

The Ecology of Ontologies in the Public Domain

Christopher J. O. Baker¹, Robert H. Warren², Volker Haarslev³, Greg Butler³

¹Data Mining Department,
Institute for Infocomm Research, Singapore

²School of Computer Science,
University of Waterloo, Waterloo, Ontario, Canada

³Department of Computer Science and Software Engineering,
Concordia University, Montreal, Quebec, Canada

1 Introduction

Since the announcement of the creation of a Web Ontology Working Group in 2001, two standards, DAML+OIL[1] and its successor, the Web Ontology Language (OWL) [2] were introduced. The latter is now the agreed-upon standard for formally specifying knowledge in the web and its introduction has renewed the focus on ontologies. Categories of researchers interested in ontologies include philosophers, software developers, computational linguists, biologists and logicians, and their motives are varied. Much debate now focuses on the criteria by which ontologies should be evaluated [3] and this debate is often driven by the question regarding the usefulness of ontologies for applications such as support for natural language processing [4] text mining [5], annotation of genes for gene expression analysis [6], inference or knowledge discovery in general, the facilitation of data integration [7], providing cornerstones of the Semantic Web [8], and as educational resources [9] There is however general agreement that Ontologies can serve as portable and easily exploitable conceptualizations for use in a multitude of computational applications. This has lead to an increase in the creation of ontologies and the emergence of a new generation of computational ontologists.

In order to characterize the proliferation of ontologies in the public domain we conducted a study to assess and review the general characteristics of existing OWL ontologies paying attention to features such as their abundance, continuing availability, originating sources and domain covered. Our long-term goal is to evaluate ontologies in light of their suitability for reasoning that yields non-obvious insight or new knowledge in the corresponding domain. The next section discusses criteria and related work for ontology evaluation. We then report on the findings from our survey of ontologies in the public domain and on their domain dependent and domain independent features. The paper concludes with a summary and an outlook on future research.

1.1 Ontology Evaluation

Criteria for ontology evaluation can be divided into domain *dependent* and domain *independent* criteria. Domain dependent criteria focus on the content of a given ontology. This is a criterion fundamental to the Semantic Web which is based on the vision of ontologies being made available across the World Wide Web. In this case alignment and integration of domain conceptualizations in areas where ontologies overlap is a critical step. Domain content is also the primary criterion when considering a given ontology for a specific application or for evaluating the on-

tology for reuse in a new knowledge engineering project. Though some existing tools simplify the identification of domain content within an ontology, such approaches require significant contributions of ontology meta-data from the developer [10] though some approaches [11] automatically tag concepts with appropriate meta-properties. In many cases ontologies are poorly characterized with respect to domain content, and require much user interaction using ontology visualization tools to produce a thorough assessment. Furthermore, the rate at which ontologies are being self-published makes a thorough evaluation of domain coverage and content particularly difficult.

Ontology content is one of many criteria that ontology users are interested in. After having identified an ontology with relevant and sufficient domain content, a user may consider applying this ontology for a specific task. For instance, in the area of natural language processing (NLP) ontologies are often used as hierarchically controlled vocabularies [4]. In such applications a series of metrics are commonly derived which seek to quantify the success of extraction of structured information content from corpora of unstructured text. Tools also exist which seek to measure the compatibility of an ontology's vocabulary with that of the text corpus under examination. For instance, the approach reported in [12] uses NLP techniques to compare the content of ontologies but does not consider their conceptual structure.

Beyond the content and application features of ontologies, the philosophical correctness of the conceptualization of knowledge in ontologies is an important consideration. Philosophical rules of thumb should not be violated, though frequently they are, especially the rules of univocity, positivity, objectivity, single inheritance, exhaustiveness, and intelligibility [13]. These rules provide important guidelines to ensure that conceptualizations have clarity and avoid ambiguity. Such criteria are however difficult to evaluate and enforce computationally, since they require a thorough understanding of the domain content and the axioms and relationships being employed in a conceptualization.

Taxonomic structures of formal ontological representations are also the subject of evaluation approaches [14]. Notions of rigidity, unity, identity, and dependence have been employed to provide useful controls on the correctness of hierarchies – thus a child (subtype) term in an ontology should not have identity criteria incompatible with those of its parent term. Formal ontologies such as those written in OWL can be further evaluated with respect to OWL tags used to build the conceptualization. Such investigations provide insight into the logic architecture underlying the conceptualization. To do this specific tags, such as those for declaring subclass relationships, can be used as inputs to Description Logic (DL) reasoners [15] in order to derive explicit or implicit knowledge. This can include knowledge about the internal correctness of an ontology and the presence of gaps. For further reference, a comprehensive study of ontology evaluation methodology is provided by [16].

Metrics of the afore mentioned evaluation criteria are of interest to ontology engineers and the Semantic Web community at large. To facilitate coordinated access to such metrics a number of initiatives have focussed on establishing resources and algorithms for searching, ranking and classifying ontologies. The OntoSelect library [17] is a collection of ontologies obtained by monitoring and trawling the web for new ontologies available in standard RDFS, DAML and OWL formats. The ontologies are marked up for human users according to a series of usability criteria. OntoKhoj [18] is a similar portal providing access to both human and computational agents to retrieve ontology classification information. OntoKohj classifies ontologies according to domain content using a naïve Bayesian approach. Swoogle's search services such as the *Ontology Dictionary* offer algorithms for ranking the importance, in terms of frequency and communal acceptance, of Semantic Web components at three levels of granularity: documents, terms and RDF graphs [19]. AKTiveRank [20] is a prototype system for ranking ontologies based on the analysis of their structures using a series of combined metrics. These resources

make it easier for knowledge engineers to access the ontologies and Semantic Web resources relevant to their needs.

2 Ontology Survey

One of the questions raised in most ontology projects concerns usefulness. Specifically, are ontologies valuable research tools, or are they an academic curiosity that cannot support tangible productive work? To address this multi-faceted issue, we first estimated ontology adoption by enumerating ontologies publicly available on the internet and assessing the rate at which they are modified or updated. We used this as a metric to indicate the level of acceptance or adoption of OWL ontologies as vehicles for knowledge representation. Additionally we identified the originating sources of these ontologies according to the institutions that published them and used this to illustrate the extent to which public organizations have endorsed ontology use. We then sought to classify ontologies according to domain content in order to identify the intended application domains and scale of their conceptualizations by enumerating the defined entities in the ontologies. Our focus then switched to the structures of these ontologies and specifically to the deployment of OWL tags. By identifying tag profiles within ontologies we were able to determine the extent to which support for DL-based reasoning has been incorporated into ontologies. Finally we reviewed features of ontologies that can be enumerated using DL reasoning tools, such as the numbers of concepts, relations, average number of child classes and occurrences of multiple-inheritance. This specifically pertains to OWL-DL (decidable First Order Logic) and OWL Lite (not expressive) but not to OWL full (undecidable). Note: the range of knowledge that can be captured in the language (i.e. its expressiveness) limits the means to reason with the captured knowledge (i.e. its decidability).

2.1 Ontology Adoption

We set out to create a data set of publicly available ontologies and study their contents and evolution over time. The ontologies were found using Google queries and a free-running web crawler [21] that searched over 2 million web pages on a daily basis. Figure 1 plots the number of ontologies observed to be available over time on the World Wide Web. The first official release of the DAML+OIL ontology standard was in December 2001, while the OWL standard was released in February 2004. Interestingly, the rise in the number of ontologies available is similar to a power curve; a behavior that is consistent with the adoption of most Internet technologies. One might ask whether the rate of increase will be sustained and whether the ontologies created are small experimental ‘me too’ ventures or robust applications of the OWL technology. One way to resolve this is to review the rate of ontology change and update.

To assess this we retrieved and stored OWL and DAML+OIL ontologies along with a time-stamp. Each day, we re-visited these ontologies and noted any changes, including suspension of availability. In each case, we attempted to query the Internet Way-Back Archive [22] for previously archived copies of the ontologies. When available, this data gave us two interesting pieces of information: the earliest point in time when the ontology was made available and any observed changes made to the ontologies. Hence we were capable, not only of tracking the current activity on selected ontologies, but also of looking backwards in time for changes that have since been made over time.

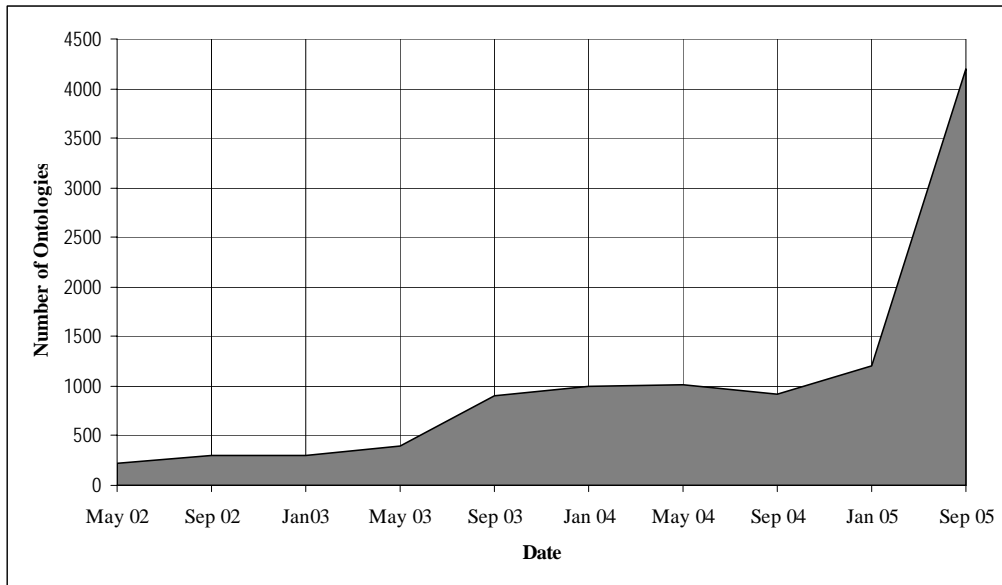


Fig. 1. Number of ontologies available on the internet

Figure 2 plots the number of ontologies which changed each month over a time period of 4 years from September 2002 onwards. The figure must be interpreted carefully as a small bias exists in the graph due to the creation of new ontologies. Also change data for post-January 2005 period was not available. An analysis of the graph reveals that ontological development is undergoing considerable expansion. On average 200 ontologies were updated each month. It is not possible to establish if this is a reflection of their use in a production setting or in ongoing experimentation. It is, however, an indication that not all of the available ontologies are stale, static remnants of a technological fad. It is to be expected that many more ontologies are available, and frequently updated, either commercially or internally within organizations. We concern ourselves only with ontologies available on the Internet, therefore our data set may not be a completely representative sample of all ontologies in actual use.

Overall we found references to over 5,100 ontologies, 60% of which were in OWL format. A similar magnitude of ontologies was reported by [19], who found 4000 Semantic Web Ontologies (SWO). We selected a set of 1,600 ontologies for further analysis based on their relative complexity and an encoding which allowed ontologies to function with all of our tools and software libraries without error. A similar position was adopted by other researchers selecting DAML ontologies for group analysis [23] These authors found coding errors in 50% of ontologies and were able to use 38% of their selected ontologies for subsequent evaluation and analysis.

2.2 Accessing Ontology Content

OWL files were parsed for Universal Resource Identifier tags [24] which were then deposited in a relational database, allowing for simple keyword searches and a series of ontology evaluations. As a means of exploring the content of the available ontologies, we developed a simple query web interface [25] that allowed us to search for strings, i.e. ordered sequences of sym-

bols, letters or words within the ontologies. In this way, we were able to query all retrieved ontologies for similar class names, irrespective of the structure of the ontologies. As a result it became possible to search for commonality between ontologies and potential semantic junctions for integration between ontologies. Ontology alignment and matching are important research areas [26] with great significance to Semantic Web initiatives. Our analysis involved simple clustering experiments based on the labels assigned to properties, classes and individuals. For example, the term ‘service_description’ occurred as a unique concept in 11 different ontologies, as determined from the ontology name and its publishing source. Other terms such as ‘enzyme’ found in 14 different ontologies, occurred in 249 entities but not as simple concept name, instead in entities which contained some usage of the word e.g. Enzyme Purification or Acetyl co-enzyme_A. The results showed that while there are a number of distinct integration opportunities, in the great majority of cases they involve very specific portions of ontologies, primarily those which described web-based services and transaction processes. This clearly suggests that we are still in the ‘early adoption’ phase of this technology and not yet able to expect ‘global’ interoperability of ontologies in any one domain. A systematic and coordinated approach is necessary to focus and coordinate the interoperability of ontologies, particularly if important information systems which are critical in our daily lives begin to depend on ontology compatibility.

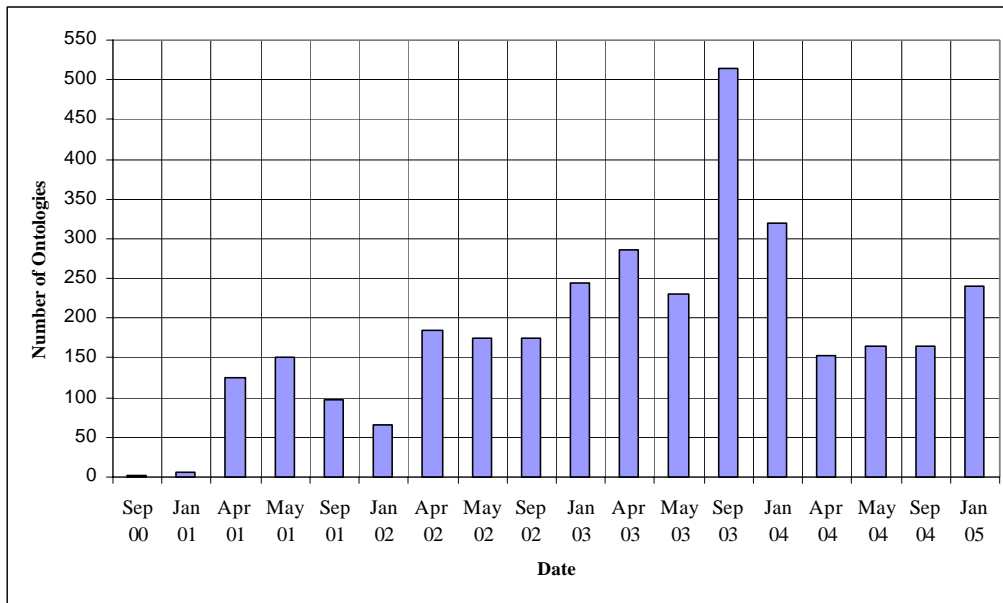


Fig 2. Number of ontologies observed to have been updated or changed

2.3 Ontology Contributors

We also sought to review the contributors to ontology development. We classified the organizations publishing ontologies into several categories as depicted in (Figure 3). The categories were large and small corporations, government institutions, individuals, standards bodies, universities and non-academic research institutions. Predictably, universities and research establishments represent the major organizations publishing ontologies on the web. We thought it

interesting to add the ‘Individuals’ category to illustrate the fact that a number of ontologies are being published by people who are contributing out of personal interest. Government organizations were the most prolific. This is primarily due to several health related ontologies released by US government institutions.

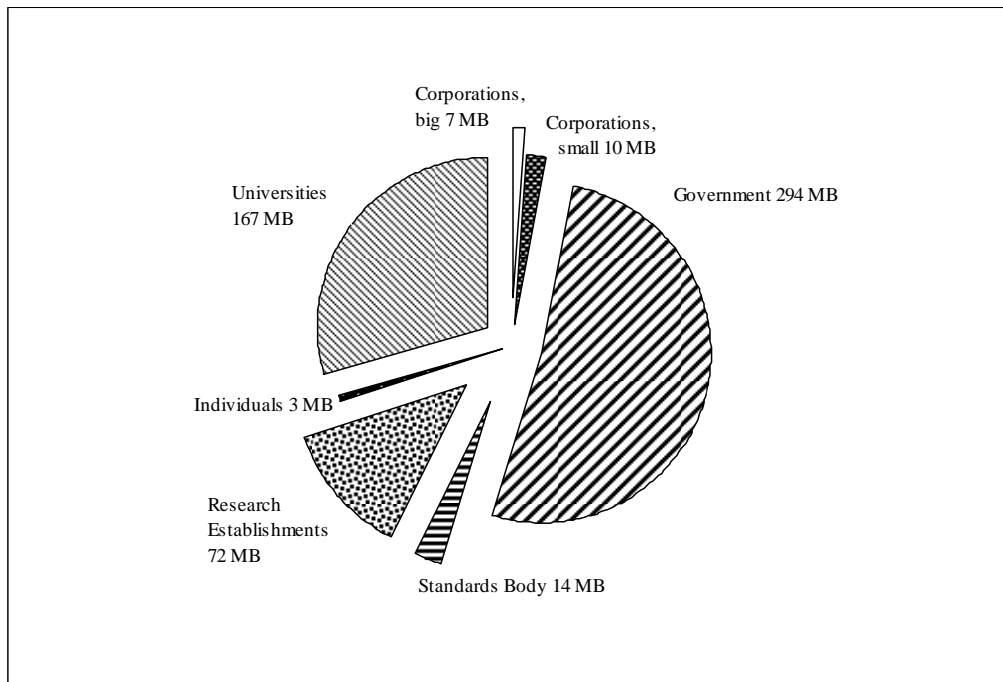


Fig. 3. Classification of ontology contributors according MB of data published

2.4 Domain Content

Using our query tool, we manually reviewed for content about 1,000 ontologies selected for size and adequate encoding. These were then hand-classified into an ad-hoc set of topics using the pooled entities in the ontologies, namely the labels on the concepts relations and instances (see Figure 4). For example, the inspection of several ontologies revealed the use of terms representing organizational roles, and associated titles and relationships that led us to create the “organization” category. The resulting classifications are not perfect, but they proved satisfactory for purposes of gross aggregation. We created a small generic category called ‘toy’ used to classify example ontologies created to demonstrate specific ontological constructs.

Again the categories with the most ontology coverage were those of web services and organizations, followed by biomedicine. These had a three-fold higher occurrence than the next distinguishable categories of Transportation and ontologies for ‘Toy’ domains (see Figure 4). The domain of biomedicine could be further subdivided into biology (40) cancer (5) and genetics (114). The largest of such ontologies include the OWL versions of the Gene Ontology. Within the classified categories the numbers of entities (classes, properties, and individuals) was enumerated. The result of our evaluation identified the domain of genetics as the domain

most richly represented with ontological entities. Clearly the domain of biomedicine is a significant source of ontology content. In the group of unclassified ontologies most of these ontologies had one or less parent (super) class and hence, a low number of entities, and it is this which made it impossible to classify them. There would appear to be a significant number of immature ontologies in the public domain, suggesting considerable low-level experimentation with OWL. This may reflect strong interest in the technology, albeit a preliminary step in the full-scale adoption of this knowledge representation vehicle.

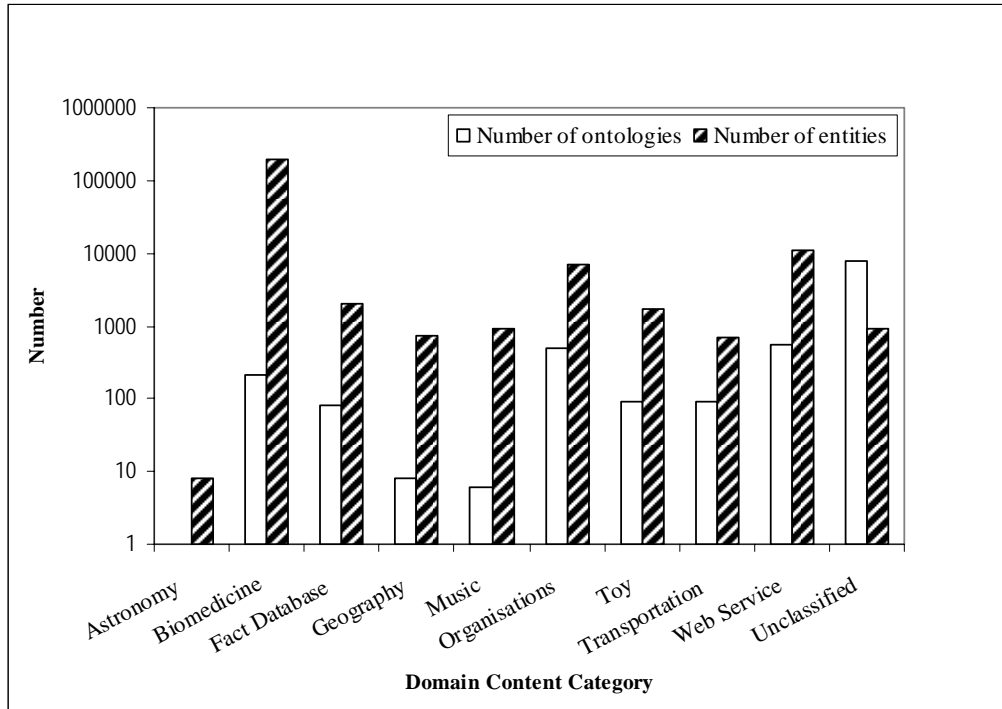


Fig. 4. Domain content classification of ontologies

2.5 OWL Deployment

Having established that OWL ontologies are indeed broadly accepted in the public domain by major institutions as knowledge representation vehicles which cover a variety of subject domains, we then considered domain independent features of these ontologies. We are interested in the degree to which the OWL language has been deployed. Specifically we wish to evaluate whether the OWL language elements in contemporary ontologies have captured ‘implicit knowledge’. We describe implicit knowledge as information represented in such a way that a reasoner can derive additional information beyond the explicit description. For example in the case of the concept ‘malignant melanoma’, this medical term can be described as equivalent to an aggregate of explicitly modeled OWL elements declaring a ‘cancer’ of the ‘skin’ that is an ‘aggressive and life-threatening manifestation’ etc. Such a conceptualization, along with declarations of the properties of, and constraints on, these concepts, facilitate reasoning tools to logically derive further non-obvious insights into the domain of the given ontology. Our ongoing

interest is to identify those sufficiently well-developed ontologies which could support reasoning.

Wolstencroft et al. [27] illustrate an example where, in conjunction with domain context, they use reasoning to enhance a discovery process. They used classification-based reasoning over OWL tags in the ‘Phosphatase ontology’, specifying necessary and sufficient conditions on membership for protein instances to belong to a protein class, primarily according to the specific phosphatase domain architecture of the protein instance. They also used closure axioms to limit the domains allowable for a specific class membership.

Furthermore we wished to determine whether in the context of the currently available ontologies, the expressiveness of DL-based reasoning tools and query languages is a limiting factor in knowledge mining from ontologies. To investigate this set of issues we analyzed more than 800 ontologies randomly gathered from the internet and counted the occurrence and frequency of several tags (called ‘DL-interesting’ in the following), which are candidates for being able to support the derivation of implicit knowledge, described above. Figure 5 shows a histogram of the OWL and RDFS tags we studied in 845 ontologies. RDF Schema tags used in the RDF knowledge representation language are also employed in OWL ontologies. Our results show the most common tags to be RDFS tags, namely the tags for declaring sub classes, relationship-axioms and hierarchies of relations. These tags also predate the introduction of OWL tags and are presumably more frequently occurring for this reason. The most frequently occurring OWL tags were those denoting cardinality and equivalence of classes. With almost equal usage constructs for declaring hierarchies of relations cardinality, class equivalence, intersection, disjointness, inversion and uniqueness. Tags declaring functional and transitive relations were used in fewer than 100 ontologies. Likewise inverse functional and symmetric relations were infrequently used.

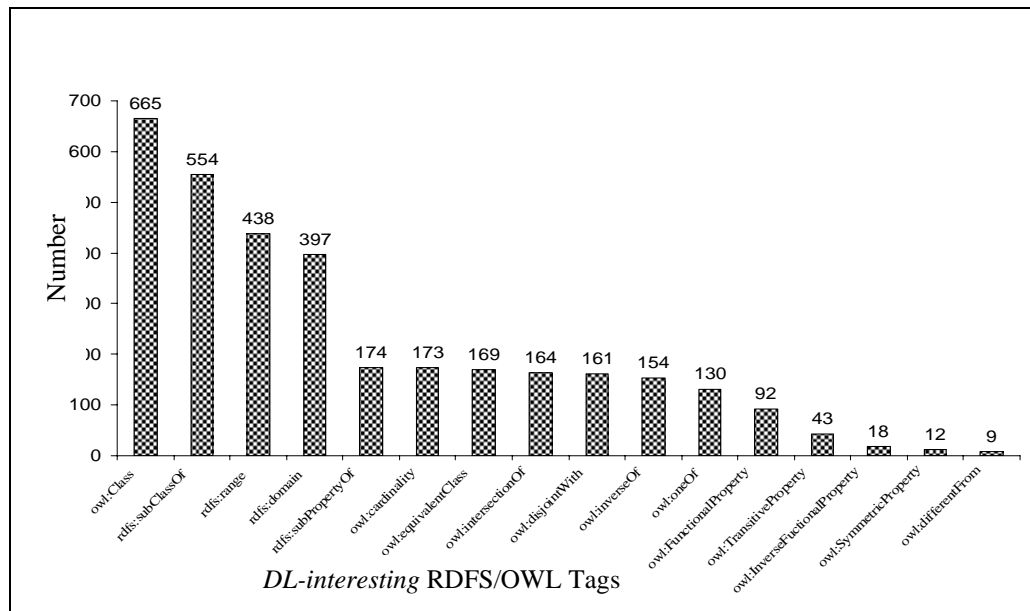


Fig. 5. Occurrence of OWL/RDFS tags in more than 800 ontologies

In order to reduce noise introduced into our analysis by the low level occurrence of tags we normalized our data using a similar approach to other authors [23,28]. The resulting analysis

revealed 63 non-redundant patterns of tag-usage [29] and their abundance. Clearly RDFS tags are found in the majority of ontologies. Table 1 [29] details 30 of the 63 usage patterns identified in our analysis. We comment briefly on the most heavily used patterns which serves to characterize the status quo with respect to the usage of DL-interesting tags in the modeling of domain knowledge.

Firstly the most deployed tag usage, the RDFS subclass tag alone - Pattern 32, results in simple hierarchical ‘*is a*’ ontologies. Secondly a significant number of ontologies contained none of the DL-interesting tags investigated in this study, Pattern 1. Most characteristic was the occurrence of the combination of tags `rdfs:subclass`, `rdfs:domain` and `rdfs:range`, Pattern 23, for specifying hierarchies and relationship-axioms. Furthermore they were the most significant tags across all usage-profiles occurring in some combination that included at least one of these three tags. OWL tags for representing cardinality and disjointness were used frequently in conjunction with the RDFS tags but in general OWL tags occurred in less than a third of the patterns. Taken together our one-time snapshot demonstrates the heavier usage of RDFS over OWL presumably because OWL was introduced more recently. Ontology Engineers in general are likely to be more accustomed to database modeling, which involves some use of cardinality, or object modeling using subclasses and axioms yet less with logical or mathematical notions such as inverse of functional relations. This would explain the less frequent deployment of OWL tags.

Our current assessment is that implicit knowledge recorded in OWL tags is relatively sparse in most ontologies and that reasoning over the majority of conceptualisations is limited to the use of subsumption taxonomies and axioms. This level of complexity is trivial for a reasoner and we came across only 9 ontologies with 6 or more different *DL-interesting* tags. Few if any of the ontologies were currently mature enough to challenge the capabilities of existing reasoning tools like RACER, suggesting we are well-tooled to interrogate OWL knowledgebases but that in general the community’s skills in knowledge representation to support new discoveries through the disclosure of implicit knowledge remains immature. This appears somewhat paradoxical, however our analysis-snapshot did not ascertain the speed at which OWL is being adopted and the time period of our analysis was curtailed. Moreover, a growing interest in ontology engineering is fuelled by the availability of industrial strength commercially available editors [30] and a widespread concern about the quality of ontologies has emerged [3]. In the last two years the Semantic Web paradigm has received considerable fanfare and established itself in on the agenda of many scientific communities and permeated research work presented at their top conferences. This is particularly true in the life science domain [31]. It is therefore realistic to anticipate in the near future the emergence of a larger subset of advanced ontologies which can be reasoned over for new insights.

2.6 DL-Derived Ontology Metrics

Ontologies have distinct compositions and structures which are relevant domain-independent criteria for ontology assessment. The OWL-DL and OWL Lite species of OWL make it possible to use description logics to rapidly access structural and organizational features of ontologies. The added value of using DL-based evaluation goes beyond the rapid enumeration of elemental components of ontologies, such as classes, properties and instances. It lies also in being able to compute and enumerate internal features of the ontology’s architecture. For our investigations we used the OWL reasoner [15] and its high performance query language nRQL [32] which is comprehensive enough for users to custom design new queries and evaluate ontologies for distinctive but relevant structural attributes such as multiple-inheritance. Multiple-inheritance, the presence in an ontology of terms with multiple parents, is sometimes considered the result of poor conceptualization of domain knowledge. It can also present obstacles to ontology integra-

tion, since in such cases `is_a` relations are permitted to mean different things in different contexts [13]. Consider: `dog is_a mammal`; `dog is_a pet`; `nurse is_a person`, `nurse is_a hospital role`.

Our DL-based analyses of more than 100 randomly selected OWL ontologies and almost 70 ontologies manually selected for enriched numbers of RDFS tags disclosed that on average the depth of most ontologies was relatively shallow, most were only 3-4 levels deep, though depths of up to 17 levels were encountered. On average each class had around 0.4 subclasses, suggesting relatively lean subclass hierarchies in most ontologies. Multiple-inheritance within class hierarchies (measured by the average number of parents) was only 0.76, although we noted significant reuse of parent classes in some ontologies. The average number of classes or concepts within ontologies was typically below 100, supported on average by a higher number of properties, commonly referred to as relations. Interestingly all selected ontologies made use of at least 18 properties but classes were instantiated by a very low average of 20 individuals per ontology. This result suggests that most ontologies are not yet designed to support large scale data mining or discovery initiatives.

3 Conclusion and Future Work

An era is emerging in which the need for evaluations and classifications of ontologies has been collectively recognized and this becomes all the more urgent where ontologies are being put forward to serve as robust knowledge representation components of the Semantic Web. While the criteria by which ontologies should be evaluated depend upon their role and purpose within an application domain, we have conducted evaluations of the ontologies available across the internet which is of quite general significance.

Firstly we observe a rapid increase in the numbers of ontologies available on the web since the publication of the Web Ontology Language (OWL) reference guidelines, though a lag phase existed before uptake was evident. We observe that changes are being made to ontologies at a relatively stable rate. This suggests active revision and maintenance by the ontology publishing community which is dominated by government and academic institutions conceptualising web service and biomedical related domains. Our studies also show that most OWL ontologies are constructed from relatively simple combinations of RDFS tags deployed in conjunction with select OWL tags.

A large portion of the results presented in this study were the product of considerable manual effort, and it would be challenging to conduct such evaluations regularly to monitor the evolution of the Semantic Web. This motivated us to take advantage of the accessibility of OWL ontologies to DL-based querying tools. These allowed us rapidly to assess the state of the art across a multitude of ontologies, highlighting elementary yet profound structural parameters relevant to ontology developers and the knowledge representation community at large. These investigations highlighted the relatively shallow depths of most publicly available ontologies, the relatively low but real existence of multiple-inheritance, and the relatively lean structure of most ontologies. We see that classes in ontologies are still associated with relatively small real world instance-data and that the ontologies are in general still quite small, with numbers of classes falling on average well below 100 per ontology.

We advocate the importance of metrics to monitor and assess the state of the art with respect to quantity and quality of ontologies. We see the need for further development of ontology evaluation tools and continued investigations of the range and appropriateness of criteria by which ontologies should be assessed. Particularly we urge for the uptake of DL-based approaches that use generic query algorithms to produce architectural information about ontologies. We foresee that the results of such analyses becoming significant in supporting integration of ontologies and thereby contributing the realization of the Semantic Web vision.

Acknowledgements

This work was financed by the Génome Québec funded project Ontologies, the semantic web and intelligent systems for genomics.

References

- [1] van Harmelen F., Patel-Schneider P.F. and Horrocks I. (Editors) DAML+OIL, Reference description of the DAML+OIL (March 2001) ontology markup language, <http://www.daml.org/2001/03/reference> (2001).
- [2] Bechhofer S., van Harmelen F., Hendler J., Horrocks I., McGuinness D. L., Patel-Schneider P. F. and Stein L. A. <http://www.w3.org/TR/owl-ref/>
- [3] Obrst L., Ceusters W., Mani I., Ray S., and Smith B., "The Evaluation of Ontologies: Toward Improved Semantic Interoperability," in *Semantic Web: Revolutionizing Knowledge Discovery in the Life Sciences*, C. J. O. Baker and K.-H. Cheung, Eds. Springer, 2007, Ch. 7, pp. 139–158.
- [4] Simon J., Dos Santos M., Fielding, J., and Smith, B., Int. Journal Med. Inform., Formal ontology for natural language processing and the integration of biomedical databases,(2005)
- [5] Doms A. and Schroeder M., Nucleic Acids Research, GoPubMed: Exploring PubMed with the Gene-Ontology, 33,(2005)
- [6] Doniger, S.W. and Salomonis, N. and Dahlquist, K.D. and Vranizan, K. and Lawlor, S.C. and Conklin, B.R., Genome Biology, 1, MAPPFinder: using Gene Ontology and GenMAPP to create a global gene-expression profile from microarray data, 4, 2003
- [7] Qi D. and Bult C.J. and Blake J.A., Kadin J.A., Richardson J.E., Ringwald M., Eppig J.T. and the Mouse Genome Informatics Group Computational Systems Bioinformatics, Aug. 8-11 2005, Stanford University, California, USA, Computational Systems Bioinformatics, Data Integration in the Mouse Genome Informatics (MGI) Database, 2005
- [8] Wolstencroft K. and Brass A. and Horrocks I. and Lord P., Sattler U., Stevens R. and Turi D., 4th International Semantic Web Conference (ISWC), November 6-10, 2005, Galway, Ireland, A Little Semantic Web Goes a Long Way in Biology 2005
- [9] Mitrovic A. and Devedzic V., Int. J. of Continuous Engineering Education and Life-long Learning (IJCEELL) 3, A Model of Multitutor Ontology-Based Learning Environments 14 2004
- [10] Lopez-Perez A., Facultad de Informatica, Universidad Politecnica de Madrid, Integracion de la aplicacion OntoMetric en la plataforma WebODE,(2004)
- [11] Voelker J. and Vrandečić D. and Sure Y., 4th International Semantic Web Conference (ISWC), November 6-10, 2005, Galway, Ireland, Automatic Evaluation of Ontologies AEON 2005
- [12] Maedche, A. and Staab, S., Proceedings of the European Conference on Knowledge Acquisition and Management EKAW-2002. Madrid, Spain, 251-263, Springer-Verlag Berlin, LNCS/LNAI, Measuring Similarity between Ontologies, 2473, (2002)
- [13] Smith, B. and Koehler, J. and Kumar A., Database Integration in the Life Sciences (DILS 2004), Rahm E. Springer-Verlag, Berlin, On the Application of Formal Principles to Life Science Data: A Case Study in the Gene Ontology (2004)
- [14] Guarino, N. and Welty, C., Communications of the ACM, 2, Evaluating Ontological Decisions with OntoClean, 45, (2002)
- [15] Haarslev V. and Moeller R., Proceedings of International Joint Conference on Automated Reasoning (IJCAR 2001), June, Siena, Italy, RACER System Description, (2001)
- [16] Hartmann, J., Spyns P., Giboin, A., Maynard D., Cuel R., Suarez-Figueroa, M. C. and Sure, Y., EU-IST Network of Excellence (NoE) IST-2004-507482 KWEB, Methods for ontology evaluation,(2005)
- [17] Buitelaar P., Eigner T., and Declerck T., OntoSelect: A Dynamic Ontology Library with Support for Ontology Selection. In Proceedings of the Demo Session at the International Semantic Web Conference. Hiroshima, Japan 2004.

- [18] Patel C., Supekar K., Lee Y., and Park E.K. OntoKhoj: A Semantic Web Portal for Ontology Searching, Ranking and Classification. In Proceedings of the Workshop on Web Information and Data Management, ACM 2003.
- [19] Ding L., Pan R., Finin T., Joshi A., Peng Y., and Kolari P. (2005) Finding and Ranking Knowledge on the Semantic Web. In: Proceedings of the 4th International Semantic Web Conference Date: November 07, 2005, LNCS 3729, p. 156 ff.
- [20] Alani H and Brewster C. EON2006, Evaluation of Ontologies for the Web 4th International EON Workshop May 22nd, 2006 Edinburgh International Conference Center, Edinburgh, United Kingdom
- [21] Cafarella M., Cutting D.: Building nutch: Open source search. Queue 2 (2)(2004) 54-61.
- [22] Internet Archive, Wayback Machine <http://www.archive.org/> Last accessed 22/5/2006
- [23] Tempich C and Volz R., Towards a Benchmark for Semantic We reasoners – An analysis of the DAML ontology library. 2nd International Workshop on Evaluation of Ontology-Based Tools (EON2003) at ISWC 2003 Sanibel Island 2003, Florida 2003.
- [24] <http://www.w3.org/Addressing/>
- [25] <http://www.cs.concordia.ca/FungalWeb/PublicOntologiesDatabase/>
- [26] Noy N., and Stuckenschmidt H. (2005) Ontology Alignment: An annotated Bibliography <http://drops.dagstuhl.de/opus/volltexte/2005/48/>
- [27] Wolstencroft K, Stevens R., and Haarslev V. Applying OWL Reasoning to Genomics: A Case Study. In: Baker C. J. O. and Cheung K.-H. Eds., Semantic Web: Revolutionizing Knowledge Discovery in the Life Sciences. Springer, 2007 ch. 7, pp. 139–158.
- [28] Baker C. J. O., Shaban-Nejad A., Xu S., Haarslev V. and Butler G. (2006), Semantic Web Infrastructure for Fungal Enzyme Biotechnologists, Journal of Web Semantics Special Edition on Semantics Web for the Life Sciences (2006).
- [29] <http://www.cs.concordia.ca/FungalWeb/EcologyTable1and2.html>
- [30] <http://www.topbraidcomposer.com/>
- [31] Baker C. J. O. and Cheung K.-H. Eds., Semantic Web: Revolutionizing Knowledge Discovery in the Life Sciences. Springer, 2007.
- [32] Wessel M. and Moeller R. , Proc. of the 2005 Description Logic Workshop (DL 2005), CEUR Electronic Workshop Proceedings, <http://ceur-ws.org/>, A High Performance Semantic Web Query Answering Engine, (2005).