Are We Better Off With Just One Ontology on the Web?

Armin Haller [a,*] Axel Polleres [b]
[a] *Research School of Management & Research School of Computer Science, Australian National University, Canberra, Australia*
E-mail: armin.haller@anu.edu.au
[b] *Institute for Information Business, Vienna University of Economics and Business, Vienna, Austria*
E-mail: axel.polleres@wu.ac.at

Editors: Pascal Hitzler, Wright State University, USA; Krzysztof Janowicz, University of California, Santa Barbara, USA

Abstract. Ontologies have been used on the Web to enable semantic interoperability between parties that publish information independently of each other. They have also played an important role in the emergence of Linked Data. However, many ontologies on the Web do not see much use beyond their initial deployment and purpose in one dataset and therefore should rather be called what they are – (local) schemas, which per se do not provide any interoperable semantics. Only few ontologies are truly used as a shared conceptualization between different parties, mostly in controlled environments such as the BioPortal. In this paper, we discuss open challenges relating to true re-use of ontologies on the Web and raise the question: "are we better off with just one ontology on the Web?"

Keywords: Ontology, Knowledge Representation

1. Introduction

Back in 1993, Gruber introduced "ontologies"[1] as an "*explicit specification of a conceptualization*" consisting of a "*set of objects, and the describable relationships among them*" represented in a declarative formalism [23]. Uschold and Grüninger [65] argued later that semantic interoperability between parties that exchange data is a key application of ontologies.

The use of ontologies as an approach to overcome the problem of semantic heterogeneity on the World Wide Web has since been well established. Semantic heterogeneity occurs whenever two contexts do not use the same interpretation of information. According to Goh [21] three causes for such semantic heterogeneity can be identified.

– **Confounding conflicts** refer to those arising from the confounding of concepts which are in fact distinct. An example is the maximum temperature on a given day. Due to different time-periods (e.g., calendar day vs. a 24 hour time-period) and different methods of averaging (e.g., over a minute vs. over an hour) the actual values, even when recorded by the same sensor, will often differ when published by different parties.
– **Naming conflicts** occur when naming schemes of information differ significantly, for example synonyms and homonyms among attribute values. For example, the entities *Product* and *Item* are often found to be synonyms in commerce applications.
– **Scaling and units conflicts** refer to the adoption of different units of measure or scales, e.g., imperial gallon vs US gallon vs litre.

Many ontology-based approaches that address these causes of semantic heterogeneity have been proposed

*Corresponding author. E-mail: armin.haller@anu.edu.au.
[1] The plural use of the term "ontology" in computer science quite likely still raises eyebrows for anyone with a background in ontology in philosophy.

since [47, 71]. The idea is that a shared ontology which carries a formal semantics, acts as a gold standard for the definition of information in different contexts and applications. Many kinds of ontologies have been proposed that can be classified on a spectrum from very lightweight ones that may consist of terms only, with little or no specification of the meaning of the term, to rigorously formalized logical theories [66]. In this paper we focus on the latter, i.e., formal ontologies expressed in RDFS/OWL.

The ontology engineering community has proposed ontologies with different levels of abstractions to ease reuse and to also layer ontologies upon each other. Although no agreed upon ontology hierarchy exists, adapting the ontology classification of Guarino [25], we can largely distinguish four different levels of abstraction in ontology design as shown in Figure 1.

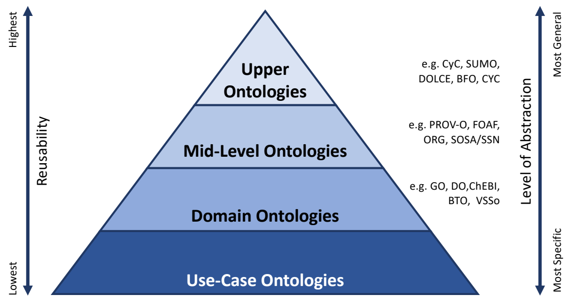

Fig. 1. Levels of Abstraction in Ontology Design

1. **Upper ontologies** that define very general terms that are common across all knowledge domains, examples of which are CYC [40], SUMO [46], DOLCE [19] and BFO [60].
2. **Mid-level ontologies** (sometimes also called *top domain ontologies* or *global domain area ontologies*) act as a bridge between the abstract content of an upper ontology and the richer detail of various domain ontologies. Space and time are two modelling aspects shared between any domain, and ontologies such as the OWL Time Ontology [9] and Geonames are widely used across domains. Other examples of mid-level ontologies are PROV-O [39], FOAF [6], ORG [56] and SOSA/SSN [30] that define concepts generally enough so that their semantics can be further narrowed by a domain ontology.
3. **Domain ontologies** define concepts and relations that belong to a specific domain. Each domain ontology typically models domain-specific definitions of terms. Examples of domain ontologies are the Gene Ontology [3], the Disease Ontology [57], ChEBI [15], the Building Topology Ontology (BTO) [55] or VSSo [38], the Vehicle Signal and Attribute Ontology. The latter is a recently developed car signal ontology that derives from the automotive standard VSS, and that builds upon a mid-level ontology pattern, i.e., from SSN/SOSA, for representing observations and actuations.
4. **Use case ontologies** include a set of detailed classes and relations highly dependent on the use case. For example, in a smart home environment for an apartment building, a use case ontology may extend terms in a domain ontology to be able to use those terms for a number of similar units in an apartment complex.

2. Challenges in Reusing Ontologies

While upper ontologies have experienced strong research interest in the early 2000's, their use on the Web has largely been confined to the biomedical domain where the community, through the OBO foundry, maintained and mandated the use of the BFO upper ontology. In fact, in an analysis of links [29] in the LOD Cloud [1] we have discovered that not a single dataset in a corpus of 430 Linked Open Datasets that were investigated for this study reuses DOLCE or SUMO, the other two main open-source upper ontologies.

This lack of adoption of upper ontologies outside the biomedical domain can mostly be attributed to the complexity and rigidity of these ontologies and the often unintended inferences that would result from importing the upper ontology in a mid-level or domain ontology. Examples of such unintended inferences are global domain and range restrictions defined in an upper ontology (e.g., DOLCE+DnS Ultralite (DUL) uses global property restrictions) that may lead to inferences in the importing domain ontology that are inconsistent in its domain of discourse. Another example is the disjointness of a set of classes defined in an upper ontology that results in an unintended restriction on the use of the domain class that is a subclass of such an upper level class. For example, in the old SSN, the `Sensor` class was defined as a subclass of a DUL `PhysicalObject`. However, users of the SSN ontology who wanted to use the `Sensor` class for computational methods, could not, because a `dul:PhysicalObject` is disjoint with a `dul:SocialObject` (which most certainly would include a computational algorithm). For this and other reasons [30], in the redesign of the SSN ontology, the working group decided to remove the dependency of the SSN ontology on the DOLCE Ultralite ontology and make its alignment optional, i.e., provide it in a separate ontology file that is not imported [30] (while at the same time relax its semantics by using higher level ontology classes from DUL). However, in terms of Linked Data principles, this optionality breaks findability through automated means, that is, solely by dereferencing links ("following your nose").

Recognising the issues with adoption of upper ontologies, the ontology engineering community has developed reusable ontology design patterns [18] that are suitable to be used as templates (i.e., guiding design principles) in lower level ontologies. These patterns bring the benefits of a traditional upper-ontology-based integration approach while avoiding its pitfalls, i.e., the need of importing the upper ontology with all its ontological commitment. Over 200 such patterns have since been submitted to the ontology design pattern initiative[2] and several of those have been reused or proposed in mid-level ontologies.

Beyond the aforementioned challenges in reusing upper ontologies, evaluating which mid-level or domain ontology is suitable for a given use case is challenging for several reasons. Gómez-Pérez [22] has proposed a criteria-based approach to ontology evaluation. Yu et al. [72] have reviewed the various criteria that have been proposed for the evaluation of ontologies. These include clarity, coherence, extendibility, minimal ontological commitment, and minimal encoding bias as proposed by Gruber [23]; competency as proposed by Grüninger and Fox [24]; consistency, completeness, conciseness, expandability, and sensitiveness as proposed by Gómez-Pérez [22] and correctness as proposed by Guarino and Welty [26].

While some of these criteria (e.g., consistency) can be verified automatically using reasoners such as Pellet [59], FaCT++ [63] or HermiT [20], others like clarity or expandability, can be difficult to evaluate as there are no means in place to determine them [72]. Other criteria require manual inspection of the ontology. For example, correctness requires a domain expert to manually verify that the definitions are correct with reference to the real world.

[2]see http://ontologydesignpatterns.org

In the following we identify a set of challenges that we have repeatedly encountered in ontology engineering consultancies with Government and industry clients. These include some of the ontology evaluation criteria above (some of which, e.g., clarity, consistency, correctness, conciseness, are combined together into one category, 'quality'), but also include other challenges that are specific to the reuse of distributed ontologies on the Web.

Availability: For ontologies to be any use in terms of serving Linked Data, they need to be highly available, preferably in perpetuity. What that means is that the file encoding the ontology needs to be permanently retrievable at the namespace URI of the ontology. Although studies have shown [7, 29] that ontologies have higher availability than Linked datasets built using these ontologies, various issues with accessing ontologies still exists. For example, purl.org, a popular service for over 15 years for creating permanent URLs on the Web that was used for many ontology namespaces including the Dublin Core Metadata initiative, ran into availability issues in 2015, as it was mostly a volunteer-driven community service. The Internet Archive has taken control of the service in the meantime and guarantees its continued support, while the W3C has since introduced w3id.org, a permanent identifier service for the Web. However, both services only offer a solution for the permanence of the URI, the ontology file itself has to still be stored persistently somewhere else. Many ontologies are now hosted on Github, but the long-term availability of this service depends on its commercial viability, and as history has shown not all such services survive: e.g., Google Code turned off its hosting services in 2016,[3] or, likewise, SourceForge, as another example, was confronted with problematic incidents like malware bundling, and changing service ownership in the past, raising doubts about its sustainability.

Discoverability: One of the main barriers for the uptake of ontologies has been the difficulty that data publishers face in discovering ontologies on the Web to describe the semantics of their data. Although, again the biomedical community has developed and maintained their own successful repository, the BioPortal [49], there has been a lack of a general-purpose ontology search engine or a central ontology library [11], beyond the relatively recently proposed Linked Open Vo-

[3]https://code.google.com/archive/

cabulary repository [67]. However, neither of the major search engine providers support the search or discovery of ontologies on the Web and therefore a non-expert ontology user has to largely rely on their social network to find and reuse existing ontologies. Ideally, in order to facilitate discoverability, search engines would need to provide a dedicated concept/property search operator, similar to "filetype" or "site" in Google. We emphasise that such services existed in the past[4], but these community-operated, academic services have in the meanwhile been discontinued.

***Completeness & Adaptability**:* Completeness of an ontology can only be evaluated against the purpose it was built for. Typically this purpose has been expressed through a number of use cases against which the ontology has been validated [24]. Often, when reusing a specific ontology, the use case may differ from the one the ontology was built for, and consequently, not all concepts and axioms that are needed, are included in the ontology for reuse. Also, ideally, the ontology should be adaptable, i.e., the ontological commitment of the ontology should not prevent the reuse of a term in a different context (e.g., through unrestricted domain and range restrictions). However, studies [36] have found that term reuse from existing ontologies is not widespread (most ontologies reuse less than 5% of their terms), while almost one in three terms overlapped in the investigated ontology corpus, i.e., they could have been reused. While the study itself did not present findings on why these terms were not reused, the ontological commitment and semantic completeness of a term often influences its potential reuse.

***Maintenance & Versioning**:* Curating and maintaining reusable ontologies is a prerequisite for their continuous relevance since the mental models of the world that the ontology has been created for may change. Just imagine a mobile phone ontology that was created in the late 90's. It would not include concepts for a 'touchscreen', 'fingerprint sensor' or even a 'wifi antenna'. These and human factors (mistakes in the ontology design) can lead to semantic drift in ontologies over time. In order to address these, ontologies need to undergo regular revision. Some of the most used ontologies on the Web [42], such as FOAF [6], SIOC [5] or SKOS [45], have undergone several revisions. Ontologies managed by the W3C, for example, do undergo regular revisions, most recently the W3C Time Ontology [9] underwent a revision more than 10 years after its first publication. When an ontology is revised, decisions have to be made on the versioning of the ontology namespace. In their seminal work on ontology versioning, Klein and Fensel [37] identified four different methods of how an ontology might be versioned; 1) the previous version is silently replaced by the new version; 2) the ontology is visibly changed, but the old version is replaced by the new version; 3) the ontology is visibly changed, and both versions are accessible at different URIs; or 4) there are two versions available at separate URIs and there is an explicit specification of the relation between terms in the new version and terms in the previous version. The authors also raise a question at what point a new URI should be minted, and recommend to change the namespace URI only in cases where the conceptualization of the ontology changes.

Ideally, every ontology should follow the guidelines proposed in Klein and Fensel [37] in combination with more recent guidelines around content negotiation [35] and use version numbers for changes in the conceptualization of the ontology in combination with a persistent URI that redirects to the most recent version of the ontology [41]. Another possible approach to versioning is to use the Memento protocol [13], or components thereof, to express temporal versioning of a dataset and to allow access to the version that was operational at a given datetime.

In many cases, however, either one of the first three approaches mentioned above is chosen instead when publishing an ontology. Even the popular FOAF ontology violates some of the proposed versioning principles. Although it uses different version numbers for the evolution of the ontology, it still uses the original namespace URI (i.e., http://xmlns.com/foaf/0.1/) for its most recent version, 0.99, and it does not make the changes from one version to the other formally explicit. In fact, many other more recent ontologies like schema.org [27] or the DBpedia ontology [4] do not adhere to the guidelines proposed in [37] and silently update the semantics of terms. Only very recent ontologies standardised in the W3C, the Time Ontology [9] and SSN/SOSA [28], make the relation to terms in the previous version of the ontology explicit through a mapping file, but then again, the Time Ontology continues to use the old URI including the old date (i.e., http://www.w3.org/2006/time#) for its most recent version, while SSN/SOSA introduces a new

[4] For instance, we used services like Sindice [64] and SWSE [33] in the past for auto-completion of ontology term search in Drupal [8].

ontology namespace URI (i.e., http://www.w3.org/ns/ssn/), while no versioned URI is linked from that new namespace.

Modularization: There are two different methods one can reuse terms from an ontology; 1) either by directly importing the source ontology using an `owl:imports` statement and therefore importing all entities, expressions, and axioms; or 2) by selectively reusing class or property URIs from an external ontology without importing its ontological commitment. While the former is the preferred approach to avoid errors in the reuse of terms, the latter is the more common in the Linked Data Web [53]. One of the reasons why using an `owl:imports` statement is often avoided, is that the importing ontology may be large and by importing all axioms, one may end up with inferences that are either hard to handle in software using the ontology or are unintended in a given domain. A solution to this problem is the splitting up of the set of axioms of an ontology into a set of modules. Largely two approaches to modularization exist [12], either at design time by the ontology designers themselves using several ontology namespace URIs for the ontology modules (e.g., DOLCE [19] has been redesigned to be available in modules), or at reuse time through segmentation [58] or traversal view extraction [48]. However, very few ontologies besides DOLCE and SSN/SOSA use a modularization architecture.

Quality: Beyond syntactic and semantic errors that can be checked by reasoners as mentioned above, the notion of the quality of an ontology is rather imprecise. Some even argue that ontologies on the Web do not need to be consistent, and systems should be able to deal with noise, different perspectives, and uncertainty [31]. In his dissertation, Vrandeĉić [69] investigates how to assess the quality of an ontology on the Web and concludes that a single measure to assess the overall quality of an ontology is elusive, and proposes ontology evaluation methods that identify shortcomings in ontologies instead. Few tools exist [54], though, that test such common shortcomings in ontologies, while no framework is available that assesses and compares the quality of ontologies available on the Web. Some ontologies are now undergoing a peer-review process in scientific conferences and journals, while others are being standardised, but still the vast majority of ontologies are not assessed for their quality. Therefore, users of ontologies need to have the expertise to assess the quality of an ontology themselves. Since most naïve users do not possess this skill and can not distinguish between high-quality and low-quality ontologies, they assess the ontology rather by its fit for a given use case.

Trust: While ontologies are built in a truly decentralised manner, companies and organisations still need to trust the publisher when reusing a digital asset on the Web, such as an ontology. Consequently, the most popular ontologies have either been developed and/or are hosted by standardisation bodies such as the W3C (e.g., PROV-O [39], ORG [56], SSN/SOSA [30]), have a long history of availability, curation and community support (e.g., FOAF [6], SIOC [5]) or are supported through a community of best practices (e.g., the OBO Foundry). While the W3C has resisted to standardise ontologies for a long time, and still does not see itself in the business of doing so, the major search engines Google, Yahoo!, and Bing have built their own ontology (schema.org [27]) while Facebook has built its own simple social profile ontology, the Open Graph Protocol[5], both of which are now the most widely used vocabularies/ontologies on the Web [42].

3. The Present and the Future

The success story of schema.org as an ontology with very lightweight semantics, that already in 2015 has been used in 31.3% of all pages on the Web [27] and that is backed by a trusted consortium of search engine providers, raises the question of whether it is an end-all solution for defining terminology on the Semantic Web [44]. Revisiting the above challenges, let us briefly discuss if and how schema.org addresses these (cf. also Table 3.1).

3.1. The Schema.org Approach

Availability: While neither the schema.org ontology itself is hosted by a publicly-funded open-access repositories nor is the namespace registered with a persistent URI service such as w3id.org, the ontology and namespace are managed by a consortia of globally operating search engines, which implies high availability and support for the ontology.

[5] see http://ogp.me/

Discoverability: Although the schema.org ontology is surprisingly hard to find on Google[6], it is a well known and highly advertised vocabulary/ontology in the Web developers community. It is also used by Google to inform their rich snippets, which gives Web developers an incentive to use the ontology to improve their search results on the Google Search Engine.

Completeness & Adaptability: With a strong focus on the eCommerce domain, schema.org is far from being a complete ontology for general human knowledge. However, a mechanism is provided where the community can propose extensions to schema.org. From personal experience (in the concrete case, a suggestion for addition to the ontology from the SOSA/SSN specification [28]), it appeared that the feedback process from outside the community is handled by a few individuals and not very dynamic. Although this is sufficient for data publishers that are mainly interested in improving the appearance of their search results on Google or the inclusion of their data in the Google Knowledge graph, it is an unsuitable process for governmental, industrial or science applications.

Maintenance & Versioning: Schema.org is continuously curated since its launch in 2011 [27]. Although the process of change in schema.org is transparent, with a release history that works through issues that have been raised on the tracker being published online, the changes to terms in the ontology are not made explicit in the term definition itself and the class or property URI is just servicing the new semantics of the term.

Modularization: While schema.org is not published in a modular fashion, each term in the ontology is being served by its own webpage and through using a Linked Data content negotiation technique a subgraph is served at the same URI.

Quality: While an ontology like schema.org that is constantly evolving may not always be consistent or correct, there is a feedback mechanism in the form of an issue tracker. Also, schema.org is using lightweight semantics with annotation properties (`schema:domainIncludes` and `schema:rangeIncludes`) instead of domain and range restrictions and no OWL constructs other than `owl:equivalentClass` and `owl:equivalentProperty`, and therefore there are only a few axioms that could be violated by additions to the ontology. On the other hand, these lightweight semantics also undermine some of the data integration benefits of fully-fledged OWL-based ontologies as discussed earlier.

Trust: Since schema.org is supported by a consortia of all major search engine providers (other than Baidu) there is little doubt that users (will) trust schema.org. While that is true for the ontology itself, the data modelled using schema.org, however, has trustworthiness/reliability issues similar to any other data that is created on the Web for a commercial benefit of the publisher.

The analysis above shows that schema.org scores well in most of the considered reuse criteria. However, although we believe that schema.org will continue to evolve and we will see an even bigger uptake of it, we believe it is not yet the end-all ontology on the Web for two reasons; 1) in terms of its *Completeness* there is little indication that it will be extended beyond the eCommerce domain (with few exceptions like the Health and Lifesciences domain) any time soon. Moreover, data providers are providing *schema.org* annotations mainly for commercial reasons, i.e., better ranking and visibility on search engines [43], while there is little to no incentive for them to annotate non-commercial knowledge with *schema.org*; 2) in regards to its *Quality*, while the lightweight semantics were deliberately chosen to make annotations on the Web easier for the average Web developer [27], they prevent the use of the ontology in environments with a requirement for stricter formal connections such as in sciences' domains or in the Governmental policy domain. Also, while community extensions are managed through an open process, the decision on additions to the ontology still sits with the providers of the ontology, i.e., the search engine companies.

The large uptake of schema.org [27, 43] and the Open Graph protocol on the Web [42], however, are signs of an emerging trend of a long tail in ontology use on the Web, with some few ontologies seeing the majority of use, while most other ontologies are only used once in the use case they were built for, a phenomena that we also observed in a recent study [29].

[6] e.g., a Google search for "product concept" or "product ontology concept" does not yield in a result to the schema.org "product" class (which is core to the ontology) within the first 10 result pages.

	schema.org	**Wikidata ontology**	**DBpedia ontology**
Availability	Highly available	Highly available	Highly available
Discoverability	Relatively easy	Relatively difficult *Linked from Wikipedia, but ontology itself hard to retrieve*	Relatively difficult *Only known in Semantic Web community*
Completeness & Adaptability	Domain specific *Community extensions available*	Generic *Combined Top-Down/Bottom-up creation process*	Generic *Top-down ontology engineering process, combined with auto-generated entities*
Maintenance & Versioning	Continuous curation Versions are not made explicit	Continuous curation Explicit entity version, and version history available through version control	Continuous curation Explicit ontology version
Modularization	Fully distributed ontology *Easy access through Linked Data content negotiation*	Fully distributed ontology *Difficult to access, through SPARQL endpoint and list pages*	Monolithic ontology *Easy access through file and SPARQL endpoint*
Quality	High quality, but lightweight semantics	Variable quality in lower parts of the ontology *No DL semantics, therefore few provable inconsistency*	Medium to Low Quality
Trust	High Trust *Developed by major search engines*	Medium Trust *Developed by community, maintained by Wikimedia Foundation*	Medium Trust *Developed and maintained by University partners*

Table 1
Evaluation of reuse criteria for schema.org, wikidata.org and dbpedia.org ontologies

3.2. The DBpedia and Wikidata Approach

There have been mainly two approaches, DBpedia[7] and Wikidata, for generic knowledge ontologies that address well our reuse criteria and could emerge as the one reference ontology on the Web. DBpedia, created in 2007 by Free University of Berlin and Leipzig University in collaboration with OpenLink Software, extracts data from Wikipedia info boxes to build an RDF graph. Wikidata [70], the “Wikipedia for data” project, established in late 2012, manages the factual information of the popular online encyclopedia. Its main goal is to provide high-quality structured data acquired and maintained collaboratively to be directly used by Wikipedia to enrich its content. In Table 3.1 and the following paragraphs we will assess these two approaches in regards to our ontology reuse criteria. For comparative studies that go beyond our focus on the ontology underlying Wikidata and DBpedia the interested reader is referred to Färber et al. [16] or Abián et al. [2].

Availability: Both ontologies are highly available. That being said, while Wikidata is run by Wikimedia, the same organisation successfully hosting Wikipedia for more than 18 years, DBpedia is run by an association affiliated with the University of Leipzig.

Discoverability: Although Wikidata does not yet have the same visibility as Wikipedia, its Alexa rank is 8,496 as of October 2019 compared to Wikipedia’s rank of 9, it can easily be reached through any page on Wikipedia. DBpedia, while extremely well known in the Semantic Web community, only ranks 158,385 on Alexa. From our own experience representing the W3C in Australia and chairing a Government Linked Data working group, it is largely unknown outside of the scientific Semantic Web community, even to people with ontology engineering skills. Assessing the discoverability of the ontology itself, Wikidata leaves a lot to be desired. To the best of our knowledge, it is impossible to download the entire ontology from the Wikidata site. There are pages listing some of the top-

[7] Yago [61] is another very similar approach to DBpedia with a stronger taxonomic backbone that ensures better quality than DBpedia. However, at the time of writing, the latest stable release of Yago is from 2017, whereas DBpedia releases a new version monthly. We therefore limit our analysis to DBpedia, while both approaches can be considered largely equivalent in the assessment of the reuse criteria other than on the quality aspect.

level concepts and relations[8], but to retrieve only the TBox statements from the Wikidata dump or SPARQL endpoint, someone would need to write sophisticated queries. DBpedia on the other hand releases its ontology as one file that is easily discoverable from its namespace URI (i.e., http://dbpedia.org/ontology/).

Completeness & Adaptability: Neither Wikidata nor DBpedia are built for a specific use case, but they are rather generic knowledge bases that aim to capture the sum of all human knowledge (as of their vision statement). Studies have compared the breadth and depth of the knowledge captured and concluded that they are comparable [2]. Comparing the ontologies themselves is difficult, as for the difficulty in obtaining the entirety of the Wikidata ontology. However, it has to be noted, though, that there is a fundamental difference in how the two ontologies are built and how they can be adapted. Anyone can add concepts or relations to the Wikidata ontology directly, whereas in DBpedia concepts and relations are added to the ontology through the "schema" of the info boxes in Wikipedia, i.e., they cannot be added to the ontology directly. Reusing and adapting specific entities of either ontology is easy, as both ontologies are served through Linked Data APIs that allow one to reference the entity by its URI (while retrieving only its subgraph). The implications of doing that with a DBpedia entity are different to a Wikidata entity, as the former is an OWL-based ontology, whereas the latter does not rely on Description Logics' (DL) semantics: the fact that Wikidata, by defining own properties and classes for relationships such as `instanceOf` (`P31`), `subclassOf` (`P279`), etc., instead of relying on RDFS' and OWL's properties such as `rdf:type` and `rdfs:subclassOf` with their standardized semantics, may be viewed – on the one hand – as lack of ontological commitment. However, on the other hand, this lack of commitment also leaves applications and users more room for contextual, maybe even collaboratively evolving interpretations of Wikidata's terminological vocabulary: we might in the future envision different sets of inference rules or semantics being defined as extensions of or within Wikidata itself, rather than remaining caught in the prescriptive semantics of the OWL and RDF(S) vocabularies. In fact, as earlier works have shown, relying on strict OWL and RDFS reasoning [32, 51], or even on strict interpretations of the RDF vocabulary (e.g., in terms of blank nodes [34]) is not suitable in all contexts when applied to collaboratively published Web data "in the wild", leading to unintended and non-intuitive inferences.

Maintenance & Versioning: While both, Wikidata and DBpedia are continuously evolving ontologies that rely on a manually developed core, the major difference is that large parts of the Wikidata ontology are generated in a collaborative, bottom-up fashion by a large number of contributors, while the DBpedia ontology is created by the maintainers of the mapping from the Wikipedia info boxes to the DBpedia data set. Each release of the DBpedia ontology corresponds to a new release of the DBpedia data set. In terms of versioning the two approaches differ too. While DBpedia continuously uses the same namespace of the ontology, the version number is made explicit by an `owl:versionInfo` annotation property. Wikidata relies on the versioning mechanism offered by the MediaWiki software and changes are made explicit through annotation properties that indicate the *timestamp*, *version* and *dateModified* of a term. There is no mechanism that allows to refer to the semantics of a term in Wikidata at a specific point in time; i.e., for each change in the conceptualization of a term, no new URI is minted that includes a reference to the old version of that term. Unfortunately, for both, DBpedia and Wikidata entities, there is no explicit mechanism to reference a specific version of an entity, i.e., if a domain ontology references an entity in either of the two ontologies, the semantics of the entity could have changed from when it was referenced. While changes in DBpedia at the instance level could be traced back to Wikipedia's built-in version control, ontological changes are somewhat hidden in DBpedia's extractor framework, with its versions being managed separately. On the contrary, terminological changes in Wikidata's properties and classes are accessible explicitly via MediaWiki's built-in version control system, as mentioned above. As earlier works, such as the DBpedia Wayback Machine [17], have demonstrated, URIs corresponding to such changes in a Wiki version control, could be minted and linked to each other exposed through the Memento protocol, allowing for references to particular versions by explicit URIs: while Fernández et al. [17] only demonstrated this approach for instance changes in Wikipedia, the same approach seems feasible for making terminological changes in Wikidata explicit.

[8] e.g., https://www.wikidata.org/wiki/Wikidata:WikiProject_Ontology/Top-level_ontology_list)

Modularization: Neither of the two ontologies is modularized. Whereas the DBpedia ontology is provided in one monolithic file, the Wikidata ontology can only be retrieved on the basis of an entity. The ontology itself can not be transparently retrieved at its namespace URI, nor can the ontology itself, to the best of our knowledge, be downloaded from a single source. The ontology is, of course, retrievable through the Wikidata SPARQL API, but even for expert users it is a challenge to just retrieve the TBox statements, given that this SPARQL endpoint gives also access to the entire Wikidata ABox.

Quality: Both, the Wikidata ontology and the DBpedia ontology are collaboratively created. While editors can directly manipulate the Wikidata ontology through the MediaWiki software, the DBpedia ontology is derived through a mapping from the Wikipedia info boxes, which themselves are created by contributors to the English Wikipedia. However, since these info boxes are created using natural language, the mapping of attributes from those info boxes to ontology relations in DBpedia leads to issues with the conciseness and minimal commitment of the DBpedia ontology. For example, the current version of the ontology includes over two dozen relations (e.g., `dbo:winsAtLAGT`, `dbo:winsAtLET` or `dbo:winsAtLPGA`) that are used to define wins of players in various sports at various events. A recent approach by Paulheim and Gangemi [50] proposes the use of an upper ontology, i.e., DOLCE, to detect such inconsistencies within DBpedia.

The Wikidata ontology does not introduce such redundancies, since the software will alert an editor if a relation already exists. It does, however, still suffer from modelling inconsistencies at lower levels of the class hierarchy. For example, in its current version as of October, a "Beef Wellington" (`wd:Q1412680`) is defined as a subclass of dish (`wd:Q746549`) and a subclass of beef dish (`wd:Q28100368`). Since beef dish is a subclass of dish itself, this is redundant information. A "Wiener Schnitzel" (`wd:Q6497852`) on the other hand is defined as an instance of veal dish (`wd:Q28100665`), itself a subclass of dish, while at the same time it is defined as a subclass of schnitzel (`wd:Q11293688`). Since Wikidata does not use DL semantics as mentioned above, i.e., neither *instance of* or *subclass of* are defined as `rdf:type` or `owl:subClassOf`, respectively, this example of (most likely) unintended punning does not introduce errors in the ontology. It is, however, just one of many examples of inconsistencies in the ontology. The Wikidata ontology, however, has a strong focus on including references to external ontologies that either informed the modelling of an entity or that are equivalent (i.e., not DL equivalent) to the entity. For example, the concept "cellular homeostasis" references the Gene Ontology entity `GO:0019725` and defines `wd:Q14881703` to be an exact match `wd:P2888` to `GO:0019725`.

Trust: Beyond a manually created core, the Wikidata ontology is created in a collaborative fashion. As such, the quality varies, similar to how the quality of Wikipedia articles varies. Still, users of Wikipedia trust that the moderation process and the many editors make sure that the information is largely correct. Similarly, Wikidatans have collaborated to create and maintain the Wikidata ontology and one can expect that the users will have a fairly high trust in the ontology. While the same applies to DBpedia to a certain extent, the ontology itself is created through a mapping process and hosted by Universities that do not have the same brand recognition as Wikipedia/Wikidata.

While DBpedia has been around since its first public release in 2007 and seen great success as a core reference ontology and dataset in the Linked Data Cloud [29], it has not become the one general knowledge reference ontology on the Web. Also, studies have shown that the Linked Data cloud itself has become rather stale, of late [14, 52, 68]. Interestingly, parts of the Wikipedia info boxes that are used to create the RDF graph in DBpedia are now created from Wikidata (with a plan to progressively create all Wikipedia info boxes from Wikidata). This should lead, in the long term, to a convergence between the Wikidata and DBpedia ontology (essentially, making the latter obsolete).

While a future of highly distributed ontologies on the Web with strong linkage between them is still possible, evidence from analysis [29] of the most successful Linked Data project, the LOD cloud [1], largely paints a different picture. We believe, however, that the Wikidata ontology, which was only introduced in late 2012 together with the Wikidata project, may have more success in becoming this "one ontology on the Web". Its strength lies in the bottom-up, collaborative development approach that strives to incorporate the source of a term. This means, for the ontology part, it reuses and references existing ontologies where possible, but mints URIs for entities in the Wikidata namespace. This clearly sets it apart from the schema.org and DBpedia approach, the former just creates entities in its namespace without an explicit reference to existing models, while the latter relies on these references being part of the Wikipedia info boxes. What that means for Wikidata is that it can incorporate existing, highly curated and high-quality ontologies. This means, that such ontologies that are built and maintained in domain portals, such as the BioPortal [49], the ETSI community building the Smart Appliances REFerence (SAREF) ontology [10] or the FiBO financial ontology[9], will be made more accessible to the wider public through its duplication and reference in the Wikidata namespace.

However, although Wikidata meets most of the reuse criteria outlined above, there are still challenges that need to be addressed for it to become a true reference ontology for general knowledge on the Web, in particular in terms of its quality assurance and better accessibility and discoverability of the TBox itself. There are efforts to improve the quality of entities by including shape expressions for entities in Wikidata [62]. This should lead, in the long term, to more consistency between similar typed entities, and as such, also in its ontology. For the latter, we are not aware of efforts to make the ontology more accessible, but we are hoping that this discussion paper may contribute to this issue being addressed.

4. Conclusion

In this paper we have asked the question if we "are better off with just one ontology on the Web?". Analysing the major challenges that publishers and users of ontologies face, and how schema.org addressed some of these challenges to become the most widely used ontology on the Web, we argue that we may indeed be better off with just one ontology on the Web. Similar to how the likes of Amazon, Google, Apple, Facebook or AirBnB benefit from the phenomena of a "winner takes all" network effect, a single winner-takes-it-all ontology would be a true boon for data interoperability on the Web. We argue that schema.org, despite its success in the eCommerce domain, is not (yet) the end-all solution to our ontology woes. We further argue that a winner-takes-it-all ontology should follow the same approach as the one taken by Wikipedia, and provide a bottom-up development of the ontology by the Web community. This bottom-up development of content on Wikipedia helped it, through a network effect, to become the only encyclopedia in use on the Web.

Wikidata as the sister project of Wikipedia to manage the factual human knowledge is building such a community-driven ontology with a strong focus on incorporating and referencing existing ontologies, while at the same time minting URIs in the Wikidata namespace. This allows it to thrive along-side specialised, high-quality domain ontology repositories, while at the same time increasing their visibility to people outside of these specialised communities.

While the Wikidata ontology still has issues with its modularization and access, only partially addresses the ontology versioning problem through metadata annotations (but not versioned URIs), and has variable quality in some knowledge domains due to its relative young age, we believe and propose that with small changes (the details of which are still in need to be worked out), its ontology could eventually become this one end-all solution to semantic interoperability on the Web.

[9] cf. https://edmcouncil.org/page/aboutfiboreview

References

[1] Andrejs Abele, John P McCrae, Paul Buitelaar, Anja Jentzsch, and Richard Cyganiak. Linking open data cloud diagram 2017. http://lod-cloud.net, Insight-Centre, 2017.

[2] D. Abián, F. Guerra, J. Martínez-Romanos, and Raquel Trillo-Lado. Wikidata and DBpedia: A Comparative Study. In Julian Szymański and Yannis Velegrakis, editors, *Semantic Keyword-Based Search on Structured Data Sources*, pages 142–154. Springer, 2018. https://doi.org/10.1007/978-3-319-74497-1_14.

[3] M.M. Ashburner, C.A.C. Ball, Judith Blake, David Botstein, Heather Butler, J.M.J. Cherry, Allan Peter Davis, Kara Dolinski, Selina Dwight, Janan Eppig, Midori Harris, D.P. Hill, Laurie Issel-Tarver, A Kasarskis, Suzanna Lewis, John Matese, J.E. Richardson, M Ringwald, G.M. Rubin, and Gavin Sherlock. Gene Ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nature Genetics*, 25:25–29, May 2000. https://doi.org/10.1038/75556.

[4] Christian Bizer, Jens Lehmann, Georgi Kobilarov, Sören Auer, Christian Becker, Richard Cyganiak, and Sebastian Hellmann. DBpedia - A crystallization point for the Web of Data. *Journal of Web Semantics*, 7(3):154–165, 2009. https://doi.org/10.1016/j.websem.2009.07.002.

[5] John G. Breslin, Stefan Decker, Andreas Harth, and Uldis Bojars. SIOC: An approach to connect web-based communities. *Int. J. Web Based Communities*, 2(2), 2006. https://doi.org/10.1504/IJWBC.2006.010305.

[6] Dan Brickley and Libby Miller. FOAF Vocabulary Specification 0.99. Technical report, http://xmlns.com/foaf/spec/, 2014.

[7] Anila Sahar Butt, Armin Haller, and Lexing Xie. Ontology Search: An Empirical Evaluation. In Peter Mika, Tania Tudorache, Abraham Bernstein, Chris Welty, Craig Knoblock, Denny Vrandečić, Paul Groth, Natasha Noy, Krzysztof Janowicz, and Carole Goble, editors, *Proc. of the 13th International Semantic Web Conference*, pages 130–147, Berlin Heidelberg, 2014. Springer. https://doi.org/10.1007/978-3-319-11915-1_9.

[8] Stéphane Corlosquet, Renaud Delbru, Tim Clark, Axel Polleres, and Stefan Decker. Produce and Consume Linked Data with Drupal! In Abraham Bernstein, David R. Karger, Tom Heath, Lee Feigenbaum, Diana Maynard, Enrico Motta, and Krishnaprasad Thirunarayan, editors, *Proc. of the 8th International Semantic Web Conference (ISWC)*, pages 763–778, Berlin, Heidelberg, October 2009. Springer. https://doi.org/10.1007/978-3-642-04930-9_48.

[9] Simon J D Cox and Chris Little. Time ontology in OWL. W3C Recommendation, W3C, October 2017. https://www.w3.org/TR/owl-time/.

[10] Laura Daniele, Frank den Hartog, and Jasper Roes. Created in Close Interaction with the Industry: The Smart Appliances REFerence (SAREF) Ontology. In Roberta Cuel and Robert Young, editors, *Formal Ontologies Meet Industry*, Berlin Heidelberg, 2015. Springer. https://doi.org/10.1007/978-3-319-21545-7_9.

[11] Mathieu d'Aquin and Natalya F. Noy. Review: Where to Publish and Find Ontologies? A Survey of Ontology Libraries. *Journal of Web Semantics*, 11:96–111, March 2012. https://doi.org/10.1016/j.websem.2011.08.005.

[12] Mathieu d'Aquin, Anne Schlicht, Heiner Stuckenschmidt, and Marta Sabou. Criteria and Evaluation for Ontology Modularization Techniques. In Heiner Stuckenschmidt, Christine Parent, and Stefano Spaccapietra, editors, *Modular Ontologies: Concepts, Theories and Techniques for Knowledge Modularization*, pages 67–89. Springer, Berlin Heidelberg, 2009. https://doi.org/10.1007/978-3-642-01907-4.

[13] H. Van de Sompel, M. Nelson, and R. Sanderson. HTTP Framework for Time-Based Access to Resource States – Memento. Rfc7089, IETF, 2013.

[14] Jeremy Debattista, Christoph Lange, Sören Auer, and Dominic Cortis. Evaluating the quality of the LOD cloud: An empirical investigation. *Semantic Web*, 9(6):859–901, 2018. https://doi.org/10.3233/SW-180306.

[15] Kirill Degtyarenko, Paula de Matos, Marcus Ennis, Janna Hastings, Martin Zbinden, Alan McNaught, Rafael AlcÃąntara, Michael Darsow, Mickael Guedj, and Michael Ashburner. ChEBI: A database and ontology for chemical entities of biological interest. *Nucleic acids research*, 36:D344–50, February 2008. https://doi.org/10.1093/nar/gkm791.

[16] Michael Färber, Frederic Bartscherer, Carsten Menne, and Achim Rettinger. Linked data quality of DBpedia, Freebase, OpenCyc, Wikidata, and YAGO. *Semantic Web*, 9(1):77–129, 2018. https://doi.org/10.3233/SW-170275.

[17] Javier D Fernández, Patrik Schneider, and Jürgen Umbrich. The DBpedia wayback machine. In *Proc. of the 11th International Conference on Semantic Systems*, pages 192–195, New York, NY, USA, 2015. ACM. https://doi.org/10.1145/2814864.2814889.

[18] Aldo Gangemi. Ontology design patterns for semantic web content. In Yolanda Gil, Enrico Motta, V. Richard Benjamins, and Mark A. Musen, editors, *Proc. of the International Semantic Web Conference*, pages 262–276, Galway, Ireland, 2005. Springer. https://doi.org/10.1007/11574620_21.

[19] Aldo Gangemi, Nicola Guarino, Claudio Masolo, Alessandro Oltramari, and Luc Schneider. Sweetening Ontologies with DOLCE. In *Knowledge Engineering and Knowledge Management: Ontologies and the Semantic Web*, pages 166–181, Berlin Heidelberg, 2002. Springer. https://doi.org/10.1007/3-540-45810-7_18.

[20] Birte Glimm, Ian Horrocks, Boris Motik, Giorgos Stoilos, and Zhe Wang. HermiT: An OWL 2 Reasoner. *Journal of Automated Reasoning*, 53(3):245–269, Oct 2014. https://doi.org/10.1007/s10817-014-9305-1.

[21] Cheng Hian Goh. *Representing and Reasoning about Semantic Conflicts in Heterogeneous Information Systems*. PhD thesis, Sloan School of Management, MIT, Boston, MA, USA, January 1997.

[22] Asunción Gómez-Pérez. Towards a framework to verify knowledge sharing technology. *Expert Systems with Applications*, 11(4):519–529, 1996. https://doi.org/10.1016/S0957-4174(96)00067-X.

[23] Tom Gruber. A translation approach to portable ontology specifications. *Knowledge Acquisition*, 5:199–220, 1993. https://doi.org/10.1145/503124.503150.

[24] Michael Grüninger and Mark Fox. Methodology for the Design and Evaluation of Ontologies. In *Proc. of the Workshop on Basic Ontological Issues in Knowledge Sharing*, July 1995.

[25] Nicola Guarino. Semantic matching: Formal ontological distinctions for information organization, extraction, and integration. In Maria Teresa Pazienza, editor, *Information Extraction A Multidisciplinary Approach to an Emerging Information Technology*, pages 139–170, Berlin Heidelberg, 1997. Springer. https://doi.org/10.1007/3-540-63438-X.

[26] Nicola Guarino and Christopher Welty. Evaluating Ontological Decisions with OntoClean. *Commununication of the ACM*, 45(2):61–65, February 2002. https://doi.org/10.1145/503124.503150.

[27] R.V. Guha, Dan Brickley, and Steve Macbeth. Schema.org: Evolution of structured data on the web. *ACM Queue*, 13(9), 2015. https://doi.org/10.1145/2857274.2857276.

[28] Armin Haller, Krzysztof Janowicz, Simon J D Cox, Danh Le Phuoc, Kerry Taylor, and Maxime Lefrançois. Semantic Sensor Network Ontology. W3C Recommendation, W3C, October 19 2017. https://www.w3.org/TR/vocab-ssn/.

[29] Armin Haller, Javier D. Fernández, Axel Polleres, and Maulik R. Kamdar. What are Links in Linked Open Data? A Characterization and Evaluation of Links between Knowledge Graphs on the Web. Technical Report 2/2019, Department für Informationsverarbeitung und Prozessmanagement, WU Vienna University of Economics and Business, 2019.

[30] Armin Haller, Krzysztof Janowicz, Simon J. D. Cox, Maxime Lefrançois, Kerry Taylor, Danh Le Phuoc, Joshua Lieberman, Raúl García-Castro, Rob Atkinson, and Claus Stadler. The modular SSN ontology: A joint W3C and OGC standard specifying the semantics of sensors, observations, sampling, and actuation. *Semantic Web*, 10(1):9–32, 2019. https://doi.org/10.3233/SW-180320.

[31] Pascal Hitzler and Frank van Harmelen. A Reasonable Semantic Web. *Semantic Web*, 1(1,2):39–44, April 2010. https://doi.org/10.3233/SW-2010-0010.

[32] Aidan Hogan. *Exploiting RDFS and OWL for Integrating Heterogeneous, Large-Scale, Linked Data Corpora*. PhD thesis, Digital Enterprise Research Institute, 2011.

[33] Aidan Hogan, Andreas Harth, Jürgen Umbrich, Sheila Kinsella, Axel Polleres, and Stefan Decker. Searching and browsing Linked Data with SWSE: The Semantic Web Search Engine. *Journal of Web Semantics*, 9(4):365–401, 2011. https://doi.org/10.1016/j.websem.2011.06.004.

[34] Aidan Hogan, Marcelo Arenas, Alejandro Mallea, and Axel Polleres. Everything You Always Wanted to Know About Blank Nodes. *Journal of Web Semantics*, 27:42–69, 2014. https://doi.org/10.1016/j.websem.2014.06.004.

[35] Bernadette Hyland, Ghislain Atemezing, and Boris VillazÃşn-Terrazas. Best Practices for Publishing Linked Data. http://www.w3.org/TR/ld-bp/, W3C, January 2014.

[36] Maulik Kamdar, Tania Tudorache, and Mark Musen. A Systematic Analysis of Term Reuse and Term Overlap across Biomedical Ontologies. *Semantic Web - Interoperability, Usability, Applicability*, 1:1–5, February 2016. https://doi.org/10.3233/SW-160238.

[37] Michel Klein and Dieter Fensel. Ontology Versioning on the Semantic Web. In *Proc. of the First International Conference on Semantic Web Working*, SWWS'01, pages 75–91, Aachen, Germany, 2001. CEUR-WS.org.

[38] Benjamin Klotz, Raphaël Troncy, Daniel Wilms, and Christian Bonnet. VSSo - A Vehicle Signal and Attribute Ontology. In *Proc. of the 9th International Semantic Sensor Networks Workshop co-located with ISWC*, volume 2213, pages 56–63, Aachen, Germany, October 2018. CEUR-WS.org.

[39] Timothy Lebo, Satya Sahoo, and Deborah McGuinness. PROV-O: The PROV Ontology. Technical report, W3C Recommendation 30 April, 2013.

[40] Douglas B. Lenat. CYC: A Large-scale Investment in Knowledge Infrastructure. *Communications of the ACM*, 38(11):33–38, November 1995. https://doi.org/10.1145/219717.219745.

[41] Bernadette Farias Lóscio, Caroline Burle, and Newton Calegari. Data on the Web Best Practices. Technical report, W3C Recommendation 31 January, 2017. https://www.w3.org/TR/dwbp/.

[42] Robert Meusel, Petar Petrovski, and Christian Bizer. The WebDataCommons Microdata, RDFa and Microformat Dataset Series. In Peter Mika, Tania Tudorache, Abraham Bernstein, Chris Welty, Craig Knoblock, Denny Vrandečić, Paul Groth, Natasha Noy, Krzysztof Janowicz, and Carole Goble, editors, *Proc. of the 13th International Semantic Web Conference: Replication, Benchmark, Data and Software Track*, pages 277–292, Berlin Heidelberg, 2014. Springer. https://doi.org/10.1007/978-3-319-11964-9_18.

[43] Robert Meusel, Christian Bizer, and Heiko Paulheim. A Web-scale Study of the Adoption and Evolution of the schema.org Vocabulary over Time. In *Proc. of the 5th International Conference on Web Intelligence, Mining and Semantics, WIMS 2015*, pages 15:1–15:11, New York, NY, USA, 2015. ACM. https://doi.org/10.1145/2797115.2797124.

[44] Peter Mika. On Schema.org and Why It Matters for the Web. *IEEE Internet Computing*, 19(4):52–55, July 2015. https://doi.org/10.1109/MIC.2015.81.

[45] Alistair Miles and Sean Bechhofer. SKOS Simple Knowledge Organization System Reference. W3C Recommendation, W3C, August 2009. http://www.w3.org/TR/skos-reference.

[46] Ian Niles and Adam Pease. Towards a Standard Upper Ontology. In *Proc. of the International Conference on Formal Ontology in Information Systems*, pages 2–9, New York, NY, USA, 2001. ACM. https://doi.org/10.1145/505168.505170.

[47] Natalya F. Noy. Semantic Integration: A Survey of Ontology-based Approaches. *SIGMOD Rec.*, 33(4):65–70, December 2004. https://doi.org/10.1145/1041410.1041421.

[48] Natasha Noy and Mark Musen. Specifying Ontology Views by Traversal. In *Proc. of the 3rd International Conference on the Semantic Web*, pages 713–725, Berlin Heidelberg, 11 2004. Springer. https://doi.org/10.1007/978-3-540-30475-3_49.

[49] Natasha Noy, Nigam Shah, Patricia L Whetzel, Benjamin Dai, Michael Dorf, Nicholas Griffith, Clement Jonquet, Daniel Rubin, Margaret-Anne D Storey, Christopher Chute, and Mark Musen. BioPortal: Ontologies and Integrated Data Resources at the Click of a Mouse. *Nucleic acids research*, 37:W170–3, June 2009. https://doi.org/10.1093/nar/gkp440.

[50] Heiko Paulheim and Aldo Gangemi. Serving DBpedia with DOLCE – More than Just Adding a Cherry on Top. In Marcelo Arenas, Oscar Corcho, Elena Simperl, Markus Strohmaier, Mathieu d'Aquin, Kavitha Srinivas, Paul Groth, Michel Dumontier, Jeff Heflin, Krishnaprasad Thirunarayan, Krishnaprasad Thirunarayan, and Steffen Staab, editors, *Proc. of the International Semantic Web Conference (ISWC)*, pages 180–196, Berlin Heidelberg, 2015. Springer. https://doi.org/10.1007/978-3-319-25007-6_11.

[51] Axel Polleres, Aidan Hogan, Renaud Delbru, and Jürgen Umbrich. RDFS & OWL reasoning for linked data. In *Proc. of Reasoning Web. Semantic Technologies for Intelligent Data Access*, pages 91–149, Mannheim, Germany, July 2013. Springer. https://doi.org/10.1007/978-3-642-39784-4_2.

[52] Axel Polleres, Maulik R. Kamdar, Javier D. Fernández, Tania Tudorache, and Mark A. Musen. A More Decentralized Vision for Linked Data. In *Proc. of the 2nd Workshop on Decentralizing the Semantic Web, co-located with ISWC*, volume 2165, Aachen, Germany, 2018. CEUR-WS.org.

[53] María Poveda-Villalón, Mari Carmen Suárez-Figueroa, and A. Gómez-Pérez. The Landscape of Ontology Reuse in Linked Data. In *Proc. of the 1st International Workshop on Ontology Engineering in a Data-driven World (OEDW 2012)*, Galway, Ireland, 2012.

[54] María Poveda-Villalón, Asunción Gómez-Pérez, and Mari Carmen Suárez-Figueroa. OOPS! (ontology pitfall scanner!): An on-line tool for ontology evaluation. *International Journal on Semantic Web and Information Systems (IJSWIS)*, 10(2), 2014. https://doi.org/10.4018/ijswis.2014040102.

[55] Mads Holten Rasmussen, Pieter Pauwels, Maxime Lefrançois, and Georg Ferdinand Schneider. Building topology ontology. Draft Community Group Report, W3C, January 2019. https://w3c-lbd-cg.github.io/bot/.

[56] Dave Reynolds. The Organization Ontology. Technical report, W3C Recommendation 16 January, 2014. https://www.w3.org/TR/vocab-org/.

[57] Lynn Marie Schriml, Cesar Arze, Suvarna Nadendla, Yu-Wei Wayne Chang, Mark Mazaitis, Victor Felix, Gang Feng, and Warren Alden Kibbe. Disease Ontology: a backbone for disease semantic integration. *Nucleic Acids Research*, 40 (D1):D940–D946, November 2011. https://doi.org/10.1093/nar/gkr972.

[58] Julian Seidenberg and Alan Rector. Web Ontology Segmentation: Analysis, Classification and Use. In Heiner Stuckenschmidt, Christine Parent, and Stefano Spaccapietra, editors, *Proc. of the 15th International Conference on World Wide Web*, pages 211–243, Berlin Heidelberg, January 2006. Springer. https://doi.org/10.1007/978-3-642-01907-4_10.

[59] Evren Sirin, Bijan Parsia, Bernardo Cuenca Grau, Aditya Kalyanpur, and Yarden Katz. Pellet: A practical OWL-DL reasoner. *Journal of Web Semantics*, 5(2):51–53, 2007. https://doi.org/10.1016/j.websem.2007.03.004.

[60] Barry Smith. Basic formal ontology. Technical report, University of Leipzig, Germany, 2003.

[61] Fabian M. Suchanek, Gjergji Kasneci, and Gerhard Weikum. YAGO: A Large Ontology from Wikipedia and WordNet. *Journal of Web Semantics*, 6(3):203–217, 2008. https://doi.org/10.1016/j.websem.2008.06.001.

[62] Katherine Thornton, Harold Solbrig, Gregory S. Stupp, Jose Emilio Labra Gayo, Daniel Mietchen, Eric Prud'hommeaux, and Andra Waagmeester. Using Shape Expressions (ShEx) to Share RDF Data Models and to Guide Curation with Rigorous Validation. In Pascal Hitzler, Miriam Fernández, Krzysztof Janowicz, Amrapali Zaveri, Alasdair J.G. Gray, Vanessa Lopez, Armin Haller, and Karl Hammar, editors, *Proc. of the European Semantic Web Conference*, pages 606–620, Berlin Heidelberg, 2019. Springer. https://doi.org/10.1007/978-3-030-21348-0_39.

[63] Dmitry Tsarkov and Ian Horrocks. FaCT++ Description Logic Reasoner: System Description. In Ulrich Furbach and Natarajan Shankar, editors, *Automated Reasoning*, pages 292–297, Berlin, Heidelberg, 2006. Springer. https://doi.org/10.1007/11814771_26.

[64] Giovanni Tummarello, Renaud Delbru, and Eyal Oren. Sindice.com: Weaving the Open Linked Data. In *Proc. of the 6th International Semantic Web Conference*, pages 552–565, Busan, South Korea, January 2007. https://doi.org/10.1007/978-3-540-76298-0_40.

[65] Michael Uschold and Michael Grüninger. Ontologies: Principles, methods and applications. *Knowledge Engineering Review*, 11:93–136, 1996.

[66] Michael Uschold and Michael Grüninger. Ontologies and Semantics for Seamless Connectivity. *SIGMOD Record*, 33:58–64, 12 2004. https://doi.org/10.1145/1041410.1041420.

[67] Pierre-Yves Vandenbussche, Ghislain Atemezing, María Poveda-Villalón, and B Vatant. Linked Open Vocabularies (LOV): A gateway to reusable semantic vocabularies on the Web. *Semantic Web*, 8:437–452, January 2017. https://doi.org/10.3233/SW-160213.

[68] Pierre-Yves Vandenbussche, Jürgen Umbrich, Luca Matteis, Aidan Hogan, and Carlos Buil Aranda. SPARQLES: Monitoring public SPARQL endpoints. *Semantic Web*, 8(6):1049–1065, 2017. https://doi.org/10.3233/SW-170254.

[69] Denny Vrandečić. *Ontology Evaluation*. PhD thesis, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany, 6 2010.

[70] Denny Vrandečić and Markus Krötzsch. Wikidata: A free collaborative knowledgebase. *Communications of the ACM*, 57 (10):78–85, 2014. https://doi.org/10.1145/2629489.

[71] Holger Wache, Thomas Vögele, Ubbo Visser, Heiner Stuckenschmidt, Gerhard Schuster, H Neumann, and Sebastian Hübner. Ontology-based integration of information — a survey of existing approaches. In *IJCAI-01 Workshop: Ontologies and Information*, pages 108–117, January 2001.

[72] Jonathan Yu, James A. Thom, and Audrey Tam. Requirements-oriented methodology for evaluating ontologies. *Information Systems*, 34(8):766–791, 2009. https://doi.org/10.1016/j.is.2009.04.002.