A systematic mapping study on combining conceptual modeling with semantic web

Cordula Eggerth [a], Syed Juned Ali [a] and Dominik Bork [a,*]
[a] *Business Informatics Group, TU Wien*

Abstract. Conceptual models aim to represent real systems at a higher abstraction level. The Semantic Web intends to add meaning to any kind of data format to arrive at linked data. Taken together, both help facilitate data processing and integration for humans and machines. In this article, we systematically analyze the research landscape at the intersection of conceptual modeling and the Semantic Web by means of a systematic mapping study (SMS). Following the SMS research methodology, first, the research scope is defined, then the search queries are designed and executed. From initially 5.107 publications, we systematically filtered out all irrelevant ones, leaving us with 484 eventually relevant ones. These publications are analyzed and mapped to a number of taxonomies. The extracted and refined data is analyzed in several analysis steps comprising bibliographical, content, combined taxonomy, and research community analyses. In this paper, we show the most active institutions and authors in the field and the topics on which they work. We moreover show, by which means (e.g., modeling language and Semantic Web standard) research predominantly combines the two areas. Besides highlighting flourishing research areas, we also point to many remaining areas for future research in which we scarcely found existing works and/or see great potential.

Keywords: Conceptual modeling, Semantic Web, Systematic Mapping Study

1. Introduction

Conceptual modeling (CM) sees the underlying reality from a more abstract perspective, which focuses on the necessary features, while leaving out the unnecessary ones [1]. Hence, a conceptual model always remains a partial excerpt or view of the underlying real system which relies on assumptions made relating to the underlying real system [2, 3]. It can also occur on even higher abstraction levels, which is represented by “meta“ models, which are basically models of models [2]. To some extent, also semantic parts are included at this perspective [4]. Key to conceptual modeling is the process of determining the right degree of abstraction from the reality, and to determine which features are essential for the model, and which ones can be left out [3, 4]. Next to this, determining the most suitable conceptual modeling language for a specific use is crucial [3]. There are, among others, languages with structural elements like “entities, relationship, and constraints“, and languages with strengths in representing behavioral aspects comprising, e.g., “states, transition, and actions“ [4]. Apart from determining a suitable abstraction level, highlighting the goal of conceptual modeling/a conceptual model as a visual support and communication tool among all kinds of stakeholders still represents a challenge [4, 5]. Furthermore, transferring the abstract view into formal structures that could be used for inference, while enabling the visual and communication support in the real perspective containing all details, is likewise a defying task [2, 4]. Generally speaking, conceptual models come in various forms and shapes, i.e., from non-formal, rather conversational to very formal, mathematically sound ones (e.g., Petri Nets) that have defined rule interlinks [2, 6]. But a formal underpinning does not add rich model semantics yet. This has to be done by “associating semantics to the language elements“, according to [2].

*Corresponding author. E-mail: dominik.bork@tuwien.ac.at.

Semantic Web (SW) in contrast focuses more on adding meaning to systems and web technologies for input processing automation, meta data generation and analysis, and formalized representations of the underlying reality [7]. Berners-Lee (2001), one of the founding personalities of the World Wide Web, expressed it in a way that the Semantic Web "brings structure to the meaningful content of web pages [...] as an extension to the usual websites" [8]. The intention was was to enhance the automated processing of websites by extending web pages with meta data and meaning – a specific structure that could be read and processed by machines, and ultimately include reasoning within the web via a "semantic markup" [8]. Overall, the Semantic Web has concentrated on integrating data in different shapes and sizes from various data sources, and to organize them more conveniently for further processing and linked data purposes [9]. It has intended a transition from "a web of data to a web of documents" based on tailored data formats such as Resource Description Framework (Schema) (RDF(S)), RDF XML, N3, and query languages, which aims to make data from various sources more interoperable [10]. The formal approach is reinforced by specified notations like RDFS and Ontology Web Language (OWL) to include and share defined structures as well as their interrelations using ontologies (with underlying description logic) or less formal vocabularies and taxonomies to provide linked data [10]. SW has become closely intertwined with several other areas, among them sensor networks, Internet of Things (IoT), and natural language processing (NLP) [10, 11]. The available ontologies and vocabularies can be matched by the means of machine learning (ML) [10–12].

Both **conceptual modeling and Semantic Web** are independent research areas, in which researchers are continuously exploring the field. However, the topics also intersect to some extent. For example with regard to ontology creation, they overlap because both use ontologies to "formally represent the conceptualization of a domain" [7]. Conceptual modeling and semantic web can both be used stand-alone, but can also be combined. Taken together, both conceptual modeling and semantic web help to facilitate processing of data objects and data integration for humans as well as for machines. When Sandkuhl et al. (2018) conceived their suggestions for extending conceptual modeling, they mentioned semantic annotation (notably with regard to assistive technologies) and further semantic web topics due to their capability of automated reasoning and inference, next to gamification, knowledge management, architectural thinking, and user-centered innovation [13, 14]. Semantic technologies in combination with conceptual modeling can range from ontology languages that describe conceptual models to rather light-weight semantic annotations, and tools that transform conceptual models to formal ontologies [13]. Conceptual modeling as well as Semantic Web use ontologies, which led to emerging topics like ontology matching, patterns, and analysis, even extended to the meta level [12, 15]. According to Storey (2015), the rise of the Semantic Web contributed to the enhanced use of ontologies in conceptual modeling to add reasoning, and semantics to CM, i.e., it "semantically enriches" it [12]. As traditional conceptual models did not comprise formal "specifications of the semantics of the terminology of the underlying models", which undermined their consistency, the use of ontologies enabled a more suitable consideration of consistency and semantic aspects [14–16]. This contributed to new opportunities within conceptual models to apply inference and logical reasoning tasks [15].

In view of the above, this paper aims to explore and analyze the research landscape at the intersection of conceptual modeling and semantic web in terms of research published in recent years by the means of a systematic mapping study. This particular intersection has neither been covered systematically nor comprehensively.

In the remainder of this paper, Section 2 reports on related works targeting the intersection of conceptual modeling and semantic web. The research questions motivating and the research methodologies guiding this paper are covered in Section 3. The findings of mapping the relevant literature to the previously introduced taxonomies are then presented in Section 4. In Section 5 we then briefly describe the Web Knowledge Base we developed to ease access to this studies data and findings. Threats to validity are discussed in Section 6 before we close this paper with a discussion of its implications for future research (Section 7) and a conclusion in Section 8.

2. Related Work

We were interested to identify if other works at the intersection of conceptual modeling and Semantic Web exist. Consequently, we used the search query as depicted in Fig. 1 and executed it on the Scopus database. We derive, that many surveys on conceptual modeling and/or Semantic Web exist, but non cover both (i.e., the intersection or combination of both). The query yielded 55 publications across both topics, which were then narrowed down by filtering out unsuitable papers in terms of research field and type, the latter being either a systematic mapping study or systematic literature review. After screening, 11 publications on conceptual modeling and 12 publications on the Semantic Web remained in the relevant subset.

(survey OR systematic mapping study OR sms OR mapping study OR systematic mapping) AND (semantic web OR semantic systems OR knowledge graph OR linked data OR linked open data OR ontology OR rdf)
OR
(survey OR systematic mapping study OR sms OR mapping study OR systematic mapping) AND (conceptual model OR modeling language OR modelling language)

Fig. 1. Scopus query for related work

The 11 publications related to **conceptual modeling** are depicted in Table 1. They range from the year 2015 to 2022. They are mainly of the research type systematic mapping study, with a minority of two publications being systematic literature reviews.

Year	First Author Name	Topic	Type
2015	Kosar	Domain-Specific Languages	SMS
2015	Verdonck	Ontology-driven Conceptual Modeling	SMS, SLR
2015	Wakil	Model Driven Web Engineering	SMS
2017	Kolukisa	Ontologies in Software Process Assessment	SLR
2017	Wortmann	Modeling for Industry 4.0	SMS
2019	Alkharabsheh	Software Design Smell Detection	SMS
2019	Rodrigues	Legal Ontologies over Time	SMS
2019	Wortmann	Modeling Languages in Industry 4.0	SMS
2020	Iung	Domain-specific Language Development Tools	SMS
2021	Harley	Data Modeling and NoSQL Databases	SMS
2022	Zahid	Formal Methods in Requirements Engineering of Industrial CPS	SMS

Table 1
Related surveys on conceptual modeling

In 2015, for instance Kosar published a SMS on *Domain-Specific Languages*, which was based on a search query limited to the time span between 2006 and 2012 [17]. It intended to provide a fine-grained understanding of the domain-specific language (DSL) field and its evolutionary trends based on a previous work from 2005 [17]. In this paper, the sequential flow from 1153 first query results from Web of Science (WoS) and ACM Digital Library to the final selected 390 publications was documented [17]. In line with the findings, it turned out that the focus in DSL research was usually rather placed on the "development of new techniques / methods rather than investigating the integration of DSL with other software engineering processes or measuring their effectiveness", according to Kosar (2015) [17]. The number of works published did not change either over the years analyzed, meaning that the research field stayed as it is, and "domain analysis, validation, and maintenance" were revealed as areas for further research [17].

Verdonck et al. (2015) rather focused on *Ontology-driven Conceptual Modeling* in their study, which was back then a cutting edge research field [18]. It covered overall 180 publications [18]. Based on the SMS, several "research gaps" were identified, among them the lack of empirical projects, of "model purpose" specification, and of "experimental, observational, and testing evaluation methods" [18]. Verdonck (2015) also recommend to direct more research effort to "how learning, interpretation, and understanding of conceptual representation" [18].

Zahid (2022) provided a SMS on *Semi-formal and Formal Methods in Requirements Engineering of Industrial Cyber-Physical Systems* comprising 93 underlying publications from 2009 to 2020 [19]. According to the authors, publications in that research area concentrate on “formal analysis and verification of safety and timing requirements“ [19]. However, semi-formal methods, privacy-considering methods, and industrial standards are not much represented in the underlying papers, which would suggest potential for further research [19].

Table 2 lists the 12 publications related to **Semantic Web** including their publication year and main author. They range from the year 2009 to 2022. They are mainly of the research type systematic mapping study, with a minority of two publications being systematic literature reviews.

Year	First Author Name	Title	Type
2009	Janev	Maturity and Applicability Assessment of SW Technologies	SMS
2016	Pauwels	Semantic Web Technologies in AEC Industry	SLR
2016	Zander	SW Techniques for Description of Robotic Components	SMS
2017	Moussallem	Machine Translation using Semantic Web Technologies	SMS
2018	De Souza Neto	Semantic Web and Human Computation	SMS
2018	Sabou	Semantic Web Services Testing	SMS
2019	Alloghani	The XML and Semantic Web	SMS
2019	Gacitua	The XML and Semantic Web	SMS
2020	Dadkhah	Semantic Web Enabled Software Testing	SLR
2020	Drury	Semantic Web Technology for Agriculture	SMS
2021	Enriquez-Reyes	Using SW Techniques in Development of Data Warehouses	SMS
2022	Senthil	SW Techniques in Healthcare	SMS

Table 2
Related surveys on Semantic Web

The SMS by Janev (2009) for instance deals with the topic *Maturity and Applicability Assessment of Semantic Web Technologies*, which includes technologies and tools used in the SW field [20]. As benefits, the authors determined “data reuse and sharing, improved search, open or incremental modeling, decreased implementation time, and customization to individual cases“, and identified that SW is typically supported by conceptual modeling (e.g. using UML) [20]. Still, ontologies have not yet been developed in a systematic way, which should be reinforced in the future [20].

Sabou (2018) published a SMS on *Semantic Web and Human Computation: The Status of an Emerging Field*, which provided insights into an intersection of two topics [21]. Based on publications from 2008 to 2018, it has matured as a research area as papers moved from conference proceedings to journal articles [21]. The most popular topics within the research area are “ontology engineering and knowledge validation“ [21]. More research would be necessary regarding “reusable tools, semantic annotation, and user interfaces“ [21].

Senthil’s (2022) SMS on *Utilizing Semantic Web Technologies in Healthcare, Virtual Communities, and Ontology-based Information Processing Systems* concentrated notably on ontology creation and reuse as well as on semantic data retrieval in the named fields [22]. Senthil (2022) noted that the “role of semantic web is becoming pervasive“ in those areas, and that publication activity has risen sharply over the last couple of years [22]. SW technologies appeared to bring integration and interoperability capabilities to software projects, and freely accessible ontologies such as *DBpedia* or *schema.org* are frequently used as role models or compatible ontologies to integrate one’s own ontology with [22].

As it could be seen from above, previous works exist on either conceptual modeling or Semantic Web, but the intersection of the two topics has not been systematically surveyed yet. Thus, this paper provides insights into an area that has not be tackled yet. The review of related literature also revealed that many studies are dedicated to specific niche topics or domains while some of them touched upon sub-topics of both CM and SW. Still, a systematic overview at a general level is lacking. This paper aims to elaborate on the research landscape covering the intersection of conceptual modeling and Semantic Web to close the identified research gap, thereby structuring the state of research and pointing to vivid future research directions.

3. Research Questions and Research Methodology

In this section, we introduce the research questions we aim to respond to and the research methodology we follow to derive the answers.

3.1. Research Questions

For the systematic mapping study, seven research questions (RQ) were defined as follows:

- **RQ1:** *How has the research area at the intersection of conceptual modeling and Semantic Web evolved over time in general, and with regard to publications, research type, contribution type, and modeling purpose?*
- **RQ2:** *What are the main contributing institutions and where is the relevant research published?*
- **RQ3:** *Who are the main contributing researchers and research communities in the field, what topics are they focusing on, and how do these research groups interact?*
- **RQ4:** *Are the contributions in the field attributed to foundational research or rather to specific industries / domains, and what kind of conceptual modeling languages are used?*
- **RQ5:** *In what kinds of semantic web technology segments and W3C main area did the contributions occur, which SW standard(s) did they use*?
- **RQ6:** *What value is generated by the combination of conceptual modeling and Semantic Web?*
- **RQ7:** *Which insights can be derived from a cross-taxonomy analysis?*

3.2. Systematic Mapping Study

The outlined research questions are answered by the means of a systematic mapping study based on [23, 24]. According to Kitchenham [24], an SMS is methodologically somewhat related to a systematic literature review (SLR), but emphasizes rather the goal of achieving a "wide overview of the research area" under concern for scientists with regard to publication activity, evolution over time and content, and involves a classification scheme [24]. An SMS covers both qualitative as well as quantitative techniques to depict the research area, and intends to classify the publications in the research area under concern according to taxonomies [25]. The research questions are therefore formulated in a way that research trends, evolution, and publication activity can be observed systematically based on the used taxonomies.

We chose to conduct an SMS to examine the research published at the intersection of conceptual modeling and Semantic Web. This is notably due to the fact that an SMS provides the opportunity to get a multi-faceted understanding of the current state of the research along the chosen dimensions and categories, and to detect content gaps that might constitute viable directions for future research [23, 24]. With the SMS, we seek to systematically present the chosen research area, classify publications, thematically analyze selected publications, and facilitate the understanding of research trends and topics for both new and experienced researchers. Our SMS followed the SMS framework elaborated by Petersen [23], which comprises the following **phases**[1]:

1. **Define research scope**
   The research questions notably influence the research scope. They seek to gather information on research and contribution types at the intersection of conceptual modeling and Semantic Web, to identify the modeling purpose and languages used in the corresponding conceptual models, and to outline the major research communities in the respective research field. This paper explores whether the publications related to a specific industry or domain, or whether they represent foundational research. The semantic web technology segments where the contributions occur and what modeling purpose they served is investigated. This paper further evaluates what benefits can be achieved by the combination of conceptual modeling with Semantic Web.
2. **Conduct search**
   As the objective is to investigate the intersection of conceptual modeling and Semantic Web, the search query contains two parts, i.e. one related to conceptual modeling, and the other related to Semantic Web which are

[1]Note: Literature source refers to all phases mentioned below.

combined using the logical operator “and“. The queries for the two areas draw partially upon knowledge from previous literature (e.g. [21, 26]). Several queries are tried in order to grasp the field best possible, and finally select the most suitable one. The query has to be refined as necessary. This SMS limits itself to publications written in English. The query is then executed in literature search engines (considering title and abstract) such as Scopus[2], IEEE Xplore[3], ACM Digital Library[4], and Web of Science[5].
This SMS considers publications in English. The search was executed in the title and abstract of publications. Fig. 2 depicts the final (refined) search query we executed.

```
TITLE-ABS-KEY ( ( ( {conceptual modeling} OR {conceptual modelling} OR {metamodel} OR {meta-model}
OR {metamodels} OR {meta-models} OR {domain specific language} OR {domain-specific language} OR
{domain specific languages} OR {domain-specific languages} OR {modeling formalism} OR {modelling
formalism} OR {modelingformalisms} OR {modelling formalisms} OR {modeling tool} OR {modelling tool}
OR {modeling tools} OR {modelling tools} OR {modeling language} OR {modelling language} OR {modeling
languages} OR {modelling languages} OR {modeling method} OR {modellingmethod} OR {modeling
methods} OR {modelling methods} OR {modeldriven} OR {model-driven} OR {mde} ) AND ( {knowledge
graph} OR{knowledge graphs} OR {linked data} OR {linked-data} OR {semanticweb} OR {ontolog} OR
{RDF} OR {OWL} OR {SPARQL} OR {SHACL} OR {semantic systems} OR {semantic system} OR
{semantic technologies} OR {semantic technology} OR {RDFS} OR {protege} OR {SKOS} OR {simple
knowledge organisation system} OR {JSON-LD} OR{rule interchange format} OR {semantic modeling} OR
{semantic modelling} OR{linked open data} OR {vocabularies}) ) ) AND
( LIMIT-TO ( SUBJAREA , "COMP" ) )
```

Fig. 2. Final search query (Scopus notation)

3. **Screen papers**
In the screening phase, the criteria with regard to which search results to include or exclude (i.e. deemed non-/relevant), are defined. For example papers from non-computer science areas are excluded, and papers below or above a specified length are excluded. The citation files of the search results are downloaded using the Scopus API, and online interface for the remaining search engines. The BibTeX files are converted to CSV format, and duplicates are removed. For each filtering step, the number of papers involved is tracked. As follows, the abstracts of the remaining publications are downloaded, and analyzed with regard to their relevance. For the documents considered relevant, the full text version is downloaded, and is prepared for reading. At this stage, the publications are ready for mapping to the elaborated classification taxonomies.
In this SMS, the following **inclusion criteria** were applied:

– IC1: Publication is written in English
– IC2: Publication is in the area of computer science
– IC3: Publication is peer-reviewed (i.e. journal article, book, conference proceeding)
– IC4: Publication length >= 4 and <150 pages
– IC5: Relevant abstract

Publications were excluded according to the following **exclusion criteria**:

– EC1: Duplicates based on DOI
– EC2: Duplicates based on title
– EC3: Published before 2005
– EC4: Publication length <4 or >150 pages
– EC5: Non-relevant abstract
– EC6: Duplicates based on manual check

[2]https://www.scopus.com/search/form.uri?display=advanced (last accessed on 19.12.2022)
[3]https://ieeexplore.ieee.org/search/advanced (last accessed on 19.12.2022)
[4]https://dl.acm.org/search/advanced (last accessed on 19.12.2022)
[5]https://www.webofscience.com/wos/woscc/advanced-search (last accessed on 19.12.2022)

The exclusion and inclusion criteria related to the language (see IC1), area (see IC2) and peer-reviewed publication type (see IC3) were already applied in the query. Based on the exported data fields, the further exclusion criteria were applied, which is illustrated in Fig. 3. The number of publications for further review was reduced from initially 5.107 potentially relevant to 484 eventually relevant papers.

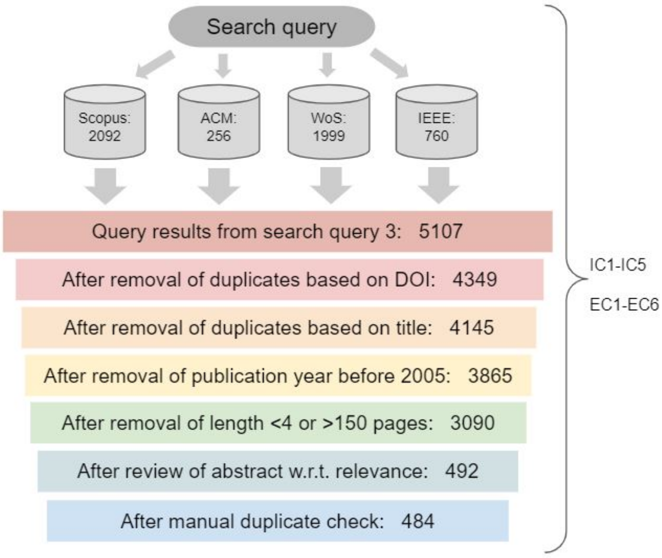


Fig. 3. Publication search and screening process

The reduced list of search results was exported in the form of BibTeX citation files using the Scopus API and the web UI in an automated way. The information on title, abstract, DOI, publication year amongst others were transformed into CSV format, and rendered using Python scripts so that a standardized format emerged.

4. **Keyword abstracts**
The abstracts are analyzed and keywords, which appear to characterize the publications' main content and contribution, are assigned to formulate the classification scheme. In line with classification, taxonomies are created to assign the publications accordingly. The taxonomies comprise several components, namely the W3C main areas of Semantic Web [26], the Semantic Web activity areas [27], the semantic technology segments [20], the Semantic Web standards [28], research types, contribution types, conceptual modeling purposes, and value added by combining CM and SW. The result of this phase is the classification scheme, i.e. the taxonomies [23]. The taxonomies forming the classification scheme of our mapping will be introduced in the following:

W3C Main Areas of Semantic Web Taxonomy

- **Linked Data**[6]:
  "Collection of interrelated datasets" which is available in a standardized format that "provides an environment where applications can query the data, draw inferences using vocabularies", where "relationships among data" are defined, and which is "fit for use" by semantic tools [26, 28].
- **Vocabularies**[7]:
  "Vocabularies define the concepts and relationships (i.e., terms) used to describe and represent an area of concern to classify the terms that can be used in an application, characterize possible relationships, and define possible constraints on using terms". Here, vocabularies also comprise ontologies, which

[6]https://www.w3.org/standards/semanticweb/data (last accessed on 24 November 2022)
[7]https://www.w3.org/standards/semanticweb/ontology (last accessed on 24 November 2022)

are more complex and formal term collections. Vocabularies in the narrower sense do not require such a strict formalism [26, 28].

- **Queries**[8]: "Technologies/protocols that can programmatically get information from linked data" [26, 28].
- **Inference**[9] Inference refers to the "automatic procedures used to generate new relationships based on the data and additional information from vocabularies (e.g., rule sets)". Inference means "reasoning to discover new relationships" [26, 28].

W3C Activity Areas Taxonomy

The activity areas comprises the categories **Foundational**[10], **Application lifecycle management**, **Arts**, **Manufacturing**, **Media**, **Cultural**, **Education**, **Government**, **Energy**, **Financial**, **Tourism**, **Geographical Information System (GIS)**, **Healthcare**, **IT**, **Legal**, **Life sciences**, **Oil and gas**, **Service management**, **Telecommunications**, **Utilities**[11].

Semantic Technology Segments Taxonomy

- **Semantic data management and integration**: "Ontology-driven information systems and server platforms that enable RDF triple storage, semantic data / service integration and management, semantic interoperability based on W3C standards" [20]. Janev's publication on the "maturity and applicability assessment of Semantic Web technologies" was used as a foundation for this taxonomy [20].
- **Semantic modeling and development**: "Tools that enable design and development of ontologies, RDF or OWL knowledge stores, and tools for semantic services applications development" [20].
- **Semantic collaboration incl. portal technologies**: "Portals based on semantic standards, semantic wiki technology; solutions that support social networking, data aggregation, dynamic publishing of contents" [20].
- **Learning and reasoning**: "(OWL) reasoners, ontology learning tools, rule engines" [20].
- **Semantic annotation**: "Technologies that support automatic semantic annotation, information extraction, text mining, other language processing tasks" [20].
- **Semantic search and retrieval**: "Semantic data access and search tools based on W3C standard query languages, semantic search engines based on NLP and technologies including content classification and clustering; fact and entity extraction, taxonomy creation, and management (tagging engines); knowledge presentation" [20].

Semantic Web Standards Taxonomy

- **RDF**: A "standard model for data interchange and linking on the web" and provides schemas to enable easier data integration and is represented in graph notation[12].
- **OWL**: A "Semantic Web language" used for representation in the form of knowledge graphs, and logical knowledge including the relationships between its parts[13].
- **SPARQL**: A query language used to extract relationship data and graph data from knowledge graphs, and semantic data formats[14].
- **RDFa**: RDFa means "RDF in Attributes" and constitutes a "specification for attributes to express structured data in HTML5, XHTML, and any XML application"[15].

[8] https://www.w3.org/standards/semanticweb/query (last accessed on 24 November 2022)
[9] https://www.w3.org/standards/semanticweb/inference (last accessed on 24 November 2022)
[10] General foundational research in the context of Semantic Web technologies, which is not specifically tailored to an application domain.
[11] https://www.w3.org/2001/sw/sweo/public/UseCases/ (last accessed on 24 November 2022)
[12] https://www.w3.org/2001/sw/wiki/RDF (last accessed on 24 November 2022)
[13] https://www.w3.org/2001/sw/wiki/OWL (last accessed on 24 November 2022)
[14] https://www.w3.org/2001/sw/wiki/SPARQL (last accessed on 24 November 2022)
[15] https://www.w3.org/2001/sw/wiki/RDFa (last accessed on 24 November 2022)

– **JSON-LD**: Refers to a JSON format which is tailored to the use in the context of linked data[16].
– **SKOS**: It means Simple Knowledge Organization System and “is a common data model for sharing and linking knowledge organization systems“ online[17].
– **RDFS**: RDFS is a linked data format, refers to the schema and “represents simple vocabularies online“ which constitutes the foundation for more complex ontologies[18].
– **GRDDL**: A “technique for obtaining RDF data from XML documents and in particular XHTML pages“, and open to integrate further algorithms and procedures necessary to handle semantic data[19].
– **POWDER**: Offers “a mechanism to describe and discover Web resources and provide a succinct way to define any number of predicates for those resource“ to more integrate data into big linked data systems[20].
– **PROV**: A “provenance specification“ which enables the “exchange of provenance information“ in linked data format[21].
– **RIF**: A rule interchange format which helps to “interchange rules between different logical-based systems“[22].
– **SAWSDL**: It “defines extension attributes for WSDL and XML schema definition language that allows description of additional semantics of WSDL components, which specifies how semantic annotation is accomplished using references to semantic models“[23].
– **RDB2RDF**: RDB2RDF is a “collection of two Recommendations to map the content of relational databases to RDF“, which uses “direct mapping and R2RML“ as mapping languages that transform linked data from one to another format[24].
– **SHACL**: It is a “standard language for describing shape of RDF data which is used for validating conditions in a linked data and graph data setting by referring to numeric ranges, string patterns, values, and the like“[25].

Research Type Taxonomy

– **Experience**: “Explain on what and how something has been done in practice, referring to personal experience of author(s)” [23].
– **Evaluation**: “Observation of how a technique is implemented to solve a research problem (solution implementation and measure consequences in terms of benefits and drawbacks)” [23, 29].
– **Solution**: New solution or extension of existing solution to a problem, whose applicability is shown by an example or a solid argumentation [23, 30].
– **Vision**: “Non-disruptive research agenda setting papers” [31].

Contribution Type Taxonomy

– **Discussions**: “Investigations without constructive contributions (e.g., reviews, comments, opinions)” [31].
– **Concepts**: “Suggestions of ways of thinking (e.g., meta-models, frameworks, taxonomies) that are constructed from a set of statements, assertions, or other concepts” (including mathematical theories) [23, 30].

[16]https://www.w3.org/2001/sw/wiki/JSON-LD (last accessed on 24 November 2022)
[17]https://www.w3.org/2001/sw/wiki/SKOS (last accessed on 24 November 2022)
[18]https://www.w3.org/2001/sw/wiki/RDFS (last accessed on 24 November 2022)
[19]https://www.w3.org/2001/sw/wiki/GRDDL (last accessed on 24 November 2022)
[20]https://www.w3.org/2001/sw/wiki/POWDER (last accessed on 24 November 2022)
[21]https://www.w3.org/2001/sw/wiki/PROV (last accessed on 24 November 2022)
[22]https://www.w3.org/2001/sw/wiki/RIF (last accessed on 24 November 2022)
[23]https://www.w3.org/2001/sw/wiki/SAWSDL (last accessed on 24 November 2022)
[24]https://www.w3.org/2001/sw/wiki/RDB2RDF (last accessed on 24 November 2022)
[25]https://www.w3.org/2001/sw/wiki/SHACL (last accessed on 24 November 2022)

- **Methods**: "Suggestions of new ways of doing things (e.g., applying existing models) by means of actionable instructions that are conceptual (not algorithmic)" [23, 30].
- **Algorithms**: "Suggestions of new automatic ways of computing (e.g., model transformation) or measuring things (e.g., metrics) by means of formal logical instructions" [30].
- **Tools**: "Presenting novel software tools (e.g., modeling tools)" [23, 31].

Modeling Purpose Taxonomy[32–39][26]

- **Representation**: "Creation of abstract representations of the system under study (descriptive modeling)".
- **Analysis**: "Analysis of properties of the system under study by means of e.g., simulations or queries".
- **(Re-)Design**: "(Re-)design of future version of the system under study".
- **Code Generation**: "Generation of code (parts) that can be executed to realize a (software) system".

Conceptual Modeling Language Taxonomy

The conceptual modeling languages used are recorded for each paper so that their popularity and application in combination with the further taxonomies could be analyzed later on.

Value Added of Combining SW and CM Taxonomy

- **Representation flexibility**: "Any extant data structure or format can be represented as RDF. RDF can readily express information contained within structured (conventional databases), semi-structured (Web page or XML data streams), or unstructured (documents and images) information sources" [40].
- **Incremental schema and modeling**: "Semantic technologies, on the other hand, allow domains to be captured and modeled in an incremental manner. As new knowledge is gained or new integrations occur, the underlying schema can be added to and modified without affecting the information that already exists in the system. It is a benefit that enables experimentation and lowers risk" [40].
- **Interoperability of multimedia metadata**: SW technologies can help to make metadata from different, otherwise not compatible sources interoperable, and W3C standards using formal semantics can be used for this purpose [41, 42].
- **Enhanced inference capabilities**: SW technologies extend the reasoning capabilities of CM using formal logic (e.g., description logic) to make inferences based on ontologies. Models that include OWL and RDF(S) have the necessary formal foundations [42, 43].

5. **Extract and map data**

The relevant publications are mapped to the classification scheme along the dimensions stated in the previous step. The mapping is subject to a feedback round, and review to capture possibly occurring gaps or errors. Based on the mapping results, a content analysis is carried out as necessary, and a series of plots, tables, and figures are created to depict the results in a visually appealing and systematic way. The output of this phase is the systematic mapping of research publications according to the taxonomies [23].

For this purpose, an online spreadsheet was created which contained the exported data in one tab, including the categorization per taxonomy with each taxonomy in one separate column respectively. One tab per taxonomy was created to give an early, short overview on the frequency per classification category.

After the data extraction, the data was cleaned, formatting was aligned, and integrity checks on the data were performed. This dataset comprises a series of attributes such as university and country where the researchers were located at the time of publication, their Scopus ID, the publication's abstract, year, document type, publication channel, DOI, number of pages, relevant documents, and all taxonomies. All in all, it comprised 30 features for 484 publications.

4. Findings

4.1. Bibliographic Analysis

The systematic mapping study considered publications ranging from the year 2005 to 2022. As Fig. 4 shows, the number of publications at the intersection of conceptual modeling and Semantic Web has grown over the last two decades from around 3 to 10 publications in the 2000s to over 50 per year in the late 2010s. The trend is an overall positive one, with small exceptions in 2007, 2011, and 2015 (see Fig. 4). In 2022, the 11 publications that are depicted have been published until May, so a larger number is expected for the entire year.

Fig. 4. Nr. of publications per year

Fig. 5 splits the publications into journal articles and conference papers, which shows that the number of conference papers has grown much faster than the number of journal articles up to the year 2019. This confirms that the field was growing but was not maturing until then. Since 2019, the number of conference papers published has come down to a level similar to the number of journal articles, which indicates that the research in the field is starting to mature in recent years (see Fig. 5).[27]

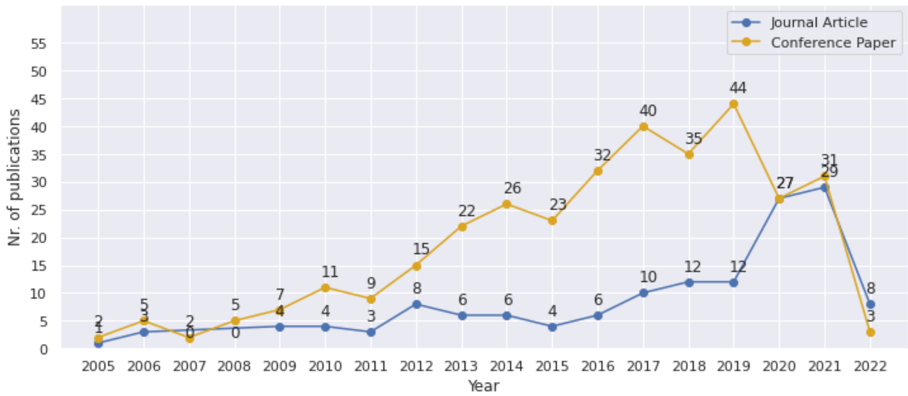


Fig. 5. Nr. of publications per year and type

The top 10 institutions in terms of number of publications count are depicted in Table 3. The Federal University of Espirito Santo from Brazil and the University of Vienna from Austria stand out with over 20 publications

[27] For reasons of simplicity and better visibility of the remaining categories, the two book chapters were not shown in Fig. 5.

Table 3

Top 10 contributing institutions based on nr. of publications

Institution	Country	Count
Federal University of Espirito Santo	Brazil	23
University of Vienna	Austria	21
Free University of Bozen-Bolzano	Italy	14
Babes-Bolyai University	Romania	10
Vienna University of Technology	Austria	8
Kaunas Institute of Technology	Lithuania	8
University of Leipzig	Germany	7
Wuhan University	China	7
Karlsruhe Institute of Technology	Germany	6
Polytechnical University of Valencia	China	6

Table 4

Most used conference publication channels (left) and Most used journal publication channels (right)

Publication channel (conference)	Count
Conceptual Modeling, ER	17
IEEE Conference on Emerging Technologies & Factory Automation	6
Winter Simulation Conference	5
Model and Data Engineering	5
Conference on Model-Driven Engineering & Software Development	5
IEEE Enterprise Distributed Object Computing Workshop	4
Conference on Knowledge Discovery, Engineering & Management	4
Semantic Web	3
Procedia Computer Science	3
IEEE Aerospace Conference	3
Federated Conference on Computer Science & Information Systems	3
International Semantic Web Conference	3

Publication channel (journal)	Count
Journal of Biomedical Informatics	6
Expert Systems with Applications	5
IEEE Transactions on Services Computing	5
Journal of Systems and Software	4
Semantic Web	4
Applied Sciences	4
Data and Knowledge Engineering	4
Information Systems	3
IEEE Access	3
Advanced Engineering Informatics	3
Advances in Production Management Systems	3

each, followed by the Free University of Bozen-Bolzano (Italy), Babes-Bolyai University (Romania), and TU Wien (Austria) with at least 10 publications each, according to Table 3.

The main publication channels were conference proceedings and journal articles. With regard to conference papers, the most found conferences, in line with which the most publications are shown in Table 4 left. *Conceptual Modeling, ER* was the leading conference among the analyzed publication channels, followed by the *IEEE International Conference on Engineering Technologies and Factory Automation*, and the *Winter Simulation Conference*. The journals in which the topic has been most prevalent are depicted in Table 4 right. The *Journal of Biomedical Informatics*, *Expert Systems with Applications*, and *IEEE Transactions on Services Computing* were among the three most used journals.

4.2. Content Analysis

This section contains analyses based on the taxonomies considered independently and over time, as well as the abstracts of the relevant publications. Fig. 6 illustrates the development of the number of publications split by research type over time from 2005 to May 2022, which indicates that the publications in each category have increased. The major part is formed by publications of the *solution* type, i.e. it presents a "new solution or extension of existing solution to a problem, whose applicability is shown by an example or a solid argumentation", which rose fast from the 2010s onwards compared to all other types [23, 29]. The *evaluation* and the *experience* type also saw rather steady, but not high growth, as shown in Fig. 6. The *vision* type, i.e., "non-disruptive research agenda setting papers" [31], grew up to around 2017, and then declined.

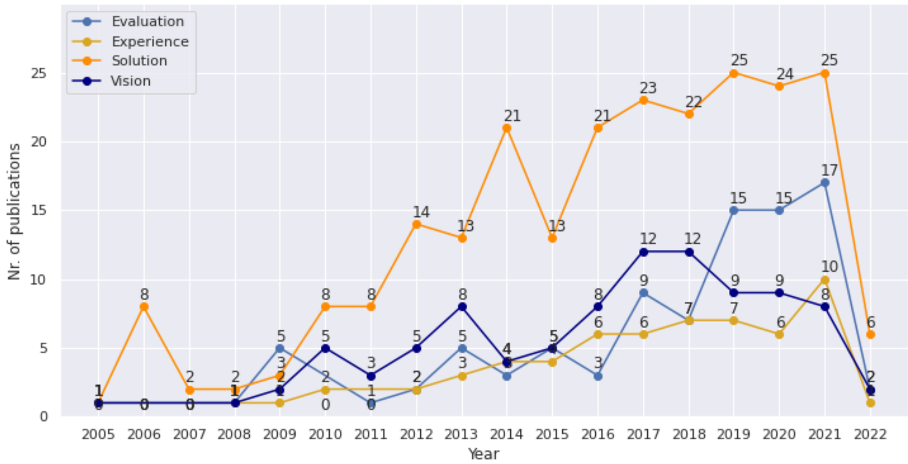


Fig. 6. Nr. of publications per year and research type

The number of publications has developed differently depending on the contribution type, which is illustrated in Fig. 7. The largest part of publications have recently appeared to provide new or adaptations of existing *methods*, which considerably increased in numbers over the last two decades. The number of *concepts* has not frequently been among the largest contribution types, but has recently become more popular, and came in second in 2021. The number of *discussions* peaked in 2017 to 2018, but has since declined again, while the publications presenting *algorithms or tools* have been growing in recent years, but at a lower level (see Fig. 7).

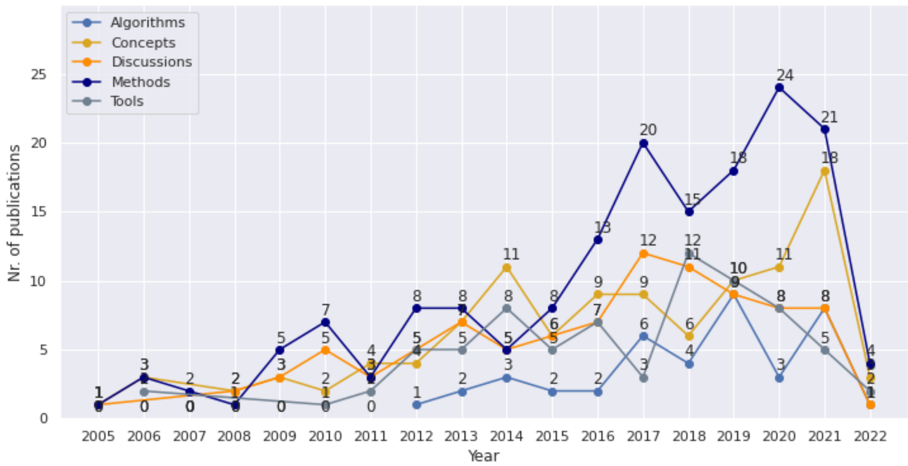


Fig. 7. Nr. of publications per year and contribution type

As for the conceptual modeling part of the intersection topic, UML (used in 231 publications) is by far the most frequently used conceptual modeling language, followed by any kind of DSL (125), and BPMN (59). Semantic Web Rule Language (SWRL, 38), OntoUML (33), SysML (31), ER (31), and OCL (30) were among the most used out of over 100 modeling languages mentioned in the publications.

In the next step, the modeling purpose taxonomy was applied to the publications. According to Fig. 8, the publications were split into four categories, which revealed *representation* and *analysis* as the most recent major modeling purposes among the publications. These two categories were leading the modeling purpose most of the time period analyzed, but not all as exceptions occurred in the early 2010s, and around 2018 to 2019 (see Fig. 8). *Code generation* as a modeling purpose rose before 2015, then shortly declined, and finally rose again up to 2019, to then decline again. The *(Re-)Design* modeling purpose occurred at a low level, but never considerably increased.

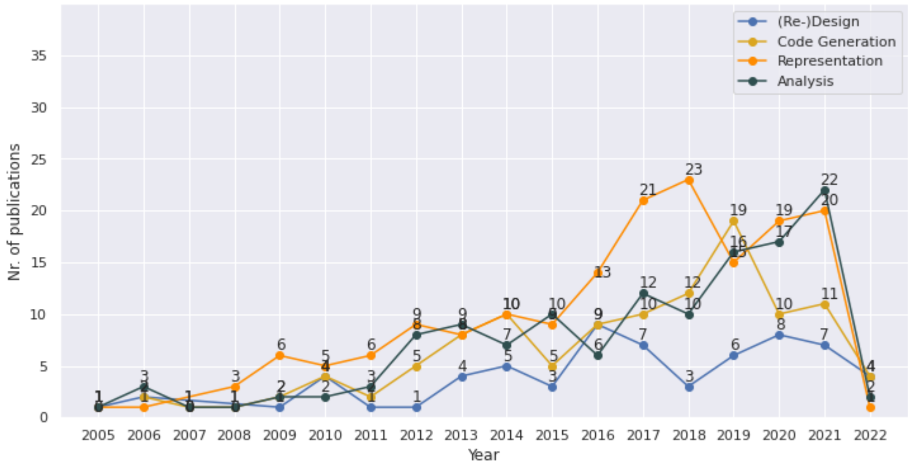


Fig. 8. Nr. of publications by modeling purpose

When considering the W3C main areas of Semantic Web overall, the largest part of the papers were related to *Linked Data* and *Vocabularies*, while only a lower share can be attributed to *Inference*, and *Queries*. The picture is still quite similar once the time component is also considered, as Fig. 9 shows. In the mid-2000s, all categories started from a low level, while the number of publications on *Linked Data* and *Vocabularies* increased considerably after 2011, the number of publications on *Inference* and *Queries* achieved merely a slightly higher level in this time period. It also has to be added at this point that the development of the number of publications in each one of the categories was not steady, but exhibited several increases and decreases (see Fig. 9).

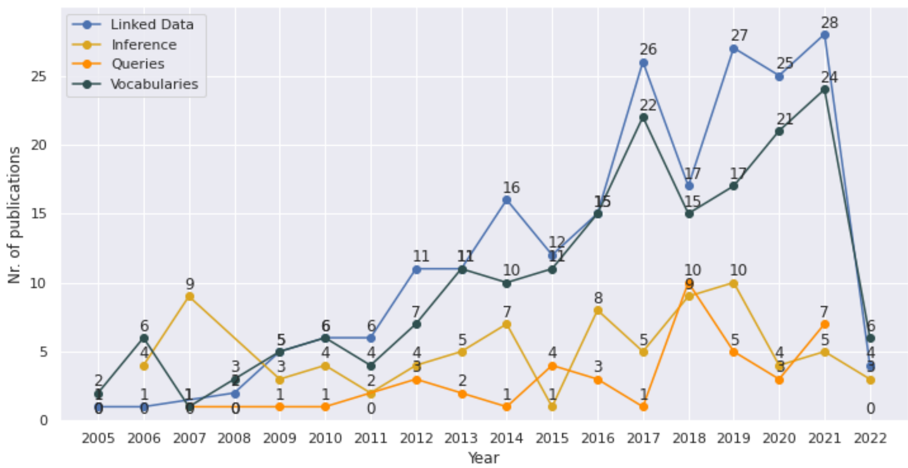


Fig. 9. Nr. of publications by W3C main area and year

The overall most frequently semantic technology segment (see Fig. 10) (i.e., in 286 publications) is *semantic modeling and development*. 121 publications dealt with *learning and reasoning* (i.e., inference theories and engines), 104 with *semantic data management and integration*, and 95 with *semantic annotation* (i.e., adding or extracting meaning from text or other underlying data) [20, 27]. The lowest number of publications, i.e., 35 papers, were related to *semantic search and retrieval*.

The publications were also categorized according to Semantic Web activity areas either in *foundational* or specific industry background works. One publication was assigned to only one Semantic Web activity area. 187 out of the 484 publications, i.e., 38.6%, were of foundational nature, while the remaining 61.4% are split across specific

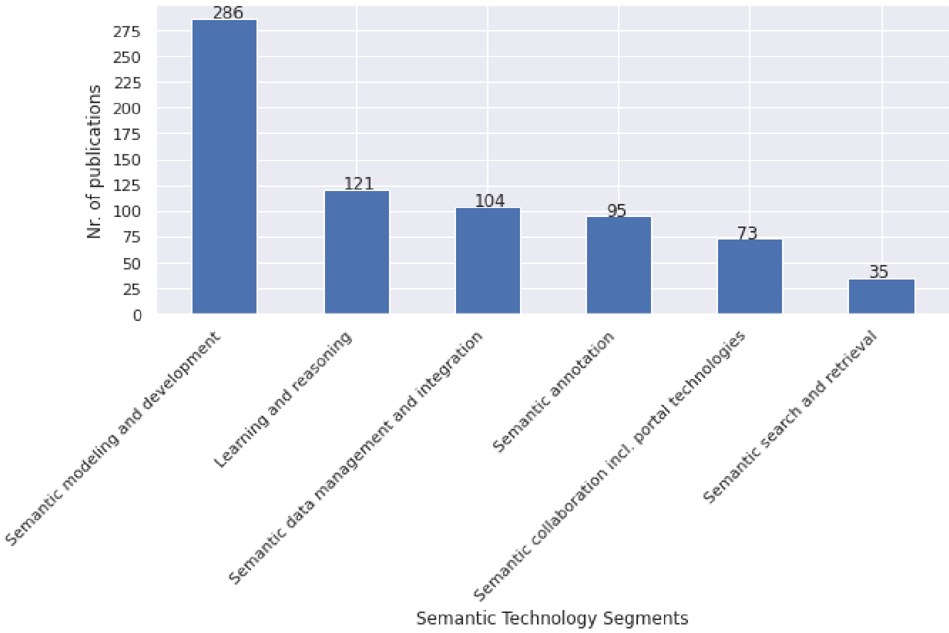


Fig. 10. Nr. of publications by technology segment

domains. Among most prevalent domains-specific Semantic Web activity areas are *manufacturing*, *IT*, *healthcare*, *education*, *GIS*, *cultural*, and *government* in this given order , according to Fig. 11.

Regarding Semantic Web standard taxonomy, the largest part of the publications were related either to *OWL* or *RDF* or both (see Fig. 12). Here, one publication could refer to either one, several, or even none of the standards. The next most occurring SW standards were *SPARQL*, *RDFS*, *JSON-LD*, and *RIF*. 111 out of the 484 publications did not contain any reference to a Semantic Web standard from the taxonomy. This does, however, not necessarily mean that they did not relate to any standard, but that just in the abstract, title, and full text no standard was mentioned.

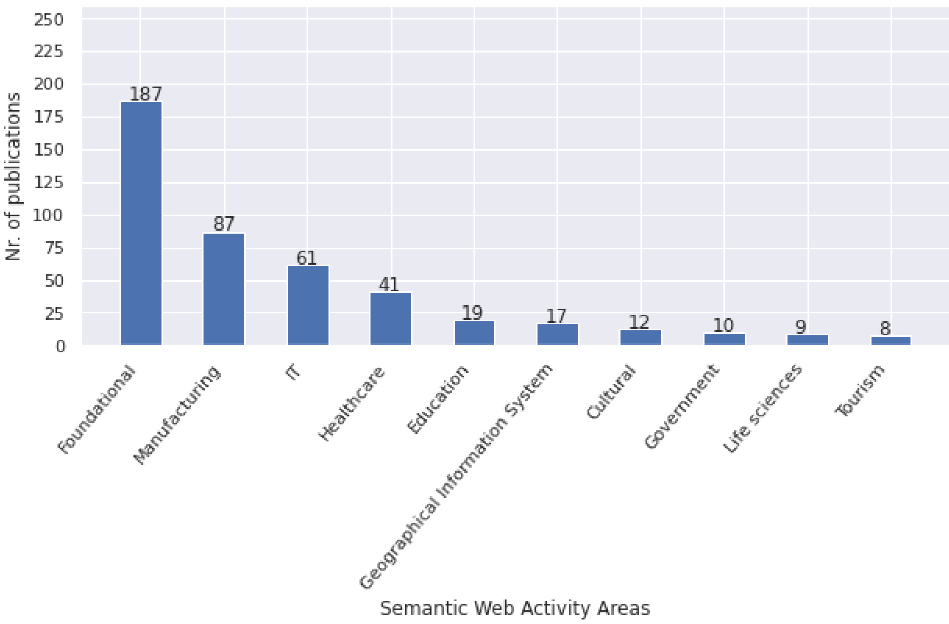


Fig. 11. Nr. of publications by Semantic Web activity area

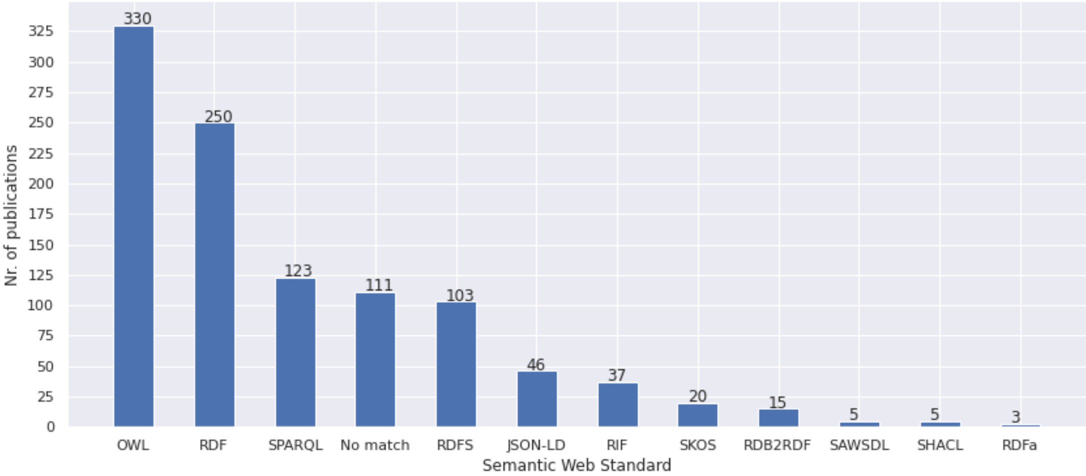


Fig. 12. Nr. of publications by Semantic Web standard

4.3. Combined Analysis

The combined analysis considers both the value of combining Semantic Web and conceptual modeling and the combination of several taxonomies.

Over time, publications mapped to *incremental schema and modeling* evolved fastest among the different value added options, and peaked in 2021 with 49 publications, as Fig. 13 depicts. The remaining three value added elements have developed over time from around two to three per year to around 15 to 20 each per year up to 2021, but their prevalence changed depending on the year.

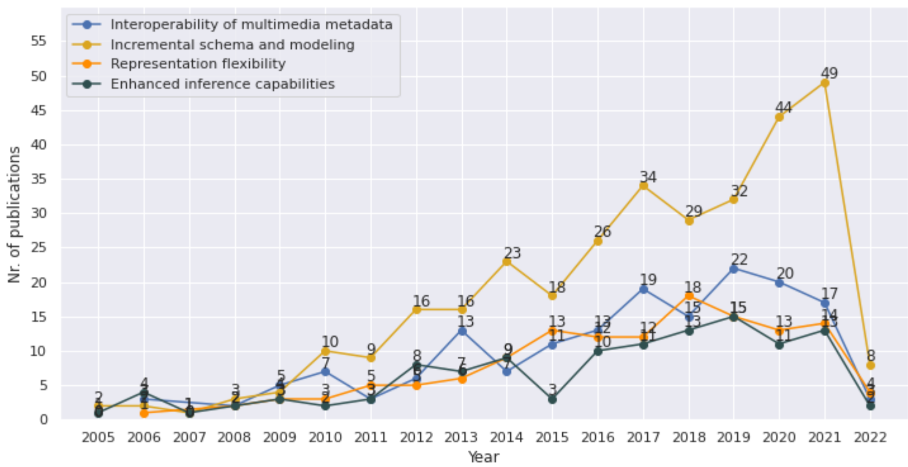


Fig. 13. Value added by combining SW and CM over time

The bubble plot in Fig. 14 shows the longitudinal change of combining the W3C main areas of Semantic Web with the modeling purpose. The plots separate publications before 2015 (on the left in orange), and since 2015 (on the right side in blue). The development over time indicates that combinations of modeling purpose with *inference* or *queries* have tendentially stayed at a similar level or decreased, but combinations with *linked data* in general and *vocabularies* have increased considerably, according to Fig. 14.

Fig. 15 illustrates the combination of the contribution type with the modeling purpose taxonomy, which indicates a concentration of papers along *representation* modeling purpose combined with *discussions* or *concepts* contribution type. The contribution type of *methods* rather appears in combination with the modeling purpose of *code generation*

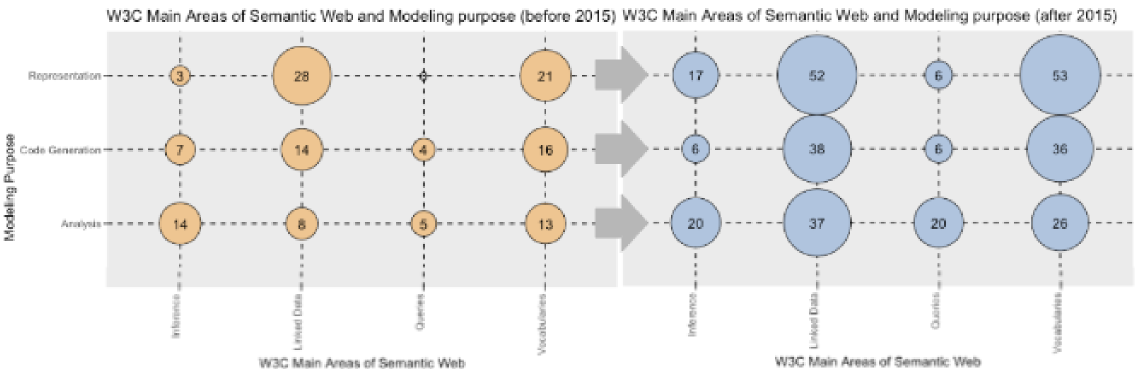


Fig. 14. Development of taxonomy combination over time

or *analysis*, and the contribution type *tools* is mostly combined with *code generation* as a modeling purpose. Over time notably the combination of *methods* with *representation* have grown considerably, as well as in general all of the largest combinations mentioned above. However, the combinations of taxonomy elements in the lower left corner exhibited a significant decrease over time.

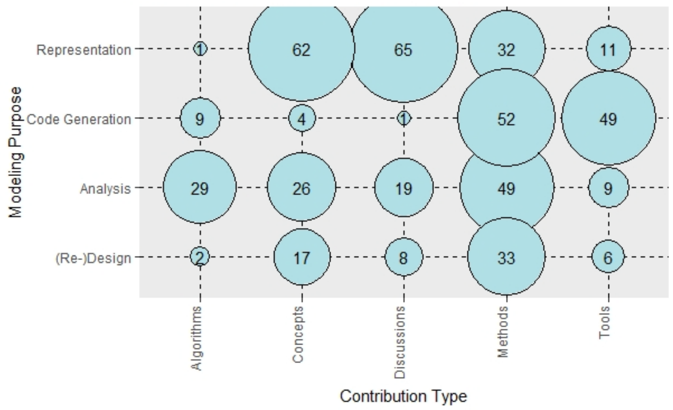


Fig. 15. Nr. of publications by contribution type and modeling purpose

When it comes to the combination of Semantic Web standards with conceptual modeling languages, it appears that UML, DSL, BPMN as modeling languages stand out in combination with the Semantic Web standards OWL, RDF(S), and SPARQL (see Fig. 16). Over time, all of those main combinations have increased considerably, and additionally the standard JSON-LD became visible more frequently in combination with UML, ER, OntoUML, DSL, BPMN. The combinations with the Semantic Web standards RIF, RDB2RDF, and SKOS slightly increased over time, whereas those with SAWSDL disappeared over time.

Most publications concentrate at the combination of firstly UML and secondly DSL with all kinds of value added opportunities (cf. Fig. 17). BPMN combined with the value added types *incremental schema and modeling*, *interoperability of multimedia metadata*, and *representation flexibility* accounts for a large part of combinations. The combinations related to OntoUML, OCL, ER, ArchiMate, and AML are at a lower range, but are quite evenly spread across the value added options. As for the development over time, the strongest growth was observed in UML and DSL combined with *incremental schema and modeling*, and UML with *interoperability* and *representation flexibility*. Comparatively low growth was recorded for Petri Nets, ArchiMate, OCL, and ER overall. The *enhanced inference capabilities* have recorded major growth combined with OCL, ER, DSL, and BPMN. As for *representation flexibility* as a value added category, it has grown for ArchiMate, AML, and OntoUML, but has stayed stable at a low level for ER, OCL, and Petri Nets.

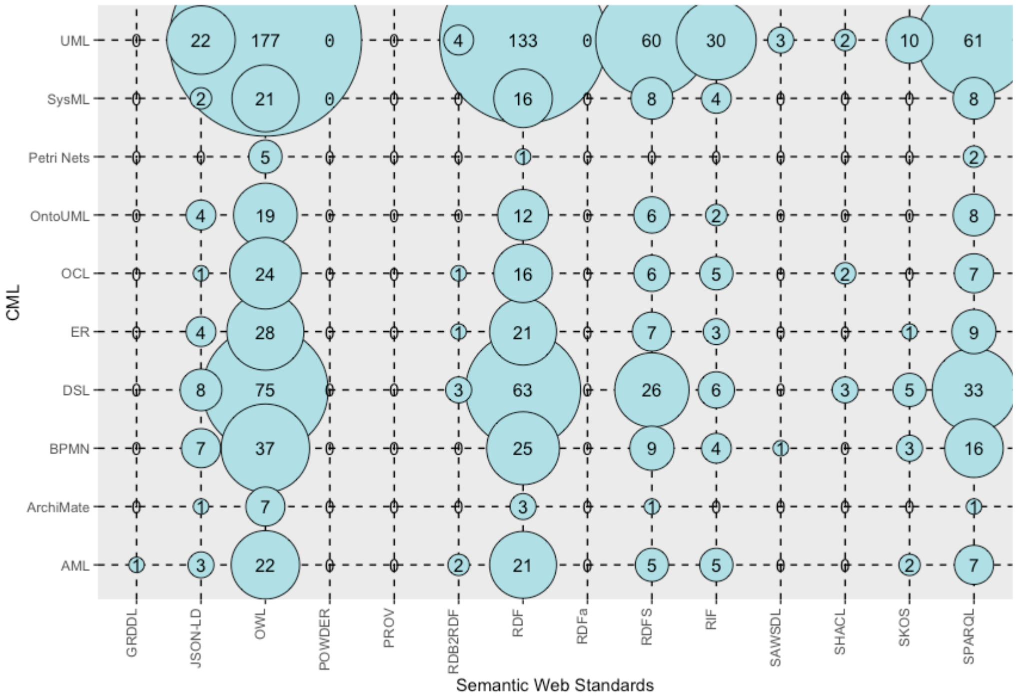


Fig. 16. Nr. of publications by W3C standard and conceptual modeling language

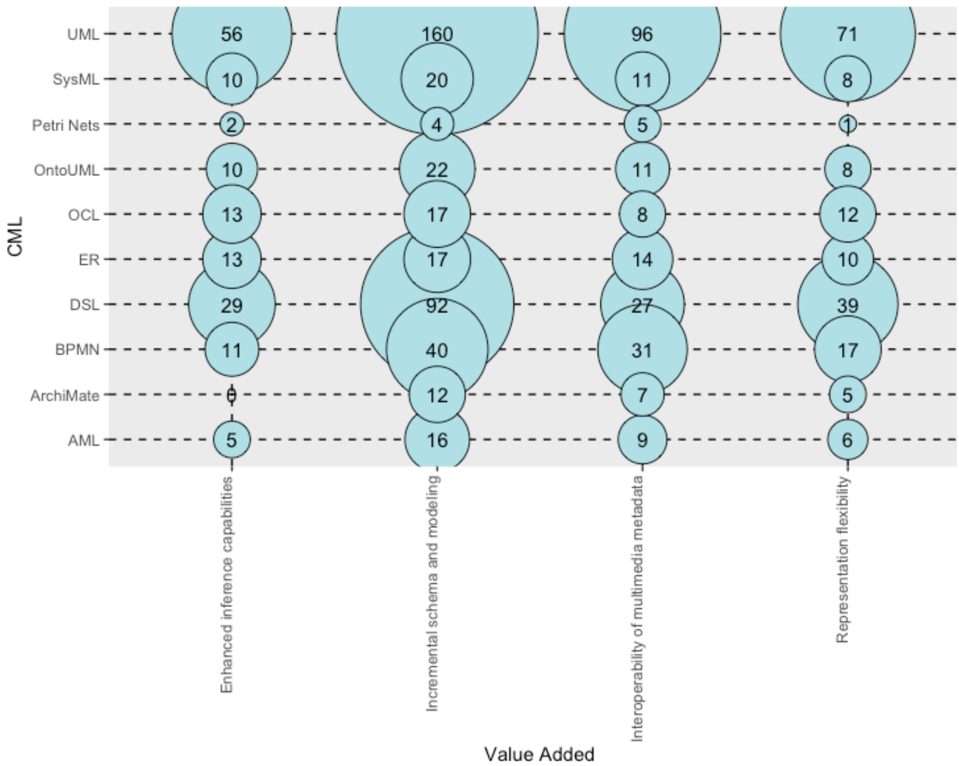


Fig. 17. Nr. of publications by value added and modeling language

4.4. Research Community Analysis

In line with the research community analysis, the relationships between researchers at the intersection of conceptual modeling and Semantic Web topics were explored and their main topics were identified. Fig. 18 depicts the whole publication landscape used for this SMS in the form of a co-authorship graph weighted by documents. The main research communities are highlighted in color and are shown in bigger font size according to their document output weighting (see Fig. 18). Some of the largest research clusters are structured around the researchers M. Wimmer, R. Verborgh, T. Walter, D. Gasevic, M. Malki, G. Guizzardi, R. A. Buchmann, J. Sun, G. Kardas, H. Paulheim, X. Zheng, and X. Wang.

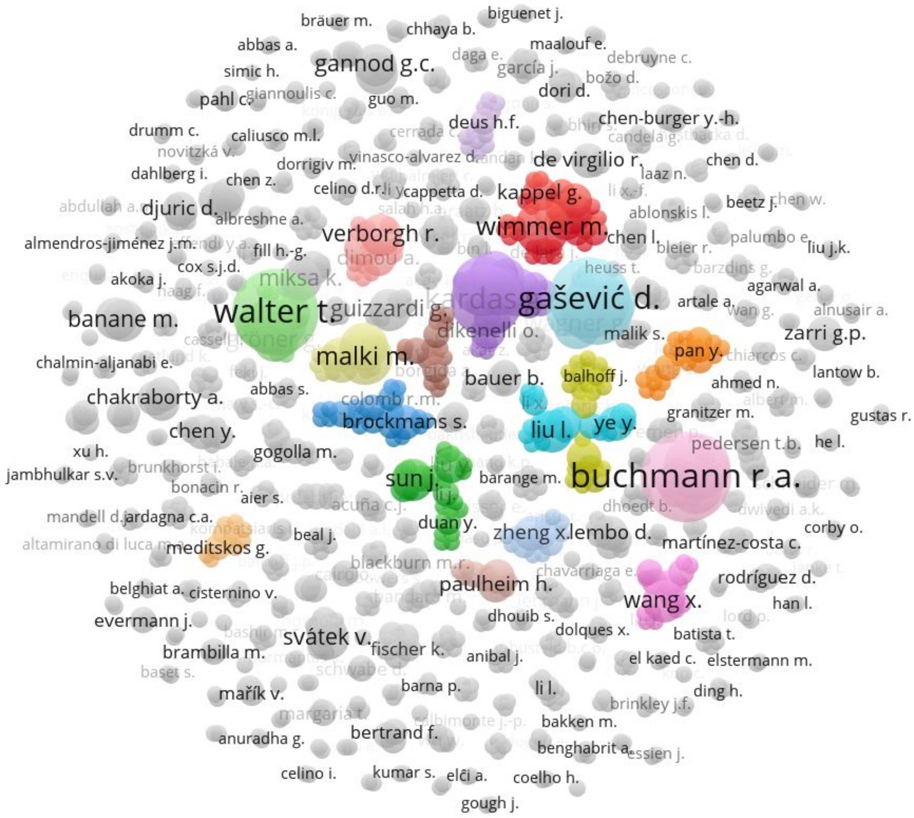


Fig. 18. Co-authorship graph weighted by documents

Next, we drilled down on some of the active research communtities (see Fig. 19). *Manuel Wimmer*, whose community is depicted in red in Fig. 19, is currently a professor at the Johannes Kepler University Linz (Austria), but was at the time of his publications employed at TU Wien (Austria). Some researchers and professors co-authoring papers with him are for instance G. Kappel, E. Kapsammer, W. Schwinger, W. Retschitzegger, J. Delara, M. Sabou, S. Biffl, F. Ekaputra, and O. Kovalenko, many of whom were colleagues at TU Wien (Austria) or Johannes Kepler University Linz (Austria). The most prevalent topics of this community has been involved were model transformations, graph grammars, and UML, all with a focus on an industrial context. *Geylani Kardas's* community is depicted in violet in Fig. 19, and is linked to M. Challenger, S. Getir, T. Kosar, M. Mernik, and A. Goknil, amongst others, who are mainly working in research at Ege University (Turkey). The main topics covered by this research community evolve around agent-oriented software engineering, multi-agent systems, and domain-specific modeling. The third cluster in Fig. 19 is highlighted in light blue and is structured around *Dragan Gasevic* from Monash University (Australia) and includes V. Devedzic, G. Wagner, A. Giurca, and H. S. Carvalho. Gasevic's community devotes to self-regulated learning and modeling in an educational context.

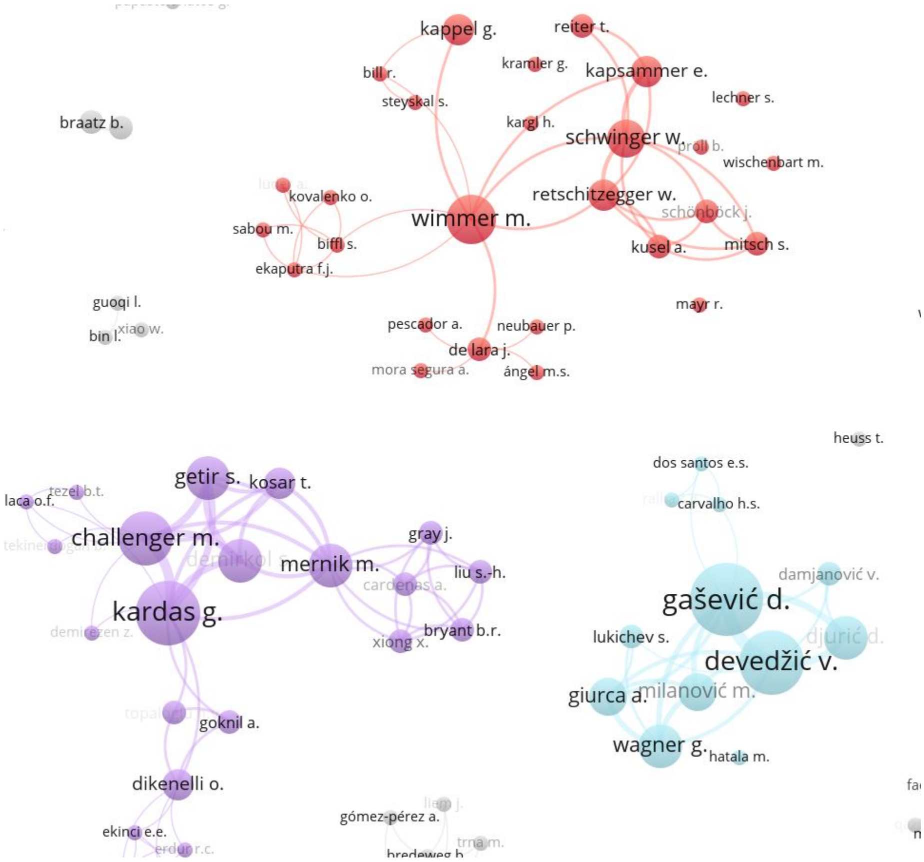


Fig. 19. Co-authorship graph: clusters Wimmer, Kardas, Gasevic

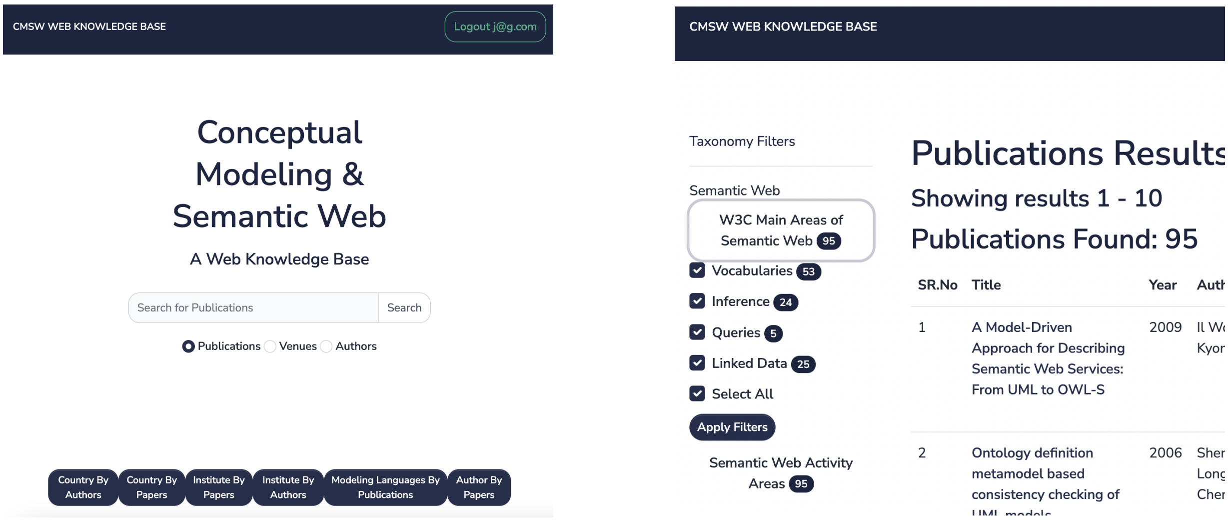
5. Web Knowledge Base

Complementary to the systematic mapping, a web knowledge base containing information on the relevant publications was created to enable researchers to retrieve an excerpt and related meta data of the results that they are interested in. The Web Knowledge Base aims to provide self-service functionalities to explore the data of the systematic mapping study. Functionalities comprise the detailed analysis results with regard to the countries and modeling languages as well as concerning the taxonomies. Moreover, it offers search opportunities so that publications can be retrieved according to search terms and taxonomies. The Web Knowledge Base is accessible via http://me.big.tuwien.ac.at/cmsw.

The landing page of the Web Knowledge Base offers users the opportunity to enter a search key word for which a list of publications that contain the search term are retrieved. The search term refers either to publications, venues, or authors. On the lower side of the page, frequency tables concerning the *country by authors*, *country by papers*, *institute by papers*, *institute by authors*, *modeling languages by publications*, and *author by papers* can be displayed based on the SMS data.

The results of an exemplary search for the search term *UML* are displayed in Fig. 21. In this case, 95 results, i.e., publications, were found in the SMS data. The title, the publication year, the authors, the abstract, and the badges (meaning the taxonomy elements assigned to the respective publications) are shown in the resulting columns. By clicking the button in the abstract column, the abstract of the selected publication is shown.

The menu on the left side also provides the *taxonomy filters*, by which some taxonomy elements can be selected or unselected as the user prefers. By clicking the button *Apply Filters*, the selection can be confirmed and the publication results are displayed according to the filters set, whose selection is shown in Fig. 20b.



(a) Landing page (b) Taxonomy filters

Fig. 20. Landing page (left) and taxonomy filters of the Web Knowledge Base

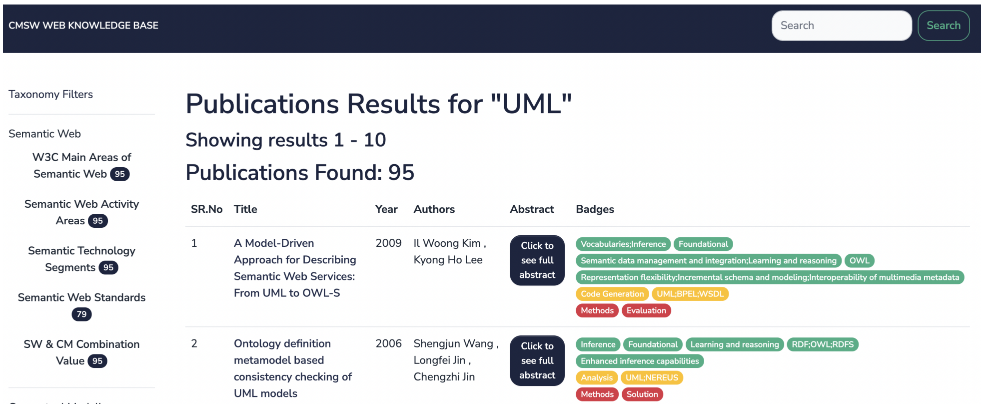


Fig. 21. Search

6. Threats to Validity

The term "*validity*" refers to degree of reliability and correctness of the results of the systematic mapping study [44, 45]. Hence, in this chapter potential limitations, i.e., "threats to validity", of the SMS are reviewed from a critical perspective. Validity can be categorized into several sub-types, namely *conclusion validity*, *internal validity*, *construct validity*, and *external validity* [44] which will be reviewed in the following.

Conclusion validity deals with the "relation of the research process to the outcomes" [44] and its replicability by the means of using an appropriate systematic research method [17]. A potential threat in this category is presented by how the search for papers was conducted in this paper. As the selection of specific inclusion criteria (IC1-IC5) or exclusion criteria (EC1-EC6) might have had a considerable impact on the number of papers and on the content of papers that finally remained in the list of publications subject to analysis. The screening phase also involved only title and abstract of the publications, and no further attributes. The SMS was based on literature search runs in the publication databases ACM Digital Library, Web of Science, Scopus, and IEEE Xplore, which is only a selection of the available publication databases, and might thus limit the generalizability of the conclusions. Precautions to ensure conclusion validity was the systematic execution of the methodologically given phases, with the goal to draw conclusions once the method allows for it.

Internal validity refers to the fact that there might be “a relationship between the treatment and the outcomes“ [44], or to the “extraction of information“ from the underlying data [45]. Threats to internal validity might have occurred in the process of evaluating the publications as relevant or non-relevant, and in the process of assigning taxonomy elements to the papers based on the abstracts and / or full text. At this point, it might have happened that relevant papers were overlooked or non-relevant papers have inadvertently been added to the final list of selected publications. Moreover, taxonomy elements might have been incorrectly or contradictorily assigned to the publications (e.g., due to skipped information or poor analysis of the content) so that the internal validity or correctness could have been hampered.

Construct validity refers to “the relationship between the theory and the observations, while reflecting the researcher’s initial expectations“ could have been negatively impacted [44]. Construct validity is affected by the research design and the formulation of adequate research questions [45]. The search queries were compiled using previous related research as a yardstick, but it could have happened that relevant terms or synonyms might have been left out, and therefore led to papers not being included. Relevant combinations of terms might have been overlooked in the search string definition process so that the RQ might not be fully answered. To prevent from this threat, the selection process of the publications, and classification as relevant or not based on abstract and title was carried out once and then reviewed in a second correction run. Likewise, the classification according to the taxonomies was carried out and reviewed to spot misclassifications and contradictory data entries. The search query was adapted and tested with different search terms to identify terms that could largely impact the number of publications in the search result.

Threats to *external validity* finally refer to issues regarding the “generalizability of the results outside the scope of the study“ [44]. The SMS might be limited in generalizability as not all relevant papers might have been captured in the selection process, and / or not sufficient or sufficiently targeted previous related work might have been analyzed upfront as a basis. The relevant papers were also selected from a time period from 2005 up to May 2022, which means that after that limit further critically relevant papers might have been published outside the chosen time frame, and are now missing in this study. Another external validity problem might be the fact that some rare case happened in which papers that looked relevant were unavailable, and therefore had to be excluded from the analysis, which means that their contribution to the research field is now missing. However, this SMS covers a very specific topic so that threats to generalizability are as such of lesser importance compared to the other threats to validity, as the SMS aims to be representative for its narrow research topic at the intersection of conceptual modeling and Semantic Web.

7. Discussion

7.1. Response to Research Questions

RQ1 aimed to explain how the research area at the intersection of conceptual modeling and semantic web evolved over time in general, and with regard to publication, research, contribution type, and modeling purpose. The data analysis indicated that the research area under concern has been subject to substantial growth since 2005. The number of publications increased from around three to five annually (in the late 2000s) to almost 60 annually (around 2020). In terms of publication types, there has been a shift from conference papers to journal articles which hinted towards a maturation process. Among the various research types, the *solution* type has prevailed over the observed time period, followed by the *evaluation* type, although it can be said that all research types including *vision* and *experience* have increased in publications until 2022. As for the contribution types, *methods* papers have surged since 2015, and *concepts* papers have started to catch up in 2019. *Discussions*, *tools*, and *algorithms* rather declined. Concerning the modeling purpose, most publications aimed to *represent* or *analyze*, but *code generation* stayed stable at a lower level.

RQ2 intended to identify the main contributing institutions, and in what outlet did they publish their research. Based on the number of researchers, the Federal University of Espirito Santo (Brazil), the Kaunas University of Technology (Lithuania) and the Free University of Bozen-Bolzano (Italy) ranked at the top of publishing institutions in this research field. Among the top journals were the Journal of Biomedical Informatics, Expert Systems with Applications, and IEEE Transactions on Services Computing. Among conferences, the most frequently occurring ones were Conceptual Modeling (ER), IEEE International Conference on Engineering Technologies and Factory Automation, and Winter Simulation Conference.

RQ3 aimed to explore the main contributing researchers and research communities in the field, what topics they are focusing on, and how do these research groups interact. As main contributing research communities, the analysis identified the clusters around T. Walter (University of Koblenz-Landau / Germany), M. Malki (Université Djillali Liabes de Sidi Bel Abbes / Algeria), M. Wimmer (Vienna University of Technology / Austria), R.A. Buchmann (University Babes-Bolyai / Romania), H. Paulheim (University of Mannheim), G. Meditskos (Aristotle University of Thessaloniki / Greece), R. Verborgh (University of Ghent), G. Guizzardi (Free University of Bozen-Bolzano / Italy and Federal University of Espirito Santo / Brazil), S. Brockmans (Karlsruhe Institute of Technology / Germany), D. Gasevic (Monash University / Australia), and G. Kardas (Ege University / Turkey)[28]. The cluster of M. Wimmer focuses on model transformation, model-driven engineering, knowledge graphs, and UML. The main topics covered by the research community around Kardas evolve around agent-oriented software engineering and multi-agent systems combined with domain-specific modeling languages. Gasevic's community devotes to modeling in an educational context.

RQ4 tried to find out whether the contributions in the CM-SW field are attributed to foundational research or rather to specific industries / domains, and what kind of conceptual modeling languages are used. It turned out that 38.6% of the publications collected were *foundational* research, whereas the remaining 61.4% were attributed to various specific industries. The most frequently occurring industries or Semantic Web activity areas, as they were called in the related taxonomy, were *manufacturing*, *IT*, *healthcare*, and *education*. Among the modeling languages 48% applied UML, and 26% any type of DSL. BPMN was used in 12% of the publications. Less than 10% used SWRL, OntoUML, SysML, ER, and OCL. Overall, more than 100 modeling languages were mentioned in the publications.

RQ5 explored in what kinds of semantic technology segments and W3C main area did the contributions occur, and what SW standard(s) they used. The data analysis revealed that 59% of the publications involved *semantic modeling and development* as a semantic technology segment. Around 25% of the publications referred to *learning and reasoning*, *semantic data management and integration*, *semantic annotation*, and *semantic collaboration incl. portal technologies* respectively. The steepest increase over the observed time period occurred in publications on *linked data* and *vocabularies*, when it comes to the W3C main area. The categories *queries* and *inference* have slightly increased in their presence, but at a lower level. With regard to the Semantic Web standards, 68% of the publications used OWL, 52% RDF, which were the most used standards. Next, around 25% of the publications related to SPARQL and / or RDFS. For 23%, no match with any of the SW standards in the taxonomy could be identified, which raised a topic for further research.

RQ6 aimed to explain what value added conceptual modeling can achieve in combination with Semantic Web. The data analysis showed that 67% of the papers achieved a benefit from *increment schema and modeling*, while 34% improved *interoperability of multimedia data*, 28% enjoyed greater *representation flexibility*, and 24% *enhanced* their *inference capabilities*. At the beginning of the observed time period, i.e., in the mid- to late 2000s up to around 2013, the number of publications per value added opportunity was approximately the same. But after 2013 the *incremental schema and modeling* as a value added has surged, while the remaining options were subject to a slower increase.

RQ7 intended to combine several taxonomies to obtain more fine-grained mapping results. At this point, selected developments of combinations are summarized. Further combinations are covered in Section 4.3. The *combination of the contribution type with the modeling purpose taxonomy* indicates a concentration of papers along *representation* modeling purpose combined with *discussions* or *concepts* contribution type. The contribution type of *methods* rather appears in combination with the modeling purpose of *code generation* or *analysis*, and the contribution type *tools* is mostly combined with *code generation* as a modeling purpose. Over time notably the combination of *methods* with *representation* have grown considerably. The taxonomy *combination of World Wide Web Consortium*

[28]The institutions indicated are the ones at the time of publication of the respective papers.

(W3C) main area with conceptual modeling language reveals that the main areas *inference*, *linked data*, and *vocabularies* are very often combined with the *UML* modeling language. In addition to this, *DSL* also appear to be used widely with regard to *linked data*, and *vocabularies*. The increase in the use of *DSL* in these combinations almost tripled over time, whereas the use of *UML* only doubled. The evolution over time hinted towards a growth in *inference* main area together with modeling languages such as OntoUML, OCL, ER, DSL, AML, and BPMN. The *combination of the Semantic Web activity areas with the modeling purpose taxonomy* showed that an integral part of the publications concentrates in the foundational activity area in combination with the modeling purposes *representation* (79 publications), *analysis* (57), and *code generation* (39). Notably the modeling purposes *representation* and *analysis* prevailed across the activity areas. The largest part was in both analyzed time periods the category of *foundational* papers. The modeling purposes *representation* and *code generation* have become more extensive in combination with the activity areas *IT*, *manufacturing*, *healthcare*, and *education* over time. The *cultural* and *education* activity area stayed very small with regard to *representation*, *code generation*, and *analysis* as modeling purposes. The publications in the *tourism* activity area has grown stronger in combination with the modeling purpose *representation*, and *government* the other around. In addition to this, the *legal* activity area performed a shift from *code generation* to *representation* and *analysis* at a low level.

7.2. Implications for Future Research

The findings showed that the number of publications has been growing annually over the last decades, and that the topics at the intersection of conceptual modeling and Semantic Web have become more and more interlinked. This could also be observed by a sharper increase in journal articles more recently compared to conference papers, which hints towards a beginning process of maturation. However, this is so far just a presumption, but should be analyzed in future research to identify where this field of research heads to.

As no other SMS exists on this topic, this paper is based on related survey s on conceptual modeling or Semantic Web. Given this fact, future research still needs to verify how appropriate the selected taxonomies are, as no yardstick had existed at the time of writing. The taxonomy elements should also be critically examined with regard to completeness and relevance. Based on this, future research should try to replicate the results achieved in order to verify the conclusions.

When it comes to the findings, it turned out that the most frequently occurring research type was the *solution* paper, while the types *vision*, *experience*, and *evaluation* were not as prevalent. The same tendency holds for the contribution type, for which *concepts* and *methods* prevailed while *tools*, *algorithms*, and *discussions* were rather rare. As for conceptual modeling languages, UML, DSL, and BPMN were most frequently occurring, but others like SWRL, OntoUML, SysML, ER were on the rise. Given this, future research should focus on inspecting why those differences exist, what factors contribute to this situation, and more closely examine the papers of the rare types.

With respect to the W3C main areas, *Linked Data* and *Vocabularies* were covered a lot compared to *Inference* and *Queries*. The same situation was observed for the taxonomy referring to the value added, which led to *Incremental schema and modeling* as the mainly appearing category, followed by *Interoperability*. Such discrepancies could constitute an interesting topic for future research as well. Furthermore, among the Semantic Web standards OWL, RDF, and SPARQL were most prevalent, followed by the category "*no match*", which would again require future research to determine why there was no match, and what kinds of publications are affected by this.

A series of combinations of two taxonomies were analyzed in both a quantitative and a visual way. They were extended by a third dimension, namely the time component. Still, future research should aim to combine more taxonomies in order to generate even more fine-grained insights into the publications data.

This SMS provides a comprehensive overview of the research landscape including their most contributed topics. In relation to this, future research should cover the links within as well as between research clusters in even greater depth to capture relationships that have previously not been revealed, and to reveal the full dynamics of research communities.

Eventually, we believe this SMS provides a basis for manifold possible drill-downs on e.g., specific taxonomies or a specific combination of taxonomies. As such, the relevant papers mapped in this SMS can form the basis for many in-depth Systematic Literature Reviews on e.g., the use of *UML* and *OWL* for *Incremental schema and modeling*.

8. Conclusion

This systematic mapping study explored the research landscape at the intersection of conceptual modeling and the Semantic Web in the form of a systematic mapping study that comprised 484 publications. It followed the research method guidance from Petersen [23] and Kitchenham [24]. In line with the systematic mapping studies, seven research questions were answered using the classification scheme.

All in all, this SMS answered the research questions raised at the outset throughout a systematic data analysis process. It showed that the research area at the intersection of conceptual modeling and Semantic Web has grown from 2005 to May 2022, and likely will grow further, with an ongoing shift from conference papers to journal articles. The single taxonomy as well as the combined analysis indicated that some parts of the research area have been covered extensively, while others remained almost untouched, and could constitute potential opportunities for further research.

We prepared an accompanying Web Knowledge Base that enables users to explore and query the SMS data. The Web Knowledg Base can be accessed freely via: http://me.big.tuwien.ac.at/cmsw.

References

[1] S.M. Sanchez and H. Wan, Work smarter, not harder: a tutorial on designing and conducting simulation experiments, in: *Proceedings of the 2015 Winter Simulation Conference, Huntington Beach, CA, USA, December 6-9, 2015*, IEEE/ACM, 2015, pp. 1795–1809. doi:10.1109/WSC.2015.7408296.

[2] H.C. Mayr and B. Thalheim, The triptych of conceptual modeling, *Softw. Syst. Model.* **20**(1) (2021), 7–24. doi:10.1007/s10270-020-00836-z.

[3] S. Robinson, Choosing the right model: conceptual modeling for simulation, in: *Winter Simulation Conference 2011, WSC'11, Phoenix, AZ, USA, December 11-14, 2011*, S. Jain, R.R.C. Jr., J. Himmelspach, K.P. White and M.C. Fu, eds, IEEE, 2011, pp. 1428–1440. doi:10.1109/WSC.2011.6147862.

[4] S. ER, ER Conceptual Modeling, 2022, https://conceptualmodeling.org/.

[5] S. Robinson, G. Arbez, L.G. Birta, A. Tolk and G. Wagner, Conceptual modeling: definition, purpose and benefits, in: *Proceedings of the 2015 Winter Simulation Conference, Huntington Beach, CA, USA, December 6-9, 2015*, IEEE/ACM, 2015, pp. 2812–2826. doi:10.1109/WSC.2015.7408386.

[6] D. Bork and H. Fill, Formal Aspects of Enterprise Modeling Methods: A Comparison Framework, in: *47th Hawaii International Conference on System Sciences, HICSS 2014, Waikoloa, HI, USA, January 6-9, 2014*, IEEE Computer Society, 2014, pp. 3400–3409. doi:10.1109/HICSS.2014.422.

[7] M. Sabou and F. Ekaputra, Introduction to Semantic Systems, 2020, TU Wien.

[8] T. Berners-Lee, J. Hendler and O. Lassila, THE SEMANTIC WEB, *Scientific American* **284**(5) (2001), 34–43. http://www.jstor.org/stable/26059207.

[9] P. Hitzler, A review of the semantic web field, *Commun. ACM* **64**(2) (2021), 76–83. doi:10.1145/3397512.

[10] A. Patel and S. Jain, Present and future of semantic web technologies: a research statement, *International Journal of Computers and Applications* **43** (2019), 413–422. doi:10.1080/1206212X.2019.1570666.

[11] J. Euzenat, A map without a legend, *Semantic Web* **11**(1) (2020), 63–68. doi:10.3233/SW-190376.

[12] V.C. Storey, J. Trujillo and S.W. Liddle, Research on conceptual modeling: Themes, topics, and introduction to the special issue, *Data Knowl. Eng.* **98** (2015), 1–7. doi:10.1016/j.datak.2015.07.002.

[13] K. Sandkuhl, H. Fill, S. Hoppenbrouwers, J. Krogstie, F. Matthes, A.L. Opdahl, G. Schwabe, Ö. Uludag and R. Winter, From Expert Discipline to Common Practice: A Vision and Research Agenda for Extending the Reach of Enterprise Modeling, *Bus. Inf. Syst. Eng.* **60**(1) (2018), 69–80. doi:10.1007/s12599-017-0516-y.

[14] G. Guizzardi, G. Wagner, R. de Almeida Falbo, R.S.S. Guizzardi and J.P.A. Almeida, Towards Ontological Foundations for the Conceptual Modeling of Events, in: *Conceptual Modeling - 32th International Conference, ER 2013, Hong-Kong, China, November 11-13, 2013. Proceedings*, W. Ng, V.C. Storey and J. Trujillo, eds, Lecture Notes in Computer Science, Vol. 8217, Springer, 2013, pp. 327–341. doi:10.1007/978-3-642-41924-9_27.

[15] M. Verdonck, F. Gailly, R. Pergl, G. Guizzardi, B.F. Martins and O. Pastor, Comparing traditional conceptual modeling with ontology-driven conceptual modeling: An empirical study, *Inf. Syst.* **81** (2019), 92–103. doi:10.1016/j.is.2018.11.009.

[16] G.G. Shanks, E. Tansley and R. Weber, Using ontology to validate conceptual models, *Commun. ACM* **46**(10) (2003), 85–89. doi:10.1145/944217.944244.

[17] T. Kosar, S. Bohra and M. Mernik, Domain-Specific Languages: A Systematic Mapping Study, *Inf. Softw. Technol.* **71** (2015), 77–91. doi:10.1016/j.infsof.2015.11.001.

[18] M. Verdonck, F. Gailly, S. de Cesare and G. Poels, Ontology-driven conceptual modeling: A systematic literature mapping and review, *Appl. Ontology* **10**(3–4) (2015), 197–227. doi:10.3233/AO-150154.

[19] F. Zahid, A. Tanveer, M.M.Y. Kuo and R. Sinha, A systematic mapping of semi-formal and formal methods in requirements engineering of industrial Cyber-Physical systems, *J. Intell. Manuf.* **33**(6) (2022), 1603–1638. doi:10.1007/s10845-021-01753-8.

[20] V. Janev and S. Vranes, Maturity and Applicability Assessment of Semantic Web Technologies, in: *5th International Conference on Semantic Systems, Graz, Austria, September 2-4, 2009. Proceedings*, A. Paschke, H. Weigand, W. Behrendt, K. Tochtermann and T. Pellegrini, eds, Verlag der Technischen Universität Graz, 2009, pp. 530–541. http://www.i-semantics.at/2009/papers/maturity_and_applicability_assessment.pdf.

[21] M. Sabou, L. Aroyo, K. Bontcheva, A. Bozzon and R.K. Qarout, Semantic Web and Human Computation: The status of an emerging field, *Semantic Web* **9**(3) (2018), 291–302. doi:10.3233/SW-180292.

[22] S. Kumar, K. Srinivasan, Y.-C. Hu, S. Masilamani and K.-Y. Huang, A Contemporary Review on Utilizing Semantic Web Technologies in Healthcare, Virtual Communities, and Ontology-Based Information Processing Systems, *Electronics* **11** (2022), 453. doi:10.3390/electronics11030453.

[23] K. Petersen, R. Feldt, S. Mujtaba and M. Mattsson, Systematic Mapping Studies in Software Engineering, in: *12th International Conference on Evaluation and Assessment in Software Engineering, EASE 2008, University of Bari, Italy, 26-27 June 2008*, G. Visaggio, M.T. Baldassarre, S.G. Linkman and M. Turner, eds, Workshops in Computing, BCS, 2008. http://ewic.bcs.org/content/ConWebDoc/19543.

[24] B.A. Kitchenham, D. Budgen and O.P. Brereton, Using mapping studies as the basis for further research - A participant-observer case study, *Inf. Softw. Technol.* **53**(6) (2011), 638–651. doi:10.1016/j.infsof.2010.12.011.

[25] F. Febrero, C. Calero and M.Á. Moraga, A Systematic Mapping Study of Software Reliability Modeling, *Inf. Softw. Technol.* **56**(8) (2014), 839–849. doi:10.1016/j.infsof.2014.03.006.

[26] W3C, The Semantic Web, 2022, W3C Community. https://www.w3.org/standards/semanticweb/.

[27] W3C, Semantic Web Case Studies and Use Cases, 2012. https://www.w3.org/2001/sw/sweo/public/UseCases/.

[28] W3C, Semantic Web Wiki - Main Page, 2019. https://www.w3.org/2001/sw/wiki/.

[29] K. Peffers, T. Tuunanen, M.A. Rothenberger and S. Chatterjee, A Design Science Research Methodology for Information Systems Research, *J. Manag. Inf. Syst.* **24**(3) (2008), 45–77. doi:10.2753/mis0742-1222240302.

[30] K. Peffers, M.A. Rothenberger and W.L.K. Jr. (eds), Design Science Research in Information Systems. Advances in Theory and Practice - 7th International Conference, DESRIST 2012, Las Vegas, NV, USA, May 14-15, 2012. Proceedings, in *Lecture Notes in Computer Science*, Vol. 7286, Springer, 2012. ISBN 978-3-642-29862-2. doi:10.1007/978-3-642-29863-9.

[31] A. Wortmann, O. Barais, B. Combemale and M. Wimmer, Modeling languages in Industry 4.0: an extended systematic mapping study, *Softw. Syst. Model.* **19**(1) (2020), 67–94. doi:10.1007/s10270-019-00757-6.

[32] D. Karagiannis, H.C. Mayr and J. Mylopoulos (eds), *Domain-Specific Conceptual Modeling, Concepts, Methods and Tools*, Springer, 2016. ISBN 978-3-319-39416-9. doi:10.1007/978-3-319-39417-6.

[33] D. Bork, *Conceptual Modeling and Artificial Intelligence: Challenges and Opportunities for Enterprise Engineering: Keynote Presentation at the 11th Enterprise Engineering Working Conference (EEWC 2021)*, 2022, pp. 3–9. ISBN 978-3-031-11519-6. doi:10.1007/978-3-031-11520-2_1.

[34] Nelson, H James and Poels, Geert and Genero, Marcela and Piattini, Mario, A conceptual modeling quality framework, *SOFTWARE QUALITY JOURNAL* **20**(1) (2012), 201–228. http://dx.doi.org/10.1007/s11219-011-9136-9.

[35] D. Bork, R.A. Buchmann, D. Karagiannis, M. Lee and E. Miron, An Open Platform for Modeling Method Conceptualization: The OMiLAB Digital Ecosystem, *Commun. Assoc. Inf. Syst.* **44** (2019), 32. https://aisel.aisnet.org/cais/vol44/iss1/32.

[36] A. Olivé, *Conceptual modeling of information systems*, Springer, 2007. doi:10.1007/978-3-540-39390-0.

[37] A.R. da Silva, Model-driven engineering: A survey supported by the unified conceptual model, *Comput. Lang. Syst. Struct.* **43** (2015), 139–155. doi:10.1016/j.cl.2015.06.001.

[38] B. Thalheim, The Science and Art of Conceptual Modelling, *Trans. Large Scale Data Knowl. Centered Syst.* **6** (2012), 76–105. doi:10.1007/978-3-642-34179-3_3.

[39] M. Smajevic, S. Hacks and D. Bork, Using Knowledge Graphs to Detect Enterprise Architecture Smells, in: *The Practice of Enterprise Modeling - 14th IFIP WG 8.1 Working Conference, PoEM 2021, Riga, Latvia, November 24-26, 2021, Proceedings*, E. Serral, J. Stirna, J. Ralyté and J. Grabis, eds, Lecture Notes in Business Information Processing, Vol. 432, Springer, 2021, pp. 48–63. doi:10.1007/978-3-030-91279-6_4.

[40] M.K. Bergman, Glossary of Semantic Technology Terms, 2012, AI3 - Adaptive Information, Adaptive Innovation, Adaptive Infrastructure. https://www.mkbergman.com/1017/glossary-of-semantic-technology-terms/.

[41] T. Sjekavica, G. Gledec and M. Horvat, Advantages of Semantic Web Technologies Usage in the Multimedia Annotation and Retrieval, 2014.

[42] A. Le Hors and S. Speicher, Semantic Web Use Cases and Case Studies - Case Study: Open Services Lifecycle Collaboration framework based on Linked Data, 2012, W3C.

[43] M.R. Blackburn and P.O. Denno, Using Semantic Web Technologies for Integrating Domain Specific Modeling and Analytical Tools, in: *Proceedings of the Complex Adaptive Systems 2015 Conference, San Jose, CA, USA, November 2-4, 2015*, C.H. Dagli, ed., Procedia Computer Science, Vol. 61, Elsevier, 2015, pp. 141–146. doi:10.1016/j.procs.2015.09.174.

[44] J.B. de Souza Neto, A.M. Moreira and M.A. Musicante, Semantic Web Services testing: A Systematic Mapping study, *Comput. Sci. Rev.* **28** (2018), 140–156. doi:10.1016/j.cosrev.2018.03.002.

[45] A. Wortmann, B. Combemale and O. Barais, A Systematic Mapping Study on Modeling for Industry 4.0, in: *20th ACM/IEEE International Conference on Model Driven Engineering Languages and Systems, MODELS 2017, Austin, TX, USA, September 17-22, 2017*, IEEE Computer Society, 2017, pp. 281–291. doi:10.1109/MODELS.2017.14.

[46] H. Hanke and D. Knees, A phase-field damage model based on evolving microstructure, *Asymptotic Analysis* **101** (2017), 149–180.

[47] E. Lefever, A hybrid approach to domain-independent taxonomy learning, *Applied Ontology* **11**(3) (2016), 255–278.
[48] P.S. Meltzer, A. Kallioniemi and J.M. Trent, Chromosome alterations in human solid tumors, in: *The Genetic Basis of Human Cancer*, B. Vogelstein and K.W. Kinzler, eds, McGraw-Hill, New York, 2002, pp. 93–113.
[49] P.R. Murray, K.S. Rosenthal, G.S. Kobayashi and M.A. Pfaller, *Medical Microbiology*, 4th edn, Mosby, St. Louis, 2002.
[50] E. Wilson, Active vibration analysis of thin-walled beams, PhD thesis, University of Virginia, 1991.
[51] A.M. Turing, On computable numbers, with an application to the Entscheidungsproblem, *J. of Math* **58** (1936), 345–363.
[52] R.M. Stallman and S. Krempl, Software muß frei sein!, 1999, Online erhältlich unter http://www.heise.de/tp/deutsch/inhalt/te/2860/1.html; abgerufen am 8. Januar 2005..
[53] E.S. Raymond, A Brief History of Hackerdom, T. O'Reilly, ed., O'Reilly, Sebastopol, CA, 1999, pp. 7–26.
[54] T. O'Reilly (ed.), The Cathedral and the Bazaar – Musings on Linux and Open Source by an Accidental Revolutionary, O'Reilly, Sebastopol, CA, 1999.
[55] B. Fitzgerald, N. Russo and J. DeGross, *Realigning Research and Practice in Information Systems Development: The Social and Organizational Perspective*, Kluwer Academic Press, Boston, 2001, Eine optionale Notiz.
[56] J. Kaye and P. Dourish, Special issue on science fiction and ubiquitous computing, *Personal and Ubiquitous Computing* **18**(4) (2014), 765–766. doi:10.1007/s00779-014-0773-4.
[57] W. Lin and C.-F. Tsai, Missing value imputation: a review and analysis of the literature (2006–2017), *Artificial Intelligence Review* **53** (2019), 1487–1509.
[58] I. Ezzine and L. Benhlima, A Study of Handling Missing Data Methods for Big Data, in: *2018 IEEE 5th International Congress on Information Science and Technology (CiSt)*, 2018, pp. 498–501. doi:10.1109/CIST.2018.8596389.
[59] S. Biffl, M. Eckhart, A. Lüder and E. Weippl, *Security and Quality in Cyber-Physical Systems Engineering*, Springer Publishing, Cham (Switzerland), 2019. ISBN 978-3-030-25311-0. doi:https://doi.org/10.1007/978-3-030-25312-7.
[60] R. Gacitúa, J. Mazón and A. Cravero, Using Semantic Web technologies in the development of data warehouses: A systematic mapping, *WIREs Data Mining Knowl. Discov.* **9**(3) (2019). doi:10.1002/widm.1293.
[61] C. Tikva and E. Tambouris, A systematic mapping study on teaching and learning Computational Thinking through programming in higher education, *Thinking Skills and Creativity* **41** (2021), 100849. doi:https://doi.org/10.1016/j.tsc.2021.100849. https://www.sciencedirect.com/science/article/pii/S187118712100064X.
[62] G. Sebastián, J.A. Gallud and R. Tesoriero, Code generation using model driven architecture: A systematic mapping study, *Journal of Computer Languages* **56** (2020), 100935. doi:https://doi.org/10.1016/j.cola.2019.100935. https://www.sciencedirect.com/science/article/pii/S2590118419300607.
[63] A. Iung, J. Carbonell, L. Marchezan, E. Rodrigues, M. Bernardino, F.P. Basso and B. Medeiros, Systematic mapping study on domain-specific language development tools, *Empir. Softw. Eng.* **25**(5) (2020), 4205–4249. doi:10.1007/s10664-020-09872-1.
[64] M. Verdonck and F. Gailly, Insights on the Use and Application of Ontology and Conceptual Modeling Languages in Ontology-Driven Conceptual Modeling, in: *Conceptual Modeling*, I. Comyn-Wattiau, K. Tanaka, I.-Y. Song, S. Yamamoto and M. Saeki, eds, Springer International Publishing, Cham, 2016, pp. 83–97. ISBN 978-3-319-46397-1.
[65] K. Wakil and D.N.A. Jawawi, Model Driven Web Engineering: A Systematic Mapping Study, *e Informatica Softw. Eng. J.* **9**(1) (2015), 87–122. doi:10.5277/E-INF150106.
[66] M. Mernik, Domain-Specific Languages: A Systematic Mapping Study, in: *SOFSEM 2017: Theory and Practice of Computer Science*, B. Steffen, C. Baier, M. van den Brand, J. Eder, M. Hinchey and T. Margaria, eds, Springer International Publishing, Cham, 2017, pp. 464–472. ISBN 978-3-319-51963-0.
[67] C. dos Santos Garcia, A. Meincheim, E.R. Faria Junior, M.R. Dallagassa, D.M.V. Sato, D.R. Carvalho, E.A.P. Santos and E.E. Scalabrin, Process mining techniques and applications – A systematic mapping study, *Expert Systems with Applications* **133** (2019), 260–295. doi:https://doi.org/10.1016/j.eswa.2019.05.003. https://www.sciencedirect.com/science/article/pii/S0957417419303161.
[68] P. Nistala, K.V. Nori and R. Reddy, Software Quality Models: A Systematic Mapping Study, in: *2019 IEEE/ACM International Conference on Software and System Processes (ICSSP)*, 2019, pp. 125–134. doi:10.1109/ICSSP.2019.00025.
[69] V.A. Pinto and F.S. Parreiras, Enterprise Linked Data: A Systematic Mapping Study, in: *Advances in Conceptual Modeling*, M. Indulska and S. Purao, eds, Springer International Publishing, Cham, 2014, pp. 253–262. ISBN 978-3-319-12256-4.
[70] C.M. de Oliveira Rodrigues, F.L.G. de Freitas, E.F.S. Barreiros, R.R. de Azevedo and A.T. de Almeida Filho, Legal ontologies over time: A systematic mapping study, *Expert Systems with Applications* **130** (2019), 12–30. doi:https://doi.org/10.1016/j.eswa.2019.04.009. https://www.sciencedirect.com/science/article/pii/S0957417419302398.
[71] A. Venceslau, R. Andrade, V. Vidal, T. Nogueira and V. Pequeno, IoT Semantic Interoperability: A Systematic Mapping Study, 2019, pp. 535–544. doi:10.5220/0007732605350544.
[72] L. Ding, P. Kolari, Z. Ding and S. Avancha, *Using Ontologies in the Semantic Web: A Survey*, in: *Ontologies: A Handbook of Principles, Concepts and Applications in Information Systems*, R. Sharman, R. Kishore and R. Ramesh, eds, Springer US, Boston, MA, 2007, pp. 79–113. ISBN 978-0-387-37022-4. doi:10.1007/978-0-387-37022-4_4.
[73] M. Alloghani, D. Al-Jumeily Obe, A. Hussain, A. Aljaaf, J. Mustafina, M. Khalaf and S.Y. Tan, THE XML AND SEMANTIC WEB: A SYSTEMATIC REVIEW ON TECHNOLOGIES, 2019, pp. 92–102. doi:10.33965/bigdaci2019_201907L012.
[74] L. Waltersdorfer, A. Breit, F.J. Ekaputra and M. Sabou, Bridging Semantic Web and Machine Learning: First Results of a Systematic Mapping Study, in: *Database and Expert Systems Applications - DEXA 2021 Workshops*, G. Kotsis, A.M. Tjoa, I. Khalil, B. Moser, A. Mashkoor, J. Sametinger, A. Fensel, J. Martinez-Gil, L. Fischer, G. Czech, F. Sobieczky and S. Khan, eds, Springer International Publishing, Cham, 2021, pp. 81–90. ISBN 978-3-030-87101-7.

[75] B. DAUM (ed.), Copyright, in: *Modeling Business Objects with XML Schema*, The Morgan Kaufmann Series in Software Engineering and Programming, Morgan Kaufmann, San Francisco, 2003, p. vi. ISBN 978-1-55860-816-0. doi:https://doi.org/10.1016/B978-1-55860-816-0.50024-0. https://www.sciencedirect.com/science/article/pii/B9781558608160500240.

[76] T.P. Nogueira, R.B. Braga, C.T. de Oliveira and H. Martin, FrameSTEP: A framework for annotating semantic trajectories based on episodes, *Expert Systems with Applications* **92** (2018), 533–545. doi:https://doi.org/10.1016/j.eswa.2017.10.004. https://www.sciencedirect.com/science/article/pii/S0957417417306796.

[77] M. Zeng, Semantic enrichment for enhancing LAM data and supporting digital humanities. Review article, *El Profesional de la Información* **28** (2019). doi:10.3145/epi.2019.ene.03.

[78] M. Salama, R. Bahsoon and N. Bencomo, Chapter 11 - Managing Trade-offs in Self-Adaptive Software Architectures: A Systematic Mapping Study, in: *Managing Trade-Offs in Adaptable Software Architectures*, I. Mistrik, N. Ali, R. Kazman, J. Grundy and B. Schmerl, eds, Morgan Kaufmann, Boston, 2017, pp. 249–297. ISBN 978-0-12-802855-1. doi:https://doi.org/10.1016/B978-0-12-802855-1.00011-3. https://www.sciencedirect.com/science/article/pii/B9780128028551000113.

[79] P. Fettke, How Conceptual Modeling Is Used, *Commun. Assoc. Inf. Syst.* **25** (2009), 43. doi:10.17705/1cais.02543.

[80] I. Davies, P.F. Green, M. Rosemann, M. Indulska and S. Gallo, How do practitioners use conceptual modeling in practice?, *Data Knowl. Eng.* **58**(3) (2006), 358–380. doi:10.1016/j.datak.2005.07.007.

[81] D. Karagiannis, R. Buchmann, P. Burzynski, U. Reimer and M. Walch, *Fundamental Conceptual Modeling Languages in OMiLAB*, 2016, pp. 3–30. ISBN 978-3-319-39416-9. doi:10.1007/978-3-319-39417-6_1.

[82] V.C. Storey and I. Song, Big data technologies and Management: What conceptual modeling can do, *Data Knowl. Eng.* **108** (2017), 50–67. doi:10.1016/j.datak.2017.01.001.

[83] L. Ding and T. Finin, Characterizing the Semantic Web on the Web, in: *The Semantic Web - ISWC 2006, 5th International Semantic Web Conference, ISWC 2006, Athens, GA, USA, November 5-9, 2006, Proceedings*, I.F. Cruz, S. Decker, D. Allemang, C. Preist, D. Schwabe, P. Mika, M. Uschold and L. Aroyo, eds, Lecture Notes in Computer Science, Vol. 4273, Springer, 2006, pp. 242–257. doi:10.1007/11926078_18.

[84] N. Shadbolt, T. Berners-Lee and W. Hall, The Semantic Web Revisited, *IEEE Intelligent Systems* **21**(3) (2006), 96–101. doi:10.1109/MIS.2006.62.

[85] A. Gyrard, M. Serrano and G.A. Atemezing, Semantic web methodologies, best practices and ontology engineering applied to Internet of Things, in: *2015 IEEE 2nd World Forum on Internet of Things (WF-IoT)*, 2015, pp. 412–417. doi:10.1109/WF-IoT.2015.7389090.

[86] S. Mehla and S. Jain, *Rule Languages for the Semantic Web: Proceedings of IEMIS 2018, Volume 1*, 2019, pp. 825–834. ISBN 978-981-13-1950-1. doi:10.1007/978-981-13-1951-8_73.

[87] E. Karan, J. Irizarry and J. Haymaker, BIM and GIS Integration and Interoperability Based on Semantic Web Technology, *Journal of Computing in Civil Engineering* **30** (2015), 1–12. doi:10.1061/(ASCE)CP.1943-5487.0000519.

[88] G. Guizzardi, C. Fonseca, A. Benevides, J. Almeida, D. Porello and T. Prince Sales, Endurant Types in Ontology-Driven Conceptual Modeling: Towards OntoUML 2.0, 2018. ISBN 978-3-030-00846-8. doi:10.1007/978-3-030-00847-5_12.

[89] K. Alkharabsheh, Y. Crespo, M.E. Manso and J.A. Taboada, Software Design Smell Detection: a systematic mapping study, *Softw. Qual. J.* **27**(3) (2019), 1069–1148. doi:10.1007/s11219-018-9424-8.

[90] A. Tarhan and G. Giray, On the Use of Ontologies in Software Process Assessment: A Systematic Literature Review, in: *Proceedings of the 21st International Conference on Evaluation and Assessment in Software Engineering, EASE 2017, Karlskrona, Sweden, June 15-16, 2017*, E. Mendes, S. Counsell and K. Petersen, eds, ACM, 2017, pp. 2–11. doi:10.1145/3084226.3084261.

[91] C.M. de Oliveira Rodrigues, F.L.G. de Freitas, E.F.S. Barreiros, R.R. de Azevedo and A.T. de Almeida Filho, Legal ontologies over time: A systematic mapping study, *Expert Syst. Appl.* **130** (2019), 12–30. doi:10.1016/j.eswa.2019.04.009.

[92] H.V. Olivera, G. RuiZhe, R.C. Huacarpuma, A.P.B. da Silva, A.M. Mariano and M. Holanda, Data Modeling and NoSQL Databases - A Systematic Mapping Review, *ACM Comput. Surv.* **54**(6) (2022), 116:1–116:26. doi:10.1145/3457608.

[93] R. Enríquez-Reyes, S. Cadena-Vela, A.F. Guilló, J. Mazón, L. Ibáñez and E. Simperl, Systematic Mapping of Open Data Studies: Classification and Trends From a Technological Perspective, *IEEE Access* **9** (2021), 12968–12988. doi:10.1109/ACCESS.2021.3052025.

[94] P. Pauwels, S. Zhang and Y. Lee, Semantic web technologies in AEC industry: A literature review, *Automation in Construction* **73** (2016). doi:10.1016/j.autcon.2016.10.003.

[95] M. Dadkhah, S. Araban and S. Paydar, A systematic literature review on semantic web enabled software testing, *J. Syst. Softw.* **162** (2020). doi:10.1016/j.jss.2019.110485.

[96] D. Moussallem, M. Wauer and A.N. Ngomo, Machine Translation Using Semantic Web Technologies: A Survey, *CoRR* **abs/1711.09476** (2017). http://arxiv.org/abs/1711.09476.

[97] S. Zander, N. Ahmed and M.T. Frank, A Survey about the Usage of Semantic Technologies for the Description of Robotic Components and Capabilities, in: *Proceedings of the 1st International Workshop on Science, Application and Methods in Industry 4.0 co-located with (i-KNOW 2016), Graz, Austria, October 19, 2016*, R. Kern, G. Reiner and O. Bluder, eds, CEUR Workshop Proceedings, Vol. 1793, CEUR-WS.org, 2016. http://ceur-ws.org/Vol-1793/paper2.pdf.

[98] B. Drury and M. Roche, A survey of the applications of text mining for agriculture, *Computional Electronics Agriculture* **163** (2019). doi:10.1016/j.compag.2019.104864.