A survey of the first 20 years of research on semantic Web and linked data

Fabien Gandon

Inria, Université Côte d'Azur, CNRS, I3S, Wimmics
2004 rt des Lucioles, 06902
Sophia Antipolis, France

fabien.gandon@inria.fr

ABSTRACT. *This paper is a survey of the research topics in the field of Semantic Web, Linked Data and Web of Data. This study looks at the contributions of this research community over its first twenty years of existence. Compiling several bibliographical sources and bibliometric indicators, we identify the main research trends and we reference some of their major publications to provide an overview of that initial period. We conclude with some perspectives for the future research challenges.*

RÉSUMÉ. *Cet article est une étude des sujets de recherche dans le domaine du Web sémantique, des données liées et du Web des données. Cette étude se penche sur les contributions de cette communauté de recherche au cours de ses vingt premières années d'existence. En compilant plusieurs sources bibliographiques et indicateurs bibliométriques, nous identifions les principales tendances de la recherche et nous référencons certaines de leurs publications majeures pour donner un aperçu de cette période initiale. Nous concluons avec une discussion sur les tendances et perspectives de recherche.*

KEYWORDS: *survey, semantic Web, linked data, Web of data.*

MOTS-CLÉS : *état de l'art, Web sémantique, données liées, Web de données.*

DOI:10.3166/ISI.23.3-4.11-56

1. Introduction: Weaving a Web of Everything

In nature, a web is a network of fine threads formed by weaving or interweaving. It is both resilient to many things and fragile. A web may be regularly rebuilt, patched or updated. Used as an abstraction, a web is a complex system of interconnected elements, with links interlaced into an intricate lattice-like structure. And from the beginning, the core idea of the World Wide Web (the Web) was to link as many things through as many links and from as many sources as possible (Berners-Lee, 1989). In fact the initial vision of the Web that Tim Berners-Lee had was already very "semantic webby".

This initial vision led us, thirty years after the birth of the Web, to have a Web linking applications, things, people, data, etc. In order to be able to scale, in terms of volume and variety, Tim Berners-Lee insisted very early on the need to provide on the Web "more machine oriented semantic information, allowing more sophisticated processing" (Berners-Lee *et al.*, 1994).

To bootstrap that evolution Tim Berners-Lee then proposed in September 1998 a "Semantic Web Road map" (Berners-Lee, 1998) giving 20 years ago the blue prints of the architecture of the Semantic Web. In 1999 the first versions of RDF and RDFS were published by the W3C and the vision of a Semantic Web was then made visible to a broad audience in 2001 with an article in the Scientific American (Berners-Lee *et al.*, 2001). This well-known article presents the Semantic Web as an extension of the existing document-based Web with a Web of structured data and formal semantics better enabling computers and people to work in cooperation. A few years later, Tim Berners-Lee will be again instrumental in pushing what can be seen as a first wave of deployment of the Semantic Web with the Linked Data principles and the Linked Open Data 5-star rules (Berners-Lee, 2006) leading to the publication and growth of linked open datasets weaving a Web of Linked Data.

But in parallel to these new developments and since the beginning of the years 2000 a research community has formed on the topic of the Semantic Web. It all started with the first international Semantic Web Working Symposium (SWWS), a workshop held in Stanford, Palo Alto, the 30th of July and 1st of August 2001. The following year the symposium became the International Semantic Web Conference (ISWC) series. Nowadays, Semantic Web not only has its conferences (e.g. ISWC, ESWC, SemTech, SemWeb.Pro) and journals (e.g. Semantic Web Journal, Journal of Web Semantics) but is also an established topic of older conferences and journals from other domains (e.g. The Web Conference WWW, VLDB, EKAW, IJCAI/ECAI, WI, etc.).

In this paper, we will survey the research topics over the first twenty years of research on Semantic Web and Linked Data and how researchers and developers are growing linked data and linked schemata on the Web to bridge natural and artificial intelligence worldwide. In rest of this article, RDF and RDFS will be used to denote any version, including the early drafts, of respectively the resource description framework to publish linked (meta)data on the Web and its schema language to publish lightweight linked ontologies - essentially taxonomies. Likewise the acronym OWL will refer to any version of the Web ontology language and its different profiles to publish and link formal ontologies on the Web. The term SPARQL will be used to refer indistinctively to both versions of the query language and protocol to access RDF triplestores over the Web.

In section 2, we introduce the method applied to identify the main research areas of the first twenty years of research on Semantic Web and structure this paper accordingly. Then, each one of the next sections identifies and explains a research area. Finally section 17 provides a discussion and concluding remarks about these research trends and their perspectives.

2. Tag Cloud Atlas: Mapping the Semantic Web Research Community

The rest of the article is structured by a study of the research tracks, sessions and calls in the Semantic Web venues over the first twenty years of its existence. In this first section, we summarize the method followed and we provide an overview of these topics. The following sections will then group research topics and major references into a set of main research areas.

As a first step, we performed a review of the session titles of the programs of ISWC and ESWC conferences since their first edition until 2017. This review suggested topical clusters and candidate labels based on the frequency of that topic and grouping over the years and the regularity in its labelling. The resulting groups suggested a first set of research area to be consolidated and extended.

The second source followed the "eat your own dog food" famous saying in the Web community and used linked open data about our community to extract major research topics.

In Figure 1 are shown the top 100 topics from 1903 topics found by a SPARQL query on the Scholarly Data end-point[1] in January 2018. The query extracts, groups, counts and orders the topics (keywords) according to the number of articles linked to them. The result was used both as a first ordered list of topics to complement the session labels and as a corpus to generate the tag cloud[2].

In Figure 2 are shown the top 200 words from 5070 words extracted from titles found by a SPARQL query again on the Scholarly Data end-point in January 2018. The query extracts the titles to generate a corpus which is then fed to the same tag cloud generator.

To validate this overview independently, Mylène Leitzelman, a colleague, extracted 2547 articles from the Web of Science selecting sources which titles contain "Semantic*" (conferences, journals, etc.) and obtained a tagcloud from n-grams in

1. http://www.scholarlydata.org/sparql/
2. https://tagcrowd.com/

Figure 1. Top 100 from 1903 topics found on scholarly data

Figure 2. Top 20 from 5070 words from titles found on scholarly data.

Gargantext (see Figure 3). In addition, her analysis on the Web of Science identified 14157 document on the topic "Semantic Web", showing a now stable scientific production (Figure 4) and a clearly international interest (Figure 5). This independent analysis also confirmed by statistics on the Web of Science that ISWC and ESWC conferences are the main sources of articles about the Semantic Web and therefore important sources for this survey.

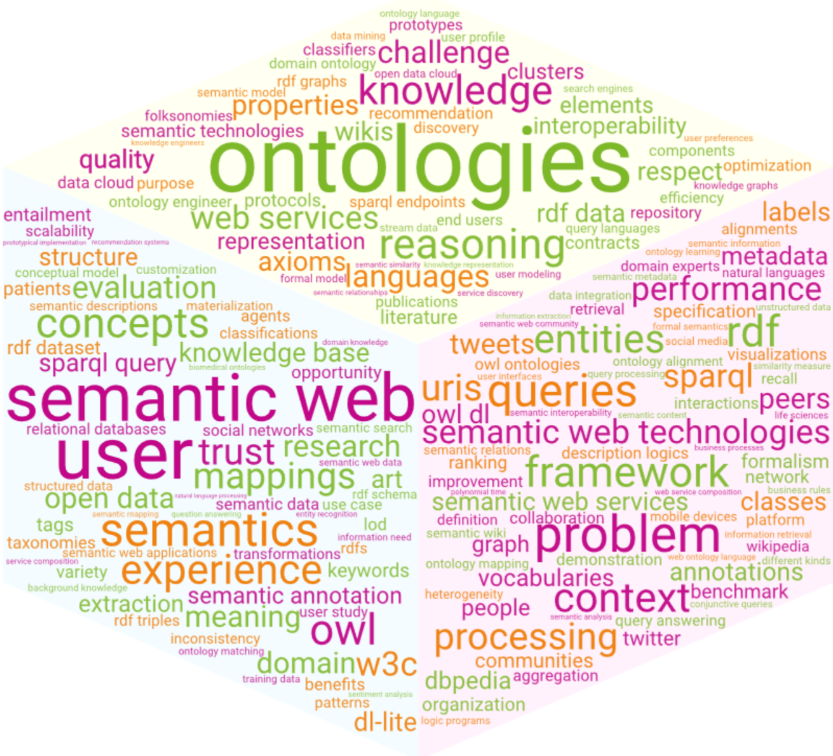

Figure 3. Tagcloud from n-grams in Gargantext extracted 2547 articles from the Web of Science by selecting sources which titles contain "Semantic"*

Comparing the three tag clouds and their statistics we validated the major topics and made sure each one of them has an identified research area. This allows us to consolidate the topic list and the topic clustering into research areas.

As a third step, to enforce the representativeness of the survey in terms both of structure and of references we identified a large number of distinguished papers and mapped them to the topics and cluster to ensure their coverage and provide relevant selected references. From a methodological point of view, the following sections include bibliographical references that have been selected by searching for the awarded papers (best papers award, test of time award) at two conferences (ISWC, ESWC) and the most cited papers (Google Scholar). In addition the keywords of each sections

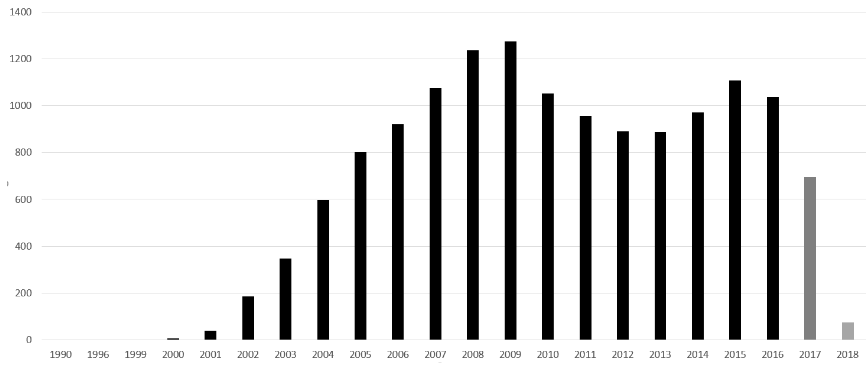

Figure 4. Yearly distribution of 14157 documents found on the Web of Science on the topic "Semantic Web", note that at the moment of the survey the years 2017 and 2018 were largely incomplete

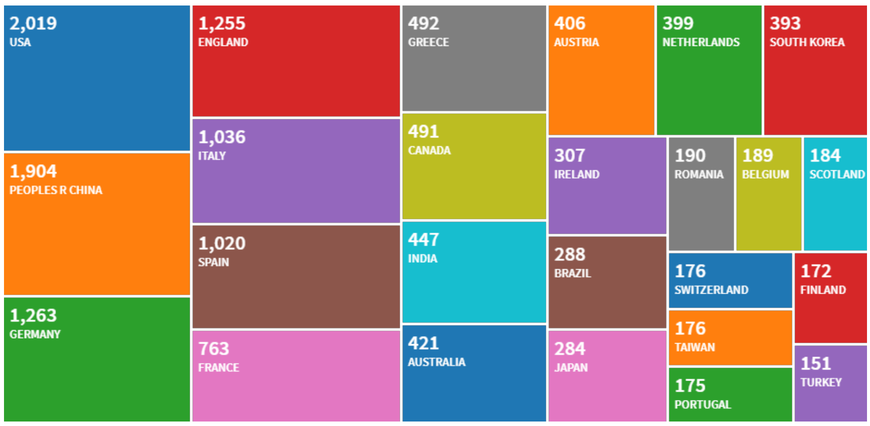

Figure 5. Distribution per country of 14157 documents found on the Web of Science on the topic "Semantic Web"

have been entered in Google Scholar with a variation of keywords for query extension for instance the keyword for ontology was searched on Google scholar as : "ontology linked data", "ontology Semantic Web", "ontology Web of data", "ontology Web". The top cited papers were considered for inclusion. Finally the last criteria for inclusion was the explicit mention of the domains in the paper (e.g. Semantic Web, linked data) or keywords of the domain (e.g. RDF, OWL, SPARQL). If a paper revealed a missing research topic, this one was added to the selection.

Applying this method, we can now distinguish the following main research areas of the Semantic Web community that we will describe in the next sections:

- Knowledge Representation, Reasoning and processing (KRR) (section 3).
- Ontologies and semantic vocabularies for machines on the Web (section 4).

– Matching, aligning and mapping the vocabularies (section 5).

– Interoperability or the Web providing a distributed semantic blackboard framework (section 6).

– Retrieving and querying formal knowledge graphs on the Web (section 7).

– Data Management and the challenges of Big (Web of Linked) Data (section 8).

– Software Architectures supported by or to support semantics (section 9).

– Linked (Open) Data and linked Schemata published on the Web (section 10).

– Semantics in ensuring security, trust and privacy (section 11).

– Machine learning and data mining to obtain knowledge but also with the help of knowledge (section 12).

– Natural Language processing with and for semantics and its applications (section 13).

– Semantics in human-computer interactions and their design (section 14).

– Semantics-based social networks and media representation and management (section 15).

– The Semantic Web in use, its applications and their returns on experience (section 16).

The next sections detail and survey each one of these areas, providing pointers to relevant and distinguished contributions, before providing some concluding remarks. The order of the sections is based on both the chronology of topics and their articulation.

3. Knowledge Representation, Reasoning and Processing (KRR)

The topic of Knowledge Representation and Reasoning (KRR) is one of the oldest and still very active core research area in Semantic Web (since 2002). It is part of the foundations of the Semantic Web because it was a key question since the initial work on RDF and RDFS and also because the Semantic Web community was bootstrapped by a number of researchers coming from the knowledge representation and management communities (e.g. EKAW, KR, DL, ICCS or KCap).

This topic can be seen as covering two families of questions inherently related: (1) the knowledge representation formalisms (e.g. logics, graphs, RDF, RDFS, OWL, Rule, RIF, semantic networks or knowledge graphs) with their possible fragments, profiles and extensions and (2) the reasoning and processing mechanisms (e.g. inferences, entailment, validation, transformation, non-standard, temporal, spatial or approximate reasoning). Both topics are still active for instance to support ever more intelligent processing on top of RDF data or to adapt to specific platforms and context such as in the case of mobile reasoning. Key research questions of the KRR papers include expressiveness, decidability, completeness and complexity.

In terms of knowledge representation a lot of work has been done since the initial idea of reducing the semantic discontinuity between Web languages and architecture and KRR languages and architecture targeting a common semantic foundation on top of which to build the Semantic Web (Patel-Schneider, Siméon, 2002). Through its research, the community contributed a lot to the work on standards (RDF, RDFS, OWL, etc.), their links to existing languages (description logics, conceptual graphs, etc.) and also in proposing extensions or alternative languages such as SWRL (O'connor *et al.*, 2005) to support rule system interoperability over the Web or STTL to support the definition of RDF transformations (Corby *et al.*, 2015).

Maybe a controversial view on the advent of the Semantic Web is to see it as a major milestone in going toward a unified theory of KR in the sense that it succeeded to provide a convergence between historical KR formalisms and schools of thoughts by starting from a standard shared data structure, the RDF graph model. For instance work has been done to combine OWL and rules in a decidable, expressive and safe manner (Motik *et al.*, 2005). But also, more recently, discussions resurfaced about alternative knowledge representations for instance based on prototypes (Cochez *et al.*, 2016) echoing old debates in knowledge representation from the previous century. In addition, the fact the Semantic Web is built on Web standards also created a new way to transfer KRR research results to industrial contexts, usages and products.

The Semantic Web formalisms also brought new KRR problems such as the issue of blank nodes in RDF that still raises theoretical questions and empirical analysis of data publicly available on the Web (Mallea *et al.*, 2011). In addition, the contact between the Web and KRR in general, and in particular in the context of the Semantic Web, gave importance and even created specific challenging research questions. When moving from traditional KRR to Semantic Web a number of approaches encountered theoretical problems such as the "open world assumption", evolution, contradictions and practical problems such as scalability. This triggered work on language fragments with limited expressiveness and complexity on one hand and on approaches for scaling the processing such as distributed reasoning technique combining local reasoning chunks (Serafini, Tamilin, 2005) or scalable distributed reasoning using MapReduce (Urbani *et al.*, 2009) on the other hand.

In terms of reasoning, the Semantic Web approaches now encompass and combine many different processing techniques including deduction but also, induction, graph matching, learning, approximation, statistical methods, etc. To the problems of soundness, completeness and complexity were added the problems of precision, recall, quality, support, etc. (Hitzler, Van Harmelen, 2010). As a result topics like statistical KRR became important to address, for instance, uncertainty and vagueness of knowledge from the Web (Lukasiewicz, Straccia, 2008). As an other example, the changing nature of the Web also pushed the community to consider scalable defeasible reasoning approaches, for instance to update the classification of OWL ontologies and support the addition and deletion of axioms (Kazakov, Klinov, 2013) or to support defeasible logic reasoning on the Semantic Web (Bassiliades *et al.*, 2006).

Finally, very early the community also worked on providing tools and implementations as prototypes, proofs of concept, for evaluations and for deployment. Among the earliest open-source tools are:

– TRIPLE (Sintek, Decker, 2002) supporting inference and transformation of RDF data.

– the Jena Semantic Web toolkit (McBride, 2002) and its continuous effort to implement and support the Semantic Web recommendations and their evolution (Carroll *et al.*, 2004).

– CORESE (Corby *et al.*, 2000) that started by mapping RDF to conceptual graphs in order to exploit querying and inferring capabilities enabled by conceptual graphs formalisms (Corby *et al.*, 2004).

– Sesame: A generic architecture for storing and querying RDF and RDF schema with efficient storage and expressive querying (Broekstra *et al.*, 2002).

Many other tools and platforms have joined them since, and a number of them can be found on the Wiki page of the W3C titled "Tools - Semantic Web Standards"[3].

4. Ontologies and Semantic Vocabularies for Machines on the Web

The Semantic Web has been strongly ontology-oriented since the beginning in the sense that its early links to KRR was through ontology-based formalisms (e.g. descriptions logics, conceptual graphs) and also RDF and RDFS were drafted together. The notion of ontology in the semantic Web was sometimes misconstrued as the quest for the "one ontology to bind them all" while in fact the standards and the research on ontologies for the Web always acknowledged their plurality and never targeted a universal ontology for the Web. The semantic Web community talks about ontologies, schemata and vocabularies and their diversity is one of the challenges.

Like KRR in general, ontologies are discussed in the Semantic Web community since the first symposium in 2001 and the work on ontologies in the Semantic Web can be divided into two large families: (1) ontology languages and reasoning and (2) specific ontologies published on the Semantic Web. Looking at the contributions, every aspect of the life-cycle of ontologies is covered (engineering, extracting, publishing, visualizing, aligning / matching, reasoning, modularizing, evaluating, maintaining, etc.) but with the additional challenges and solutions that the Web brings to them.

The early years of the Semantic Web were marked by lightweight ontology language to start from (RDFS) and an effort to merge existing more expressive languages (DAML, OIL, DAML+OIL) to provide a starting point for the W3C's Semantic Web Activity's Ontology Web Language that will lead to Web Ontology Language (OWL) (McGuinness *et al.*, 2002). With a large support and involvement from their community, Description Logics became the core ontology language for the Semantic Web and

3. https://www.w3.org/2001/sw/wiki/Tools

drove the development of the OWL recommendations (Baader *et al*., 2005; Antoniou, Van Harmelen, 2004). OWL and its profiles were the subject of many contributions, proving their characteristics and providing efficient processing over the years such as entailment from satisfiability (Horrocks, Patel-Schneider, 2003) or class subsumptions computing (Krötzsch, 2012).

Ontological engineering methodologies were both needed and fed by the Semantic Web launch, driving the development process, supervising the ontology life cycle, and proposing the tools that support them (Gómez-Pérez *et al*., 2006). With the release of different languages and fragments or profiles of languages developers needed help to find the most suitable languages for their representation needs, requiring methodologies from the beginning (Gómez-Pérez, Corcho, 2002) and still more recently (Allemang, Hendler, 2011). Design practices and design patterns were also proposed for Semantic Web content to facilitate or improve the techniques used during the ontology life-cycle (Gangemi, 2005). In addition a lot of attention was paid to help automate, and therefore scale, part of the design work especially by providing methods for ontology learning from and for the Semantic Web (Maedche, Staab, 2001; Delteil *et al*., 2001).

Another impact of the Web on ontologies comes from the social dimension and scale of the Web. It shifted the focus from domain experts and knowledge engineers to the Web crowd and all its potential in terms of collaboration space first and crowdsourcing platform later. For instance the Web ontologies were immediately subject to the idea of collaborative ontology development (Sure *et al*., 2002; Tudorache *et al*., 2008) and, latter, ontology engineering tasks requiring human contributions were envisioned as micro-tasks to be crowd-sourced on online labor markets of the Web (Sarasua *et al*., 2012). Again as ontology started to change the Web, the Web started to change ontologies.

Just like for KRR, tools are often associated to results and advances on Web ontology languages. FaCT++ provides reasoning services for ontology tools supporting the OWL DL ontology language (Tsarkov, Horrocks, 2006). Pellet was created as a practical OWL-DL reasoner (Sirin *et al*., 2007). One of the most well known open-source tool is Protégé and its OWL support to provide an open development environment for Semantic Web application and schemata edition (Knublauch *et al*., 2004). The OWL Plugin is often used to edit ontologies in OWL and it allows its users to access description logic reasoners and to edit Semantic Web content.

Finally an important impact of the deployment of ontologies and the Semantic Web was the evolution toward the notion of vocabularies encompassing ontologies but also thesauri, lexicon, and other types of more or less formal vocabulary answering different needs and requiring different languages such as SKOS (Miles *et al*., 2005).

All this work also supported the second family of contributions on ontologies we mentioned at the beginning of this section: the actual publishing of vocabularies on the Web in large numbers to the point we needed, and now have, gateways to find these reusable semantic vocabularies on the Web such as the Linked Open Vocabularies

(LOV) (Vandenbussche *et al.*, 2017) and the service prefix.cc to manage the usual prefixes associated to their namespaces.

There would be no way to be exhaustive here, but some of the most well-known ontologies published on the Semantic Web include:

– The Semantic Web version of the early Dublin Core Metadata proposal for resource discovery (Weibel *et al.*, 1998) and in particular to represent documentary resources.

– The FOAF ontology (Brickley, Miller, 2007) to represent social networks of acquaintance and user profiles.

– The Creative Commons ontology describing copyright in RDF[4].

– The ontology provided by several Web giants on schema.org in particular to improve Web experience in searching and interacting with content exchanged over the Internet (e.g. Web search, browsing, email).

– The WGS84 Geo Positioning ontology for representing latitude, longitude and other information about spatially-located things[5].

– The Event Ontology (Raimond, Abdallah, 2007).

– The Time Ontology (Cox *et al.*, 2017).

– . . . and many more can be found on the LOV[6].

5. Matching, Aligning and Mapping the Vocabularies

In continuity with the previous section, a very immediate result of the contact of ontologies with the Web is the importance of being able to align or match different ontologies. This is a shared problem with the domain of ontology-based interoperability: the resulting alignments are useful for cross-enriching ontologies and increasing the interoperability on the Semantic Web (see section 6).

Ontology matching rapidly became a major research trend (Le *et al.*, 2004) and is still very active (Euzenat *et al.*, 2007) with a strong state of the art and renewed challenges (Shvaiko, Euzenat, 2013). The community compared, benchmarked but also very rapidly combined different approaches integrating, for instance, various similarity methods (Ehrig, Sure, 2004), formalizing ontology alignment and its operations (Zimmermann *et al.*, 2006) and also considering the quality of the mappings obtained from different techniques and proposing formal semantics to weight the ontology mappings (Atencia *et al.*, 2012). With the growing number of datasets published on the Web (see section 10), ontology alignment also relied on these linked open data to generate schema-level links between datasets (Jain *et al.*, 2010) and to find concept

4. https://creativecommons.org/ns
5. http://www.w3.org/2003/01/geo/wgs84_pos#
6. http://lov.okfn.org/

coverings and alignments between concepts in ontologies from multiple Linked Data sources (Parundekar *et al.*, 2012).

As we will see in the next section the problem of matching, aligning and mapping now concerns (RDF) resources in general with approaches relying on both the schemata and the data to detect mappings between identifiers in general, beyond vocabularies.

6. Interoperability or the Web as a Distributed Semantic Blackboard

The research contributions on interoperability are closely related to ontologies as these shared conceptualizations are a cornerstone of formal knowledge exchange and semantic integration. By coupling the results in ontology-oriented approaches and the standardization offered by Web languages, the Semantic Web was rapidly identified as supporting new approaches to semantic integration developed by researchers in the ontology community. It provided testbed for scalability and a common ground for hybrid approaches (N. F. Noy, 2004).

Interoperability advances also heavily relied on methods from ontology matching (section 5). Indeed data integration benefits not only from aligned ontologies, but also adapts the ontology alignment methods to resource alignment, entity resolution and key generation. Research topics included: semantics for distributed systems and their relations with alignment composition (Zimmermann, Euzenat, 2006); scalable and distributed methods for entity consolidation to detect identifiers that correspond to the same entity (Hogan, Zimmermann *et al.*, 2012); or more recently on the problem of unsupervised entity resolution on multi-type graphs (Zhu *et al.*, 2016).

Another topic related to interoperability is the integration of legacy systems, heterogeneous systems and in particular the interface between the Semantic Web and traditional relational databases. Very early we saw propositions of approaches for mapping relational databases to RDF (Sahoo *et al.*, 2009) either for linked data publication from relational databases (Auer *et al.*, 2009) or for querying relational databases with Semantic Web languages for example through SQL views (Sequeda, Miranker, 2013). Bridging the gap between relational databases and the Semantic Web also covers the problems of the creation of an ontology from an existing database instance and the discovery of mappings between an existing database instance and an existing ontology (Spanos *et al.*, 2012).

The pursue for interoperability can also be seen as viewing the semantic Web as a way to turn the Web into a universal distributed semantic blackboard where very different kinds of intelligence can co-operate.

7. Retrieving and Querying Knowledge Graphs on the Web

Starting from the publication and access to RDF through HTTP the question of retrieving knowledge pieces and querying the Semantic Web sources quickly focused on the need for a query language that became SPARQL. Then this trend continued with studies on one hand on extensions, scalability and optimization of the query language and on the other hand on alternative access and query mechanisms and architectures.

In 2005, we had several proposals of query languages (SquishQL, RDQL, TriQL, RQL, SeRQL, etc.) (Bailey *et al.*, 2005) and although SPARQL became the standard, extensions, variants and alternatives are still studied nowadays. As a standard, many aspects of SPARQL are studied such as the semantics and complexity of the query language (Pérez *et al.*, 2006) and its patterns (Angles, Gutierrez, 2016). Another important question is the query answering in the presence of ontologies with different strategies to materialize, rewrite or filter query answers or different techniques to improve efficiency in terms of computation time or memory space (Lutz *et al.*, 2013). This research trend is also called Ontology-Based Data Access (OBDA).

Rapidly the questions of performance and scalability became an active research area in its own. Optimization techniques have been proposed, for instance, using selectivity estimation of SPARQL basic graph patterns and heuristics for static optimization (Stocker *et al.*, 2008). On the other hand benchmarks have been designed and made available in particular to assess the performance systems for real queries on real data (Morsey *et al.*, 2011).

In terms of extension, a special case is the support for approximate query processing based on ontologies metrics (Corby *et al.*, 2006) or RDF Query relaxation based on failure causes (Fokou *et al.*, 2016). In both cases the core idea is to extend the query mechanism with algorithms and operators to retrieve close results as an alternative to no answers.

In terms of alternatives to SPARQL endpoints, one of the most well-known initiatives is the use of Triple Pattern Fragments recommending client-side querying for single knowledge graphs and federations, to reduce server load and increase caching effectiveness (Verborgh *et al.*, 2016). This alternative is based on the Linked Data Fragments framework to analyze Web interfaces to Linked Data and to compare them (Hartig *et al.*, 2017).

Finally, the Semantic Web is not a centralized data warehouse but a network of distributed and linked datasources. As a result a very active domain of research on query mechanisms is the study of distributed and federated querying.

A first problem in that domain was to manage federated repository for querying graph structured data with distributed indexing methods and parallel query evaluation methods for instance on a cluster of computers (Harth *et al.*, 2007). One motivation for this is to have storage and querying architectures that can scale. Query rewriting and cost-based query optimization were proposed to speed-up federated query execution (Quilitz, Leser, 2008) and optimization techniques for federated query processing on linked data became a research topic (Schwarte *et al.*, 2011). SPARQL 1.1 federation extension syntax, semantics and possible optimization techniques when dealing with large amounts of intermediate and final results is another example of research in that area (Buil-Aranda *et al.*, 2013).

A second problem is the case of executing SPARQL queries over the distributed sources of the Web of linked data and discover data that might be relevant for answering a query during the query execution itself. In the work of (Hartig *et al*., 2009) the discovery is based on the "follow your nose" principle of the Web[7] by following RDF links between data sources based on URIs in the query and in partial results. The discovered URIs are resolved over HTTP to obtain new RDF data continuously added to the queried dataset. This problem can also be specialized for specific parts of the query such as property path patterns looking at their query semantics and how it can be coupled with navigating the Web graph (Hartig, Pirrò, 2015).

Finally, the special case of hybrid search queries in a federation of multiple data sources looks at extending the SPARQL algebra to incorporate keyword search to express queries on distributed and heterogeneous data sources on the Web (Nikolov *et al*., 2013).

8. Data Management and the Big (Web of) Data

To some extend, the topic of data management generalizes the two previous ones: (1) interoperability and the links to databases for persistence and legacy reasons for instance, and (2) querying and accessing data. The contributions to data management consider all the steps of the life-cycle of data and datasets with an emphasis on: volume and scalability, infrastructure and robustness, availability and performance.

Very early, this led to the evaluation of knowledge base systems for OWL datasets that are getting larger and larger every year (Guo *et al*., 2004). Benchmarks were designed for OWL knowledge base systems together with methods for benchmarking Semantic Web knowledge base systems with respect to their use in large OWL applications (Guo *et al*., 2005). Competitions and challenges were also proposed such as, for instance, the Billion triple challenge.

Concerning databases, and besides query rewriting and data transformation we mentioned before, contributions were made on indexing RDF data efficiently with regard to its triple structure (Weiss *et al*., 2008) and on different partitioning approaches (Abadi *et al*., 2007). Ideas have also been adapted from traditional database and data management, for instance for defining views to create virtual resource descriptions and schemas customized to the needs of specific applications (Magkanaraki *et al*., 2003). This domain topic also generally considered heterogeneity and the integration and management of different kinds of data just like, in parallel, the data activity at W3C generalized the Semantic Web activity and the focus moved from established core standards to bridging them and other data models.

Another important aspect are the temporal dimension and dynamic nature of data with problems of versioning, streaming, updating and propagating changes. Examples of research questions here include the computation of differences that exist be-

7. https://www.w3.org/wiki/FollowYourNose

tween two RDF models to reduce the amount of data that needs to be exchanged and managed over the network and hence build advanced synchronization and versioning services (Zeginis *et al.*, 2007). Efficient formats for publication and exchange were also proposed such as a binary RDF representation (Fernández *et al.*, 2013). A last example are the approaches proposed to support stream reasoning and to bridge the gap between reasoning and stream processing (Margara *et al.*, 2014).

9. Software Architectures for and by Semantics

The research on software architecture in the Semantic Web community comes in two flavours: software architecture to support the life-cycle of the Semantic Web and Semantic Web approaches for metadata and their processing in software architectures. Again, there is a link to previous sections on finding the adequate architectures for federation or distribution of data and on distributing or parallelizing processing.

An early example of software architecture for and by Semantic Web are the Semantic Web-Services which started as a hot topic in 2001. In this domain, semantic metadata can be used to characterize services, and service architecture can be built to support meta-data life-cycle. An early problem addressed was the semantic matching of Web Services Capabilities since the first step toward interoperation and composition is the location of other services and the semantic match between a declarative description of a service being sought, and a description of the services being offered (Paolucci *et al.*, 2002). Ontologies were developed to describe services at the application layer (Ankolekar *et al.*, 2002) such as DAML-S and OWL-S. The next step was to propose automated composition approaches (Wu *et al.*, 2003) and formal models for that (Lécué, Léger, 2006). Later the Semantic Web supported representing and taking into account other non functional aspects of services in their management. For instance (Kuter, Golbeck, 2009) generate OWL-S compositions of Semantic Web services using social trust information from user ratings of the services relying on a taxonomy of features, such as interoperability, availability, privacy, security, etc.

Close to Semantic Web services, an alternative architecture comes from Distributed Artificial Intelligence: autonomous agents and multi-agent systems (MAS). The distributed multi-agent architecture can be leveraged to address the distributed Semantic Web data sources and vice-versa ontology-oriented knowledge representation can be used to formalize agent profiles, messages, protocols, knowledge and in general metadata in societies of autonomous agents (F. Gandon, 2002). This approach was, for instance, applied to knowledge management in corporate Semantic Web (F. Gandon *et al.*, 2002).

Two other architectures that met the Semantic Web are grids and clouds. Again they can be used to provide processing and storage approaches to the Semantic Web or benefit from the Semantic Web and its KRR approaches for their own metadata needs. For instance, (Tangmunarunkit *et al.*, 2003) solve resource matching in the Grid using Semantic Web technologies. On the other hand, (Ngomo *et al.*, 2013) assesses cloud computing and different parallel processing paradigms hardware, including the use of GPUs and MapReduce platforms, for link discovery.

Last, but certainly not least, a number of contributions have looked at peer-to-peer (P2P) architectures in particular to support fully decentralized storage, querying and reasoning. This research trend started very early and is still very active. Peer-to-peer Semantic Web started by looking at distributed environments for sharing semantic knowledge on the Web (Arumugam *et al.*, 2002) and specific application such as a semantics-based bibliographic peer-to-peer system (Haase *et al.*, 2004). It evolved from decentralized management and exchange of knowledge and information (Staab, Stuckenschmidt, 2006) to distributed reasoning in a peer-to-peer setting (Adjiman *et al.*, 2006). It also now touches other very specific aspects of the linked data architecture for instance to provide decentralized caches for triple pattern fragments based on an overlay network weaved from linked data fragments similarity. These P2P caches are then used to answer queries efficiently (Folz *et al.*, 2016).

The P2P architecture is also the occasion to note that some works consider mixing several architectures for instance to provide a scalable P2P infrastructure of registries for semantic publication and discovery of Web services (Verma *et al.*, 2005).

As a final note, currently the software architecture witnesses a lot of interest for Web APIs and RESTful architecture (Fielding, Taylor, 2000) with bridges for instance at the language level (e.g. JSON-LD) or at the architecture level (e.g. Linked Data Platform LDP) and with the idea of fully investigating the capabilities of the distributed hypermedia software architecture or HATEOAS[8]. This application-centric view is also supported by the provision of dedicated programming languages for linked data (Corby *et al.*, 2017) and software architecture to integrate APIs and linked data (Michel *et al.*, 2018). Coming back to the general topic of software architectures for and by semantics, although the topic of Semantic Web-Services is no longer active at the moment of writing this article, it does share a lot of challenges with other currently very active topics such as Web APIs and Web of things. These topics include in particular the annotation of programming interfaces, of software and hardware capabilities, of offered operations and services, and of composition, for instance.

10. Linked (Open) Data and Schemata on the Web

In *circa* 2005 the topic of linked data and Web of data appeared as a sub-domain in itself focusing on methods, approaches and practices for publishing and connecting structured data on the Web. In particular Linked Open Data (LOD) and the LOD cloud (Bizer *et al.*, 2007) started with pioneering initiatives such as DBpedia, one of the first seeds for a Web of data (Auer *et al.*, 2007; Bizer, Lehmann *et al.*, 2009) based on a method for revealing Wikipedia structured content by extracting information from template instances (Auer, Lehmann, 2007). It was followed by Yago (Suchanek *et*

8. Hypermedia As The Engine Of Application State

al., 2007) and later Wikidata (Vrandečić, Krötzsch, 2014) also exposing, cleaning and structuring data from the Wikimedia projects. Other seeds from more specific domains included DBLP, Geonames or Music-brainz for instance. Conferences and journals also started to have resource tracks and special issues to publish descriptions and characteristics of new ontologies and datasets made available on the Web of Data.

The whole linked data trend is based on principles and methodologies for weaving a Web of data. The concepts, technical principles and progresses of Linked Data on the Web were studied and documented over the years (Bizer, Heath, Berners-Lee, 2009). Best practices and handbooks were provided to help the adoption (Heath, Bizer, 2011) as well as common errors in RDF publishing on the Web, their consequences for applications and approaches to improve the quality of structured, machine-readable and open data on the Web (Hogan *et al.*, 2010). To support the Web of data publishing activity data-lifting approaches and tools were also proposed to facilitate the contributions to the Web of data in terms of transforming, linking and publishing linked data (Scharffe *et al.*, 2012).

This last point introduces a subsequent research topic of the LOD: the study and fostering of the quantity and quality of data and the provision of crawling, indexing, selecting, filtering and ranking algorithms. As an example, to guide users among data sources, (Franz *et al.*, 2009) provide a relevance ranking of the available data. The available data can also be leveraged to provide new metrics for instance to find the most relevant entity type based on statistics and on the graph structure interconnecting entities and types (Tonon *et al.*, 2013). With this growing volume of data available, the community also gained the resources to conduct empirical studies such as surveys of linked data conformance and quantitative empirical analyses of crawled data with regard to guidelines and best practices (Hogan, Umbrich *et al.*, 2012).

As linked data grew, they also provided new challenges and material to benchmark the solutions proposed by the community. The problem of discovering links between data published on the Web called for frameworks such as SILK (Volz *et al.*, 2009) and time-efficient approaches for large-scale link discovery (Ngomo, Auer, 2011). Experiments at LOD scale in terms of volumes and variety were proposed to evaluate performances in the wild (Rietveld *et al.*, 2015).

11. Semantics in Security, Trust and Privacy

The coupling with the Web also pushed the consideration of quality, transparency, security, uncertainty and trust in all aspects of KRR. As soon as the Semantic Web started to advertise the publication of data the problems of security & privacy and provenance & quality were raised under the common concern of ensuring the trust of practitioners and users. The community also soon demonstrated that while semantics can be used to improve and enrich access to knowledge, it can also be leveraged to finely control and restrict the access to personal or confidential data.

The publishing of data very early raised the need to have privacy, confidentiality, access control and security enforcing mechanisms. Languages for policy-based security were proposed for the Semantic Web (Kagal *et al.*, 2003) as well as approaches to enforce access control with context-awareness relying on semantics to reason about access rights, access contexts and levels of details to expose (F. L. Gandon, Sadeh, 2004). The question again had two sides: adapting access control approaches to the Semantic Web but also using semantics to provide new ways to capture and declare privacy and security policies. For instance the question of representing role-based access control and more generally attribute-based access control in OWL and performing security analysis in a trust-management framework was studied in (Finin *et al.*, 2008) as well as context-aware access control to RDF triple stores (Costabello *et al.*, 2012) and this trend of works led to various access control models, standards and policy languages, and different access control enforcement strategies (Kirrane *et al.*, 2017).

Once the data have been released comes the problem of representing metadata to characterize datasources and the data they contain. Provenance metadata are needed to support data traceability. This need was a major motivation for the introduction of named graphs in RDF (Carroll *et al.*, 2005) providing a generic metadata structure for RDF that was standardized in RDF 1.1. When coupled with ontologies such as PROV-O (Lebo *et al.*, 2013) or the open provenance model vocabulary defining a lightweight provenance vocabulary (Moreau *et al.*, 2011) this approach to RDF meta-annotation provides means to enable data producers to publish their data responsibly. Contributions were also made at the Web architecture level and linked data practices level. As an example, trusty URIs support verifiable, immutable, and permanent digital artifacts identified in linked data by URIs including cryptographic hash values (Kuhn, Dumontier, 2014).

The next step is to use the metadata about RDF pieces when querying and reasoning, and propagate data annotations to the results. With the increased amount of inconsistent and non-reliable data on the Web, representing and reasoning with annotated data was studied (Zimmermann *et al.*, 2012) as well as robust and scalable linked data reasoning incorporating provenance and trust annotations (Bonatti *et al.*, 2011). An example of specific problem is querying probabilistic instance data in the presence of OWL ontology and computing answer probabilities (Jung, Lutz, 2012). The metadata may also require specific type of reasoning such as in the case of licenses attached to data requiring deontic reasoning (Governatori *et al.*, 2013). In addition, the data processing itself is likely to generate additional metadata (Hasan, Gandon, 2012) such as justification for an entailment in an OWL ontology in the form of a minimal subset of the ontology that is sufficient for that entailment to hold (Horridge *et al.*, 2008). These metadata can also be published as linked data and form linked justifications or linked explanations that can be used by data consumers and when interacting with the users (Hasan, 2014).

Finally, researchers studied the notion of trust in the context of the Semantic Web as a metric of how much credence to give each source and as a network structure (Web of trust) in which each member maintains trusts in a small number of other members. They studied how these trusts can be composed and personalized (Richardson *et al.*, 2003) and how they propagate or not on social networks and which algorithms can be used to infer trust relationships (Golbeck *et al.*, 2003), leading to a whole new area at the cross-road of trust research in computer science and the Semantic Web (Artz, Gil, 2007).

12. Learning and Mining with and for Semantics

The Semantic Web frameworks provide standardized data structures, linked data principles and formalisms for ontology-oriented knowledge representation above these. Each layer can support different kinds of artificial intelligence processing including reasoning but also: mining, clustering, classifying, learning, estimating, extracting, checking, etc.

Discovering, mining and extracting knowledge and semantics is a topic of high interest as it contributes to feed the linked datasets. Many approaches have been proposed in this area that combine Semantic Web data with the data mining and knowledge discovery process (Ristoski, Paulheim, 2016). The research topic covers a spectrum from unstructured data mining to formal knowledge mining as the mined input becomes richer in terms of structure and semantics. Very early, in (Berendt *et al.*, 2002), Semantic Web mining is presented as combining Semantic Web and mining with the double goal to improve, on the one hand, mining by exploiting semantics and to make use, on the other hand, of mining to feed the Semantic Web data sources (Berendt *et al.*, 2002).

Machine learning approaches have been integrated to the intelligent techniques needed and supported by the Semantic Web since the beginning. And the latest challenges in the field of machine learning still have their echos in the domain such as, recently, representing and exchanging embeddings and deep learning on the Semantic Web. Learning was first used to help address challenges of the Semantic Web. A first example is to learn to map between ontologies on the Semantic Web. The problem is to learn a mapping which, for each concept in one ontology, gives the most similar concept in the other ontology (Doan *et al.*, 2002). Another example is ontology learning i.e. machine learning techniques that supports semiautomatic ontology construction tools, encompassing ontology import, extraction, pruning, refinement and evaluation (Maedche, Staab, 2001). And the research on these topics continued with now methods for (semi-)automatically building and enriching ontologies by inductive learning from existing sources of information such as Linked Data, tagged data, social networks, ontologies (d'Amato *et al.*, 2010). As mentioned about data mining, machine learning approaches can provide a variety of methods applicable to different expressivity levels of Semantic Web knowledge bases with a range statistical inferences applicable from just bare RDF graphs up to rich semantic representations (Rettinger *et al.*, 2012). The challenge of adapting machine learning approaches to the RDF graph data model also led to interesting specific problems such as kernel-based machine learning algorithms tailored to be applied to instances represented as RDF graphs (Lösch *et al.*, 2012).

Here again, the contact with the Web led to the study of scalable machine learning for Linked Data. For instance (Nickel *et al.*, 2012) made a contribution based on the distributed computation and factorization of a sparse tensor that scales and are also able to incorporate ontological knowledge to improve learning results. Concerning the topic of link discovery mentioned before, (Ngomo, Lyko, 2012) proposed an active learning approach based on genetic programming to generate link specifications i.e. a specification of the conditions under which a link is to be built.

13. Natural Language processing with and for Semantics

One of the first motivations for the use of natural language processing (NLP) in the Semantic Web is shared with the previous section: knowledge extraction. Except in this case the input data are exclusively texts. The tasks involving natural language processing include: ontology learning, linked data bases population, entity resolution, text annotation, natural language querying and question answering, etc.

The idea of bootstrapping the Semantic Web via automated semantic extraction and annotation started very early with the goal to provide platforms for large-scale text analytics and automated semantic tagging of large corpora (Dill *et al.*, 2003). Approaches combined unsupervised pattern-based approach to categorize instances with regard to an ontology and to identify certain ontological relations with the idea of using the enormous corpus of the Web to overcome data sparseness (Cimiano *et al.*, 2004) really taking the best of both domains.

A specific sub-topic emerged with the need to recognize named-entity (NER) in text data as the key first step towards extracting RDF data. The challenge is to disambiguate and detect the correct URIs for a given set of named entities within an input text. Different methods started to be proposed and compared such as unsupervised extraction (Etzioni *et al.*, 2005), disambiguation of named entities using linked data (Usbeck *et al.*, 2014), and approaches that combine the state-of-the art from named entity recognition in the natural language processing domain and named entity linking from the Semantic Web community (Rizzo *et al.*, 2014). This research led to the provision of famous NER services such as DBpedia spotlight for automatically annotating text documents with DBpedia URIs (Mendes *et al.*, 2011).

Methods from natural language processing can also be used for specific tasks such as ontology design from natural language texts by combining Discourse Representation Theory, linguistic frame semantics, and ontology design patterns (Presutti *et al.*, 2012). Inversely, specific structures or relations can be targeted such as detecting arguments in natural language in social platforms or medias and returning the relations among them to provide an overview view of the argumentative discussion (Cabrio *et al.*, 2013). And the state of the art of existing knowledge extraction methods for the

Semantic Web is growing to cover different tasks of the Semantic Web (Gangemi, 2013).

As it was the case in other sections, the cross-fertilization between natural language and Semantic Web goes both ways. Indeed, natural language services and resources can benefit from Semantic Web and linked data for integration and reuse purposes. For instance, the NLP Interchange Format (NIF) relies on URIs to identifying textual elements and an ontology of common NLP primitives to support the creation of heterogeneous, distributed and loosely coupled NLP pipelines over the Web (Hellmann *et al.*, 2013) and one can even publish the NLP results as linked data (Rizzo *et al.*, 2012).

This research area also raised specific new questions such as the ones of a multilingual Semantic Web dealing with data expressed or extracted from different natural languages and cultures. One research topic then is to provide methods ensuring that data expressed in a certain language are accessible to speakers of other languages (Gracia *et al.*, 2012).

Natural language is also needed to support natural language based human-machine interactions. Natural language interfaces to Semantic Web resources hide the complexity of the linked data, ontologies and formal languages from the user behind an natural language interaction (Lopez *et al.*, 2013). Different tasks can benefit from that user-friendliness with more or less freedom in the natural language accepted. Controlled natural languages may be used, for instance, to guide users' input when editing ontologies (Bernstein, Kaufmann, 2006). Ontology-based question answering on the Semantic Web (Unger, Cimiano, 2011) and multilingual question answering over linked data (Cimiano *et al.*, 2013) target the generation of formal queries from natural language questions with a growing complexity. Inversely, Natural Language Generation from linked data is concerned with transforming some formal content input into a natural language output (Bouayad-Agha *et al.*, 2014) such as a text document, an answer, a question for a quiz, etc.

Finally, and in direct link with the next section, natural language interfaces have been shown to offer a user-friendly option to query ontology-based knowledge sources. Their usability and the alternative options have been studied (Kaufmann, Bernstein, 2007).

14. Semantics in Interactions and their Design

Although many Semantic Web languages and algorithms naturally fit at the backend of software architecture, the need for editors and interfaces even for developers quickly appeared to be vital for adoption and tools like Protégé (Knublauch *et al.*, 2004) were rapidly proposed.

More generally, the need to design interactions and visualizations and the opportunity to support context adaptation, user personalizing and semantically-driven interactions are example of key topics in this area.

Since we are on the Web, a first important interaction task that was considered was browsing. First as an existing interaction that can be augmented by Semantic Web data. For instance, the Magpie browser (Dzbor *et al.*, 2003) was designed to offer complementary knowledge to the user to support the interpretation of Web pages viewed. On the other hand the new resources published on the Web by the Semantic Web called for new browsing techniques. The Tabulator is an example of RDF browser to give humans access means to the Web of data and really experience the linked data paradigm (Berners-Lee *et al.*, 2006). Other established Web interaction approaches were also combined with Semantic Web such as facet-based approaches (Hildebrand *et al.*, 2006) and the generation of Web portals above linked data (Corby *et al.*, 2015).

Related to the notion of browsing, yet independent of a specific browser or browsing technique, is the sub-question of resource-centered visualization and the notion of presentation lenses and semantic style-sheets. Fresnel is a browser-independent presentation vocabulary containing core RDF display concepts to promote the exchange of presentation knowledge (Pietriga *et al.*, 2006). It was extended later for instance to integrate context-aware adaption of the presentation of RDF data to a user (Costabello, Gandon, 2014).

Another immediate need, besides browsing, is visualizing linked data and schemata. This is a special case of information visualization for semantic annotation, and of data visualization for linked data with the questions of what visualization techniques can do, where they must adapt and inversely what the Web can bring to the table. Visualizing Semantic Web structural information and ontology-based data was studied from an information visualization and graphical representation perspective (Geroimenko, Chen, 2006). Researchers also designed interaction between users and semantic data and proposed visualization techniques for semantic data and dynamic queries based on graspable dimensions, such as space and time to support sense-making (Petrelli *et al.*, 2009). The issue is also to assist both non-domain and non-technical users in reaching a good understanding and querying abilities and to support knowledge making (Dadzie, Rowe, 2011). Beyond visualization, exploratory browsing and exploratory search over linked data have been studied as a specific category of interactions (Marie, Gandon, 2014).

As we saw with the special case of enriched browsing, this research area is not limited to interactions with the Semantic Web resources and it also includes the potential of the Semantic Web to improve user experience in general. For instance the task of personalizing and enriching educational and learning resources may benefit from Semantic Web methods and resources, optimizing recommendations and adaptation of pedagogical materials (Dolog *et al.*, 2004). Another case for having special adaptation mechanisms is when taking into account the context and the device (e.g. mobile access) for instance handheld—multimodal interaction with ontological knowledge bases and Semantic Web services (Sonntag *et al.*, 2007). Many usage scenarios of applications mentioned in section 16 actually have to design domain-specific or task-specific interactions and interfaces.

So, to open the conclusion on this research area, from the interaction design point of view a whole domain that can be explored is the one considering how the Semantic Web resources and services can be leveraged to improve the users' experience with the devices surrounding them (F. Gandon, Giboin, 2017).

15. Semantic social networks and media

Moving from the individual level of the previous section to the collective dimension, this last topic is extremely important as it re-emphasizes the social dimension that the Web brings into the Semantic Web. It will no longer come as a surprise that the contributions to the topic of semantic social networks and media go both way: from linked data to support and model social media structure, activity and content, and from social platforms and sciences to feed, enrich, improve, etc. linked data. To some extent, 2005 is the social year for Semantic Web with important papers opening two research directions: the semantic representation and interlinking of online communities, and the semantic-based analysis of social networks and social media.

The motivation for semantically-interlinked online communities (Breslin *et al.*, 2005) is to enable connectivity and interoperability across communities and platforms by providing a lightweight ontologies (e.g. SIOC) supporting access, linking, querying and transfer of data and accounts from one social application to another (Breslin *et al.*, 2006). The FOAF ontology (Brickley, Miller, 2007) to represent user profiles, accounts and social relations was an important piece of the puzzle and these pioneering contributions started the idea of a social Semantic Web (Breslin *et al.*, 2009).

Complementary, the work of (Mika, 2004) proposes to feed methods from social network analysis by relying on ontology-oriented representations of social network data and mining for online data acquisition. The author then proposed to extend the bipartite model of ontologies with social aspects creating a tripartite model composed of concepts, instances and actors thus merging ontological models and social network models in a community-based ontology model (Mika, 2005b). Social medias also come with their structures and models such as folksonomies. These can be combined with Semantic Web models and used to find communities and improve search (Hotho *et al.*, 2006), they can be enriched with semantics (Specia, Motta, 2007) to improve community exchanges (Limpens *et al.*, 2013).

From this point, several research directions were opened starting with semantics-based social media analysis. In that context, the Semantic Web framework is used to extract, aggregate and visualize online social networks, reasoning with personal knowledge acquired from a number of sources and used for social network analysis and community presentation (Mika, 2005a). The semantics can then be leveraged to formally define and extend social network analysis metrics using Semantic Web frameworks for reasoning, querying and analyzing the communities and their activity (Erétéo *et al.*, 2009). The social Semantic Web provides a powerful framework to jointly analyze the network structure, the communication behaviors and the content of the communication, for instance to measure the dynamic bi-directional influence between content and social networks (Wang, Groth, 2010). Conversely, this joint representation can help enrich the content and interaction. As an example, approaches have been proposed for the semantic modeling of Twitter users based on their posts and for linking posts with related content to contextualize the activities (Abel *et al.*, 2011). Even the structure of the discussions can be enriched using for instance argument graphs (Cabrio *et al.*, 2013).

This research direction is also related to research on trust network and social propagation of trust mentioned in section 11. For instance, authors of (Golbeck, Hendler, 2004) proposed an approach for calculating locally the reputation ratings from a Semantic Web Social Network and applied it to rate emails.

These models and methods combining Semantic Web and social network also supported the emergence of social semantic applications. In (Gruber, 2008), "collective knowledge systems" are defined as a class of applications supporting collective intelligence on the Social Web with KRR techniques of the Semantic Web. An important special case of collaborative Semantic Web applications are semantic wikis. Reconciling Semantic Web and social Web in one application, semantic wikis allow every user to be an active provider and consumer of information. For instance (Buffa *et al.*, 2008) makes heavy use of Semantic Web concepts and languages, and demonstrates how the use of such paradigms can improve navigation, search, and usability in a wiki. One of the most important example was the semantic extension to be integrated in the Wikipedia open-source engine allowing the typing of links and entities directly inside the articles (Völkel *et al.*, 2006). This prefigured the idea of Wikipedia becoming a rich source of data for the Semantic Web.

Finally the social applications also provided new solutions to existing problems of the Semantic Web. We can mention, for instance, the collaborative edition of ontologies and datasets. An innovative example is the use of crowd-sourcing to contribute to the acquisition or curating of knowledge. For instance (Waitelonis *et al.*, 2011) designed a game with a purpose to evaluate linked data heuristics with a quiz that cleans up DBpedia, detects inconsistencies in Linked Data and scores properties for semantic search. But, beyond this example there is a whole research area at the cross-road of Semantic Web and Human Computation (Sabou *et al.*, 2018).

16. Semantic Web in Use: application scenarios and domains

Since the beginning of the Semantic Web research community there has always been conference tracks and journal special issues for its applications, its industry adoption and its benchmarks and challenges competitions.

Looking at publications in Semantic Web venues, the application domains include at least:

– Knowledge Management and Content Management systems.

– Information retrieval and Search engines, semantic searching, ranking and filtering.

– Information systems and their integration, enterprise applications, intranets, private Webs.

– Multi-media systems, multi-media, annotation, video annotation, music collections annotation.

– Education and e-Learning.

– Cultural data and cultural heritage, cultural events and programs.

– Life Science, Healthcare Medical and Biomedical Applications.

– Scientific applications and e-Science.

– Publishing industry, libraries.

– Public Sector, Government and e-Government.

– Legal systems, risk and compliance.

– Software and systems engineering.

– Industry, Manufacturing and Automation, and Process Models.

– Environmental Data.

– Sensors and data streams.

– Internet of Things and smart thing, smart homes, smart cities, smart planet.

– Mobile platforms and Mobile Web.

Knowledge management and information management have always been application domains of KRR and the advent of the Semantic Web made them an application domain for it too. In fact Web approaches in general are popular options to provide standard-based intranet applications and interoperability between legacy systems. Just like intrawebs are based on open Web technologies, corporate Semantic Web applications apply on intranets, behind firewalls, the Semantic Web and linked data approaches (F. Gandon, 2002) for standard based ontology-driven knowledge management (Davies *et al.*, 2003).

A special case of knowledge management is educational knowledge management for e-Learning. This was also identified as a promising application scenario of the Semantic Web very early, in order to modularize, open up, share and reuse semantically annotated pedagogical resources and services of educational systems (Aroyo, Dicheva, 2004) and support ontology-based reasoning and personalizing in e-Learning for instance by adapting adaptive educational hypermedia methods (Henze *et al.*, 2004). More recently new directions have been investigated such as the use of available linked open data on the Web to automatically generate educative and customized quizzes (Rodríguez Rocha, Faron Zucker, 2018).

Another important special case of knowledge management is in the scientific domain with e-science and the need to share knowledge, data and services among research scientists. In this domain we had a lot of contributions in terms of ontologies and datasets produced and published on the Web of data. The Gene Ontology (Ashburner *et al.*, 2000), Bio2RDF (Belleau *et al.*, 2008) and BioPortal (N. Noy *et al.*, 2009) are examples in the pioneer domain of bioinformatics. This domain adopted Semantic Web approaches to mashup, integrate and build scientific knowledge systems. In addition, the availability of the Web of data is also supporting new innovative ways of doing science for instance by generating hypotheses for possible interpretations of statistical results from Linked Open Data graphs (Paulheim, 2012). This is especially interesting at a time where we are looking for explainable systems and automated explanation generation.

The multimedia collections and content publishers also soon identified the potential of the standard annotation framework offered by the Semantic Web to represent, exchange, reason and query on the indexes of their resources. Semantic Web annotation of images and videos supports multimedia indexing and analysis for instance by linking low level MPEG-7 visual descriptions to ontology-based Semantic Web annotations (Bloehdorn *et al.*, 2005) and applying linked data principles to multimedia fragments (Hausenblas *et al.*, 2009). In this domain, BBC was a pioneer integrating data and linking documents across domains (Kobilarov *et al.*, 2009). In turn, the multimedia metadata support new usages and interactions such as the exploratory search of videos (Waitelonis, Sack, 2012).

In a similar way, cultural institutions saw the Semantic Web as a new way to index their collections and support exchanges between cultural actors. They also rapidly identified the opportunities to support semantics-driven recommendations and museum tour generation (Aroyo *et al.*, 2007) via semantic annotation and search of cultural-heritage collection (Schreiber *et al.*, 2008).

The case of museum tours also touches the domain of mobile applications, geo-located usages and data and dynamic information about occurring events. A growing application domain of the Semantic Web is the integration of APIs and data streams coming from connected objects, sensors, smart devices and smart places for instance to support spatial ontology-mediated query answering over mobility streams (Eiter *et al.*, 2017) or predict the severity of road traffic congestion (Lécué *et al.*, 2014).

Last but not least, search engines, information retrieval systems and their APIs are a key component of the Web and they were among the first services to be revisited by the Semantic Web community. In particular, researchers worked on providing new search engines for the Semantic Web or on improving Web search engines and information retrieval with Semantic Web technologies. On the first topic, Swoogle supports the search for metadata on the Semantic Web (Ding *et al.*, 2004) while Sindice was trying to foster the weaving of linked open data providing a search engine and a lookup index over the resources it crawled (Tummarello *et al.*, 2007). Watson and its applications extended this idea with APIs for other applications to find, select, exploit, and combine the knowledge available on the Semantic Web (d'Aquin *et al.*, 2008). On the second topic, classical information retrieval search models were extended, for instance, to integrate ontology-based semantic search capabilities in searching for documents (Fernández *et al.*, 2011).

Let us close that section by remembering one of the motivations of linked data is that "applications pass but data remain".

17. Concluding remarks

Calling the Semantic Web what it is — The Semantic Web is now an established domain with courses and handbooks. But even the differences in the titles of these handbooks show the different names of the domain we surveyed, stressing different aspects and points of interest: "Foundations of Semantic Web Technologies" (Hitzler *et al.*, 2009), "Semantic Web for the Working Ontologist: effective modeling in RDFS and OWL" (Allemang, Hendler, 2011), "Linked Data: Evolving the Web into a Global Data Space" (Heath, Bizer, 2011), and "A Semantic Web Primer" (Antoniou, vanHarmelen, 2004) are some examples. In fact, one difficulty for newcomers to enter the domain of linked data on the Web is that the initiative is presented under different names, each name insisting on a different facet of this evolution of the Web. The term "Web of data" stresses the idea of a Web where to open silos of data of all sizes, from the small data of your car maintenance to immense databases of astronomy, and to exchange them on the Web according to our needs. The names "linked data" and "linked open data" or LOD empathizes three things: the added value of linking data on the Web to integrate different sources; the wealth of having open data as commons available to everyone's applications; and the fact that all the approaches of the domain can be used in private spaces (intranets, intrawebs, extranets, etc.). The expression "giant global graph" insists on the larger perspective of the growing amount of links between data distributed on the Web and which weave a giant graph. Finally, the historical name of the "Semantic Web" reminds us of the ability we have to also exchange our data schemata, in addition to datasets, in order to enrich the range of automatic processing that can be performed on them. However, I believe in the end, all these names are just different facets of a specific evolution of the Web to make the Web more machine-friendly and to support ever more automation.

Standard stack…overflow? — Another way to look at the evolution and topics in the Semantic Web is to consider the fact that we went from two standards drafted in 1999 (RDF and RDFS) for metadata and lightweight schemata publishing on the Web, to a stack of standard depicted by the W3C in Figure 6 during the first half of the years 2000, and to an even larger number of additional standards still growing in the last years and depicted in an updated view of the pile in Figure 7. There are other recommendations about the Semantic Web that are not included in the figure like best practices recommendations to publish data. And as I write this article, future recommendations are being prepared such as more precise ontologies to describe and exchange datasets. From the multiplication of applications of the Semantic Web we saw in section 16, the community gathered lessons learned which led to the idea and the development of best practices guidelines and some of them were even turned into W3C recommendations. And here we come full circle, as we started from *a priori* standards to create new practices (RDF, RDFS) and ended-up recommending *a posteriori* standards compiling effective best practices. This growing stack of standards and literature about the Semantic Web is both an indication of the interest it generates and of the complexity it has grown to.

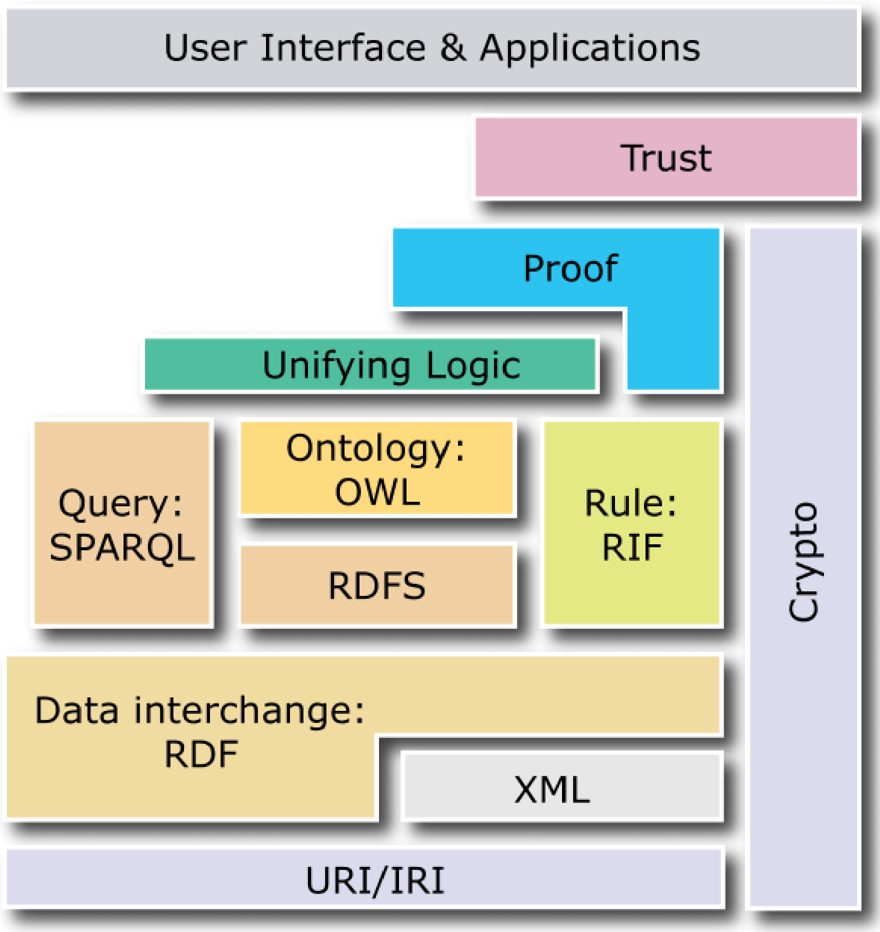

Figure 6. W3C Semantic Web Stack or Layer Cake for the years 2000

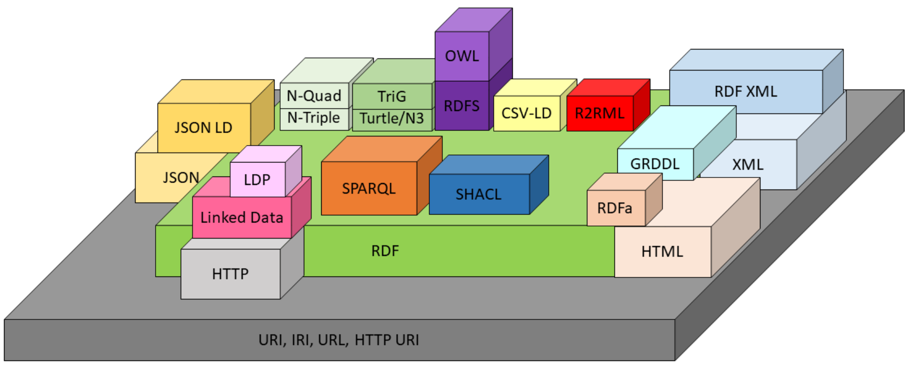

Figure 7. A new version of the Semantic Web Stack in 2018

Under construction Web. . . base — At the beginning of the Web it was very frequent to find pages with the mention "under construction" with a variety of icons to indicate a work in progress. The Web is a never-ending project and the Semantic Web is no exception to that. Over the first 20 years, we moved from solved to new challenges, starting with a small number of key topics (e.g. RDF, Query Languages, Web Ontologies) and ending-up with many more as we saw. The Semantic Web has achieved many advances as shown by the articles referenced in this paper but, as always in research, it also keeps opening new perspectives for the development and deployment of a Web of structured data and formal knowledge. For that reason, questions we sometimes hear like "when will the Semantic Web happen?" do not really make sense to me because the semantic Web has happened, is happening and will happen as a part of the never-ending project that we call "The Web". In fact, many of the Semantic Web sources could, for ever, be labeled with the mention and the icons for "Under construction Web base". Moreover, because we all use the Web, we all have a responsibility to defend it, and because the Web, and the semantic Web, are never-ending projects, we can never stop defending them.

Mind the gap. . . divide, ditch, ravine — One of the most difficult challenges of the Semantic Web is summarized in its name: the gap between formal semantics for machines and a Web for humans. As the Web grows and encompasses more persons and cultures, it also has to face cultural gaps, digital divides, accessibility issues, thin files, data poors, etc. The Semantic Web inherits these issues as we saw in the previous sections and it adds to them the growing divide between ever more formal and complex models and methods on one hand, and an ever-wider range of user profiles, usages and use contexts on the other hand. There is also a growing concern of the cost of the Web in terms of energy, infrastructure and resources in general and of who can afford that cost. The questions of identifying the "Web we want" and the "Web we can afford" can be specialized to the case of the semantic Web and translated into research challenges such as designing and optimizing the architectures and processes to improve our impact on society, environment and the world in general. For instance, the actual need to re-decentralize the Web and give everybody his Web site back again, translates into research questions for the semantic Web to find new architectures and methods supporting this re-empowerment of the Web users. For many years now, I have been concluding my talks insisting on the fact that "He who controls metadata controls the Web and, through the Web, many things of our world". A corollary of that saying is that we must ensure, by every means, especially open research, open development and open standards, that the Web in general and the semantic Web and (linked) data in particular, do not end up being centralized in one silo, one hand. The need for a constructive design of the Web and the semantic Web is at the heart of the agenda of Web Science.

This is for everything. . . but no 42 — Because most of the approaches for the Web and for the Semantic Web are domain-independent, the results may be applied and reused in many different application scenarios, as shown by the list of domains in section 16. Where Tim Berners-Lee said about the Web "This is for everyone"[9], it can also be said that "This is for everything". However because it may be used for everything, it does not mean it is the "answer to everything". An important question in building the research agenda of the Semantic Web is to systematically identify when it is actually a good option and why it can make a difference compared to alternative options (Bernstein, Noy, 2014).

From semantic checkpoint Charlie. . . to a shared blackboard — The Semantic Web provided pivot languages and frameworks at different levels. We already mentioned data integration and alignment at the data level. Going beyond, the Semantic Web has the potential of crossing: the walls of formal semantics (using RDF as a pivot data model between different formalisms) and the walls of schools of thoughts (with hybrid approaches, and the Web as an integration architecture). From the previous sections it was clear that every time the Semantic Web is combined with another research domain there is a systematic double-way cross-fertilizing. From that perspective, a first challenge is to ensure that every time cross-fertilizing is possible we avoid setting up an asymmetric relation with the other domains and that we fully investigate the three following aspects: what the other domains can bring to the Semantic Web; what the Semantic Web can bring to the other domains; and what the combined Semantic Web and domains can do better.

From an augmented world. . . to a Web-Wide World — By nature, the Semantic Web finds itself at the intersection of knowledge-based interactions and Web-augmented interactions, at all social scales from personal interactions to crowd interactions. This is raising many challenges and research questions in particular in terms of conceiving intelligent interaction design, of augmenting human intelligence and abilities and revisiting our experience of the world. From a more general perspective, and in relation to the previous point, the Web provides a distributed and shared blackboard where we can bring together very different contributions and start to link them. One of the challenges will be to use semantics and the computing intelligence to foster linkage, interactions and convergence when possible and avoid polarization and radicalization.

Linking all forms of intelligence. . . and the rest — the Web in general and the Semantic Web in particular have the potential to link all forms of intelligence. We are already seeing how it can provide a common ground between software agents and human agents. Different kinds of artificial intelligence are now leveraged at every step of the knowledge life cycle and for every parts of the components we loosely couple on the Web: to extract, import, interpret, recognize,. . . the inputs; to process, query, reason, decide,. . . from it; and to export, express, customize, adapt,. . . the outputs. But the full potential of the Semantic Web is to support a world-wide collaboration of intelligence in every forms: natural and physical intelligence and not only humans but also animals and plants; connected objects in a Web of things; and also artificial intel-

9. @timberners_lee : "This is for everyone #london2012 #oneweb #openingceremony @webfoundation @w3c" https://twitter.com/timberners_lee/status/228960085672599552

ligence in a broad sense with every approach to simulate different forms of intelligent behaviors such as learning, deducting, inducing, identifying, classifying, communicating, reading, imagining, feeling, etc. This is not science-fiction, and you just have to consider what we could achieve if the Web could seamlessly connect herdsourcing (e.g. existing projects on connected animals to predict earthquakes), sensors (Web of things), experts (social Web) and AI (e.g. different kinds of classifiers, experts systems, etc.) to create a collective intelligent decision system.

The Web was initially for homo sapiens but this has changed in many ways (services, Web bots, connected objects, connected AI, connected animals, connected plants, etc.). Likewise the Semantic Web was often presented as the Web for machines. In fact, the Semantic Web is the association, on one hand, of formal semantics and methods as in KRR and, on the other hand, one of the largest social application of the Internet that is the Web. As such the Semantic Web is a descendant of at least two fields born in the 50s: AI for Artificial Intelligence (McCarthy *et al.*, 1955) and IA for Intelligence Amplification (Ashby, 1956) and Intelligence Augmentation (Engelbart, 1962). It builds on and is a continuation of the AI and IA research programs but with the worldwide dimension that the Web brings. Augmenting and linking all kinds of intelligence, there is the long term potential of the Semantic Web.

Acknowledgements

To Mylène Leitzelman for her independent bibliometric validation of the topics and articles.
To my reviewers: Max Chevalier, Olivier Corby, Sébastien Laborie, Elodie Thieblin, Cassia Trojahn and Antoine Zimmermann

References

Abadi D. J., Marcus A., Madden S. R., Hollenbach K. (2007). Scalable semantic web data management using vertical partitioning. In *Proceedings of the 33rd international conference on very large data bases*, pp. 411–422. VLDB Endowment. Retrieved from http://dl.acm .org/citation.cfm?id=1325851.1325900

Abel F., Gao Q., Houben G.-J., Tao K. (2011). Semantic enrichment of twitter posts for user profile construction on the social web. In *Extended semantic web conference*, pp. 375–389. Heraklion, Greece, Springer.

Adjiman P., Chatalic P., Goasdoué F., Rousset M.-C., Simon L. (2006). Distributed reasoning in a peer-to-peer setting: Application to the semantic web. *Journal of Artificial Intelligence Research (JAIR)*, Vol. 25, pp. 269–314.

Allemang D., Hendler J. (2011). *Semantic web for the working ontologist: effective modeling in rdfs and owl*. Amsterdam, The Netherlands, Elsevier.

Angles R., Gutierrez C. (2016). The multiset semantics of sparql patterns. In *International semantic web conference*, pp. 20–36. Kobe, Japan, Springer.

Ankolekar A., Burstein M., Hobbs J. R., Lassila O., Martin D., McDermott D. *et al.* (2002). Daml-s: Web service description for the semantic web. In *International semantic web conference*, pp. 348–363. Sardinia, Italy, Springer.

Antoniou G., vanHarmelen F. (2004). *A semantic web primer*. Cambridge, MA, USA, MIT Press.

Antoniou G., Van Harmelen F. (2004). Web ontology language: Owl. In *Handbook on ontologies*, pp. 67–92. Springer.

Aroyo L., Dicheva D. (2004). The new challenges for e-learning: The educational semantic web. *Journal of Educational Technology & Society*, Vol. 7, No. 4.

Aroyo L., Stash N., Wang Y., Gorgels P., Rutledge L. (2007). Chip demonstrator: Semantics-driven recommendations and museum tour generation. In *The semantic web: 6th international semantic web conference, 2nd asian semantic web conference*, pp. 879–886. Busan, Korea, Springer.

Artz D., Gil Y. (2007). A survey of trust in computer science and the semantic web. *Web Semantics: Science, Services and Agents on the World Wide Web*, Vol. 5, No. 2, pp. 58–71.

Arumugam M., Sheth A., Arpinar B. (2002). Towards peer-to-peer semantic web: A distributed environment for sharing semantic knowledge on the web. In *Workshop on real world rdf and semantic web applications.* Honolulu, Springer.

Ashburner M., Ball C. A., Blake J. A., Botstein D., Butler H., Cherry J. M. *et al.* (2000). Gene ontology: tool for the unification of biology. *Nature genetics*, Vol. 25, No. 1, pp. 25.

Ashby R. (1956). Design for an intelligence-amplifier. *Automata studies*, Vol. 400, pp. 215–233.

Atencia M., Borgida A., Euzenat J., Ghidini C., Serafini L. (2012). A formal semantics for weighted ontology mappings. In *International semantic web conference*, pp. 17–33. Boston, USA, Springer.

Auer S., Bizer C., Kobilarov G., Lehmann J., Cyganiak R., Ives Z. (2007). Dbpedia: A nucleus for a web of open data. In *The semantic web: 6th international semantic web conference, 2nd asian semantic web conference*, pp. 722–735. Busan, Korea, Springer.

Auer S., Dietzold S., Lehmann J., Hellmann S., Aumueller D. (2009). Triplify: light-weight linked data publication from relational databases. In *Proceedings of the 18th international conference on world wide web*, pp. 621–630.

Auer S., Lehmann J. (2007). What have innsbruck and leipzig in common? extracting semantics from wiki content. In *European semantic web conference*, pp. 503–517. Innsbruck, Austria, Springer.

Baader F., Horrocks I., Sattler U. (2005). Description logics as ontology languages for the semantic web. In *Mechanizing mathematical reasoning*, pp. 228–248. Springer.

Bailey J., Bry F., Furche T., Schaffert S. (2005). Web and semantic web query languages: A survey. In *Proceedings of the first international conference on reasoning web*, pp. 35–133. Berlin, Heidelberg, Springer-Verlag. Retrieved from http://dx.doi.org/10.1007/11526988_3

Bassiliades N., Antoniou G., Vlahavas I. (2006). A defeasible logic reasoner for the semantic web. *International Journal on Semantic Web and Information Systems (IJSWIS)*, Vol. 2, No. 1, pp. 1–41.

Belleau F., Nolin M.-A., Tourigny N., Rigault P., Morissette J. (2008). Bio2rdf: towards a mashup to build bioinformatics knowledge systems. *Journal of biomedical informatics*, Vol. 41, No. 5, pp. 706–716.

Berendt B., Hotho A., Stumme G. (2002). Towards semantic web mining. In *International semantic web conference*, pp. 264–278. Sardinia, Italy, Springer.

Berners-Lee T. (1989). *Information management: A proposal*. Technical report. Genève, CERN. Retrieved from http://cds.cern.ch/record/1405411/files/ARCH-WWW-4-010.pdf

Berners-Lee T. (1998). *Semantic web road map*. Design Issues, Notes. Boston, USA, W3C. Retrieved from https://www.w3.org/DesignIssues/Semantic.html

Berners-Lee T. (2006). *Linked data*. Design Issues, Notes. Boston, USA, W3C. Retrieved from https://www.w3.org/DesignIssues/LinkedData.html

Berners-Lee T., Cailliau R., Luotonen A., Nielsen H. F., Secret A. (1994, August). The world-wide web. *Communications of the ACM*, Vol. 37, No. 8, pp. 76–82. Retrieved from http://doi.acm.org/10.1145/179606.179671

Berners-Lee T., Chen Y., Chilton L., Connolly D., Dhanaraj R., Hollenbach J. *et al.* (2006). Tabulator: Exploring and analyzing linked data on the semantic web. In *Proceedings of the 3rd international semantic web user interaction workshop*, Vol. 2006, p. 159. Athens, USA, SWUI2006.

Berners-Lee T., Hendler J., Lassila O. (2001). The semantic web. *Scientific american*, Vol. 284, No. 5, pp. 28–37.

Bernstein A., Kaufmann E. (2006). Gino–a guided input natural language ontology editor. In *International semantic web conference*, pp. 144–157. Budva, Montenegro, Springer.

Bernstein A., Noy N. (2014). *Is this really science? the semantic webber's guide to evaluating research contributions*. Technical report. Tech. Rep. No. IFI-2014.02). Department of Informatics, University of Zurich. Retrieved from https://www. merlin. uzh. ch/contributionDocument/download/6915.

Bizer C., Heath T., Ayers D., Raimond Y. (2007). Interlinking open data on the web. In *Demonstrations track, european semantic web conference.* Innsbruck, Austria, Springer.

Bizer C., Heath T., Berners-Lee T. (2009). Linked data-the story so far. *International journal on semantic web and information systems*, Vol. 5, No. 3, pp. 1–22.

Bizer C., Lehmann J., Kobilarov G., Auer S., Becker C., Cyganiak R. *et al.* (2009). Dbpedia - a crystallization point for the web of data. *Web Semantics: Science, Services and Agents on the World Wide Web*, Vol. 7, No. 3, pp. 154 - 165. Retrieved from http://www.sciencedirect.com/science/article/pii/S1570826809000225 (The Web of Data)

Bloehdorn S., Petridis K., Saathoff C., Simou N., Tzouvaras V., Avrithis Y. *et al.* (2005). Semantic annotation of images and videos for multimedia analysis. In *European semantic web conference*, pp. 592–607. Heraklion, Crete, Springer.

Bonatti P. A., Hogan A., Polleres A., Sauro L. (2011). Robust and scalable linked data reasoning incorporating provenance and trust annotations. *Web Semantics: Science, Services and Agents on the World Wide Web*, Vol. 9, No. 2, pp. 165–201.

Bouayad-Agha N., Casamayor G., Wanner L. (2014). Natural language generation in the context of the semantic web. *Semantic Web*, Vol. 5, No. 6, pp. 493–513.

Breslin J. G., Decker S., Harth A., Bojars U. (2006). Sioc: an approach to connect web-based communities. *International Journal of Web Based Communities*, Vol. 2, No. 2, pp. 133–142.

Breslin J. G., Harth A., Bojars U., Decker S. (2005). Towards semantically-interlinked online communities. In *European semantic web conference*, pp. 500–514. Heraklion, Crete, Springer.

Breslin J. G., Passant A., Decker S. (2009). *The social semantic web*. Springer Science & Business Media.

Brickley D., Miller L. (2007). *Foaf vocabulary specification 0.91*. Technical report, ILRT.

Broekstra J., Kampman A., Van Harmelen F. (2002). Sesame: A generic architecture for storing and querying rdf and rdf schema. In *International semantic web conference*, pp. 54–68. Sardinia, Italy, Springer.

Buffa M., Gandon F., Ereteo G., Sander P., Faron C. (2008). Sweetwiki: A semantic wiki. *Web Semantics: Science, Services and Agents on the World Wide Web*, Vol. 6, No. 1, pp. 84–97.

Buil-Aranda C., Arenas M., Corcho O., Polleres A. (2013). Federating queries in sparql 1.1: Syntax, semantics and evaluation. *Journal of Web Semantics*, Vol. 18, pp. 1-17. Retrieved from https://www.scopus.com/inward/record.uri?eid=2-s2.0-84873310631&doi=10.1016%2fj.websem.2012.10.001&partnerID=40&md5=1af2977d5f947e2e4ba8cbedacaf23e0

Cabrio E., Villata S., Gandon F. (2013). A support framework for argumentative discussions management in the web. In *European semantic web conference*, pp. 412–426. Montpellier, France, Springer.

Carroll J. J., Bizer C., Hayes P., Stickler P. (2005). Named graphs, provenance and trust. In *Proceedings of the 14th international conference on world wide web*, pp. 613–622. Chiba, Japan, ACM.

Carroll J. J., Dickinson I., Dollin C., Reynolds D., Seaborne A., Wilkinson K. (2004). Jena: implementing the semantic web recommendations. In *Proceedings of the 13th international world wide web conference on alternate track papers & posters*, pp. 74–83.

Cimiano P., Handschuh S., Staab S. (2004). Towards the self-annotating web. In *Proceedings of the 13th international conference on world wide web*, pp. 462–471. New York, USA, ACM.

Cimiano P., Lopez V., Unger C., Cabrio E., Ngomo A.-C. N., Walter S. (2013). Multilingual question answering over linked data (qald-3): Lab overview. In *International conference of the cross-language evaluation forum for european languages*, pp. 321–332. Valencia, Spain, Springer.

Cochez M., Decker S., Prud'hommeaux E. (2016). Knowledge representation on the web revisited: the case for prototypes. In *International semantic web conference*, pp. 151–166. Kobe, Japan, Springer.

Corby O., Dieng R., Hébert C. (2000). A conceptual graph model for w3c resource description framework. In *International conference on conceptual structures*, pp. 468–482. Darmstadt, Germany, Springer.

Corby O., Dieng-Kuntz R., Faron-Zucker C. (2004). Querying the semantic web with corese search engine. In *Proceedings of the 16th eureopean conference on artificial intelligence, ecai'2004, including prestigious applicants of intelligent systems, PAIS 2004, valencia, spain, august 22-27, 2004*, pp. 705–709. Valencia, Spain, IOS Press.

Corby O., Dieng-Kuntz R., Gandon F., Faron-Zucker C. (2006). Searching the semantic web: Approximate query processing based on ontologies. *IEEE Intelligent Systems*, Vol. 21, No. 1, pp. 20–27.

Corby O., Faron Zucker C., Gandon F. (2015, October). A Generic RDF Transformation Software and its Application to an Online Translation Service for Common Languages of Linked Data. In *The 14th International Semantic Web Conference.* Bethlehem, United States, Springer. Retrieved from https://hal.inria.fr/hal-01186047

Corby O., Faron Zucker C., Gandon F. (2017, October). LDScript: a Linked Data Script Language. In *International Semantic Web Conference.* Vienna, Austria, Springer. Retrieved from https://hal.inria.fr/hal-01589162

Costabello L., Gandon F. (2014). Context-Aware Presentation of Linked Data on Mobile. *International Journal On Semantic Web and Information Systems (IJSWIS)*, Vol. 10, No. 4, pp. 32. Retrieved from https://hal.inria.fr/hal-01188210

Costabello L., Villata S., Gandon F. (2012, August). Context-Aware Access Control for RDF Graph Stores. In *ECAI - 20th European Conference on Artificial Intelligence - 2012.* Montpellier, France. Retrieved from https://hal.inria.fr/hal-00724041

Cox S., Little C., Hobbs J. R., Pan F. (2017). *Time ontology in owl*. Recommendation. Boston, USA, W3C. Retrieved from https://www.w3.org/TR/owl-time/

Dadzie A.-S., Rowe M. (2011). Approaches to visualising linked data: A survey. *Semantic Web*, Vol. 2, No. 2, pp. 89–124.

d'Amato C., Fanizzi N., Esposito F. (2010). Inductive learning for the semantic web: What does it buy? *Semantic Web*, Vol. 1, No. 1, 2, pp. 53–59.

d'Aquin M., Motta E., Sabou M., Angeletou S., Gridinoc L., Lopez V. *et al*. (2008). Toward a new generation of semantic web applications. *IEEE Intelligent Systems*, Vol. 23, No. 3, pp. 20–28.

Davies J., Fensel D., Van Harmelen F. (2003). *Towards the semantic web: ontology-driven knowledge management*. John Wiley & Sons.

Delteil A., Faron-Zucker C., Dieng R. (2001). Learning ontologies from rdf annotation. In *Workshop on ontology learning*, pp. 7–12. Seattle, USA, CEUR-WS.org.

Dill S., Eiron N., Gibson D., Gruhl D., Guha R., Jhingran A. *et al*. (2003). Semtag and seeker: Bootstrapping the semantic web via automated semantic annotation. In *Proceedings of the 12th international conference on world wide web*, pp. 178–186.

Ding L., Finin T., Joshi A., Pan R., Cost R. S., Peng Y. *et al*. (2004). Swoogle: a search and metadata engine for the semantic web. In *Proceedings of the thirteenth acm international conference on information and knowledge management cikm*, pp. 652–659. Washington, USA, ACM.

Doan A., Madhavan J., Domingos P., Halevy A. (2002). Learning to map between ontologies on the semantic web. In *Proceedings of the 11th international conference on world wide web*, pp. 662–673.

Dolog P., Henze N., Nejdl W., Sintek M. (2004). The personal reader: Personalizing and enriching learning resources using semantic web technologies. In *International conference on adaptive hypermedia and adaptive web-based systems*, pp. 85–94.

Dzbor M., Domingue J., Motta E. (2003). Magpie – towards a semantic web browser. In *International semantic web conference*, pp. 690–705. Sanibel Island, USA, Springer.

Ehrig M., Sure Y. (2004). Ontology mapping–an integrated approach. In *European semantic web symposium*, pp. 76–91. Heraklion, Greece, Springer.

Eiter T., Parreira J. X., Schneider P. (2017). Spatial ontology-mediated query answering over mobility streams. In *European semantic web conference*, pp. 219–237. Portorož, Slovenia, Springer.

Engelbart D. C. (1962). *Augmenting human intellect: A conceptual framework*. SRI Summary Report No. AFOSR-3223. Stanford Research Institute. Internet Requests for Comments.

Erétéo G., Buffa M., Gandon F., Corby O. (2009). Analysis of a real online social network using semantic web frameworks. In *International semantic web conference*, pp. 180–195. Chantilly, VA, USA, Springer.

Etzioni O., Cafarella M., Downey D., Popescu A.-M., Shaked T., Soderland S. *et al.* (2005). Unsupervised named-entity extraction from the web: An experimental study. *Artificial intelligence*, Vol. 165, No. 1, pp. 91–134.

Euzenat J., Shvaiko P. *et al.* (2007). *Ontology matching* (Vol. 18). Germany, Springer.

Fernández M., Cantador I., López V., Vallet D., Castells P., Motta E. (2011). Semantically enhanced information retrieval: An ontology-based approach. *Web semantics: Science, services and agents on the world wide web*, Vol. 9, No. 4, pp. 434–452.

Fernández J., Martínez-Prieto M., Gutiérrez C., Polleres A., Arias M. (2013). Binary rdf representation for publication and exchange (hdt). *Journal of Web Semantics*, Vol. 19, pp. 22-41.

Fielding R. T., Taylor R. N. (2000). *Architectural styles and the design of network-based software architectures* (Vol. 7). University of California, Irvine Doctoral dissertation.

Finin T., Joshi A., Kagal L., Niu J., Sandhu R., Winsborough W. *et al.* (2008). R owl bac: representing role based access control in owl. In *Proceedings of the 13th acm symposium on access control models and technologies*, pp. 73–82. Estes Park, USA, ACM.

Fokou G., Jean S., Hadjali A., Baron M. (2016). Rdf query relaxation strategies based on failure causes. In *European semantic web conference*, pp. 439–454. Heraklion, Greece, Springer.

Folz P., Skaf-Molli H., Molli P. (2016). Cyclades: a decentralized cache for triple pattern fragments. In *European semantic web conference*, pp. 455–469. Heraklion, Greece, Springer.

Franz T., Schultz A., Sizov S., Staab S. (2009). Triplerank: Ranking semantic web data by tensor decomposition. In *International semantic web conference*, pp. 213–228. Chantilly, VA, USA, Springer.

Gandon F. (2002). *Distributed artificial intelligence and knowledge management: ontologies and multi-agent systems for a corporate semantic web*. Unpublished doctoral dissertation, Université Nice Sophia Antipolis.

Gandon F., Dieng-Kuntz R., Corby O., Giboin A. (2002). Semantic web and multi-agents approach to corporate memory management. In *International conference on intelligent information processing*, pp. 103–115. Montréal, Québec, Springer.

Gandon F., Giboin A. (2017, June). Paving the WAI: Defining Web-Augmented Interactions. In *Web Science 2017 (WebSci17)*, p. 381 - 382. Troy, NY, United States. Retrieved from https://hal.inria.fr/hal-01560180

Gandon F. L., Sadeh N. M. (2004). Semantic web technologies to reconcile privacy and context awareness. *Web Semantics: Science, Services and Agents on the World Wide Web*, Vol. 1, No. 3, pp. 241–260.

Gangemi A. (2005). Ontology design patterns for semantic web content. In *International semantic web conference*, pp. 262–276. Galway, Ireland, Springer.

Gangemi A. (2013). A comparison of knowledge extraction tools for the semantic web. In *Extended semantic web conference*, pp. 351–366.

Geroimenko V., Chen C. (2006). *Visualizing the semantic web: Xml-based internet and information visualization*. Springer Science & Business Media.

Golbeck J., Hendler J. (2004). Accuracy of metrics for inferring trust and reputation in semantic web-based social networks. In *International conference on knowledge engineering and knowledge management*, pp. 116–131.

Golbeck J., Parsia B., Hendler J. (2003). Trust networks on the semantic web. In *International workshop on cooperative information agents*, pp. 238–249.

Gómez-Pérez A., Corcho O. (2002). Ontology languages for the semantic web. *IEEE Intelligent systems*, Vol. 17, No. 1, pp. 54–60.

Gómez-Pérez A., Fernández-López M., Corcho O. (2006). *Ontological engineering: with examples from the areas of knowledge management, e-commerce and the semantic web*. Germany, Springer Science & Business Media.

Governatori G., Rotolo A., Villata S., Gandon F. (2013). One license to compose them all. In *International semantic web conference*, pp. 151–166. Sydney, Australia, Springer.

Gracia J., Montiel-Ponsoda E., Cimiano P., Gómez-Pérez A., Buitelaar P., McCrae J. (2012). Challenges for the multilingual web of data. *Web Semantics: Science, Services and Agents on the World Wide Web*, Vol. 11, pp. 63–71.

Gruber T. (2008). Collective knowledge systems: Where the social web meets the semantic web. *Web Semantics: Science, Services and Agents on the World Wide Web*, Vol. 6, No. 1, pp. 4 - 13. Retrieved from http://www.sciencedirect.com/science/article/pii/S1570826807000583 (Semantic Web and Web 2.0)

Guo Y., Pan Z., Heflin J. (2004). An evaluation of knowledge base systems for large owl datasets. In *International semantic web conference*, pp. 274–288. Hiroshima, Japan, Springer.

Guo Y., Pan Z., Heflin J. (2005). Lubm: A benchmark for owl knowledge base systems. *Web Semantics: Science, Services and Agents on the World Wide Web*, Vol. 3, No. 2-3, pp. 158–182.

Haase P., Broekstra J., Ehrig M., Menken M., Mika P., Olko M. *et al.* (2004). Bibster–a semantics-based bibliographic peer-to-peer system. In *International semantic web conference*, pp. 122–136. Hiroshima, Japan, Springer.

Harth A., Umbrich J., Hogan A., Decker S. (2007). Yars2: A federated repository for querying graph structured data from the web. In *The semantic web : 6th international semantic web conference, 2nd asian semantic web conference*, pp. 211–224. Busan, Korea, Springer.

Hartig O., Bizer C., Freytag J.-C. (2009). Executing sparql queries over the web of linked data. In *International semantic web conference*, pp. 293–309. Chantilly, VA, USA, Springer.

Hartig O., Letter I., Pérez J. (2017). A formal framework for comparing linked data fragments. In *International semantic web conference*, pp. 364–382. Vienna, Austria, Springer.

Hartig O., Pirrò G. (2015). A context-based semantics for sparql property paths over the web. In *European semantic web conference*, pp. 71–87. Portoroz, Slovenia, Springer.

Hasan R. (2014). *Predicting query performance and explaining results to assist Linked Data consumption*. Theses, Université Nice Sophia Antipolis. Retrieved from https://tel.archives-ouvertes.fr/tel-01127124

Hasan R., Gandon F. (2012, August). A brief review of explanation in the Semantic Web. In *7th Workshop on Explanation-aware Computing (ExaCt 2012), 20th European Conference on Artificial Intelligence (ECAI-2012).* Montpellier, France. Retrieved from https://hal.inria.fr/hal-00729037

Hausenblas M., Troncy R., Raimond Y., Bürger T. (2009). Interlinking multimedia: How to apply linked data principles to multimedia fragments. In *Linked data on the web workshop (ldow 09).*

Heath T., Bizer C. (2011). *Linked data: Evolving the wb into a global data space* (Vol. 1) No. 1. USA, Morgan & Claypool Publishers.

Hellmann S., Lehmann J., Auer S., Brümmer M. (2013). Integrating nlp using linked data. In *International semantic web conference*, pp. 98–113. Sydney, Australia, Springer.

Henze N., Dolog P., Nejdl W. (2004). Reasoning and ontologies for personalized e-learning in the semantic web. *Journal of Educational Technology & Society*, Vol. 7, No. 4.

Hildebrand M., Ossenbruggen J. van, Hardman L. (2006). /facet: A browser for heterogeneous semantic web repositories. In *International semantic web conference*, pp. 272–285. Athens, USA, Springer.

Hitzler P., Krotzsch M., Rudolph S. (2009). *Foundations of semantic web technologies*. USA, CRC press.

Hitzler P., Van Harmelen F. (2010). A reasonable semantic web. *Semantic Web*, Vol. 1, No. 1, 2, pp. 39–44.

Hogan A., Harth A., Passant A., Decker S., Polleres A. (2010). Weaving the pedantic web. In *Linked data on the web, ldow.* Raleigh, USA, CEUR, Workshop Proceedings.

Hogan A., Umbrich J., Harth A., Cyganiak R., Polleres A., Decker S. (2012). An empirical survey of linked data conformance. *Web Semantics: Science, Services and Agents on the World Wide Web*, Vol. 14, pp. 14–44.

Hogan A., Zimmermann A., Umbrich J., Polleres A., Decker S. (2012). Scalable and distributed methods for entity matching, consolidation and disambiguation over linked data corpora. *Web Semantics: Science, Services and Agents on the World Wide Web*, Vol. 10, pp. 76–110.

Horridge M., Parsia B., Sattler U. (2008). Laconic and precise justifications in owl. In *International semantic web conference*, pp. 323–338. Karlsruhe, Germany, Springer.

Horrocks I., Patel-Schneider P. (2003). Reducing owl entailment to description logic satisfiability. In *International semantic web conference*, pp. 17–29. Sanibel Island, USA, Springer.

Hotho A., Jäschke R., Schmitz C., Stumme G. (2006). Information retrieval in folksonomies: Search and ranking. In *European semantic web conference*, pp. 411–426. Budva, Montenegro, Springer.

Jain P., Hitzler P., Sheth A. P., Verma K., Yeh P. Z. (2010). Ontology alignment for linked open data. In *International semantic web conference*, pp. 402–417. Shanghai, China, Springer.

Jung J. C., Lutz C. (2012). Ontology-based access to probabilistic data with owl ql. In *International semantic web conference*, pp. 182–197. Boston, USA, Springer.

Kagal L., Finin T., Joshi A. (2003). A policy based approach to security for the semantic web. In *International semantic web conference*, pp. 402–418. Sanibel Island, USA, Springer.

Kaufmann E., Bernstein A. (2007). How useful are natural language interfaces to the semantic web for casual end-users? In *The semantic web: 6th international semantic web conference, 2nd asian semantic web conference*, pp. 281–294. Busan, Korea, Springer.

Kazakov Y., Klinov P. (2013). Incremental reasoning in owl el without bookkeeping. In *International semantic web conference*, pp. 232–247. Sydney, Australia, Springer.

Kirrane S., Mileo A., Decker S. (2017). Access control and the resource description framework: A survey. *Semantic Web*, Vol. 8, No. 2, pp. 311–352.

Knublauch H., Fergerson R., Noy N., Musen M. (2004). The protégé owl plugin: An open development environment for semantic web applications. In *International semantic web conference*, pp. 229–243. Hiroshima, Japan, Springer.

Kobilarov G., Scott T., Raimond Y., Oliver S., Sizemore C., Smethurst M. *et al.* (2009). Media meets semantic web – how the bbc uses dbpedia and linked data to make connections. In *European semantic web conference*, pp. 723–737. Heraklion, Greece, Springer.

Krötzsch M. (2012). The not-so-easy task of computing class subsumptions in owl rl. In *International semantic web conference*, pp. 279–294. Boston, USA, Springer.

Kuhn T., Dumontier M. (2014). Trusty uris: Verifiable, immutable, and permanent digital artifacts for linked data. In *European semantic web conference*, pp. 395–410. Anissaras, Greece, Springer.

Kuter U., Golbeck J. (2009). Semantic web service composition in social environments. In *International semantic web conference*, pp. 344–358. Chantilly, VA, USA, Springer.

Le B. T., Dieng-Kuntz R., Gandon F. (2004). On ontology matching problems. In *International conference on enterprise information systems*, pp. 236–243. Porto, Portugal, ICEIS.

Lebo T., Sahoo S., McGuinness D., Belhajjame K., Cheney J., Corsar D. *et al.* (2013). Prov-o: the prov ontology. *World Wide Web Consortium, Recommendation*.

Lécué F., Léger A. (2006). A formal model for semantic web service composition. In *International semantic web conference*, pp. 385–398. Athens, USA, Springer.

Lécué F., Tucker R., Bicer V., Tommasi P., Tallevi-Diotallevi S., Sbodio M. (2014). Predicting severity of road traffic congestion using semantic web technologies. In *European semantic web conference*, pp. 611–627. Anissaras, Greece, Springer.

Limpens F., Gandon F., Buffa M. (2013). A Complete Life-Cycle for the Semantic Enrichment of Folksonomies. In Springer (Ed.), *Advances In Knowledge Discovery and Management.* Springer. Retrieved from https://hal.inria.fr/hal-01171250

Lopez V., Unger C., Cimiano P., Motta E. (2013). Evaluating question answering over linked data. *Journal of Web Semantics*, Vol. 21, pp. 3-13.

Lösch U., Bloehdorn S., Rettinger A. (2012). Graph kernels for rdf data. In *European semantic web conference*, pp. 134–148. Heraklion, Greece, Springer.

Lukasiewicz T., Straccia U. (2008). Managing uncertainty and vagueness in description logics for the semantic web. *Web Semantics: Science, Services and Agents on the World Wide Web*, Vol. 6, No. 4, pp. 291–308.

Lutz C., Seylan I., Toman D., Wolter F. (2013). The combined approach to obda: Taming role hierarchies using filters. In *International semantic web conference*, pp. 314–330. Sydney, Australia, Springer.

Maedche A., Staab S. (2001). Ontology learning for the semantic web. *IEEE Intelligent systems*, Vol. 16, No. 2, pp. 72–79.

Magkanaraki A., Tannen V., Christophides V., Plexousakis D. (2003). Viewing the semantic web through rvl lenses. In *International semantic web conference*, pp. 96–112. Sanibel Island, USA, Springer.

Mallea A., Arenas M., Hogan A., Polleres A. (2011). On blank nodes. In *International semantic web conference*, pp. 421–437. Bonn, Germany, Springer.

Margara A., Urbani J., Van Harmelen F., Bal H. (2014). Streaming the web: Reasoning over dynamic data. *Journal of Web Semantics*, Vol. 25, pp. 24-44.

Marie N., Gandon F. (2014, October). Survey of linked data based exploration systems. In *IESD 2014 - Intelligent Exploitation of Semantic Data.* Riva Del Garda, Italy. Retrieved from https://hal.inria.fr/hal-01057035

McBride B. (2002). Jena: A semantic web toolkit. *IEEE Internet computing*, Vol. 6, No. 6, pp. 55–59.

McCarthy J., Minsky M., Rochester N., Shannon C. (1955, august 31). *A proposal for the dartmouth summer research project on artificial intelligence.* Hanover, New Hampshire, USA, Dartmouth College. Retrieved from http://raysolomonoff.com/dartmouth/boxa/dart564props.pdf

McGuinness D. L., Fikes R., Hendler J., Stein L. A. (2002). Daml+ oil: an ontology language for the semantic web. *IEEE Intelligent Systems*, Vol. 17, No. 5, pp. 72–80.

Mendes P. N., Jakob M., García-Silva A., Bizer C. (2011). Dbpedia spotlight: shedding light on the web of documents. In *Proceedings of the 7th international conference on semantic systems*, pp. 1–8.

Michel F., Faron Zucker C., Gandon F. (2018, April). SPARQL Micro-Services: Lightweight Integration of Web APIs and Linked Data. In *LDOW 2018 - Linked Data on the Web*, p. 1-10. Lyon, France, CEUR. Retrieved from https://hal.archives-ouvertes.fr/hal-01722792

Mika P. (2004). Social networks and the semantic web. In *Proceedings of the 2004 ieee/wic/acm international conference on web intelligence*, pp. 285–291. Washington, DC, USA, IEEE Computer Society. Retrieved from http://dx.doi.org/10.1109/WI.2004.128

Mika P. (2005a). Flink: Semantic web technology for the extraction and analysis of social networks. *Web Semantics: Science, Services and Agents on the World Wide Web*, Vol. 3, No. 2-3, pp. 211–223.

Mika P. (2005b). Ontologies are us: A unified model of social networks and semantics. In *International semantic web conference*, pp. 522–536. Galway, Ireland, Springer.

Miles A., Matthews B., Wilson M., Brickley D. (2005). Skos core: simple knowledge organisation for the web. In *International conference on dublin core and metadata applications*, pp. 3–10.

Moreau L., Clifford B., Freire J., Futrelle J., Gil Y., Groth P. *et al.* (2011). The open provenance model core specification (v1. 1). *Future generation computer systems*, Vol. 27, No. 6, pp. 743–756.

Morsey M., Lehmann J., Auer S., Ngomo A.-C. N. (2011). Dbpedia sparql benchmark–performance assessment with real queries on real data. In *International semantic web conference*, pp. 454–469. Bonn, Germany, Springer.

Motik B., Sattler U., Studer R. (2005). Query answering for owl-dl with rules. *Web Semantics: Science, Services and Agents on the World Wide Web*, Vol. 3, No. 1, pp. 41–60.

Ngomo A.-C. N., Auer S. (2011). Limes - a time-efficient approach for large-scale link discovery on the web of data. In *International joint conference on artificial intelligence.* Barcelona, Spain, AAAI Press.

Ngomo A.-C. N., Kolb L., Heino N., Hartung M., Auer S., Rahm E. (2013). When to reach for the cloud: Using parallel hardware for link discovery. In *European semantic web conference*, pp. 275–289. Montpellier, France, Springer.

Ngomo A.-C. N., Lyko K. (2012). Eagle: Efficient active learning of link specifications using genetic programming. In *European semantic web conference*, pp. 149–163. Heraklion, Greece, Springer.

Nickel M., Tresp V., Kriegel H.-P. (2012). Factorizing yago: Scalable machine learning for linked data. In *Proceedings of the 21st international conference on world wide web*, pp. 271–280. New York, NY, USA, ACM. Retrieved from http://doi.acm.org/10.1145/2187836.2187874

Nikolov A., Schwarte A., Hütter C. (2013). Fedsearch: Efficiently combining structured queries and full-text search in a sparql federation. In *International semantic web conference*, pp. 427–443. Sydney, Australia, Springer.

Noy N., Shah N., Whetzel P., Dai B., Dorf M., Griffith N. *et al.* (2009). Bioportal: ontologies and integrated data resources at the click of a mouse. *Nucleic acids research*, Vol. 37, No. suppl_2, pp. W170–W173.

Noy N. F. (2004). Semantic integration: a survey of ontology-based approaches. *ACM Sigmod Record*, Vol. 33, No. 4, pp. 65–70.

O'connor M., Knublauch H., Tu S., Grosof B., Dean M., Grosso W. *et al.* (2005). Supporting rule system interoperability on the semantic web with swrl. In *International semantic web conference*, pp. 974–986. Galway, Ireland, Springer.

Paolucci M., Kawamura T., Payne T. R., Sycara K. (2002). Semantic matching of web services capabilities. In *International semantic web conference*, pp. 333–347. Sardinia, Italy, Springer.

Parundekar R., Knoblock C. A., Ambite J. L. (2012). Discovering concept coverings in ontologies of linked data sources. In *International semantic web conference*, pp. 427–443. Boston, USA, Springer.

Patel-Schneider P., Siméon J. (2002). Building the semantic web on xml. In *International semantic web conference*, pp. 147–161. Sardinia, Italy, Springer.

Paulheim H. (2012). Generating possible interpretations for statistics from linked open data. In *European semantic web conference*, pp. 560–574. Heraklion, Greece, Springer.

Pérez J., Arenas M., Gutierrez C. (2006). Semantics and complexity of sparql. In *International semantic web conference*, pp. 30–43. Athens, USA, Springer.

Petrelli D., Mazumdar S., Dadzie A.-S., Ciravegna F. (2009). Multi visualization and dynamic query for effective exploration of semantic data. In *International semantic web conference*, pp. 505–520. Chantilly, VA, USA, Springer.

Pietriga E., Bizer C., Karger D., Lee R. (2006). Fresnel: A browser-independent presentation vocabulary for rdf. In *International semantic web conference*, pp. 158–171. Athens, USA, Springer.

Presutti V., Draicchio F., Gangemi A. (2012). Knowledge extraction based on discourse representation theory and linguistic frames. In *International conference on knowledge engineering and knowledge management*, pp. 114–129. Galway, Ireland, Springer.

Quilitz B., Leser U. (2008). Querying distributed rdf data sources with sparql. In *European semantic web conference*, pp. 524–538.

Raimond Y., Abdallah S. (2007). *The event ontology*. Technical report. Centre for Digital Music, Queen Mary, University of London.

Rettinger A., Lösch U., Tresp V., d'Amato C., Fanizzi N. (2012). Mining the semantic web. *Data Mining and Knowledge Discovery*, Vol. 24, No. 3, pp. 613–662.

Richardson M., Agrawal R., Domingos P. (2003). Trust management for the semantic web. In *International semantic web conference*, pp. 351–368. Sanibel Island, USA, Springer.

Rietveld L., Beek W., Schlobach S. (2015). Lod lab: Experiments at lod scale. In *International semantic web conference*, pp. 339–355. Bethlehem, USA, Springer.

Ristoski P., Paulheim H. (2016). Semantic web in data mining and knowledge discovery: A comprehensive survey. *Journal of Web Semantics*, Vol. 36, pp. 1-22.

Rizzo G., Erp M. van, Troncy R. (2014). Benchmarking the extraction and disambiguation of named entities on the semantic web. In *Lrec*, pp. 4593–4600.

Rizzo G., Troncy R., Hellmann S., Bruemmer M. (2012). Nerd meets nif: Lifting nlp extraction results to the linked data cloud. *LDOW*, Vol. 937.

Rodríguez Rocha O., Faron Zucker C. (2018, April). Automatic Generation of Quizzes from DBpedia According to Educational Standards. In *The 3rd Educational Knowledge Management Workshop (EKM 2018).* Lyon, France. Retrieved from https://hal.inria.fr/hal-01758737

Sabou M., Aroyo L., Bontcheva K., Bozzon A., Qarout R. K. (2018). Semantic web and human computation: The status of an emerging field. *Semantic Web*, pp. 1–12.

Sahoo S. S., Halb W., Hellmann S., Idehen K., Thibodeau Jr T., Auer S. *et al.* (2009). *A survey of current approaches for mapping of relational databases to rdf*. RDB2RDF Incubator Group Report. USA, W3C.

Sarasua C., Simperl E., Noy N. F. (2012). Crowdmap: Crowdsourcing ontology alignment with microtasks. In *International semantic web conference*, pp. 525–541. Boston, USA, Springer.

Scharffe F., Atemezing G., Troncy R., Gandon F., Villata S., Bucher B. *et al.* (2012). Enabling linked data publication with the datalift platform. In *Semantic cities.* Toronto, Canada, Proceedings AAAI Workshop.

Schreiber G., Amin A., Aroyo L., Assem M. van, Boer V. de, Hardman L. *et al.* (2008). Semantic annotation and search of cultural-heritage collections: The multimedian e-culture demonstrator. *Web Semantics: Science, Services and Agents on the World Wide Web*, Vol. 6, No. 4, pp. 243–249.

Schwarte A., Haase P., Hose K., Schenkel R., Schmidt M. (2011). Fedx: Optimization techniques for federated query processing on linked data. In *International semantic web conference*, pp. 601–616. Bonn, Germany, Springer.

Sequeda J., Miranker D. (2013). Ultrawrap: Sparql execution on relational data. *Web Semantics: Science, Services and Agents on the World Wide Web*, Vol. 22, pp. 19–39.

Serafini L., Tamilin A. (2005). Drago: Distributed reasoning architecture for the semantic web. In *European semantic web conference*, pp. 361–376. Heraklion, Crete, Springer.

Shvaiko P., Euzenat J. (2013). Ontology matching: state of the art and future challenges. *IEEE Transactions on knowledge and data engineering*, Vol. 25, No. 1, pp. 158–176.

Sintek M., Decker S. (2002). Triple—a query, inference, and transformation language for the semantic web. In *International semantic web conference*, pp. 364–378. Sardinia, Italy, Springer.

Sirin E., Parsia B., Grau B. C., Kalyanpur A., Katz Y. (2007). Pellet: A practical owl-dl reasoner. *Web Semantics: science, services and agents on the World Wide Web*, Vol. 5, No. 2, pp. 51–53.

Sonntag D., Engel R., Herzog G., Pfalzgraf A., Pfleger N., Romanelli M. *et al.* (2007). Smartweb handheld—multimodal interaction with ontological knowledge bases and semantic web services. In *Artifical intelligence for human computing*, pp. 272–295. Springer.

Spanos D.-E., Stavrou P., Mitrou N. (2012). Bringing relational databases into the semantic web: A survey. *Semantic Web*, Vol. 3, No. 2, pp. 169–209.

Specia L., Motta E. (2007). Integrating folksonomies with the semantic web. In *European semantic web conference*, pp. 624–639. Innsbruck, Austria, Springer.

Staab S., Stuckenschmidt H. (2006). *Semantic web and peer-to-peer: decentralized management and exchange of knowledge and information*. Springer.

Stocker M., Seaborne A., Bernstein A., Kiefer C., Reynolds D. (2008). Sparql basic graph pattern optimization using selectivity estimation. In *Proceedings of the 17th international conference on world wide web*, pp. 595–604. Beijing, China, ACM.

Suchanek F., Kasneci G., Weikum G. (2007). Yago: a core of semantic knowledge. In *Proceedings of the 16th international conference on world wide web*, pp. 697–706. Banff, Canada, ACM.

Sure Y., Erdmann M., Angele J., Staab S., Studer R., Wenke D. (2002). Ontoedit: Collaborative ontology development for the semantic web. In *International semantic web conference*, pp. 221–235. Sardinia, Italy, Springer.

Tangmunarunkit H., Decker S., Kesselman C. (2003). Ontology-based resource matching in the grid – the grid meets the semantic web. In *International semantic web conference*, pp. 706–721. Sanibel Island, USA, Springer.

Tonon A., Catasta M., Demartini G., Cudré-Mauroux P., Aberer K. (2013). Trank: Ranking entity types using the web of data. In *International semantic web conference*, pp. 640–656. Sydney, Australia, Springer.

Tsarkov D., Horrocks I. (2006). Fact++ description logic reasoner: System description. In *International joint conference on automated reasoning*, pp. 292–297. Seattle, USA, Springer.

Tudorache T., Vendetti J., Noy N. F. (2008). Web-protege: A lightweight owl ontology editor for the web. In *Owled*, Vol. 432.

Tummarello G., Delbru R., Oren E. (2007). Sindice.com: Weaving the open linked data. In *The semantic web : 6th international semantic web conference, 2nd asian semantic web conference*, pp. 552–565. Busan, Korea, Springer.

Unger C., Cimiano P. (2011). Pythia: Compositional meaning construction for ontology-based question answering on the semantic web. In *International conference on application of natural language to information systems*, pp. 153–160.

Urbani J., Kotoulas S., Oren E., Van Harmelen F. (2009). Scalable distributed reasoning using mapreduce. In *International semantic web conference*, pp. 634–649. Chantilly, VA, USA, Springer.

Usbeck R., Ngomo A.-C. N., Röder M., Gerber D., Coelho S. A., Auer S. *et al*. (2014). Agdistis-graph-based disambiguation of named entities using linked data. In *International semantic web conference*, pp. 457–471. Riva del Garda, Italy, Springer.

Vandenbussche P.-Y., Atemezing G. A., Poveda-Villalón M., Vatant B. (2017). Linked open vocabularies (lov): a gateway to reusable semantic vocabularies on the web. *Semantic Web*, Vol. 8, No. 3, pp. 437–452.

Verborgh R., Vander Sande M., Hartig O., Van Herwegen J., De Vocht L., De Meester B. *et al*. (2016). Triple pattern fragments: a low-cost knowledge graph interface for the web. *Web Semantics: Science, Services and Agents on the World Wide Web*, Vol. 37, pp. 184–206.

Verma K., Sivashanmugam K., Sheth A., Patil A., Oundhakar S., Miller J. (2005). Meteor-s wsdi: A scalable p2p infrastructure of registries for semantic publication and discovery of web services. *Information Technology and Management*, Vol. 6, No. 1, pp. 17–39.

Völkel M., Krötzsch M., Vrandecic D., Haller H., Studer R. (2006). Semantic wikipedia. In *Proceedings of the 15th international conference on world wide web*, pp. 585–594. Edinburgh, Scotland, ACM.

Volz J., Bizer C., Gaedke M., Kobilarov G. (2009). Silk - a link discovery framework for the web of data. In *Proceedings of the www2009 workshop on linked data on the web*, Vol. 538. Madrid, Spain, CEUR.

Vrandečić D., Krötzsch M. (2014). Wikidata: a free collaborative knowledgebase. *Communications of the ACM*, Vol. 57, No. 10, pp. 78–85.

Waitelonis J., Ludwig N., Knuth M., Sack H. (2011). Whoknows? evaluating linked data heuristics with a quiz that cleans up dbpedia. *Interactive Technology and Smart Education*, Vol. 8, No. 4, pp. 236–248.

Waitelonis J., Sack H. (2012). Towards exploratory video search using linked data. *Multimedia Tools and Applications*, Vol. 59, No. 2, pp. 645–672.

Wang S., Groth P. (2010). Measuring the dynamic bi-directional influence between content and social networks. In *International semantic web conference*, pp. 814–829. Shanghai, China, Springer.

Weibel S., Kunze J., Lagoze C., Wolf M. (1998, September). *Dublin core metadata for resource discovery*. RFC No. 2413. RFC Editor. Internet Requests for Comments.

Weiss C., Karras P., Bernstein A. (2008). Hexastore: sextuple indexing for semantic web data management. *Proceedings of the VLDB Endowment*, Vol. 1, No. 1, pp. 1008–1019.

Wu D., Parsia B., Sirin E., Hendler J., Nau D. (2003). Automating daml-s web services composition using shop2. In *International semantic web conference*, pp. 195–210. Sanibel Island, USA, Springer.

Zeginis D., Tzitzikas Y., Christophides V. (2007). On the foundations of computing deltas between rdf models. In *The semantic web : 6th international semantic web conference, 2nd asian semantic web conference*, pp. 637–651. Busan, Korea, Springer.

Zhu L., Ghasemi-Gol M., Szekely P., Galstyan A., Knoblock C. A. (2016). Unsupervised entity resolution on multi-type graphs. In *International semantic web conference*, pp. 649–667. Kobe, Japan, Springer.

Zimmermann A., Euzenat J. (2006). Three semantics for distributed systems and their relations with alignment composition. In *International semantic web conference*, pp. 16–29. Athens, USA, Springer.

Zimmermann A., Krotzsch M., Euzenat J., Hitzler P. (2006). Formalizing ontology alignment and its operations with category theory. In *Formal ontology in information systems: Proceedings of the fourth international conference*, pp. 277–288. Baltimore, USA, IOS Press.

Zimmermann A., Lopes N., Polleres A., Straccia U. (2012). A general framework for representing, reasoning and querying with annotated semantic web data. *Web Semantics: Science, Services and Agents on the World Wide Web*, Vol. 11, pp. 72–95.