Semantic Web - (2019) 0
IOS Press

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: axel.polleres@wu.ac.at

[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 conceptualisation 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

1. Introduction

Back in 1993, Gruber introduced "ontologies"[1] as an "*explicit specifications of a conceptualization*" consisting of a "*set of objects, and the describable relationships among them*" represented in a declarative formalism [15]. Uschold and Grüninger [46] argued later that semantic interoperability between parties that want to 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 [14] 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.

[*] Corresponding author. E-mail: axel.polleres@wu.ac.at.

[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.

Many ontology-based approaches that address these causes of semantic heterogeneity have been proposed since [36, 50]. 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. 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 [17], we can largely distinguish four different levels of abstraction in ontology design as shown in Figure 1.

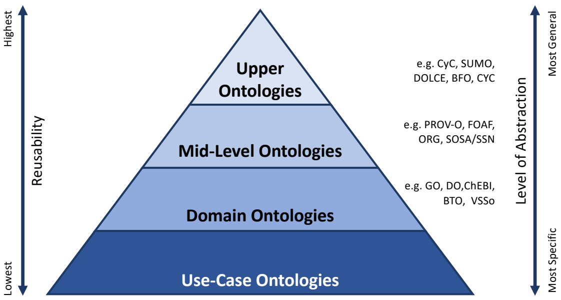

Fig. 1. Levels of Abstraction in Ontology Design

1. **Upper ontologies** that define very general terms that are common across all knowledge domains. Examples of such upper ontologies are CYC [29], SUMO [33], DOLCE [13] and BFO [44].
2. **Mid-level ontologies or Top domain ontologies** act as a bridge between the abstract content of an upper ontology and the richer detail of various domain ontologies. PROV-O [28], FOAF [4], ORG [41], SOSA/SSN [21] can be considered examples of such mid-level ontologies that define concepts generally enough so that their semantics can be further narrowed by a domain ontology or a use case 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 [1], the Disease Ontology [42], ChEBI [11], the Building Topology Ontology (BTO) [40] or VSSo [27], 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.
   Domain ontologies often also contain named individuals, for example, for a currency ontology to be useful it should include all currencies in use globally, their subunits and abbreviations as named individuals.
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. A recent exception is the proposal by Gangemi and Paulheim to use DOLCE to detect inconsistencies within DBpedia [37]. In fact, in an analysis of links in the Linked Open Data cloud [19] we have discovered that the two main open-source upper ontologies other than BFO, DOLCE and SUMO, are not used in any of 430 Linked Open Datasets that were investigated for the study. 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 result from importing the upper ontology in a mid-level or domain ontology. In the redesign of the SSN ontology, for example, the working group decided based on community feedback, 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 [21]. 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 [12] 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.

Still, beyond the lack of uptake of upper ontologies several other issues with the reuse of ontologies on the Web have been identified, which we discuss hereafter.

***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 [5, 19] that ontologies have higher availability than Linked Data instance 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 still to 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 examples, was confronted with problematic incidents like malware bundling, and changing service ownership in the past, raising doubts on 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 [35], there has been a lack of a general-purpose ontology search engine or a central ontology library [8], beyond the relatively recently proposed Linked Open Vocabulary repository [47]. 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 operators, 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 [16]. 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 have found that term reuse from existing ontologies is not widespread (most ontologies reuse less than 5% of their terms) [25], 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.

***Versioning**:* In their seminal work on ontology versioning Klein and Fensel [26] 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 two 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 rec-

[2] see http://ontologydesignpatterns.org
[3] https://code.google.com/archive/

[4] For instance, we used services like Sindice [45] and SWSE [23] in the past for auto-completion of ontology term search in Drupal [6].

ommend to change the namespace URI only in cases where the conceptualisation of the ontology changes.

Ideally, every ontology should follow the guidelines proposed in Klein and Fensel [26] in combination with more recent guidelines around content negotiation [24] and use version numbers for changes in the conceptualisation of the ontology in combination with a persistent URI that redirects to the most recent version of the ontology [30]. Another possible approach to versioning is to use the Memento protocol [10], 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 [4] 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 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 [18] or the DBpedia ontology [2] do not adhere to the guidelines proposed in [26] and silently update the semantics of terms. Only very recent ontologies standardised in the W3C, the Time Ontology [7] and SSN/SOSA [20], make the relation to terms in the previous version of the ontology explicit, 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 ontology namespace URI (i.e. http://www.w3.org/ns/ssn/), while no versioned URI is linked from that new namespace.

Modularisation: 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 [39]. 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 modularisation exist [9], either at design time by the ontology designers themselves using several ontology namespace URIs for the ontology modules (e.g. DOLCE [13] has been redesigned to be available in modules), or at reuse time through segmentation [43] or traversal view extraction [34]. Very few ontologies besides DOCLE and SSN/SOSA use a modularisation architecture, while even the large DBpedia ontology is published in just one big file using one namespace (i.e. http://dbpedia.org/ontology/).

Quality: Beyond syntactic and semantic errors that can be checked by reasoners, the notion of the quality of an ontology is rather imprecise. Some even argue that ontologies on the Web do not need to be computed in a sound and complete way, and systems should be able to deal with noise, different perspectives, and uncertainty [22]. In his dissertation, Vrandeĉić [48] 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 exists [38], though, that tests 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, a skill that itself requires considerable experience with engineering ontologies.

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 [28], ORG [41], SSN/-SOSA [21]), have a long history of availability and community support (e.g. FOAF [4], SIOC [3]) 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 [18]) 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 [31].

3. The Present and the Future

The success story of schema.org [18] as an ontology with very lightweight semantics, 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 [32]. Revisiting the above challenges, let us briefly discuss if and how schema.org addresses these.

- **Availability**: While neither the schema.org ontology itself is hosted by a publicly-funded open-access repositories such as Zenodo 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 companies, which implies high availability and support for the ontology at least for the foreseeable future.
- **Discoverability**: Although the schema.org vocabulary 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 [20]), 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 science, governmental or industrial applications.
- **Versioning**: 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 of the term itself and the class or property URI is just servicing the new semantics of the term.
- **Modularisation**: Schema.org is fully modularised with each term in the ontology being served by its own webpage and through using a Linked Data Content Negotiation technique also by providing its own graph at the same URI.
- **Quality**: While an ontology like schema.org that is constantly evolving may not always be consistent or free of dispute, 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 in the Western world there is little doubt that users (will) trust schema.org.

Although we believe that schema.org will continue to evolve and we will see an even bigger uptake of it on the general commercial Web, we believe it is not yet the end-all ontology on the Web. 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 and IoT domain) any time soon. In regards to its *Quality*, while the lightweight semantics were deliberately chosen to make annotations on the Web easier for the average Web developer [18], 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

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

[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.

extensions are managed through an open community process, the decision on additions to the ontology still sits with the providers of the ontology, i.e. the search engine companies.

The success of schema.org and the Open Graph protocol on the general Web [31], 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. In fact, we argue that we could be "better off with just one ontology on the Web" that addresses the issues identified above.

And we believe we are seeing the emergence of one such ontology, the ontology underlying Wikidata [49], the "Wikipedia for data" project that manages the factual information of the popular online encyclopedia. The data and the ontology underlying Wikidata is built bottom-up by the contributors to Wikidata, while also striving to incorporate the source of a term. This means, for the ontology part, it reuses existing ontologies where possible, but mints URIs for terms in the Wikidata namespace. It already meets most of the requirements that we outlined above. Similar to Wikipedia, it is highly available and terms in the ontology are easily discoverable, it has already more terms than any other ontology other than the DBpedia ontology, and it will enjoy similar trust as Wikipedia. However, at the moment, the ontology itself can not be transparently retrieved at its namespace URI and while it is modularised into wiki pages for each term, they are not retrievable through a Linked Data Content Negotiation technique[7] nor can the ontology itself, to the best of our knowledge, be downloaded from a single source. While 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 conceptualisation of a term, no new URI is minted that includes a reference to the old version of that term nor are the terms accessible through a Memento API. However, its greatest strength is the bottom-up engineering process with a strong moderation process similar to Wikipedia that allows terms in the ontology to be added and if necessary, be changed.

[7]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

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 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; and while it still has issues with its modularisation and access, and only partially addresses the ontology versioning problem through metadata annotations (but not versioned URIs), 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.

References

[1] M. Ashburner, C. Ball, J. Blake, D. Botstein, H. Butler, J. Cherry, A. P. Davis, K. Dolinski, S. Dwight, J. Eppig, M. Harris, D. Hill, L. Issel-Tarver, A. Kasarskis, S. Lewis, J. Matese, J. Richardson, M. Ringwald, G. Rubin, and G. Sherlock. Gene Ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nature Genetics*, 25:25–29, 05 2000.

[2] C. Bizer, J. Lehmann, G. Kobilarov, S. Auer, C. Becker, R. Cyganiak, and S. Hellmann. Dbpedia - a crystallization point for the web of data. *Journal of Web Semantics*, 7(3):154–165, 2009.

[3] J. G. Breslin, S. Decker, A. Harth, and U. Bojars. SIOC: An approach to connect web-based communities. *Int. J. Web Based Communities*, 2, 2006.

[4] D. Brickley and L. Miller. FOAF Vocabulary Specification 0.99. Technical report, http://xmlns.com/foaf/spec/, 2014.

[5] A. S. Butt, A. Haller, and L. Xie. Ontology search: An empirical evaluation. In *Proceedings of the International Semantic Web Conference*, pages 130–147, Riva del Garda, Italy, 2014. Springer International Publishing.

[6] S. Corlosquet, R. Delbru, T. Clark, A. Polleres, and S. Decker. Produce and consume linked data with drupal! In A. Bernstein, D. R. Karger, T. Heath, L. Feigenbaum, D. Maynard, E. Motta, and K. Thirunarayan, editors, *Proceedings of the 8th International Semantic Web Conference (ISWC 2009)*, volume 5823, pages 763–778, Washington DC, USA, Oct. 2009. Springer.

[7] S. J. D. Cox and C. Little. Time ontology in OWL. W3C Recommendation, World Wide Web Consortium, Oct. 2017. https://www.w3.org/TR/owl-time/.

[8] M. d'Aquin and N. F. Noy. Review: Where to publish and find ontologies? a survey of ontology libraries. *Journal of Web Semantics*, 11:96–111, Mar. 2012.

[9] M. d'Aquin, A. Schlicht, H. Stuckenschmidt, and M. Sabou. *Criteria and Evaluation for Ontology Modularization Techniques*, pages 67–89. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009.

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

[11] K. Degtyarenko, P. de Matos, M. Ennis, J. Hastings, M. Zbinden, A. McNaught, R. AlcÃąntara, M. Darsow, M. Guedj, and M. Ashburner. Chebi: A database and ontology for chemical entities of biological interest. *Nucleic acids research*, 36:D344–50, 02 2008.

[12] A. Gangemi. Ontology design patterns for semantic web content. In Y. Gil, E. Motta, V. R. Benjamins, and M. A. Musen, editors, *Proceedings of the International Semantic Web Conference*, pages 262–276, Galway, Ireland, 2005. Springer Berlin Heidelberg.

[13] A. Gangemi, N. Guarino, C. Masolo, A. Oltramari, and L. Schneider. Sweetening Ontologies with DOLCE. In *Knowledge Engineering and Knowledge Management: Ontologies and the Semantic Web*, pages 166–181, Berlin, Heidelberg, 2002. Springer Berlin Heidelberg.

[14] C. H. Goh. *Representing and Reasoning about Semantic Conflicts in Heterogeneous Information Systems*. PhD thesis, Sloan School of Management, MIT, Boston, MA, USA, 1 1997.

[15] T. Gruber. A translation approach to portable ontology specifications. *Knowledge Acquisition*, 5:199–220, 1993.

[16] M. Grüninger and M. Fox. Methodology for the design and evaluation of ontologies. In *Proc. of the Workshop on Basic Ontological Issues in Knowledge Sharing*, 07 1995.

[17] N. Guarino. Semantic matching: Formal ontological distinctions for information organization, extraction, and integration. In M. T. Pazienza, editor, *Information Extraction A Multidisciplinary Approach to an Emerging Information Technology*, pages 139–170, Berlin, Heidelberg, 1997. Springer Berlin Heidelberg.

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

[19] A. Haller, J. D. Fernandez, A. Polleres, and M. R. Kamdar. Deliverable 4.2: Data Integration (v.1). Technical report, Vienna, Austria, 2019.

[20] A. Haller, K. Janowicz, S. J. D. Cox, D. Le Phuoc, K. Taylor, and M. Lefrançois. Semantic Sensor Network Ontology. W3C Recommendation, World Wide Web Consortium, Oct. 19 2017. https://www.w3.org/TR/vocab-ssn/.

[21] A. Haller, K. Janowicz, S. J. D. Cox, M. Lefrançois, K. Taylor, D. L. Phuoc, J. Lieberman, R. García-Castro, R. Atkinson, and C. 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.

[22] P. Hitzler and F. van Harmelen. A reasonable semantic web. *Semantic Web*, 1(1,2):39–44, Apr. 2010.

[23] A. Hogan, A. Harth, J. Umbrich, S. Kinsella, A. Polleres, and S. Decker. Searching and browsing linked data with swse: The semantic web search engine. *Journal of Web Semantics*, 9(4):365–401, 2011.

[24] B. Hyland, G. Atemezing, and B. VillazÃşn-Terrazas. Best Practices for Publishing Linked Data. http://www.w3.org/tr/ld-bp/, W3C, 1 2014.

[25] M. Kamdar, T. Tudorache, and M. Musen. A systematic analysis of term reuse and term overlap across biomedical ontologies. *Semantic Web âĂŞ Interoperability, Usability, Applicability*, 1:1–5, 02 2016.

[26] M. Klein and D. Fensel. Ontology versioning on the semantic web. In *Proceedings of the First International Conference on Semantic Web Working*, SWWS'01, pages 75–91, Stanford University, CA, USA, 2001. CEUR-WS.org.

[27] B. Klotz, R. Troncy, D. Wilms, and C. Bonnet. Vsso - a vehicle signal and attribute ontology. In *Proceedings of the 9th International Semantic Sensor Networks Workshop co-located with 17th International Semantic Web Conference (ISWC 2018)*, pages 56–63, Monterey, CA, USA, 10 2018.

[28] T. Lebo, S. Sahoo, and D. McGuinness. PROV-O: The PROV Ontology. Technical report, W3C Recommendation 30 April, 2013.

[29] D. B. Lenat. Cyc: A large-scale investment in knowledge infrastructure. *Communications of the ACM*, 38(11):33–38, Nov. 1995.

[30] B. F. Lóscio, C. Burle, and N. Calegari. Data on the Web Best Practices. Technical report, W3C Recommendation 31 January, 2017.

[31] R. Meusel, P. Petrovski, and C. Bizer. The webdatacommons microdata, rdfa and microformat dataset series. In *Proceedings of the 13th International Semantic Web Conference: Replication, Benchmark, Data and Software Track*, Riva del Garda, Italy, 2014. Springer Berlin Heidelberg.

[32] P. Mika. On schema.org and why it matters for the web. *IEEE Internet Computing*, 19(4):52–55, July 2015.

[33] I. Niles and A. Pease. Towards a standard upper ontology. In *Proceedings of the International Conference on Formal Ontology in Information Systems - Volume 2001*, pages 2–9, New York, NY, USA, 2001. ACM.

[34] N. Noy and M. Musen. Specifying ontology views by traversal. In *Proceedings of the 3rd International Conference on the Semantic Web*, pages 713–725, 11 2004.

[35] N. Noy, N. Shah, P. L Whetzel, B. Dai, M. Dorf, N. Griffith, C. Jonquet, D. Rubin, M.-A. D Storey, C. Chute, and M. Musen. Bioportal: Ontologies and integrated data resources at the click of a mouse. *Nucleic acids research*, 37:W170–3, 06 2009.

[36] N. F. Noy. Semantic integration: A survey of ontology-based approaches. *SIGMOD Rec.*, 33(4):65–70, Dec. 2004.

[37] H. Paulheim and A. Gangemi. Serving dbpedia with dolce – more than just adding a cherry on top. In *Proceedings of the International Semantic Web Conference (ISWC)*, pages 180–196, Bethlehem, PA, USA,, 2015. Springer.

[38] M. Poveda-Villalón, A. Gómez-Pérez, and M. C. Suárez-Figueroa. OOPS! (ontology pitfall scanner!): An on-line tool for ontology evaluation. *International Journal on Semantic Web and Information Systems (IJSWIS)*, 10, 2014.

[39] M. Poveda Villalón, M. C. Suárez-Figueroa, and A. Gómez-Pérez. The landscape of ontology reuse in linked data. In *Proceedings of the First International Workshop on Ontology Engineering in a Data-driven World (OEDW 2012)*, Galway, Ireland, 2012.

[40] M. H. Rasmussen, P. Pauwels, M. Lefrançois, and G. F. Schneider. Building topology ontology. Draft Community Group Report, World Wide Web Consortium, Jan. 2019. https://w3c-lbd-cg.github.io/bot/.

[41] D. Reynolds. The Organization Ontology. Technical report, W3C Recommendation 16 January, 2014.

[42] L. M. Schriml, C. Arze, S. Nadendla, Y.-W. W. Chang, M. Mazaitis, V. Felix, G. Feng, and W. A. Kibbe. Disease Ontology: a backbone for disease semantic integration. *Nucleic Acids Research*, 40(D1):D940–D946, 11 2011.

[43] J. Seidenberg and A. Rector. Web ontology segmentation: Analysis, classification and use. In *Proceedings of the 15th International Conference on World Wide Web*, pages 13–22, 01 2006.

[44] B. Smith. Basic formal ontology. Technical report, University of Leipzig, Germany, 2003.

[45] G. Tummarello, R. Delbru, and E. Oren. Sindice.com: Weaving the open linked data. pages 552–565, Busan, South Korea, 01 2007.

[46] M. Uschold and M. Gruninger. Ontologies: Principles, methods and applications. *Knowledge Engineering Review*, 11:93–136, 1996.

[47] P.-Y. Vandenbussche, G. Atemezing, M. Poveda-Villalón, and B. Vatant. Linked open vocabularies (lov): A gateway to reusable semantic vocabularies on the web. 8:437–452, 01 2017.

[48] D. Vrandeĉić. *Ontology Evaluation*. PhD thesis, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany, 6 2010.

[49] D. Vrandeĉić and M. Krötzsch. Wikidata: A free collaborative knowledgebase. *Communications of the ACM*, 57(10):78–85, 2014.

[50] H. Wache, T. Vögele, U. Visser, H. Stuckenschmidt, G. Schuster, H. Neumann, and S. Hübner. Ontology-based integration of information — a survey of existing approaches. In *IJCAI-01 Workshop: Ontologies and Information*, pages 108–117, 01 2001.