. 5, no. 1, p. 14, Dec. 2014.
- [16] I. Niles and A. Pease, "Towards a standard upper ontology," in *Proc. Int. Conf. Formal Ontol. Inf. Syst.*, Oct. 2001, pp. 2–9.
- [17] A. Atkin, "Peirce's theory of signs," in *The Stanford Encyclopedia of Philosophy*, E. N. Zalta, Ed. Stanford, CA, USA: Stanford Univ., 2013.
- [18] B.-M. Roh, S. R. T. Kumara, T. W. Simpson, P. Michaleris, P. Witherell, and I. Assourocko, "Ontology-based laser and thermal metamodelling for metal-based additive manufacturing," in *Proc. 36th Comput. Inf. Eng. Conf.*, Aug. 2016.
- [19] M. Dinar and D. W. Rosen, "A design for additive manufacturing ontology," *J. Comput. Inf. Sci. Eng.*, vol. 17, no. 2, pp. 1–8, Jun. 2017.
- [20] S. Kim, D. W. Rosen, P. Witherell, and H. Ko, "A design for additive manufacturing ontology to support manufacturability analysis," *J. Comput. Inf. Sci. Eng.*, vol. 19, no. 4, pp. 1–38, Dec. 2019.
- [21] P. Valluru, J. Karlapudi, K. Menzel, T. Mätäsniemi, and J. Shemeikka, "A semantic data model to represent building material data in AEC collaborative workflows," in *Boosting Collaborative Networks 4.0*. Cham, Switzerland: Springer, 2020, pp. 133–142.
- [22] International Organization for Standardization. *ISO 16739-1:2018 Industry Foundation Classes (IFC) for Data Sharing in the Construction and Facility Management Industries—Part 1: Data Schema*. Accessed: Sep. 9, 2023. [Online]. Available: <https://www.iso.org/standard/70303.html>

- [23] International Organization for Standardization. *ISO 19650-1:2018 Organization and Digitization of Information About Buildings and Civil Engineering Works, Including Building Information Modelling (BIM)—Information Management Using Building Information Modelling—Part 1: Concepts and Principles*. Accessed: Sep. 9, 2023. [Online]. Available: <https://www.iso.org/standard/68078.html>
- [24] O. Hakimi, J. L. Gelpi, M. Krallinger, F. Curi, D. Repchevsky, and M. Ginebra, “The devices, experimental scaffolds, and biomaterials ontology (DEB): A tool for mapping, annotation, and analysis of biomaterials data,” *Adv. Funct. Mater.*, vol. 30, no. 16, Apr. 2020, Art. no. 1909910.
- [25] K. Cheung, J. Hunter, and J. Drennan, “MatSeek: An ontology-based federated search interface for materials scientists,” *IEEE Intell. Syst.*, vol. 24, no. 1, pp. 47–56, Jan. 2009.
- [26] H. Li, R. Armiento, and P. Lambrix, “An ontology for the materials design domain,” in *The Semantic Web—ISWC 2020*. Cham, Switzerland: Springer, 2020, pp. 212–227.
- [27] M. Alam, F. Dittmann, M. Niebel, J. Lehmann, D. Dessi, J. F. Morgado, P. von Hartrott, C. Eberl, P. Gumbsch, and H. Sack, “Towards digitizing physical entities in materials science,” FIZ Karlsruhe, Leibniz Inst. Inf. Infrastruct., Germany, Tech. Rep., 2020.
- [28] M. Alam, H. Birkholz, D. Dessi, C. Eberl, H. Fliegl, P. Gumbsch, P. von Hartrott, L. Mädler, M. Niebel, and H. Sack, “Ontology modelling for materials science experiments,” in *Proc. Poster&Demo Track Workshop Ontology-Driven Conceptual Modelling Digit. Twins Located Semantics*, vol. 2941. Aachen, Germany: RWTH Aachen, 2021, p. 11.
- [29] Y. Chen. (2022). *MatoLab Brinell Test Ontology*. [Online]. Available: https://matportal.org/ontologies/MOL_BRINELL
- [30] M. Poveda-Villalón and S. Chávez-Feria. *Material Properties Ontology*. Accessed: Jun. 6, 2023. [Online]. Available: <http://bimerr.iot.linkeddata.es/def/material-properties>
- [31] M. Schilling, “FAIR data in PMD: Development of MSE mid-level and standard-compliant application ontologies,” in *Proc. World Congr. Integr. Comput. Mater. Eng. (ICME)*, Orlando, FL, USA, 2023.
- [32] P. Strömert, J. Hunold, A. Castro, S. Neumann, and O. Koepler, “Ontologies4Chem: The landscape of ontologies in chemistry,” *Pure Appl. Chem.*, vol. 94, no. 6, pp. 605–622, Jun. 2022.
- [33] J. Hastings, L. Chepelev, E. Willighagen, N. Adams, C. Steinbeck, and M. Dumontier, “The chemical information ontology: Provenance and disambiguation for chemical data on the biological semantic web,” *PLoS ONE*, vol. 6, no. 10, Oct. 2011, Art. no. e25513.
- [34] N. Jeliakova, C. Chomenidis, P. Doganis, B. Fadeel, R. Grafström, B. Hardy, J. Hastings, M. Hegi, V. Jeliakov, N. Kochev, P. Kohonen, C. R. Munteanu, H. Sarimveis, B. Smeets, P. Sopsakis, G. Tsiliki, D. Vorgrimmler, and E. Willighagen, “The eNanoMapper database for nanomaterial safety information,” *Beilstein J. Nanotechnol.*, vol. 6, pp. 1609–1634, Jul. 2015.
- [35] A. F. de Baas, “Mat-O-lab/AluTrace-data-and-documentation,” Fraunhofer, Germany, Tech. Rep., 2017.
- [36] C. Schweizer, M. Huschka, P. von Hartrott, and S. Stier. *Material Science and Engineering Ontology*. Accessed: Jun. 18, 2023. [Online]. Available: <https://github.com/Mat-O-Lab/MSEO>
- [37] B. Bayerlein, T. Hanke, T. Muth, J. Riedel, M. Schilling, C. Schweizer, B. Skrotzki, A. Todor, B. Moreno Torres, J. F. Unger, C. Völker, and J. Olbricht, “A perspective on digital knowledge representation in materials science and engineering,” *Adv. Eng. Mater.*, vol. 24, no. 6, pp. 1–14, Jun. 2022.
- [38] M. Schilling. *MOL_TENSILE Ontology*. Accessed: Sep. 9, 2023. [Online]. Available: https://matportal.org/ontologies/MOL_TENSILE/?p=summary
- [39] N. Baker. *Nano Particle Ontology*. Aug. 17, 2012. [Online]. Available: <https://bioportal.bioontology.org/ontologies/NPO>
- [40] N. Schneider, D. M. Lowe, R. A. Sayle, M. A. Tarselli, and G. A. Landrum, “Big data from pharmaceutical patents: A computational analysis of medicinal chemists’ bread and butter,” *J. Medicinal Chem.*, vol. 59, no. 9, pp. 4385–4402, 2016.
- [41] C. Pahl, N. Frank, J. Breitbart, and S. Bräse, “Overview of chemical ontologies,” 2020, *arXiv:2002.03842*.
- [42] S. Clark, F. L. Bleken, S. Stier, E. Flores, C. W. Andersen, M. Marcinek, A. Szczesna-Chrzan, M. Gaberscek, M. R. Palacin, M. Uhrin, and J. Friis, “Toward a unified description of battery data,” *Adv. Energy Mater.*, vol. 12, no. 17, May 2022, Art. no. 2102702.
- [43] S. Clark, C. W. Andersen, E. Flores, F. L. Bleken, and J. Friis, “(Digital presentation) a battery interface ontology for data interoperability and semantic knowledge representation,” *ECS Meeting Abstr.*, vol. 2022, no. 7, p. 2582, Oct. 2022.
- [44] P. D. Nostro, G. Goldbeck, and D. Toti, “CHAMEO: An ontology for the harmonisation of materials characterisation methodologies,” *Appl. Ontol.*, vol. 17, no. 3, pp. 401–421, Aug. 2022.
- [45] P. D. Nostro, G. Goldbeck, and D. Toti, “The CHAMEO ontology: Exploiting EMMO’s multiperspective versatility for capturing materials characterization procedures,” in *Proc. 12th Int. Workshop Formal Ontologies Meet Ind.*, vol. 3240, 2022, pp. 1–6.
- [46] *CIF Ontology*. Accessed: Sep. 9, 2023. [Online]. Available: <https://github.com/emmo-repo/CIF-ontology>
- [47] *EMMO Disciplines*. Accessed: Sep. 9, 2023. [Online]. Available: <https://github.com/emmo-repo/EMMO/tree/added-disciplines>
- [48] *Metal-Alloy*. Accessed: Sep. 9, 2023. [Online]. Available: <https://github.com/arsarkar/ontorepo/tree/master/metal-alloy>
- [49] *Microstructure Domain Ontology*. Accessed: Sep. 9, 2023. [Online]. Available: <https://github.com/emmo-repo/domain-microstructure>
- [50] *EMMO Mechanical Testing Ontology*. Accessed: Sep. 9, 2023. [Online]. Available: <https://github.com/emmo-repo/domain-mechanical-testing>
- [51] *OIE Ontologies*. Accessed: Sep. 9, 2023. [Online]. Available: <https://github.com/emmo-repo/OIE-Ontologies>
- [52] OYSTER. (2022). *Open Characterisation and Modelling Environment to Drive Innovation in Advanced Nano-Architected and Bio-Inspired Hard/Soft Interfaces, Grant Agreement N. 760827*. [Online]. Available: <https://cordis.europa.eu/project/id/760827>
- [53] NanoMECommons. (2025). *Harmonisation of EU-Wide Nanomechanics Protocols and Relevant Data Exchange Procedures, Across Representative Cases; Standardisation, Interoperability, Data Workflow, Grant Agreement N. 952869*. [Online]. Available: <https://cordis.europa.eu/project/id/952869>
- [54] P. Del Nostro, G. Goldbeck, A. Pozzi, and D. Toti, “Modeling experts, knowledge providers and expertise in materials modeling: MAEO as an application ontology of EMMO’s ecosystem,” *Appl. Ontol.*, vol. 18, no. 2, pp. 99–118, Aug. 2023.
- [55] *Phovoltaics Domain Ontology*. Accessed: Sep. 9, 2023. [Online]. Available: <https://github.com/emmo-repo/domain-photovoltaics>
- [56] *Precipitation Model*. Accessed: Sep. 9, 2023. [Online]. Available: <https://github.com/arsarkar/ontorepo/blob/master/metal-alloy/precipitation-model.owl>
- [57] M. T. Horsch, S. Chiacchiera, M. A. Seaton, I. T. Todorov, K. Šindelka, M. Lísal, B. Andreon, E. B. Kaiser, G. Moggi, G. Goldbeck, R. Kunze, G. Summer, A. Fiseni, H. Brüning, P. Schifffels, and W. L. Cavalcanti, “Ontologies for the virtual materials marketplace,” *Künstliche Intelligenz*, vol. 34, no. 3, pp. 423–428, Sep. 2020.
- [58] M. E. Deagen, J. P. McCusker, T. Fateye, S. Stouffer, L. C. Brinson, D. L. McGuinness, and L. S. Schadler, “FAIR and interactive data graphics from a scientific knowledge graph,” *Sci. Data*, vol. 9, no. 1, p. 239, May 2022.
- [59] H. Zhao, Y. Wang, A. Lin, B. Hu, R. Yan, J. McCusker, W. Chen, D. L. McGuinness, L. Schadler, and L. C. Brinson, “NanoMine schema: An extensible data representation for polymer nanocomposites,” *APL Mater.*, vol. 6, no. 11, Nov. 2018, Art. no. 111108.
- [60] P. Kügler, M. Marian, B. Schleich, S. Tremmel, and S. Wartzack, “TribAlIn—Towards an explicit specification of shared tribological understanding,” *Appl. Sci.*, vol. 10, no. 13, p. 4421, Jun. 2020.
- [61] P. Shvaiko and J. Euzenat, “Ontology matching: State of the art and future challenges,” *IEEE Trans. Knowl. Data Eng.*, vol. 25, no. 1, pp. 158–176, Jan. 2013.
- [62] P. Atzeni, F. Polticelli, and D. Toti, “A framework for semi-automatic identification, disambiguation and storage of protein-related abbreviations in scientific literature,” in *Proc. IEEE 27th Int. Conf. Data Eng. Workshops*, Apr. 2011, pp. 59–61.
- [63] P. Atzeni, F. Polticelli, and D. Toti, “An automatic identification and resolution system for protein-related abbreviations in scientific papers,” in *Proc. Eur. Conf. Evol. Comput., Mach. Learn. Data Mining Bioinf.*, in Lecture Notes in Computer Science, vol. 6623, 2011, pp. 171–176.
- [64] D. Toti, P. Atzeni, and F. Polticelli, “Automatic protein abbreviations discovery and resolution from full-text scientific papers: The PRAISED framework,” *Bio-Algorithms Med-Syst.*, vol. 8, no. 1, p. 13, 2012.

- [65] A. Poggi, M. Rodriguez, and M. Ruzzi, "Ontology-based database access with DIG-Mastro and the OBDA Plugin for Protégé," in *Proc. 4th Int. Workshop OWL, Exper. Direction*, vol. 496, 2008, pp. 1–4.
- [66] E. Kim, K. Huang, O. Kononova, G. Ceder, and E. Olivetti, "Distilling a materials synthesis ontology," *Matter*, vol. 1, no. 1, pp. 8–12, Jul. 2019.
- [67] S. Gražulis, A. Daškevič, A. Merkys, D. Chateigner, L. Lutterotti, M. Quirós, N. R. Serebryanaya, P. Moeck, R. T. Downs, and A. Le Bail, "Crystallography open database (COD): An open-access collection of crystal structures and platform for world-wide collaboration," *Nucleic Acids Res.*, vol. 40, no. D1, pp. D420–D427, Jan. 2012.
- [68] The Gene Ontology Consortium, "The gene ontology resource: 20 years and still going strong," *Nucleic Acids Res.*, vol. 47, no. D1, pp. D330–D338, 2018.
- [69] W. D. Callister and D. G. Rethwisch, *Callister's Materials Science and Engineering*, 10th ed. Hoboken, NJ, USA: Wiley, 2020.
- [70] CEN Workshop Agreement. (2018). CWA 17284. [Online]. Available: https://www.cencenelec.eu/media/CEN-CENELEC/CWAs/RI/cwa17284_2018.pdf
- [71] A. F. de Baas, "Review of materials modelling: 'What makes a material function? Let me compute the ways..', 6th version," Publications Office Eur. Union, Eur. Commission, Tech. Rep., 2017.
- [72] *The IUPAC Compendium of Chemical Terminology: The Gold Book*, International Union of Pure and Applied Chemistry (IUPAC), Research Triangle Park, NC, USA, 2019. [Online]. Available: <https://goldbook.iupac.org/>



ANNE DE BAAS received the joint Diploma degree in physics from the Technical University of Eindhoven, in 1978, and University of Utrecht, in 1982, and the Ph.D. degree in mechanical engineering from the University of Delft, in 1986. Her career encompasses research in geophysics, meteorology, and computational fluid dynamics done with the Risoe National Laboratory, Roskilde, Denmark, and ESN, Nantes, France.

In 1992, she joined the European Commission in the research and ICT departments, where she first set up benchmarking activities Europort I and II for the use of parallel computers for simulations in a wide scope of industrial applications. Secondly, she set up a European-wide network for technology transfer called FUSE, promoting the first use of electronics by industry. Later, she joined the Production and Materials Unit, where she set up the European Materials Modeling Council and wrote the review on materials modeling (RoMM, 7 editions), the start of the multi-perspective approach of the EMMO ontology. After retiring from EC, she started her own research company AnneFdeBaasMaterialsResearch, in 2022, and joined several EU projects via Goldbeck Consulting Ltd. Furthermore, she became a Strategy Advisor of EMMC.



PIERLUIGI DEL NOSTRO received the Graduate and Ph.D. degrees in computer engineering from Roma Tre University, in 2001 and 2008, respectively. His research interests include semantic web technologies, knowledge management, databases, and translations between heterogeneous data models, which led to publications in international conferences and journals. Additionally, he has more than 15 years of professional experience in IT, starting with software design and development

and later progressing to include the role of the project/product manager and a business developer. He has participated in several European and national projects in the fields of ontology modeling and software architecture. He is currently a Scientific Consultant with Goldbeck Consulting Ltd., where he is involved in data science ontology and software architecture aspects of projects.



JESPER FRIIS received the Master of Science degree in applied physics and electrical engineering from Linköping University, Sweden, in 1999, and the Ph.D. degree in physics and in transmission electron microscopy from NTNU, Norway, in 2003. He is currently a Senior Scientist with the Materials Physics Group, Trondheim. He did his postdoctoral research on microstructure modeling in aluminum before, he started in SINTEF, in 2007, where he has worked with multiscale modeling

from atomistic to microstructure level and digitalization. He has a long experience with EU projects and serves on the board of directors for the European Materials Modeling Council (EMMC). His research interests include ontologies, interoperability, multiscale modeling, atomistic modeling, microstructure modeling, thermodynamics, electron diffraction, data-driven materials science, aluminum, solar cells, batteries, and metal dusting.



EMANUELE GHEDINI has been a Full Professor with the University of Bologna, since 2022. His research interests include plasma transport theory and the development of physical mathematical models for industrial applications of plasmas, calculation of transport properties of plasma mixtures, modeling of plasma sources devices, multi-scale modeling nanomaterial synthesis, and development of formal ontologies for plasma applications in the field of materials and modeling.

He has participated in several national and international research projects. From 2015 to 2018, he was the Project Coordinator of H2020-NMP-20-2014 NanoDome, Nanomaterials via Gas-Phase Synthesis: A Design-Oriented Modelling and Engineering Approach (three years, 4M). Since January 2019, he has been the Project Coordinator of H2020-DT-NMBP-09-2018 SimDOME, Digital Ontology-Based Modeling Environment for Simulation of Materials (four years, 4.6M). He is the author of more than 70 articles in international journals. He is a member of the European Materials Modeling Council and the main Developer of the European Materials Modelling Ontology (EMMO).



GERHARD GOLDBECK received the Diploma degree in physics from RWTH Aachen University, in 1986, and the Ph.D. degree in polymer physics from Bristol University, in 1992. His career encompasses research in solid state physics and polymer materials with Forschungszentrum Jülich, H. H. Wills Physics Laboratory, Bristol, and Department of Materials Science and Metallurgy, Cambridge, by means of a range of modeling and characterization techniques, software development, product management, and marketing at Molecular Simulations/Accelrys/Biovia. During that time, he set up and ran the nanotechnology consortium which fostered close collaboration and technology transfer between academia, software, and materials industries. In addition to his role as the Managing Director of Goldbeck Consulting, he acts as the Executive Secretary of the European Materials Modelling Council and a member of the Operational Management Board of the European Materials Characterization Council.

ware development, product management, and marketing at Molecular Simulations/Accelrys/Biovia. During that time, he set up and ran the nanotechnology consortium which fostered close collaboration and technology transfer between academia, software, and materials industries. In addition to his role as the Managing Director of Goldbeck Consulting, he acts as the Executive Secretary of the European Materials Modelling Council and a member of the Operational Management Board of the European Materials Characterization Council.



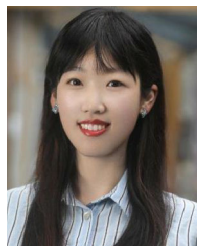
ILARIA MARIA PAPONETTI received the master's degree in energetical and nuclear engineering, in 2019. She is currently pursuing the Ph.D. degree with the Department of Industrial Engineering, Alma Mater Studiorum, Bologna. Her master's thesis in mechanical deformation in IV generation nuclear reactor. Her research interests include the domain of materials science, energy, and industry, having first pursued a path in the multinational automation industry and then started her research activities first in ENEA, then in Unibo. Her Ph.D. program concerns the development of formal knowledge frameworks at the domain level, with the aim of increasing interoperability among systems and connecting available data in the specific field.



ANDREA POZZI (Member, IEEE) received the bachelor's degree in industrial engineering, the master's degree in electrical engineering, and the Ph.D. degree in electronics, informatics, and electrical engineering from the University of Pavia, in 2015, 2017, and 2021, respectively. He has held visiting scholar positions with TU Braunschweig, in 2016, and UC Berkeley, in 2019. After working as a Postdoctoral Researcher with the University of Pavia, he joined the Faculty of Mathematical, Physical and Natural Sciences, Catholic University of the Sacred Heart, Brescia, Italy, as an Assistant Professor in machine learning, in January 2022. His research interests encompass reinforcement learning, imitation learning, machine learning, approximate dynamic programming, and advanced control theory.



ARKOPAUL SARKAR received the Ph.D. degree in mechanical and systems engineering from the Russ College of Engineering, Ohio University, with active research in semantic knowledge modeling for distributed resource and process planning in decentralized manufacturing. He is currently an architect, a designer, and a developer of large-scale business and industrial software solutions in the domain of enterprise resource planning, product design, production planning, and scheduling. He is currently a Postdoctoral Researcher with Ecole Nationale d'Ingénieurs de Tarbes. He was a business consultant for large organizations for over six years in the past and developed custom solutions based on SAP products in sales and distribution, product life-cycle, customer relationships, and supply chain management. He regularly publishes observations and findings from master's and Ph.D. research in peer-reviewed journals, conference proceedings, and book chapters. He participates in both paid and pro bono projects, funded by the university, government organizations, industries, and venture capitalists. He launched two start-ups in the mobile application and entertainment industry. He is a regular speaker at engineering conferences and a member of academic organizations.



LAN YANG received the B.E., M.E., and M.Eng.Sc. degrees and the joint Ph.D. degree from the University of Galway, in September 2020, funded by Hardiman Scholarship, co-supervised by Tsinghua University. She is currently a Postdoctoral Researcher with the Unit for Social Semantics at Insight, University of Galway. Before joining Insight, she delivered lectures with the College of Science and Engineering and Centre for Adult Learning and Professional Development and conducted research with the Enterprise Research Centre, University of Galway. Her research interests include formal ontologies, knowledge graphs, semantic web, big data, healthcare analytics, medical informatics. She has constantly been cooperating with reputed universities and organizations, such as Tsinghua University, the University of Oxford, the Norwegian University of Science and Technology, and the International Council on Systems Engineering and publishing papers in leading journals. Her research achievements have been recognized through plenary presentations at international conferences and symposia. She is a reviewer of many peer-reviewed journals in her field and a scientific committee member of many international conferences. Her current research projects at Insight include OntoCommons—Ontology-Driven Data Documentation for Industry Commons, funded by the European Union's Horizon 2020 Research and Innovation Program (2021–2023) and the DXA (Dual-Energy X-Ray Absorptiometry) Management Application Project (MAP), funded by the HRB Awards Secondary Data Analysis Projects (2021–2023).



FRANCESCO ANTONIO ZACCARINI received the B.A. degree in philosophy and the M.A. degree in philosophical science from the University of Bologna, in 2016 and 2019, respectively. He is currently a Research Fellow with the University of Bologna. His research interests include the foundations of analytic metaphysics, the formal constraints of structures, and conceptual engineering. He is also involved in the development of the elementary multiperspective material ontology (EMMO).



DANIELE TOTI (Member, IEEE) received the bachelor's and master's degrees in computer engineering and the Ph.D. degree in computer science and engineering from Roma Tre University, in 2005, 2008, and 2012, respectively. His key research interests include information extraction, natural language processing, and semantic knowledge discovery and modeling. Recently, he has been involved in ontology building, modeling, and alignment for materials science within the context of the elementary multiperspective material ontology (EMMO). His career spans both worlds of academia and industry, with appointments in consultancy companies, research centers, universities, and software houses. He also acted as a scientific consultant of a number of international IT companies and research organizations. Since 2020, he has been a member of the Faculty of Mathematical, Physical and Natural Sciences, Catholic University of the Sacred Heart, Brescia, Italy, where he is currently an Associate Professor in computer engineering. He participated in several EU-funded international and national projects, including: ARISTOTELE and LASIE (FP7); eJRM, MODERN, SIRET, and Wheesbee (PON); and SURVANT, DANTE, VIMMP, MarketPlace, OntoTrans, OntoCommons, OYSTER, OpenModel, and NanoMECommons (H2020). He is an oracle and sun certified professional with seven industry-level certifications.

...