

The Intersection Between Semantic Web and Materials Science

Andre Valdestilhas,* Bernd Bayerlein,* Benjamin Moreno Torres, Ghezal Ahmad Jan Zia, and Thilo Muth

The application and benefits of Semantic Web Technologies (SWT) for managing, sharing, and (re-)using of research data are demonstrated in implementations in the field of Materials Science and Engineering (MSE). However, a compilation and classification are needed to fully recognize the scattered published works with its unique added values. Here, the primary use of SWT at the interface with MSE is identified using specifically created categories. This overview highlights promising opportunities for the application of SWT to MSE, such as enhancing the quality of experimental processes, enriching data with contextual information in knowledge graphs, or using ontologies to perform specific queries on semantically structured data. While interdisciplinary work between the two fields is still in its early stages, a great need is identified to facilitate access for non-experts and develop and provide user-friendly tools and workflows. The full potential of SWT can best be achieved in the long term by the broad acceptance and active participation of the MSE community. In perspective, these technological solutions will advance the field of MSE by making data FAIR. Data-driven approaches will benefit from these data structures and their connections to catalyze knowledge generation in MSE.


1. Main

Innovative research often relies on advanced technological solutions to exploit (data) potential. However, the diversity of scientific data formats and structures leads to compatibility issues and slows down progress. Data-driven science benefits from consistent data generation, organization, storage, and sharing. In this context, systems need to become interoperable to enable better automation in data access and analysis. Semantic Web Technologies (SWT) are able to efficiently address these requirements.

A. Valdestilhas, B. Bayerlein, B. Moreno Torres, G. A. J. Zia, T. Muth
BAM

Federal Institute for Materials Research and Testing
12205 Berlin, Germany

E-mail: andre.valdestilhas@bam.de; bernd.bayerlein@bam.de

 The ORCID identification number(s) for the author(s) of this article can be found under <https://doi.org/10.1002/aisy.202300051>.

© 2023 The Authors. Advanced Intelligent Systems published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

DOI: 10.1002/aisy.202300051

Over the last decades, SWT have been widely applied with great success in various disciplines, especially in life sciences. Prominent resources are the Gene Ontology (GO),^[1] a knowledge base containing the world's largest source of information on genes, the Disease Ontology^[2] that integrates disease and medical vocabularies, and the recently created knowledge graph COVIDPUBGRAPH^[3] that provides a comprehensive and researcher-oriented data source for COVID-19 publications.

The extension and adaption of these technologies to other fields hold big potential, for example, in the possibility of uniformly considering, integrating, and analyzing heterogeneous research data that originate from a wide variety of sources and come in a diverse range of formats. Their application helps to improve the traceability and reproducibility of scientific procedures, experiments, and simulations.

Related technologies can also be used for

automatable data management solutions and the answering scientific questions by using semantic queries on given data. In addition, data-driven approaches such as machine learning methods benefit from increased data completeness and coherence.

Historically, there have been first points of contact between these technologies and the field of Materials Science and Engineering (MSE) even before the concept for the Semantic Web (SW) was formulated by Tim Berners-Lee et al.^[4] in 2001, for example, with the creation of the Plinius ontology^[5] for ceramic materials in 1994. It was language independent and offered a conceptual construction kit for the chemical composition of materials. Most important SW standards, such as formats, languages, or applications, however, have only been made available later (Figure 1a). For example, the Web Ontology Language (OWL) extended existing vocabularies and thereby enabled enhanced machine interoperability. Based on OWL, important ontological resources, such as the DOLCE^[6] or the BFO,^[7] could then be developed (Figure 1b). 14 years after the creation of Plinius ontology, ontologies and technical approaches relevant to materials and processes emerged, such as MatOnto^[8] for materials data management tasks, ONTORULE^[9] for business applications in the steel industry, or Materials Ontology by Ashino^[10] for integrating thermal property databases and related resources. However, it was not until launching the

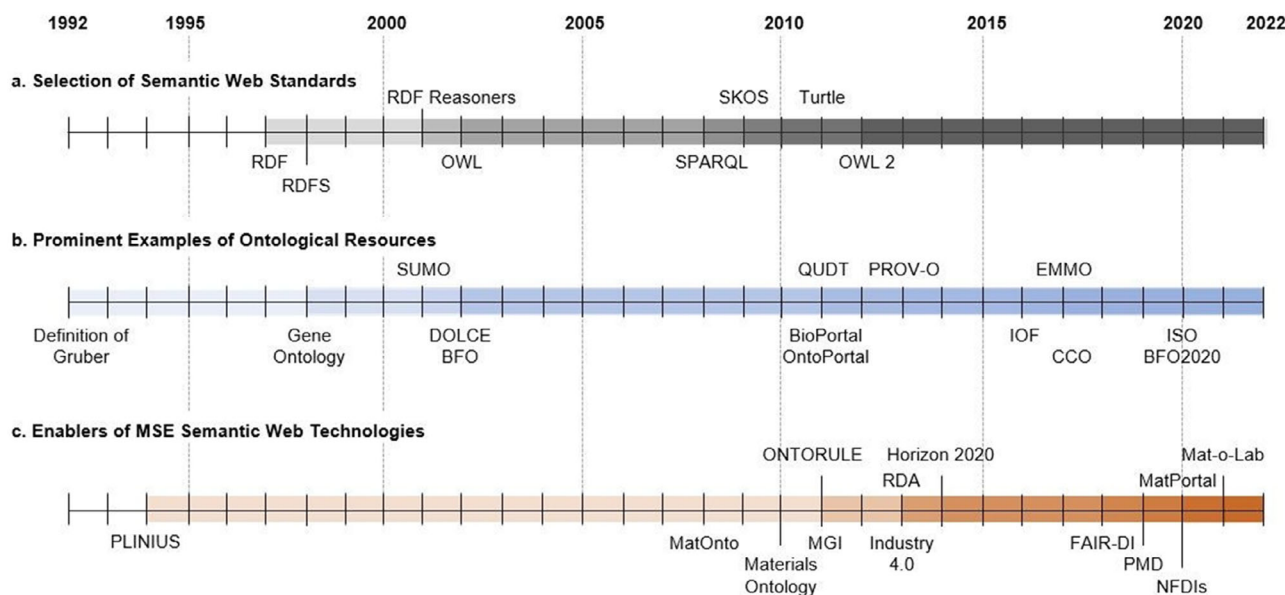


Figure 1. Temporal connection between SW standards, ontological resources, and the advent of SWT in MSE. a) Evolution of important SW standards provided by the W3C.^[49] b) Chronology of ontology development with important milestones: The agreement on Gruber's definition,^[18] a global successful applied example in the life sciences (GO^[1]), the appearance of upper-level ontologies (SUMO,^[61] DOLCE,^[6] BFO,^[7] PROV-O^[43]), the creation of the QUDT,^[47] a rich ontology for quantities, units, and dimensions, or the availability of ontology repositories (BioPortal,^[62] OntoPortal^[63]). Advanced mid-level ontologies with worldwide scope have also been included (IOF,^[64] CCO,^[65] EMMO^[45]), as well as the conversion of BFO^[7] not as a standard de facto but de iure as an ISO document.^[66] c) Increasing interest in SWT implementation in MSE with the MGI^[11] as an enabler. In addition to the globally active Research Data Alliance (RDA),^[67] which aims at open data exchange, initiatives from the European region are shown here. This is to illustrate how development in a limited geographic scope in particular is increasing rapidly. The section "Explanations of technical terms" contains terms not described here, e.g., RDF, OWL, or SPARQL, which are also used in the remaining text.

Materials Genome Initiative (MGI) in 2011,^[11] which aims to shorten typical development and innovation cycles, that SWT became relevant and found higher acceptance in the broader MSE domain (Figure 1c). Inspired by this new paradigm of a data-driven research approach, more initiatives for open science and collaborative material data spaces were funded.^[12]

The European framework provides such initiatives (We exclude initiatives not related to Semantic Web and Materials Science) that are involved mainly with Horizon 2020 (<https://www.horizont2020.de/>), a research funding program that enables the building of a knowledge- and innovation-based society and a competitive economy, while contributing to sustainable development in Europe. Concerning digital infrastructure building, the FAIR-DI (<https://www.fair-di.eu/fair-di/>)^[13,14] association also promotes an infrastructure for data from materials science, engineering, and astronomy with its European member institutes. In addition, GAIA-X (<https://www.data-infrastructure.eu/GAIA-X/Navigation/EN/Home/home.html/>) aims to build a competitive, secure, and trustworthy data infrastructure for Europe. Another example at the national level is the Platform Industry 4.0 (<https://www.bmbf.de/bmbf/de/forschung/digitale-wirtschaft-und-gesellschaft/industrie-4-0/industrie-4-0/>), a German government project with the mission to better connect machines and processes in the manufacturing industry using digital technologies. Further German initiatives such as the Platform MaterialDigital (<https://www.materialdigital.de/>) have been established to exploit SWT and especially ontologies in the MSE domain. MatPortal (<https://matportal.org/>), a repository

for MSE ontologies, was also recently set up. Beyond this, various consortia of the German national research data infrastructure (NFDI) (<https://www.nfdi.de/>) have committed to the goal of systematically providing access to valuable data from science, which is also being attempted with a close connection to the European Open Science Cloud (EOSC) (<https://digital-strategy.ec.europa.eu/de/policies/open-science-cloud/>). All these research- and industry-driven efforts aim to trigger increased community use and harmonizing appropriate digital tools and accepted standards.

So far, no detailed study has yet addressed the implementation of SWT in the MSE in a classified and structured way, detailing both the benefits and the challenges at this intersection. In this regard, it is necessary to make use of these technologies comprehensible and explicit to a larger community, as SWT has the potential to change the conventional MSE landscape through appropriate extrapolation. Therefore, it is not surprising that the literature describes the benefits, implicitly calling for the need for an overview of the state of the art. For example, the potential of using ontologies in materials science was reported before.^[15] An ontology that maps a standardized characterization method and is shared publicly could be used by anyone for organizing resulting data and metadata.^[12] Moreover, the next step here could be for manufacturers of testing machines to prescribe this data structure already as an output format. The organized formal knowledge in ontologies could also be used as a teaching tool to provide a basic understanding of the characterization method itself, the structure of the data and metadata, as well

as the application of important technological tools already in education.^[16]

Our article aims to provide stakeholders from both fields with an overview of existing approaches and implementations at this intersection. Relevant works are identified (Section 2) and summarized according to created categories (Section 3) providing insights based on current examples. The reader is guided through a distillation of challenges and opportunities that arise from approaching these overlapping areas (Section 4).

1.1. Explanations of Technical Terms

API: Computer programs can communicate with each other via an Application Programming Interface (<https://en.wikipedia.org/wiki/API>).

FAIR Data: The application of the FAIR principles^[17] (<https://www.go-fair.org/fair-principles/>) is aimed at increasing the Findability, Accessibility, Interoperability, and Reusability of (research) data.

IRI: Internationalized Resource Identifier (<https://www.rfc-editor.org/rfc/rfc3987>) can be used instead of a URI to uniquely identify resources.

MSE: Quoting “Materials Science and Engineering combine engineering, physics, and chemistry principles to solve real-world problems associated with nanotechnology, biotechnology, information technology, energy, manufacturing, and other major engineering disciplines” (Quoting from <https://mse.umd.edu/about/what-is-mse/>).

Ontology: Gruber defines “An ontology is an explicit specification of a conceptualization”.^[18] Ontology creation involves the description of knowledge through specified terms (formalized vocabularies) and their relationships to other terms, usually covering a specific domain represented by a community of users. The semantic web languages OWL (<https://www.w3.org/OWL/>) (Web Ontology Language) and its revised extension OWL2 (<https://www.w3.org/TR/2012/REC-owl2-overview-20121211/>) enable the creation and sharing of ontologies over the Web aiming at making Web content more accessible to machines.

Provenance: Traceability of chronology of MSE experiments and simulations according to the concept described in the works.^[19–22]

RDF: The Resource Description Framework (<https://www.w3.org/RDF/>) is a standard model for data interchange on the Web. In the RDF model, each statement consists of three components, subject, predicate, and object, and is represented by URIs. Statements are made in the form of RDF triples (<https://www.w3.org/TR/rdf-concepts/#section-triples/>), such as,

```
http://example.org/#specimen
<http://www.perceive.net/schemas/relationship/isMadeOf>
<http://example.org/#material>
```

An RDF graph (<https://www.w3.org/TR/rdf-concepts/#section-rdf-graph/>) is defined by a set of RDF triples, which can be stored in RDF Triple Stores (https://www.w3.org/2001/sw/wiki/Category:Triple_Store). Serialization of RDF graphs is enabled by syntaxes such as Turtle (<https://www.w3.org/TR/turtle/>). The query language SPARQL (<https://www.w3.org/TR/rdf-sparql-query/>) allows information retrieval from RDF

graphs based on Turtle syntax. RDF itself does not provide any vocabulary related to schema. RDFS (<https://www.w3.org/2001/sw/wiki/RDFS>) (RDF Schema) can be used to represent simple RDF vocabularies on the Web. There are also other vocabulary definition technologies, such as OWL, OWL2, or SKOS (<https://www.w3.org/TR/skos-reference/>), providing language for defining structured, Web-based ontologies. In this regard, a reasoner is software that can draw logical conclusions from a collection of axioms or claimed facts on ontologies.

SWT: Refers to Semantic Web Technologies that apply linked data and RDF concepts from^[4] aiming to provide RDF graphs from MSE knowledge data.

URI: Universal Resource Identifier (<https://www.w3.org/wiki/URI/>) is used to identify resources on the SW.

W3C: The World Wide Web Consortium (<https://www.w3.org/>) provides standards for the World Wide Web.

2. Literature Search Strategy and Selection Process

For the literature research, we used a specific search strategy to identify recent and relevant publications applying SWT in MSE approaches. This literature search strategy and the arising research questions are described in this section.

2.1. Selection Methodology

The merits of our literature search strategy are to provide insights into the existing research relevant to MSE approaches using SWT and to present this knowledge compactly in a written report. The literature analysis will help in understanding the benefits provided by SWT. The search strategy used in this work (**Figure 2**) follows a formal systematic literature review process. In particular, this study builds on the guidelines proposed in refs. [23–26]. Other surveys, for example,^[15,27] of relevant journals as well as on related topics, such as MSE and SWT, are also considered.

2.2. Research Questions

This work aims to provide an overview of the application of SWT in different areas of MSE, highlighting challenges and opportunities. To achieve this goal, we aimed to answer the following general research question:

Which benefits do SWT provide to the MSE domain?

This question gives rise to the following four specific research questions (RQ).RQ1. What are state-of-the-art approaches in MSE using SWT?

RQ2. Which SWTs are applied in MSE?

RQ3. How can SWT influence the completeness and coherence of materials data obtained from an MSE experiment? RQ4. What are the open challenges at the intersection between MSE and SWT?

RQ1 intends to collect available research works to retrieve knowledge from SWT resources for MSE experiments and simulations. RQ2 aims to provide an explicit comparison among

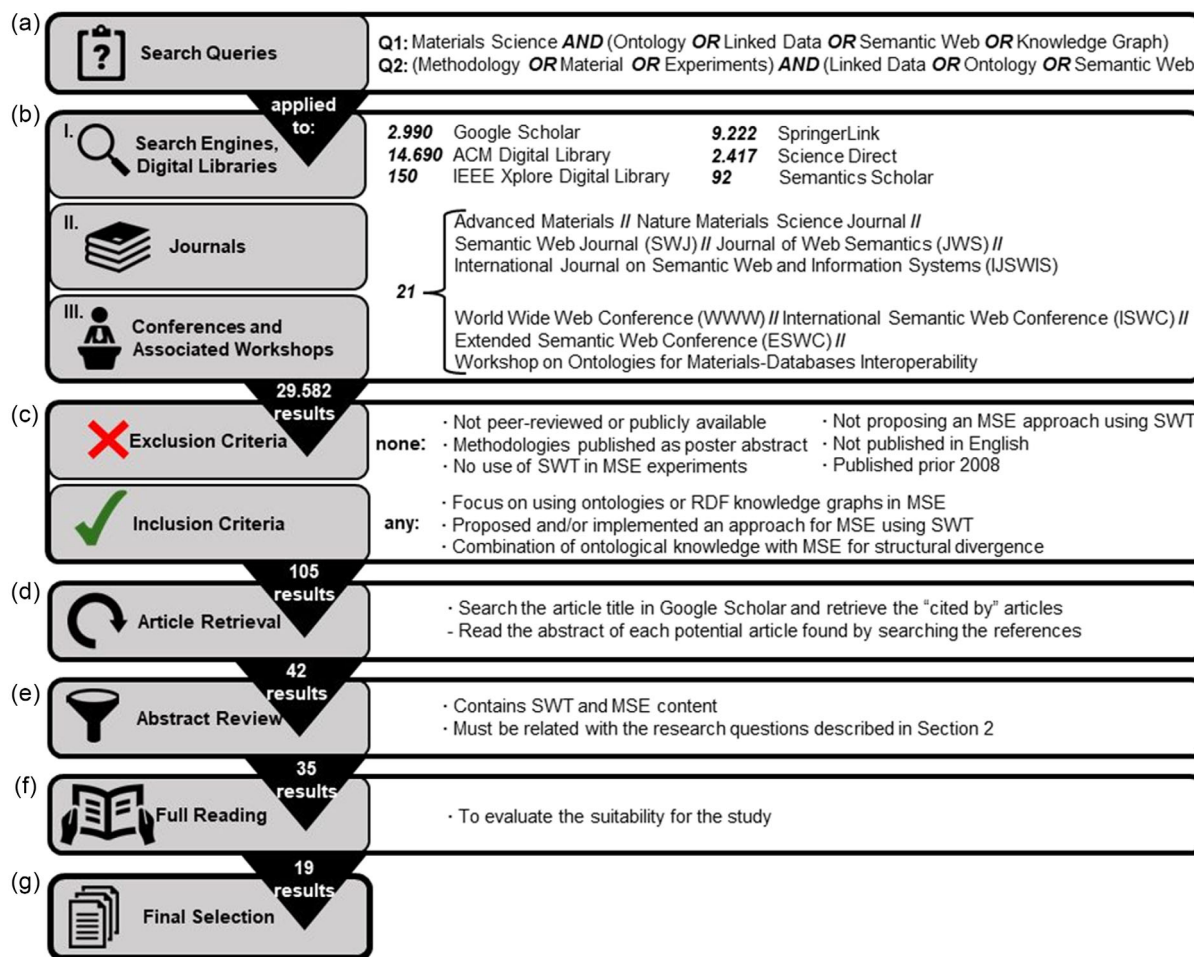


Figure 2. Selection Methodology. a) Defined a set of important keywords to find relevant SWT and MSE publications meeting the research questions and criteria. Q1 and Q2 both need to be fulfilled. The combination of two keywords, namely, “materials science” and “methodology,” fulfilled the inclusion criteria (The selection method is a manual human activity, and it was assumed that the search engines automatically consider the inflections and synonyms of the keywords, not considering duplication of entries). b) Overview of search engines and digital libraries (Google Scholar (<http://scholar.google.com/>), ACM Digital Library (<http://dl.acm.org/>), IEEE Xplore Digital Library (<http://ieeexplore.ieee.org/>), SpringerLink (<http://link.springer.com/>), ScienceDirect (<http://www.sciencedirect.com/>), journals (Advanced Materials (<https://onlinelibrary.wiley.com/journal/15214095/>), Nature—Materials Science Journal (<https://www.nature.com/subjects/materials-science/srep/>), Semantic Web Journal (SWJ) (<http://www.semantic-web-journal.net/>), Journal of Web Semantics (JWS) (<http://www.websemanticsjournal.org/>), International Journal on Semantic Web and Information Systems (IJSWIS) (<http://www.ijswis.org/>), conferences, and associated workshops (World Wide Web Conference (WWW),^[68–84] International Semantic Web Conference (ISWC),^[85–100] Extended Semantic Web Conference (ESWC),^[101–114] Workshop on Ontologies for Materials-Databases Interoperability^[115,116] used to identify relevant papers. c) The selection criteria (The use of SWT in MSE includes the development of new methodologies, not limited to only experiments) is a preselection step involving exclusion and inclusion criteria, which were defined based on best practices.^[23–25] In the Exclusion Criteria, the word “none” means that an article will be excluded if do not satisfy one of the exclusion criteria items. Similarly, this definition applies to the word “any”. d) The abstracts of the 105 articles retrieved in step c) were subjected to review based on the four research questions outlined in Section 2.2. Additional articles were obtained by following steps e) and f): e) Article titles were searched in Google Scholar to retrieve the corresponding “Cited By” articles, and f) the abstracts of the 42 obtained articles were analyzed to identify potential references. Subsequently, a full-text analysis was conducted, which led to the identification of 35 articles that were re-evaluated for relevance to the present study. g) A meticulous evaluation led to the selection of 19 articles published between 2008 and 2022, listed in Table 1.

SWT used in each respective MSE approach or system. RQ3 determines whether SWT improves the experimental process of a given MSE system. RQ4 asks for a descriptive list of MSE approaches using SWT to identify the open challenges in those works. To address the four research questions, a total of 19 papers were selected which are listed in Table 1.

3. Semantic Web Applied to Materials Science and Engineering

In the following five sections, the selected works are structured and discussed according to the SW categories introduced in Table 1 on Areas of Intersection. The intention is to provide

Table 1. Selected Publications with Authors, Year, and Title. The matrix illustrates the Areas of Intersection by the following categories: The first includes works that deal with Ontology Creation (e.g., for the representation of MSE methods and knowledge). RDF Ontology Application includes approaches that use ontologies for specific tasks. RDF Instance Creation involves the conversion of data into RDF triples. RDF Information Retrieval includes approaches for querying triples via APIs with SPARQL. Provenance covers the collection of metadata in terms of full traceability and reproducibility of generated MSE experiment and simulation data according to FAIR data principles. Ontology Reasoning involves automatable inference and reasoning, in this way obtaining logical conclusions from a collection of axioms or claimed facts on ontologies.

Ref.	Authors	Year	Title	Ontology Creation	RDF Ontology Application	RDF Instance Creation	RDF Information Retrieval	Provenance	Ontology Reasoning
[28]	Deagen et al.	2022	FAIR and interactive data graphics from a scientific knowledge graph	✓	✓	✓	✓	✓	
[29]	Zhao and Qian	2017	Ontology-based heterogeneous materials database integration and semantic query	✓	✓	✓	✓		
[30]	Zhao and Wang et al.	2018	NanoMine schema: An extensible data representation for polymer nanocomposites	✓	✓	✓	✓	✓	✓
[31]	Hakimi et al.	2020	A tool for mapping, annotation, and analysis of biomaterials data	✓					
[32]	Vardeman et al.	2017	An ontology design pattern and its use case for modeling material transformation	✓	✓	✓		✓	✓
[33]	Li et al.	2019	A method for extending ontologies with application to the materials science domain	✓	✓	✓		✓	
[34]	Nikooie et al.	2021	A first step towards a tool for extending ontologies	✓	✓			✓	
[35]	Zhang et al.	2016	MMOY: Towards deriving a metallic materials ontology from Yago	✓	✓	✓	✓		✓
[36]	Moreno et al.	2021	An ontology-based approach to enable data-driven research in the field of NDT in civil engineering	✓				✓	
[37]	Bayerlein et al.	2022	A perspective on digital knowledge representation in materials science and engineering	✓	✓	✓		✓	
[38]	Garabedian et al.	2022	Generating FAIR research data in experimental tribology	✓	✓	✓	✓	✓	✓
[39]	Romanos et al.	2019	Innovative data management in advanced characterization: Implications for materials design	✓				✓	
[8]	Cheung et al.	2008	Towards an ontology for data-driven discovery of new materials	✓					
[40]	Li et al.	2020	An ontology for the materials design domain	✓	✓	✓	✓	✓	✓
[10]	Ashino	2010	Materials ontology: An infrastructure for exchanging materials information and knowledge	✓					
[50]	An et al.	2021	OTMapOnto: Optimal transport-based ontology matching		✓				
[51]	Greenberg et al.	2020	HIVE-4-MAT: Advancing the ontology infrastructure for materials science		✓	✓	✓		
[41]	Sadigh et al.	2017	An ontology-based semantic machine tool selection for multi-scale wire EDM processes	✓	✓	✓	✓	✓	✓
[52]	Lambrix et al.	2018	Big semantic data processing in the materials design domain		✓				✓

the reader with a detailed overview of the current status and the benefits of using SWT in MSE.

3.1. Ontology Creation

Ontologies define structured and standardized dictionaries of concepts formally. These ontologies are applied to describe machine-interpretable information within an SW framework, map concepts between heterogeneous data resources, enable

advanced querying of (meta-)data, and retrieving new knowledge. In 16 of the selected papers, ontologies are created (Figure 3a) for specific MSE tasks.^[8,10,28–41] Ontologies are formulated in OWL to express knowledge in a structured, flexible, and extensible way.

Here we give an overview of how different the various techniques and approaches to ontology creation are in the selected works. Zhao and Qian^[29] proposed a semantic integration method that extracts a material database schema semiautomatically,

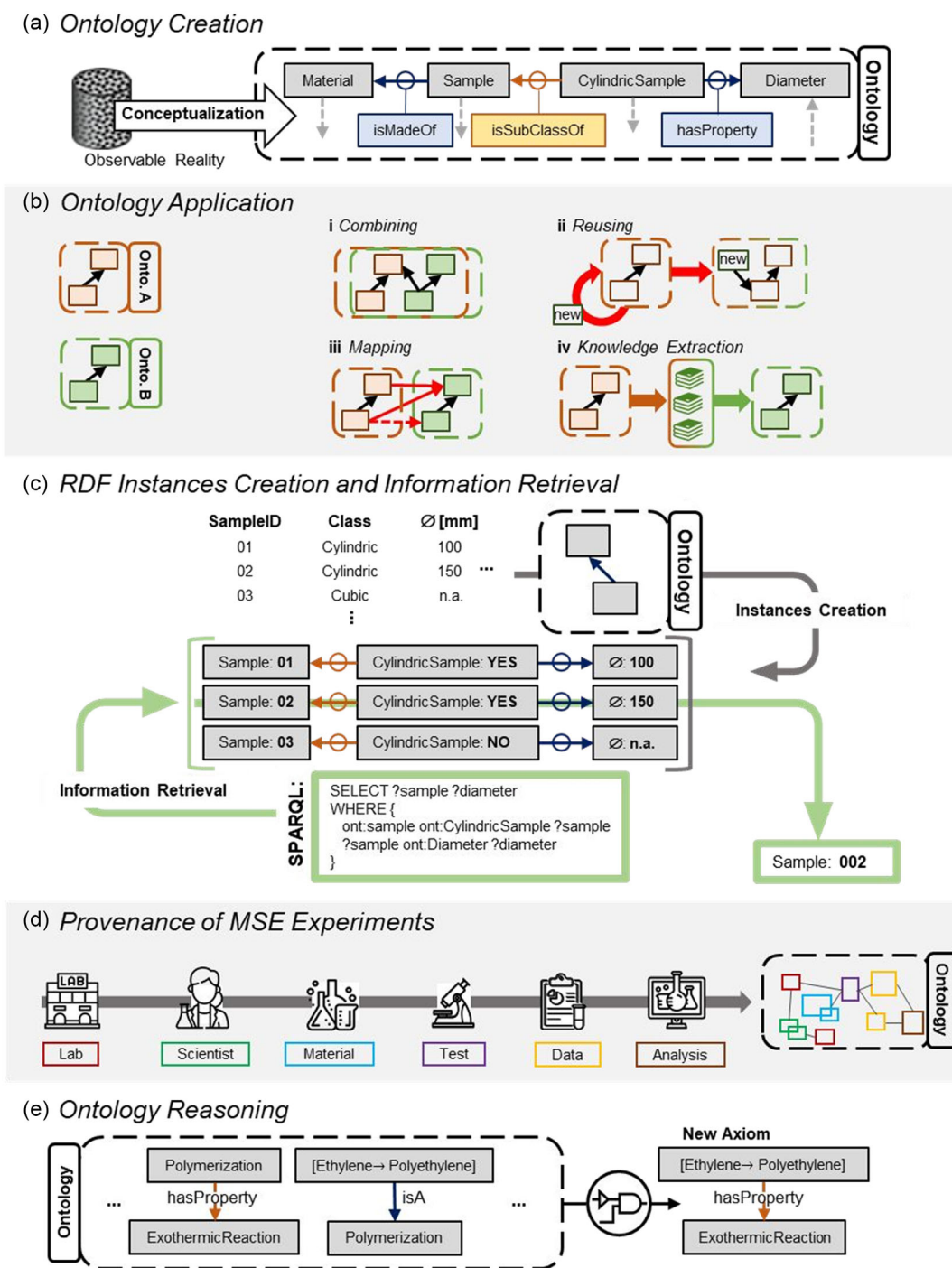


Figure 3. (preceding page). Illustration of the introduced SW categories. a) *Ontology Creation*. The process encodes a particular subset of the reality into an ontology, considering parameters like initial requirements, range of applicability, or methodology used. b) *Ontology Application*. The type of applications of ontologies in MSE is far from being fully explored. However, here we present the most common in our selection: i) *Combining*. This is the creation of a new ontology from a previous set of ontologies to be used together. ii) *Reusing*. An ontology's maintenance also includes reviewing its components to be adapted to new requirements. iii) *Mapping*. When the same reality is described independently with different approaches, it is helpful to map entities to extend their use. iv) *Knowledge Extraction*. This describes the process of using ontology as a framework for terms and relationships extraction. c) *RDF Instances Creation and Information Retrieval*. The process of creating RDF instances from tabular data using an ontology is shown. Once the data is in a triple-store format, the query language SPARQL makes information retrieval in federated data systems possible. d) *Provenance of MSE Experiments*. A structured way to encode the many steps of an experiment is to use an ontology that allows many degrees of modularization. e) *Ontology Reasoning*. The use of reasoners is possible with OWL. Different reasoners use different logic paradigms that allow different degrees of deduction.

providing an effective way to merge heterogeneous data from different sources. They used this approach to build the ontology of the material comprising five classes (material, structure, properties, processing, and performance) and subclasses. Zhao and Wang et al.^[30] provided in their work the design of an ontology following the NanoMine XML data schema and its underlying principles. This ontology represents the processing–structure–property knowledge about polymer nanocomposites with over 350 parameters. Here, a more general material data vocabulary is integrated, targeting interoperability and a wider field of application. Hakimi et al.^[31] applied a text analysis approach to a gold standard biomaterials literature set to identify the key terms for the creation of the Devices, Experimental scaffolds, and Biomaterials (DEB) ontology. The DEB ontology, which represents the field’s terminology (and rules), allows unique naming, classifying, and organizing of manufactured biomaterials such as implants and medical devices.

Another way of creating ontologies is to extend and re-use existing semantic artifacts, that is, build on already available vocabularies. Vardeman et al.^[32] presented an ontology design pattern that facilitates ontological reuse and modeling of material transformations. This allows additional capturing relationships between raw materials, intermediate components, and final products. Li et al.^[33] used a phrase-based topic model approach and a formal topical concept analysis on 600 abstracts to find additional concepts and axioms for extending and improving NanoParticle and eNanoMapper, two nanotechnology ontologies. Nikooie et al.^[34] presented the tool-based extension of another ontology. The user is guided step by step through an iterative development process. By extending the Materials Design Ontology (MDO) (by 29 concepts and 27 additional axioms), the effectiveness of this phrase-based approach is demonstrated. In the rule-based approach by Zhang et al.,^[35] ontologies were built on the structure of the open knowledge base YAGO by defining appropriate keywords. YAGO combines facts from Wikidata with a standard ontology schema and a string-matching algorithm to obtain metallic materials concepts. The concepts were used to extract domain materials knowledge and to generate the Metallic Materials Ontology (MMOY).

For such an approach, it is particularly important to use fundamental vocabularies and defined terminologies, which are rich, abstract, and complete at best.

Ontologies and their associated terminology are ideally created and harmonized in collaborative environments. Here ontology engineers and MSE domain experts can reach an agreement and achieve a shared conceptualization. Moreno et al.^[36] illustrated the experts’ interaction within this creation process, where the ontology engineer is at the center of the test description process to constructively guide the semantic transformation. Bayerlein et al.^[37] highlighted the collaborative environment and its importance for digitizing MSE research. In particular, agile processes concerning data-driven materials research and development were pointed out in the context of the Mat-o-Lab framework. The work of Garabedian and Schreiber et al.^[38] describes another example of collaborative work. Here, the domain experts derived a controlled vocabulary describing tribological processes and objects with basic semantics in a MediaWiki-based database. This is the template for the creation

of the FAIR tribological experiments ontology (TriboDataFAIR Ontology).

To make data reproducible and reusable on a larger scale, additional information needs to be collected in the form of metadata. Romanos et al.^[39] present CHADA, a transferable approach to structure generic material characterization data and metadata. Captured information, for example, about the sample, operator, laboratory conditions, calibration procedures, etc., increases quality and provides specific insights into the experiment and data evaluation.

Ontological representation can also improve automation by facilitating human–machine communication through the creation of processable knowledge structures. In addition, programming language or data structuring standards can be created. Note that already standardized processes are particularly well suited for digital representation since the sequence of individual steps and actions and specific terminology are well-defined and agreed on.^[42] The dissemination of ontologies helps to manifest these emerging standards and to create data interoperability in communities according to the FAIR principles.

Furthermore, using well-structured data, interoperable systems will improve knowledge generation from data-driven research, for example, by increasing the effectiveness of machine learning methods. For this, it is necessary to create connections to upper-level ontologies. Although it is not the aim of this study, a trend is observed in the use of specific top-level ontologies in the works reviewed. For example, the most frequently used ontologies by^[30,33,34,36,37,40] include the W3C Provenance Ontology (PROV-O),^[43] the Basic Formal Ontology (BFO),^[7] and the Chemical Entities of Biological Interest (ChEBI).^[44] Other ontologies used several times in^[8,28,30,32,34,38,40] were DOLCE,^[6] EMMO,^[45] EXPO,^[46] QUDT,^[47] and SIO.^[48]

A necessary future work would be to conduct a detailed survey on relevant ontologies for MSE. Besides the suitability of certain upper-level ontologies, existing material-specific representations could also be highlighted.

3.2. Ontology Application

Ontologies and their creation described above provide the basis for the applications described below. Note that in this section we mainly focus on describing work and approaches that go beyond organizing MSE data according to modern standards from W3C.^[49] Therefore, the RDF ontology application of the studied works is discussed together with RDF instance creation and information retrieval in the next section.

In total, ontology-based applications were observed in 14 of the selected papers.^[28–30,32–35,37,38,40,41,50–52] These applications include the reuse, combination, and extension of existing ontologies and semantic resources (Figure 3b,i,ii). An example is given by Li et al.^[33] demonstrating the extension of two nanotechnology ontologies with new concepts and axioms in a two-step procedure. The tool-based extension of the MDO was shown by Nikooie et al.,^[34] where the user is supported and guided throughout the process.

An additional important application is the mapping or alignment of one ontology to another (Figure 3b,iii). An et al.^[50]

illustrated this by introducing the two-component system OTMapOnto. In this process, terms in different ontologies are identified and mapped with each other using ontology embedding and an optimal transport approach. The approach was applied in the MSE domain bringing improvements such as gain in precision and recall.

Great potential for the application of ontologies also lies in knowledge extraction (Figure 3b,iv) and discovery through natural language processing (NLP) techniques, such as named entity recognition. These methods particularly benefit from detailed and rich semantic vocabularies. In this context, Greenberg et al.^[51] presented the “automatic” linked data ontology application HIVE-4-MAT (Helping Interdisciplinary Vocabulary Engineering for Materials discovery). The application demonstrates the power of the combined use of ontological resources with NLP by extracting knowledge and rules from unstructured text in a structured way.

The possible applications for ontologies and well-founded vocabularies are manifold. In this respect, the development of ontology infrastructures for MSE is essential for the next generation of approaches involving the convergence between MSE and SWT, bringing benefits such as the description of materials science knowledge, interoperability on heterogeneous data resources, and knowledge extraction with NLP.

Further important criteria for the application and reuse of (MSE-relevant) ontologies are precise and domain-appropriate term definitions, user-friendly documentation, referenceability via persistent IRIs, and ongoing maintenance and curation of the ontology. Midlevel ontologies are the most likely candidates to succeed in applying these criteria. A prominent example of applying these criteria is the PMD Core Ontology (PMDco) (^[26]<https://github.com/materialdigital/core-ontology>). This mid-level ontology is maintained and curated by the Platform MaterialDigital (PMD), based on continuous MSE community interactions that serve as a semantic intermediate layer and amplifier for future application ontologies.

3.3. RDF Ontology Application, Instances Creation, and Information Retrieval

The motivation for RDF instance creation is to retrieve information from heterogeneous sources, for example, tabular or non-structured data, having the benefits of an RDF triple store and SPARQL queries (Figure 3c).

The main application of ontologies concerns FAIR data management and the creation of RDF instances. In this context, Bayerlein et al.^[37] emphasized the crucial role of ontologies for data and metadata annotation and structuring. In their work, data and contextual information on aluminum alloys for high-temperature applications were semantically organized and transferred to RDF triple stores. The authors used for this purpose-specific MSE domain ontologies. For such methods and experiments, the rate of data incorporation is greatly increased. Furthermore, data is more consistent and complete.

One method applied to other areas is described by Zhao and Qian.^[29] Here, an ontology is not developed by MSE domain experts, as in the previously presented work, but is based on the conversion of a schema of a relational material database

by a set of rules. This ontology enables the conversion and integration of further heterogeneous databases by mapping their data to created individuals. Zhao and Wang et al.^[30] also aimed at providing a unified and well-structured data representation. Using the developed NanoMine schema and ontology, they curated 182 articles to share polymer nanocomposite material data. With this work, they intended to advance the development of new materials. Therefore, the extraction of knowledge from semantic relations defined in a directed and labeled RDF graph, understood as sets of triples, is introduced. This shows that SWT allow researchers and larger research teams to efficiently obtain valuable, complete, and traceable data. RDF information retrieval tasks also enable the querying of certain information that cannot be easily found due to the complexity of data’s non-explicit semantic relations. In this regard, in the work of Li et al.,^[33] data from a material database is mapped into RDF using the MDO. The query functionality with SPARQL was demonstrated here based on its terminology. Zhang et al.^[35] showed the creation and representation of RDF triples with the MMOY they developed. A prototype gave users access to the linked knowledge structure about metallic materials, their fields of applications, their properties, and other information. Another example of the potential of information retrieval is provided in the work of Sadigh et al.^[41] In this work, the authors created an ontology for the experiments workflow of wire electrical discharge machining (WEDM). Thus, they aimed the engineering vision of selecting the most suitable machine tool based on manufacturing criteria. Based on this ontology, a platform enables the storage and analysis of information about machine parts and machine tools. This can provide results of capable machine tools to produce desired parts with multiscale features.

Regarding FAIR (research) data management, TriboDataFAIR Ontology was designed and deployed by Garabedian and Schreiber et al.^[38] with a specific focus on data interoperability and reusability. In this work, a knowledge graph based on collected experimental tribological data and metadata is created in a scalable environment, enabling the targeted retrieval of information.

We conclude this section by referring to other related works that present the wide potential of ontology application, RDF instance creation, and RDF information retrieval, for example, for domain knowledge representation and FAIR data management.^[31,36,39,52]

3.4. Provenance

The objective of tracking data provenance (also referred to as “data lineage”) is to know where the data comes from and where it was modified (Figure 3d). Data provenance is used to find errors within data and to attribute them to the origin. This way, it helps to assess authenticity, reproduce MSE experiments, and facilitate data reuse. Provenance aims to add another layer of credibility and quality to the data and is the subject of 13 of the selected works which aim to trace the data with the overall goal of improving MSE experiments and their reproducibility. In this context, Romanos et al.^[39] presented an approach to organize provenance by describing ontologies that facilitate the traceability of the data. Zhao and Wang et al.^[30] specified with their method

NanoMine a Semantic Extract, Transform, and Load (SETL) script, in which the data, together with the ontology, enriches the knowledge graph with entity definitions, indicating the provenance and the source of curation of the knowledge. The work from Sadigh et al.^[41] provided an ontology for a wire electrical discharge machining (WEDM) experimental workflow, which allowed for obtaining data provenance of MSE experiments. The ontology includes the definition of materials, machines, and tools represented by classes, relationships, axioms, and constraints. They also provided a use case showing the provenance of WEDM experiments. Li et al.^[40] developed the materials design ontology (MDO), where the provenance was given using a design pattern that provides information in the repository of ontologies, used together with entities from PROV-O.^[43] MSE-related data provenance was approached by Bayerlein et al.,^[37] Garabedian and Schreiber et al.,^[38] and Nikooie et al.,^[34] using the concept of FAIR data, which implies the ability to provide the provenance of the data. Vardeman et al.^[32] introduced a work where an ontology design pattern was created for modeling material transformation. The work offered a model of relationships between products, resources, and catalysts in the transformation process and the spatial and temporal constraints necessary for a transformation to occur. The authors used this model for reasoning and provenance. They presented a use case from the sustainable construction area that leveraged the material transformation pattern in combination with the preexisting semantic trajectory ontology design pattern. Moreno et al.^[36] provided an approach for data interoperability using ontologies explicitly developed to be applied to MSE tests. They obtained the provenance of experiments with the use of ontologies also to address interoperability, providing the possibility to be traced back, furthermore allowing to visualize relevant information.

Using RDF knowledge graphs to organize MSE data brings benefits, such as data provenance, which allows tracking the information even on heterogeneous datasets. The provenance of data plays an essential role in reproducing experiments and data, which, for instance, helps when applying FAIR data principles.

3.5. Ontology Reasoning

Ontology reasoning is applied to infer logical conclusions showing new relations between the concepts using asserted facts or axioms defined in the ontology (Figure 3e). It can also be understood, from a higher perspective, as deriving new axioms from a linked data structure to create a new ontology, as Zhang et al.^[35] did to build the ontology MMOY from the YAGO knowledge base. Zhao and Wang et al.^[30] provided an ontological framework called NanoMine, where the role of the ontology and knowledge graph is to encode concept relationships that provide bindings of equation variables to specific properties. Thus, “utilizing this resource with coded inferences, existing knowledge, and inferred knowledge can be automatically incorporated rather than manually curated into the knowledge graph”. This automation is also part of the aim of Sadigh et al.^[41] when they introduce a semantic reasoning process. This process creates appropriate relations between the desired characteristics of a part of a manufactured

product and the corresponding machine tools to produce it. The system runs this reasoning using embedded semantic rules in the designed ontology model. These semantic rules, also known as axioms in the context of SWT, are explicitly shown by Vardeman et al.^[32] The work proposed an ontology design pattern for material transformation, and the potential of this pattern is better revealed when formalizing the related axioms using description logic. All the papers discussed in this section coincide in their “ontology reasoning” perspective with the conclusions of Lambrix et al.^[52] Here it is described as “work to be done” using reasoning “to debug and complete different resources, leading to higher-quality resources” and in the process of querying distributed databases. In summary, it is observed that ontological reasoning is perceived as a tool for improvement and refinement to exploit the full potential of SWT. It could be paraphrased as ontological reasoning is not a basic but an advanced task of SWT. This, together with the fact that the maturity of SWT implementation in the MSE is limited, explains the relatively low amount of publications on the subject.

4. Challenges and Perspectives

In the previous sections, the state-of-the-art of SWT implementations, differentiated by specific categories, was highlighted for various MSE use cases. While different promising approaches and solutions could be identified, still many open challenges remain in dealing with and accessing long-established SWT standards, tools, and approaches. In the following, prominent challenges are explained and discussed.

Surveyed work approaches in the field of ontology creation vary widely. Key terms and relations of specific MSE subdomains are extracted manually, semiautomatically, as well as fully automatically from multiple sources, such as text bodies, database schemata and ontological resources. At best, developed ontologies are validated by consensus in interdisciplinary collaborative environments between MSE experts and ontology engineers. In this process, it is particularly challenging to establish a fundamental understanding of how ontology creation works and which are its best practices. Depending on the target application, the quality of the created ontologies depends on various factors, such as expressiveness, richness, completeness, and degree of abstraction. For example, quality-controlled ontologies improve the results of semantic searches. Or they already contain the metadata structure, which is of increased relevance, especially in the context of reproducibility and consequently reusability of MSE experiment and simulation data. Furthermore, qualitative statements can be made about the repeatability of experiments, data completeness, and data reliability.

In the evaluation of the reviewed works, we found that essential tools for various tasks are either missing or are not sufficiently known. Especially in the creation, collection, and extension of terminology databases, collaboratively used tools could potentially enable a coordinated, standardized, internationalized, and cross-domain effort. Graphical tools, such as MatVis (Available at <https://github.com/Mat-O-Lab/MatVis/>) and Ontopanel,^[53] can help MSE domain experts to become familiar with the process of ontology creation more easily while fostering interdisciplinary communication.

Critical for the reuse and further development of existing ontologies are their findability, availability, documentation, formatting, and ultimately their sustainable maintenance and curation. An ontology with high reuse potential is, for example, the QUDT,^[47] which can be used not only, but especially for quantity and unit standards. The use of standardized top-level ontologies, such as the BFO^[7] or even the PROV-O,^[43] could be observed in several works. Nevertheless, a uniform connection and aggregation via an MSE domain ontology and as a result, crosssubdomain interoperability has not yet been achieved. The resulting benefits of linked data as a basis for advanced techniques, such as machine learning, are still far from being exploited at this stage. For example, the performance of machine learning algorithms increases with the provision of coherent and well-annotated training data, in particular for small-to-medium-sized data sets. Likewise, the NOMAD Artificial Intelligence (AI) Toolkit,^[54] which offers interactive AI-based analysis of materials science data from the NOMAD archive, benefits from well-annotated data. Curated repositories have a critical role to fulfill in publishing and organizing ontologies consistently so that, in the long run, the diversity of data can be addressed in a workable way.

The central topic of several papers is the annotation of data and metadata (including provenance) using ontologies. This involves creating RDF instances and making them findable and available in triple stores using URIs and IRIs. **Figure 4** illustrates how, in the context of FAIR research data management, for example, unstructured data and metadata from MSE experiments and simulations are migrated through ontologies into machine-processable RDF triples with graph functionality to enable information retrieval and knowledge extraction. This becomes particularly important concerning methods with high-data throughput. It will also be exciting to see how this practice is applied to the processing and analysis of image data from imaging techniques, which were not covered in the selected papers.

The retrieval of semantically structured data bundled in triple stores is mainly performed by the query language SPARQL. Challenging here is the reliable access to data, consistent ontology structure, and stable SPARQL endpoints. The difficulty of formulating SPARQL queries must also be taken care of by tools with appropriate user interfaces, such as Sparklis,^[55] leveraging the potential promise of SWT. Another approach is the query formulation in natural language to facilitate the RDF data access for the broad MSE community.^[56,57] In the long run, systematic modeling based on recurrent upper-level ontologies will induce self-similar patterns and unify queries through recurrent property paths.

A more challenging technique, although only really addressed in a few of the papers surveyed, is reasoning. Here, only perspectives are given on the quality-enhancing potential for completion, refinement, and consistency checks for resources such as ontologies. The small number of publications on this topic signals the existing complexity and low level of maturity of SWT implementation in MSE as well.

In summary, by focusing on a selection of papers from the MSE-SWT intersection, we have already been able to identify research and technology gaps. This provides the opportunity to derive future research topics and directions. We also observed many isolated approaches and ideas. Fueled by the existing complexity of the ever-evolving SWT, on the one hand, and the creative application of domain experts in the MSE, on the other hand, this range of technological implementations has been a major factor. Nevertheless, certain intersecting themes are already crystallizing. For example, combinations of specific SW modules in the form of tool chains can be observed concisely, which could become established with an increasing adoption in the MSE community.

Beyond the selected works studied in detail, there are other areas in which SWT can be applied that require detailed analysis. One such area is, for example, the enhancement of existing

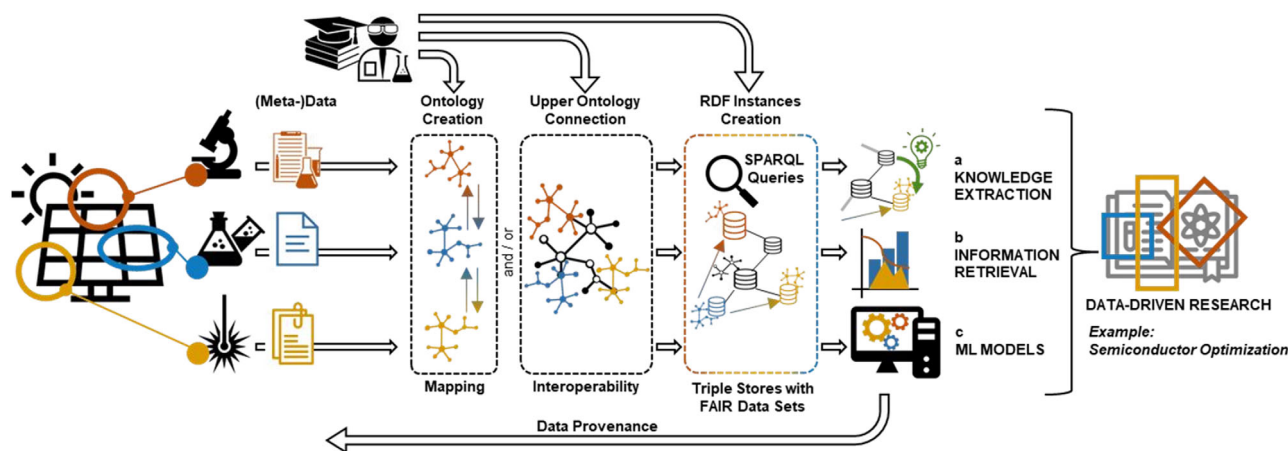


Figure 4. Schematic representation of the beneficial application of SWT in MSE. The combination of SW standards enables the comparison of FAIR data sets from different sources with multiple formats. For this purpose, unstructured (meta-)data from MSE experiments and simulations are transformed into a machine-processable, unified format in triples using ontologies. The RDF triples form the basis for: a) Knowledge extraction, where new and/or complex relationships are derived to optimize experiments and procedures. b) Information retrieval, i.e., mainly comparison and correlation with data sets according to desired criteria. c) Machine learning models that become more effective using coherent and interoperable data sets. The data provenance increases reliability, as the data can be verified at every stage. Overall, data-driven materials research benefits from SWT applications, for instance in semiconductor optimization.

material databases.^[12] These databases themselves are highly valuable resources of versatile information. Nevertheless, there is a lack of unified access to and interoperability between these heterogeneous structures. A solution in this area is provided by Andersen et al.^[58] with OPTIMADE. This API allows data to be retrieved across different databases using a uniform URL path. Material databases (<https://www.optimade.org/providers-dashboard/>) that support OPTIMADE API include Materials Project (<https://materialsproject.org/>), AFLOW (<https://www.aflowlib.org/>), COD (<http://www.crystallography.net/cod/>), TCOD (<http://www.crystallography.net/tcod/>), NOMAD (<https://cms.nomad-lab.eu/services/repo-arch>), to name a few. Another important topic, already discussed in Section 3.4, and essential for data reproducibility and reliability, involves tracking and providing complete knowledge of data provenance. In this context, Merkys et al.^[59] demonstrated how prospective materials data and corresponding databases can be automatically enriched with provenance information. The simulation results generated with an ontology-based crystal database are refined and made available with AiiDA's provenance tracker (<https://aiida.readthedocs.io/projects/aiida-core/en/v1.0.0b1/concepts/provenance.html>). Ghiringhelli et al.^[60] highlight the major necessity of a (meta)data mapping, which has to meet the requirements of complex (meta)data structures. Prospectively, ontologies are seen as the key to interoperability. The adaptation of the FAIR principles, supported by the use of SWT, also has considerable advantages for material databases. In detail, however, it is necessary to evaluate to what extent existing databases will benefit, the effort involved, and in the end, what impact it will have on the user.

Fundamentally, user-friendly ways must be established to ensure the proper application of cutting-edge technological solutions for MSE research and its advancement. This can be done via predefined data and metadata input templates in electronic lab notebooks, emulated by guidance with scalable examples, and especially in the context of teaching, in educational institutions and universities, create an early understanding of, for example, the RDF data model and ontologies.

5. Conclusion and Future Work

With this work, we provide a systematic review of publications from various MSE fields that use and integrate SWT to establish beneficial environments for FAIR research data management, data analysis, and data publication. For this purpose, we identified 19 relevant publications through a search strategy and addressed four research questions (see Section 2). We answered the first question, RQ1, by listing and discussing all surveyed articles. We answer RQ2 by providing a list of all applied SW methods and respective SW resources. For RQ3, in Sections 3 and 4, we discuss the impact of creating and applying ontologies in the context of FAIR data, which is the basic framework for data-driven materials research approaches in MSE. We conclude that ontological knowledge generally provides a comprehensible presentation of key terminologies and relations enabling result classification, where enrichment with contextual information makes experiments and simulations reproducible and therefore increases the reuse of the resulting, more complete data sets. We

could identify the following specific challenges and perspectives at the MSE–SWT intersection:

Competency Building: Creating a fundamental understanding and competencies for data management and computation in MSE education and teaching, for example, using best practice solutions as educational materials demonstrating benefits.

Reusability Increase: Capturing MSE metadata necessary for reproducibility of experiments and simulations covered by ontologies to ensure reliable, interoperable data for data-driven research.

Technology Use Facilitation: Fostering the development of user-friendly, collaborative ontology and vocabulary tools that enable, for example, the creation and consensus-building process.

Ontology Reuse: Increasing the reuse of quality ontologies by organizing them in a comprehensible and sustainable way in recognized MSE repositories. Quality-enhancing criteria include the use of domain-appropriate, precise definitions of terms, ease of use, and continuous maintenance and curation.

Common Community Standards: Aiming for agreement of the MSE community on well-defined standards, for example, a limited number of upper-level ontologies that are compatible with each other, to achieve subdomain interoperability.

SWT-Driven FAIR Data: Facilitating the generation of FAIR data by more user-friendly and robust SWT, implementing MSE use cases also with more advanced techniques such as reasoning.

We aimed to highlight the intersection between SWT and MSE for experts in both fields. The SWT experts are informed which technologies are used for which purpose and the MSE community is provided with an introductory overview of standard SWT implementations and their benefits and added value. However, the limited number of studies found also shows that this interdisciplinary area is still in its beginning. Technological implementations are mainly used for Big and FAIR data tasks, with ontology creation and its use for data and metadata structuring standing out. At the same time, it is clear from this work that challenges still exist and prevent exploiting the full potential of SWT. This is demonstrated, for example, by a lack of generic tools for converting tabular data into RDF knowledge graphs due to the heterogeneity of the data and other peculiarities of the MSE discipline.

SWT have the potential to become a key driver for data-driven materials science by making materials data available, discoverable, interoperable, and, eventually, reusable. Driven by existing and new project initiatives, we foresee increased development and use of ontologies, workflows, and databases in the future. In summary, with this work, we showcase the application of SWT, identify challenges, and thus signal for future MSE work on this intersection to be oriented and aligned.

Acknowledgements

Funding provided by Bundesanstalt für Materialforschung und-prüfung (BAM) is gratefully acknowledged. B.B. and T.M. thank the German Federal Ministry of Education and Research (BMBF) for financial support of the project Innovations-Plattform MaterialDigital (www.materialdigital.de) through project funding FKZ No: 13XP5094E (BAM). Special thanks to Thomas Hanke and Jürgen Olbricht in the context of the Mat-o-Lab project is given. The authors thank the PMD colleagues for numerous discussions

in the working group meetings. The authors would like to thank Rukeia El-Athman, Pedro Dolabella Portella, Robert Maaß, and Tilmann Hicel for their particularly valuable comments and discussions.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Conceptualization was done by A.V. and B.B. Search methodology was done by A.V. Writing the manuscript was done by A.V., B.B., and T.M. Visualizations were done by B.M.T., B.B., and A.V. Writing the review and editing were done by all authors. A.V. and B.B. contributed equally to this work.

Keywords

linked open data, materials science, ontology, semantic web

Received: January 27, 2023

Revised: March 23, 2023

Published online:

- [1] G. O. Consortium, *Nucl. Acids Res.* **2019**, *47*, D330.
- [2] L. M. Schriml, C. Arze, S. Nadendla, Y.-W. Chang, M. Mazaitis, V. Felix, G. Feng, W. A. Kibbe, *Nucl. Acids Res.* **2012**, *40*, D940.
- [3] S. Pestryakova, D. Vollmers, M. A. Sherif, S. Heindorf, M. Saleem, D. Moussallem, A.-C. N. Ngomo, *Sci. Data* **2022**, *9*, 389.
- [4] T. Berners-Lee, J. Hendler, O. Lassila, *Sci. Am.* **2001**, *284*, 34.
- [5] P. E. van der Vet, P.-H. Speel, N. J. Mars, in *Eleventh European Conf. on Artificial Intelligence (ECAI'94) Workshop on Comparison of Implemented Ontologies*, Citeseer, **1994**, pp. 8–12.
- [6] A. Gangemi, N. Guarino, C. Masolo, A. Oltramari, L. Schneider, in *Int. Conf. on Knowledge Eng. and Knowledge Management*, Springer, Berlin **2002**, pp. 166–181.
- [7] R. Arp, B. Smith, A. D. Spear, in *Building Ontologies with Basic Formal Ontology*, MIT Press, Cambridge, MA **2015**.
- [8] K. Cheung, J. Drennan, J. Hunter, in *AAAI Spring Symp.: Semantic Scientific Knowledge Integration* **2008**, p. 9–14.
- [9] C. D. Sainte Marie, M. Iglesias Escudero, P. Rosina, in *Int. Conf. on Web Reasoning and Rule Systems*, Springer, Berlin **2011**, pp. 24–29.
- [10] T. Ashino, *Data Sci. J.* **2010**, *9*, 54.
- [11] A. White, *MRS Bull.* **2012**, *37*, 715.
- [12] L. Himanen, A. Geurts, A. S. Foster, P. Rinke, *Adv. Sci.* **2019**, *6*, 1900808.
- [13] C. Draxl, M. Scheffler, *J. Physics: Mater.* **2019**, *2*, 036001.
- [14] L. M. Ghiringhelli, C. Carbogno, S. Levchenko, F. Mohamed, G. Huhs, M. Lüders, M. Oliveira, M. Scheffler, *NPJ Comput. Mater.* **2017**, *3*, 46.
- [15] X. Zhang, C. Zhao X. Wang, *Comput. Ind.* **2015**, *73*, 8.
- [16] National Academies of Sciences Engineering and Medicine, in *Ontologies in the Behavioral Sciences: Accelerating Research and the Spread of Knowledge: Digest Version*, The National Academies Press, Washington DC **2022**.
- [17] M. D. Wilkinson, M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.-W. Boiten, L. B. da Silva Santos, Bourne, *Sci. Data* **2016**, *3*, 160018.
- [18] T. R. Gruber, *Knowl. Acquis.* **1993**, *5*, 199.
- [19] L. Ding, Y. Peng, P. P. da Silva, McGuinness, TR-CS-05-06 **2005**.
- [20] R. Dividino, S. Sizov, S. Staab, B. Schueler, *J. Web Semant.* **2009**, *7*, 204.
- [21] G. Flouris, I. Fundulaki, P. Padiaditis, Y. Theoharis, V. Christophides, in *Int. Semantic Web Conf.*, Springer **2009**, pp. 196–212.
- [22] S. S. Sahoo, V. Nguyen, O. Bodenreider, P. Parikh, T. Minning, A. P. Sheth, *BMC Bioinf.* **2011**, *12*, 461.
- [23] T. Dyba, T. Dingsoyr, G. K. Hanssen in *Proc. of the First Int. Symp. on Empirical Software Engineering and Measurement, ESEM '07*, IEEE Computer Society, Washington, DC, **2007**, pp. 225–234.
- [24] B. Kitchenham, *Keele Univ.* **2004**, *33*, 1.
- [25] D. Moher, A. Liberati, J. Tetzlaff, D. G. Altman, *PLoS Med.* **2009**, *6*, 1000097.
- [26] D. Moussallem, M. Wauer, A.-C. N. Ngomo, *J. Web Semant.* **2018**, *51*, 1.
- [27] M. Fiorelli, A. Stellato, *Metadata Semant. Res.* **2020**, *1355*, 85.
- [28] M. E. Deagen, J. P. Mccusker, T. Fateye, S. Stouffer, L. C. Brinson, D. L. McGuinness, L. S. Schadler, *Sci. Data* **2022**, *9*, 239.
- [29] S. Zhao, Q. Qian, *AIP Adv.* **2017**, *7*, 105325.
- [30] H. Zhao, Y. Wang, A. Lin, B. Hu, R. Yan, J. Mccusker, W. Chen, D. L. McGuinness, L. Schadler, L. C. Brinson, *APL Mater.* **2018**, *6*, 111108.
- [31] O. Hakimi, J. L. Gelpi, M. Krallinger, F. Curi, D. Repchevsky, M.-P. Ginebra, *Adv. Funct. Mater.* **2020**, *30*, 1909910.
- [32] C. F. A. A. Krisnadhivardeman II, M. Cheatham, K. Janowicz, H. Ferguson, P. Hitzler, A. P. C. Buccellato, V. De Boer, A. Ławrynowicz, *Semant. Web* **2017**, *8*, 719.
- [33] H. Li, R. Armiento, P. Lambrix, *Data Sci. J.* **2019**, *18*, 50.
- [34] M. Abd Nikooie Pour, H. Li, R. Armiento, P. Lambrix, in *Sixth Int. Workshop on the Visualization and Interaction for Ontologies and Linked Data*, CEUR Workshop Proc. **2021**, pp. 1–12.
- [35] X. Zhang, D. Pan, C. Zhao, K. Li, *Adv. Eng. Inf.* **2016**, *30*, 687.
- [36] B. Moreno Torres, C. Völker, S. M. Nagel, T. Hanke, S. Kruschwitz, *Remote Sens.* **2021**, *13*, 2426.
- [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, J. Olbricht, *Adv. Eng. Mater.* **2022**, *24*, 2101176.
- [38] N. T. Garabedian, P. J. Schreiber, N. Brandt, P. Zschumme, I. L. Blatter, A. Dollmann, C. Haug, D. Kümmel, Y. Li, F. Meyer, C. E. Morstein, J. S. Rau, M. Weber, J. Schneider, P. Gumbsch, M. Selzer, C. Greiner, *Sci. Data* **2022**, *9*, 315.
- [39] N. Romanos, M. Kalogerini, E. P. Koumoulos, A. K. Morozinis, M. Sebastiani, C. Charitidis, *Mater. Today Commun.* **2019**, *20*, 100541.
- [40] H. Li, R. Armiento, P. Lambrix, in *Int. Semantic Web Conf.*, Springer **2020**, pp. 212–227.
- [41] B. L. Sadigh, S. N. B. Oliaei, S. Dadvandipour, in *Solid State Phenomena*, Vol. 261, Trans Tech Publ, **2017**, pp. 470–477.
- [42] M. Schilling, <https://opus4.kobv.de/opus4-bam/frontdoor/index/index/docId/52949> (accessed: September 2022).
- [43] T. Lebo, S. Sahoo, D. McGuinness, K. Belhajjame, J. Cheney, D. Corsar, D. Garijo, S. Soiland-Reyes, S. Zednik, J. Zhao, World Wide Web Consortium **2013**.
- [44] J. Hastings, G. Owen, A. Dekker, M. Ennis, N. Kale, V. Muthukrishnan, S. Turner, N. Swainston, P. Mendes, C. Steinbeck, *Nucleic Acids Res.* **2016**, *44*, D1214.
- [45] N. Adamovic, P. Asinari, G. Goldbeck, A. Hashibon, K. Hermansson, D. Hristova-Bogaerds, R. Koopmans, T. Verbrugge, E. Wimmer, in *Proc. of the 4th World Congress on Integrated Computational Materials Engineering (ICME 2017)*, Springer, Berlin **2017**, pp. 79–92.
- [46] L. N. Soldatova, R. King, *R. Soc. Interface* **2006**, *3*, 795.
- [47] D. Dauga, in *Qudt; Quantities, Units, Dimensions and Types*, **2022**.
- [48] M. Dumontier, C. J. O. 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, R. Hoehndorf, J. *Biomed. Semant.* **2014**, 5, 14.
- [49] W3C. The world wide web consortium (w3c). <https://www.w3.org/> **2022**. <https://www.w3.org/>.
- [50] Y. An, A. Kalinowski, J. Greenberg, *Ontol. Alignment Eval. Initiative* **2021**, 3063, 62.
- [51] J. Greenberg, X. Zhao, J. Adair, J. Boone, X. T. Hu, in *Research Conf. on Metadata and Semantics Research*, Springer **2020**, pp. 297–307.
- [52] P. Lambrix, R. Armiento, A. Delin, H. Li, J.: *Encycl. Big Data* **2018**, 1, XLVI, 1820.
- [53] Y. Chen, M. Schilling, P. von Hartrott, H. Beygi Nasrabadi, B. Skrotzki, J. Olbricht, *Integr. Mater. Manuf. Innov.* **2022**, 11, 545.
- [54] L. Sbailò, A. Fekete, L. M. Ghiringhelli, M. Scheffler, *NPJ Comput. Mater.* **2022**, 8, 250.
- [55] S. Ferré, *Semant. Web* **2017**, 8, 405.
- [56] K. S. Aggour, A. Detor, A. Gabaldon, V. Mulwad, A. Moitra, P. Cuddihy, V. S. Kumar, *Integrating Mater. Manuf. Innov.* **2022**, 11, 467.
- [57] A. Valdestilhas, T. Hanke, NaturalMSEQueries—A Natural Way to Query Materials Science Engineering Data Experiments, **2023**, <https://doi.org/10.5281/zenodo.7744532>.
- [58] R. Armiento, W. Andersen, P. Lambrix, G. Conduit, C. Toher, L. Ghiringhelli, *Sci. Data* **2021**, 8, 217.
- [59] A. Merkys, N. Mounet, A. Cepellotti, N. Marzari, S. Gražulis, G. Pizzi, *J. Cheminform.* **2017**, 9, 56.
- [60] L. M. Ghiringhelli, C. Baldauf, T. Bereau, S. Brockhauser, C. Carbogno, J. Chamanara, S. Cozzini, S. Curtarolo, C. Draxl, Dwaraknath, arXiv:2205.14774, **2022**.
- [61] M. Eid, R. Liscano, A. El Saddik, in *2007 IEEE Int. Conf. on Computational Intelligence for Measurement Systems and Applications*, IEEE, Piscataway, NJ **2007**, pp. 59–62.
- [62] L. Aroyo, C. Welty, H. Alani, J. Taylor, A. Bernstein, L. Kagal, N. Noy, E. Blomqvist, *Nucl. Acids Res.* **2009**, 37, W170.
- [63] S.-Y. Yang, *Expert. Syst. Appl.* **2009**, 36, 10148.
- [64] M. Karray, N. Otte, R. Rai, F. Ameri, B. Kulvatunyou, B. Smith, D. Kiritsis, C. Will, R. Arista, in *Int. Conf. on Interoperability for Enterprise Systems and Applications*, Tarbes, **2021**.
- [65] R. Moten, in *The 11th Int. Conf. on Semantic Technology for Intelligence (Defense, and Security (STIDS 2016))*, **2016**.
- [66] iso.org. Iso/iec 21838-2:2021 Information Technology — Top-Level Ontologies (TLO) — Part 2: Basic Formal Ontology (BFO), **2021**.
- [67] A. Treloar, *Learn. Publ.* **2014**, 27, S9.
- [68] WWW '17 *Proc. of the 26th Inter. Conf. on World Wide Web* (Eds: R. Barrett, R. Cummings, E. Agichtein, E. Gabrilovich) ACM, Perth, Australia **2017**.
- [69] WWW '16 *Proc. of the 25th Int. Conf. on World Wide Web* (Eds: J. Bourdeau, J. Hendler, R. Nkambou, I. Horrocks, B. Y. Zhao), ACM, Montreal, Canada **2016**.
- [70] WWW '11: *Proc. of the 20th Int. Conf. on World Wide Web* (Eds: S. Srinivasan, K. Ramamritham, A. Kumar, M. P. Ravindra, E. Bertino, R. Kumar), ACM, New York, NY **2011**.
- [71] WWW '12: *Proc. of the 21st Int. Conf. on World Wide Web* (Eds: A. Mille, F. L. Gandon, J. Misselis, M. Rabinovich, S. Staab), ACM, New York, NY **2012**.
- [72] WWW '13 *Companion: Proc. of the 22Nd Int. Conf. on World Wide Web Companion* (Eds: D. Schwabe, V. A. F. Almeida, H. Glaser, R. A. Baeza-Yates, S. B. Moon), *Inter. World Wide Web Conf.s Steering Committee*, Rio de Janeiro, Brazil **2013**.
- [73] WWW '14: *Proc. of the 23rd Inter. Conf. on World Wide Web* (Eds: C. Chung, A. Z. Broder, K. Shim, T. Suel), ACM, New York, NY **2014**.
- [74] WWW '15 *Companion: Proc. of the 24th Int. Conf. on World Wide Web Companion* (Eds: A. Gangemi, S. Leonardi, A. Panconesi), ACM, Republic and Canton of Geneva, Switzerland **2015**.
- [75] *Proc. of the 1st Int. Workshop on the Multilingual Semantic Web* (Eds: P. Buitelaar, P. Cimiano, E. Montiel-Ponsoda), CEUR-WS.org, Raleigh, NC **2010**.
- [76] *Proc. of the WWW2009 Workshop on Linked Data on the Web, LDOW* (Eds: C. Bizer, T. Heath, T. Berners-Lee, K. Idehen), CEUR-WS.org, Madrid, Spain **2009**.
- [77] *Proc. of the WWW2008 Workshop on Linked Data on the Web, LDOW* (Eds: C. Bizer, T. Heath, K. Idehen, T. Berners-Lee), CEUR-WS.org, Beijing, China **2008**.
- [78] *Proc. of the WWW2007 Workshop 13: Identity, Identifiers, Identification, Entity-Centric Approaches to Information and Knowledge Management on the Web* (Eds: P. Bouquet, H. Stoermer, G. Tummarello, H. Halpin), CEUR-WS.org, Banff, Canada **2007**.
- [79] WWW '06: *Proc. of the 15th Inter. Conf. on World Wide Web* (Eds: L. Carr, D. D. Roure, A. Iyengar, C. A. Goble, M. Dahlin) ACM, Edinburgh, Scotland, UK **2006**.
- [80] WWW '05: *Proc. of the 14th Int. Conf. on World Wide Web* (Eds: A. Ellis, T. Hagino), ACM, Chiba, Japan **2005**.
- [81] WWW '04: *Proc. of the 13th Int. Conf. on World Wide Web* (Eds: S. I. Feldman, M. Uretsky, M. Najork, C. E. Wills), ACM, New York, NY **2004**.
- [82] WWW '03: *Proc. Of The Twelfth International Conf. On World Wide Web* (Eds: G. Hencsey, B. White, Y. R. Chen, L. Kovács, S. Lawrence), ACM, Budapest, Hungary **2003**.
- [83] WWW '02: *Proc. Of The Eleventh International Conf. On World Wide Web* (Eds: D. Lassner, D. D. Roure, A. Iyengar), ACM, Honolulu, Hawaii **2002**.
- [84] WWW '01: *Proc. of the Tenth Int. Conf. on World Wide Web* (Eds: V. Y. Shen, Saito, N., Lyu, M. R., Zurko, M. E.), ACM, Hong Kong, China **2001**.
- [85] *The Semantic Web – ISWC 2002* (Eds: I. Horrocks, J. A. Hendler), Springer, Sardinia, Italy **2002**.
- [86] *The Semantic Web – ISWC 2003* (Eds: D. Fensel, K. P. Sycara, J. Mylopoulos), Springer, Sanibel Island, FL **2003**.
- [87] *The Semantic Web – ISWC 2004* (Eds: S. A. McIlraith, D. Plexousakis, F. van Harmelen), Springer, Hiroshima, Japan **2004**.
- [88] *The Semantic Web – ISWC 2005* (Eds: Y. Gil, E. Motta, V. R. Benjamins, M. A. Musen), Springer, Galway, Ireland **2005**.
- [89] *The Semantic Web – ISWC 2006* (Eds: I. F. Cruz, S. Decker, D. Allemang, C. Preist, D. Schwabe, P. Mika, M. Uschold, L. Aroyo), Vol. 4273, Springer, Athens, GA **2006**.
- [90] *The Semantic Web – ISWC 2007* (Eds: K. Aberer, K.-S. Choi, N. F. Noy, D. Allemang, K.-I Lee, L. J. B. Nixon, J. Golbeck, P. Mika, D. Maynard, R. Mizoguchi, G. Schreiber, P. Cudré-Mauroux), Vol. 4825, Springer, Busan, Korea **2007**.
- [91] *The Semantic Web – ISWC 2008* (Eds: A. P. Sheth, S. Staab, M. Dean, M. Paolucci, D. Maynard, T. W. Finin, K. Thirunarayan), Springer, Karlsruhe, Germany **2008**.
- [92] *The Semantic Web – ISWC 2009* (Eds: A. Bernstein, D. R. Karger, T. Heath, L. Feigenbaum, D. Maynard, E. Motta, K. Thirunarayan), Springer, Chantilly, VA **2009**.
- [93] *The Semantic Web–ISWC 2010* (Eds: P. F. Patel-Schneider, Y. Pan, P. Hitzler, P. Mika, L. Zhang, J. Z. Pan, I. Horrocks, B. Glimm), Springer-Verlag, Berlin Heidelberg, Germany **2010**.
- [94] *The Semantic Web–ISWC 2011* (Eds: L. Aroyo, C. Welty, H. Alani, J. Taylor, A. Bernstein, L. Kagal, N. Noy, E. Blomqvist), Springer-Verlag, Berlin Heidelberg, Germany **2011**.
- [95] *The Semantic Web–ISWC 2012* (Eds: P. Cudré-Mauroux, J. Heflin, E. Sirin, T. Tudorache, J. Euzenat, M. Hauswirth, J. X. Parreira, J. Hendler, G. Schreiber, A. Bernstein, E. Blomqvist), Springer-Verlag, Berlin Heidelberg, Germany **2012**.
- [96] *The Semantic Web–ISWC 2013* (Eds: H. Alani, L. Kagal, A. Fokoue, P. Groth, C. Biemann, J. X. Parreira, L. Aroyo, N. Noy,

- C. Welty, K. Janowicz), Springer-Verlag, Berlin Heidelberg, Germany **2013**.
- [97] *The Semantic Web—ISWC 2014, Riva Del Garda, Italy, October 19-23, 2014. Proc.* (Ed: P. Mika), Vol. 8796, Springer **2014**.
- [98] *The Semantic Web – ISWC 2015* (Eds: M. Arenas, Ó. Corcho, E. Simperl, M. Strohmaier, M. d'Aquin, K. Srinivas, P. T. Groth, M. Dumontier, J. Heflin, K. Thirunarayan, S. Staab), Springer, Bethlehem, PA **2015**.
- [99] *The Semantic Web—ISWC 2016—15th Int. Semantic Web Conf.* (Ed: P. T. Groth), Springer, Kobe, Japan **2016**.
- [100] *The Semantic Web—ISWC 2017—16th Int. Semantic Web Conf.*, Lecture Notes in Computer Science (Eds: C. d'Amato, M. Fernández, V. A. M. Tamma, F. Lécué, P. Cudré-Mauroux, J. F. Sequeda, C. Lange, J. Heflin) Springer, Vienna, Austria, **2017**.
- [101] *The Semantic Web—14th Int. Conf., ESWC*, Lecture Notes in Computer Science, (Eds: E. Blomqvist, D. Maynard, A. Gangemi, R. Hoekstra, P. Hitzler, O. Hartig) Portorož, Slovenia, **2017**.
- [102] *The Semantic Web. Latest Advances and New Domains*, vol. 9678 of Lecture Notes in Computer Science (Eds: H. Sack, et al.), Springer, Berlin Heidelberg, Germany, **2016**.
- [103] *The Semantic Web: Research and Applications*, vol. 6643 of Lecture Notes in Computer Science. Semantic Technology Institute International (STI2) (Eds: G. Antoniou, M. Grobelnik, E. Simperl, B. Parsia, D. Plexousakis, P. de Leenheer, J. Z. Pan), Springer-Verlag, Berlin Heidelberg, Germany, **2011**.
- [104] *The Semantic Web: Research and Applications*, vol. 7295 of Lecture Notes in Computer Science. Semantic Technology Institute International (STI2) (Eds: E. Simperl, P. Cimiano, A. Polleres, O. Corcho, V. Presutti), Springer-Verlag, Berlin Heidelberg, Germany, **2012**.
- [105] *The Semantic Web: Semantics and Big Data*, vol. 7882 of Lecture Notes in Computer Science. Semantic Technology Institute International (STI2), (Eds: P. Cimiano, O. Corcho, V. Presutti, L. Hollink, S. Rudolph), Springer-Verlag, Berlin Heidelberg, Germany, **2013**.
- [106] *The Semantic Web: Trends and Challenges* vol. 8465 of Lecture Notes in Computer Science. Semantic Technology Institute International (STI2) (Eds: V. Presutti, et al.), Springer International Publishing, Cham, Switzerland, **2014**.
- [107] *The Semantic Web. Latest Advances and New Domains*, vol. 9088 of Lecture Notes in Computer Science. Semantic Technology Institute International (STI2) (Eds: F. Gandon, M. Sabou, H. P. Sack, A. Zimmermann) Springer International Publishing, Cham, Switzerland, **2015**.
- [108] *The Semantic Web: Research and Applications, 7th Extended Semantic Web Conf., ESWC 2010* (Ed: L. Aroyo, et al.), Springer, Heraklion, Crete **2010**.
- [109] *The Semantic Web: Research and Applications, 6th European Semantic Web Conf., ESWC 2009* (Eds: L. Aroyo, P. Traverso, F. Ciravegna, P. Cimiano, T. Heath, E. Hyvönen, R. Mizoguchi, E. Oren, M. Sabou, E. P. B. Simperl), Springer, Heraklion, Crete **2009**.
- [110] *The Semantic Web: Research and Applications, 5th European Semantic Web Conf., ESWC 2008* (Eds: S. Bechhofer, M. Hauswirth, J. Hoffmann, M. Koubarakis), Springer, Tenerife, Canary Islands **2008**.
- [111] *The Semantic Web: Research and Applications, 4th European Semantic Web Conf., ESWC 2007* (Eds: E. Franconi, M. Kifer, W. May), Springer, Innsbruck, Austria **2007**.
- [112] *The Semantic Web: Research and Applications, 3rd European Semantic Web Conf., ESWC 2006* (Eds: Y. Sure, J. Domingue), Springer, Budva, Montenegro **2006**.
- [113] *The Semantic Web: Research and Applications, Second European Semantic Web Conf., ESWC 2005* (Eds: A. Gómez-Pérez, J. Euzenat), Springer, Crete, Greece **2005**.
- [114] *The Semantic Web: Research and Applications, First European Semantic Web Symp., ESWS* (Eds: C. Bussler, J. Davies, D. Fensel, R. Studer), Springer, Crete, Greece **2004**.
- [115] *Open Databases Integration for Materials Design – OMDI 2019* (Eds: R. Armiento, G. Conduit, S. Gražulis, F. Mohamed, G. Pizzi, G.-M. Rignanese, C. Toher,), CECAM-HQ-EPFL, Lausanne, Switzerland **2019**.
- [116] *Workshop on Ontologies for Materials-Databases Interoperability – OMDI 2021* (Ed: R. Armiento, et al.), Linköping University, Linköping University, Sweden **2021**.