

Using Aber-OWL for fast and scalable reasoning over BioPortal ontologies

Luke Slater^{1*}, Georgios V Gkoutos², Paul N Schofield³, Robert Hoehndorf¹

¹ Computational Bioscience Research Center, King Abdullah University of Science and Technology, 4700 KAUST, 23955-6900, Thuwal, Saudi Arabia

² Department of Computer Science, Aberystwyth University, Aberystwyth, SY23 3DB, Wales, United Kingdom

³ Department of Physiology, Development and Neuroscience, University of Cambridge, Downing Street, CB2 3EG, England, United Kingdom

ABSTRACT

Reasoning over biomedical ontologies using their OWL semantics has traditionally been a challenging task due to the high theoretical complexity of OWL-based automated reasoning. As a consequence, ontology repositories, as well as most other tools utilizing ontologies, either provide access to ontologies without use of automated reasoning, or limit the number of ontologies for which automated reasoning-based access is provided. We apply the Aber-OWL infrastructure to provide automated reasoning-based access to all accessible and consistent ontologies in BioPortal (368 ontologies). We perform an extensive performance evaluation to determine query times, both for queries of different complexity as well as for queries that are performed in parallel over the ontologies. We demonstrate that, with the exception of a few ontologies, even complex and parallel queries can now be answered in milliseconds, therefore allowing automated reasoning to be used on a large scale, to run in parallel, and with rapid response times.

1 INTRODUCTION

Major ontology repositories such as the BioPortal (Noy *et al.*, 2009), OntoBee (Xiang *et al.*, 2011), or the Ontology Lookup Service (Cote *et al.*, 2006), have existed for a number of years, and currently contain several hundred ontologies, enabling ontology creators and maintainers to publish their ontology releases and make them available to the wider community.

Besides the hosting functionality that such repositories offer, they usually also provide certain web-based features for browsing, comparing, visualising and processing ontologies. One particularly useful feature, currently missing from the major ontology repositories, is the ability to provide online access to reasoning services simultaneously over many ontologies. Such a feature would enable the use of semantics and deductive inference when processing data characterized with the ontologies these repositories contain (Hoehndorf *et al.*, 2015). Moreover, the ability to query multiple ontologies simultaneously further enables data integration across domains and data sources. For example, there is an increasing amount of RDF (Manola and Miller, 2004) data becoming available through public SPARQL (Seaborne and Prud'hommeaux, 2008) endpoints (Jupp *et al.*, 2014; The Uniprot Consortium, 2007; Belleau *et al.*, 2008; Williams *et al.*, 2012), which utilise multiple ontologies to annotate entities.

However, enabling automated reasoning over multiple ontologies is a challenging task since as automated reasoning can be highly complex and costly in terms of time and memory consumption (Tobies, 2000). In particular, ontologies formulated in the Web Ontology Language (OWL) (Grau *et al.*, 2008) can utilize statements based on highly expressive description logics (Horrocks *et al.*, 2000), and therefore queries that utilize automated reasoning cannot, in general, be guaranteed to finish in a reasonable amount of time.

Prior work on large-scale automated reasoning over biomedical ontologies has often focused on the set of ontologies in Bioportal, as it is one of the largest collections of ontologies freely available. To enable inferences over this set of ontologies, modularization techniques have been applied (Del Vescovo *et al.*, 2011) using the notion of locality-based modules, and demonstrated that, for most ontologies and applications, relatively small modules can be extracted over which queries can be answered more efficiently. Other work has focused on predicting the performance of reasoners when applied to the set of BioPortal ontologies (Sazonau *et al.*, 2013), and could demonstrate that performance of particular reasoners can reliably be predicted; at the same time, the authors have conducted an extensive evaluation of average *classification* times of each ontology.

Other approaches apply RDFS reasoning (Patel-Schneider *et al.*, 2004) for providing limited, yet fast, inference capabilities in answering queries over Bioportal's set of ontologies through a SPARQL interface (Salvadores *et al.*, 2012, 2013). Alternatively, systems such as OntoQuery (Tudose *et al.*, 2013) provide access to ontologies through automated reasoning but limit the number of ontologies.

The Aber-OWL (Hoehndorf *et al.*, 2015) system is a novel ontology repository that aims to allow access to multiple ontologies through automated reasoning utilizing the OWL semantics of the ontologies. Aber-OWL mitigates the complexity challenge by using a reasoner which supports only a subset of OWL (i.e., the OWL EL profile (Motik *et al.*, 2009)), ignoring ontology axioms and queries that do not fall within this subset. This enables the provision of polynomial-time reasoning, which is sufficiently fast for many practical uses even when applied to large ontologies. However, thus far, the Aber-OWL software is only applied to a few, manually selected, ontologies, and therefore does not have a similar coverage as other ontology repositories, nor does it cater for reasoning over large sets of ontologies such as the ones provided by the BioPortal ontology dataset (Bioportal contains, as of 9 March 2015, 428 ontologies consisting of 6,668,991 classes).

*To whom correspondence should be addressed: luke.slater@kaust.edu.sa

Here, we apply the Aber-OWL framework to reason over the majority of the available ontologies in Bioportal. We evaluate the performance of querying ontologies with Aber-OWL, utilizing 337 ontologies from BioPortal, we evaluate Aber-OWL's ability to perform different types of queries as well as its scalability in performing queries that are executed in parallel. We demonstrate that the Aber-OWL framework makes it possible to provide, at least, light-weight description logic reasoning over most of the freely accessible ontologies contained in BioPortal, with a relatively low memory footprint and high scalability in respect to the number of queries executed in parallel, using only a single medium-sized server as hardware to provide these services. Furthermore, we identify several ontologies for which querying using automated reasoning performs significantly worse than the majority of the other ontologies tested, and discuss potential explanations and solutions.

2 METHODS

2.1 Selection of ontologies

We selected all ontologies contained in BioPortal as candidate ontologies, and attempted to download the current versions of all the ontologies for which a download link was provided by BioPortal. A summary of the results is presented in Table 1.

Total	427
Loadable	368
Used	337
Unobtainable	39
Non-parseable	17
Inconsistent	3
No Labels	31

Table 1. Summary of Ontologies used in our test. The loadable ontologies are the ones obtained from BioPortal which could be parsed using the OWL API and which were found to be consistent when classified with the ELK reasoner. We exclude 31 ontologies that do not contain any labels from our analysis.

Out of 427 total ontologies listed by Bioportal, only 368 could be directly downloaded and processed by Aber-OWL. Reasons for failure to load ontologies include the absence of a download link for listed ontologies, proprietary access to ontologies or ontologies that are only available in proprietary data formats (e.g., some of the ontologies and vocabularies provided as part of the Unified Medical Language Systems (Bodenreider, 2004)). 39 ontologies were not obtainable. Furthermore, 17 ontologies that could be downloaded were not parseable with the OWL API, indicating a problem in the file format used to distribute the ontology. Three ontologies were inconsistent at the reasoning stage. Several ontologies also referred to unobtainable ontologies as imports; however, we included these ontologies in our analysis, utilizing only the classes and axioms that were accessible. As Aber-OWL currently relies on the use of labels to construct queries, we further removed 31 ontologies that did not include any labels from our test set.

Overall, we use set of 337 ontologies in our experiments consisting of 3,466,912 classes and 6,997,872 logical axioms (of which 12,721 are axioms involving relations, i.e., RBox axioms). In

comparison, BioPortal currently (9 March 2015) includes a total of 6,668,991 classes.

2.2 Use of the Aber-OWL reasoning infrastructure

Aber-OWL (Hoehndorf *et al.*, 2015) is an ontology repository and query service built on the OWLAPI (Horridge *et al.*, 2007) library, which allows access to a number of ontologies through automated reasoning. In particular, Aber-OWL allows users or software applications to query the loaded ontologies using Manchester OWL Syntax (Horridge *et al.*, 2006), using the class and property labels as short-form identifiers for classes. Aber-OWL exposes this functionality on the Internet through a JSON API as well as a web interface available on <http://aber-owl.net>. To answer queries, Aber-OWL utilizes the ELK reasoner (Kazakov *et al.*, 2014, 2011), a highly optimized reasoner that supports the OWL-EL profile. Ontologies which are not OWL-EL are automatically transmuted by the reasoner by means of ignoring all non-EL axioms, though as of 2013 50.7% of ontologies in Bioportal were natively using it (Matentzoglou *et al.*, 2013).

We extended the Aber-OWL framework to obtain a list of ontologies from the Bioportal repository, periodically checking for new ontologies as well as for new versions of existing ontologies. As a result, our testing version of Aber-OWL maintains a mirror of the accessible ontologies available in BioPortal. Furthermore, similarly to the functionality provided by BioPortal, a record of older versions of ontologies is kept within Aber-OWL, so that, in the future, the semantic difference between ontology versions could be computed.

In addition, we expanded the Aber-OWL software to count and provide statistics about:

- The ontologies which failed to load, with associated error messages;
- Axioms, axiom types, and number of classes per ontology; and
- Axioms, axiom types, and number of classes over all ontologies contained within Aber-OWL.

For each query to Aber-OWL, we also provide the query execution time within Aber-OWL and pass this information back to the client along with the result-set of the query.

All information is available through Aber-OWL's JSON API, and the source code freely available at <https://github.com/bio-ontology-research-group/AberOWL>.

2.3 Experimental setup

In order to evaluate the performance of querying single and multiple ontologies in Aber-OWL, randomly queries of different complexity were generated and executed. Since the ELK reasoner utilises a cache for answering queries that have already been computed, each of the generated query consisted of a new class expression. The following types of class expressions were used in the generated queries (for randomly generated A, B, and R):

- Primitive class: A
- Conjunctive query: A and B
- Existential query: R some A
- Conjunctive existential query: A and R some B

300 random queries for each of these type were generated for each ontology that was tested (1,200 queries in total per ontology). Each set of the 300 random queries that was generated, was subsequently

split into three sets each of which contained 100 class expressions. The random class expressions contained in the resulting sets were then utilised to perform superclass (100 queries), equivalent (100 queries) and subclass (100 queries) queries and the response time of the Aber-OWL framework was recorded for each of the query.

We further test the scalability of answering the queries by performing these queries in parallel. For this purpose, we remotely query Aber-OWL with one query at once, 100 queries in parallel, and 1,000 queries in parallel.

In our test, we record the response time of each query, based on the statistics provided by the Aber-OWL server; in particular, response time does not include network latency. All tests are performed on a server with 128GB memory and two Intel Xeon E5-2680v2 10-core 2.8GHz CPUs with hyper-threading activated (resulting in 40 virtual cores). The ELK reasoner underlying Aber-OWL is permitted to use all available (i.e., all 40) cores to perform classification and respond to queries.

3 RESULTS AND DISCUSSION

On average, when performing a single query over Aber-OWL, query results are returned in 10.8 milliseconds (standard deviation: 48.0 milliseconds). The time required to answer a query using Aber-OWL correlates linearly with the number of logical axioms in the ontologies (Pearson correlation, $\rho = 0.33$), and also strongly correlates with the number of queries performed in parallel (Pearson correlation, $\rho = 0.82$). Figure 1 shows the query times for the ontologies based on the type of query, and Figure 2 shows the query times based on different number of queries run in parallel. The maximum observed memory consumption for the Aber-OWL server while performing these tests was 66.1 GB.

We observe several ontologies for which query times are significantly higher than for the other ontologies. The most prevalent outlier is the NCI Thesaurus (Sioutos *et al.*, 2007) for which average query time is 600 ms when performing a single query over Aber-OWL. Previous analysis of NCI Thesaurus has identified axioms which heavily impact the performance of classification for the ontology using multiple description logic reasoners (Gonçalves *et al.*, 2011). The same analysis has also shown that it can significantly improve reasoning time to add inferred axioms to the ontology. To test whether this would also allow us to improve reasoning time over the NCI Thesaurus in Aber-OWL and using the ELK reasoner, we apply the Elvira modularization software (Hoehndorf *et al.*, 2011), using the Hermit reasoner to classify the NCI Thesaurus and adding all inferred axioms that fall into the OWL-EL profile to the ontology, as opposed to ELK's approach of ignoring non-EL axioms during classification. We then repeat our experiments. Figure 3 shows the different reasoning times for NCI Thesaurus before and after processing with Elvira. Query time reduces from 703 ms (standard deviation: 689 ms) before processing with Elvira to 51 ms (standard deviation: 42 ms) after processing with Elvira, demonstrating that adding inferred axioms and removing axioms that do not fall in the OWL-EL profile can be used to improve query time.

Another outlier with regard to average query time is the Natural Products Ontology (NATPRO, <http://bioportal.bioontology.org/ontologies/NATPRO>). However, as NATPRO is expressed in OWL-Full, it cannot reliably be classified with a Description Logic reasoner, and therefore we cannot apply

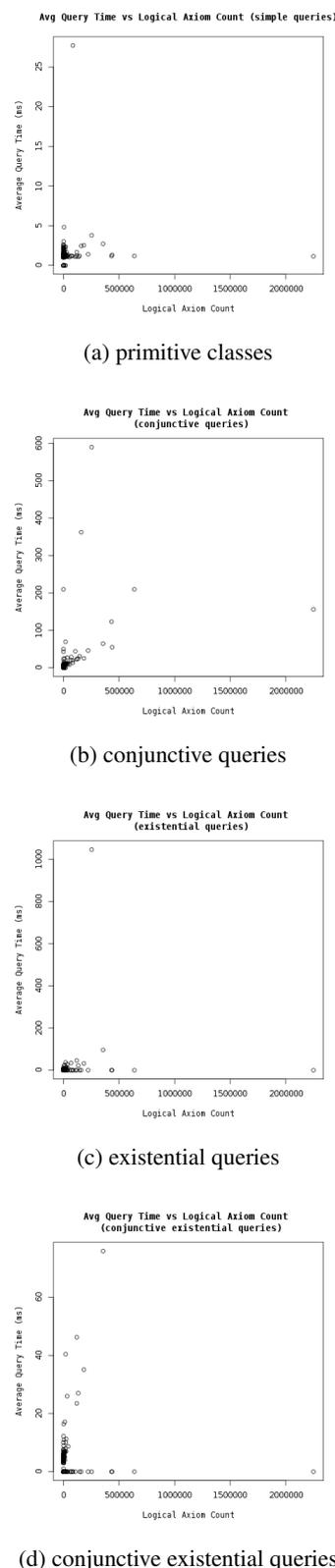


Fig. 1: Query times as function of the number of logical axioms in the ontologies, separated by the type of query.

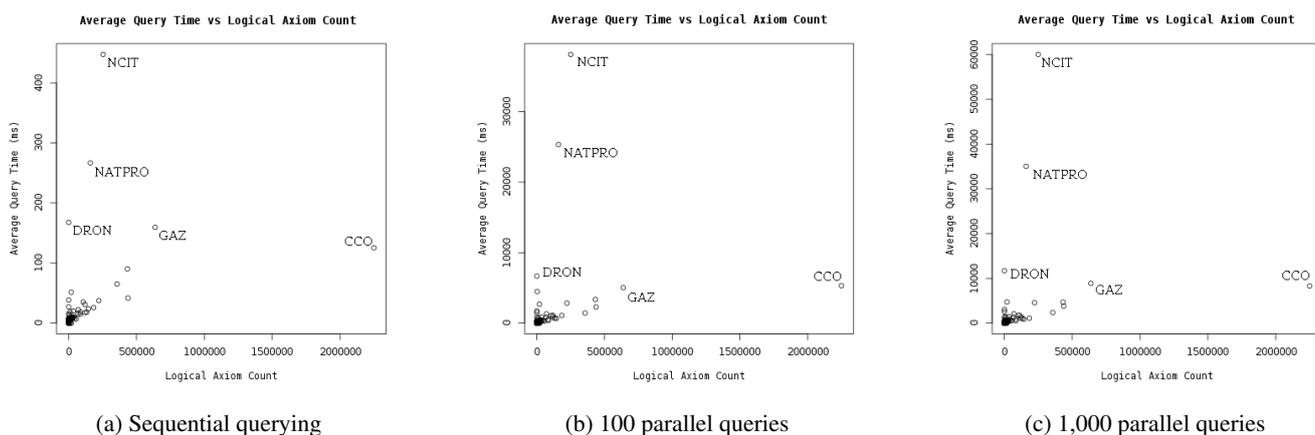


Fig. 2: Query times as function of the number of logical axioms in the ontologies, separated by the number of queries executed in parallel.

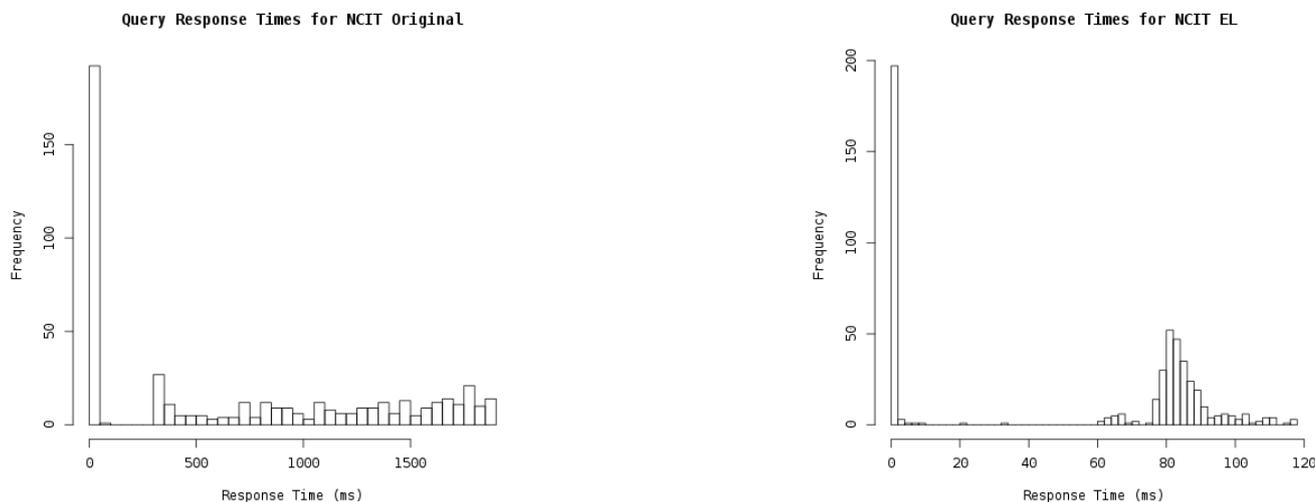


Fig. 3: Query times over the NCI Thesaurus.

the same approach to improve the performance of responding to queries.

3.1 Future Work

The performance of using automated reasoning for querying ontologies relies heavily on the type of reasoner used. We have used the ELK (Kazakov *et al.*, 2014, 2011) reasoner in our evaluation; however, it is possible to substitute ELK with any other OWLAPI-compatible reasoners. In particular, novel reasoners such as Konklude (Steigmiller *et al.*, 2014), which outperform ELK in many tasks (Bail *et al.*, 2014), may provide further improvements in performance and scalability.

We identified several ontologies as leading to performance problems, i.e., they are outliers during query time testing. For these ontologies, including the Natural Products Ontology (NATPRO),

and, to a lesser degree, the Drug Ontology (DRON) (Hanna *et al.*, 2013), similar ‘culprit-finding’ analysis methods may be applied as have previously been applied for the NCI Thesaurus (Gonçalves *et al.*, 2011). These methods may also allow the ontology maintainers to identifying possible modifications to their ontologies that would result in better reasoner performance.

4 CONCLUSION

We have demonstrated that it is feasible to reason over most of the ontologies available in BioPortal in real time, and that queries over these ontologies can be answered quickly, in real-time, and using only standard server hardware. We further tested the performance of answering queries in parallel, and show that, for the majority of cases, even highly parallel access allows quick response times.

We have also identified a number of ontologies for which performance of automated reasoning, at least when using AberOWL and the ELK reasoner, is significantly worse, which renders them particularly problematic for application that carry heavy parallel loads. At least for some of these ontologies, pre-processing ontologies using tools such as Elvira (Hoehndorf *et al.*, 2011) can mitigate these problems.

The ability to reason over a very large number of ontologies, such as all the ontologies in BioPortal, opens up the possibility to frequently use reasoning not only locally when making changes to a single ontology, but also monitor – in real time – the consequences that a change may have on other ontologies, in particular on ontologies that may import the ontologies that is being changed. Using automated reasoning over all ontologies within a domain therefore has the potential to increase interoperability between ontologies and associated data by verifying mutual consistency and enabling queries across multiple ontologies, and our results show that such a system can now be implemented with the available software tools and commonly used server hardware.

ACKNOWLEDGEMENTS

REFERENCES

- Bail, S., Glimm, B., Jiménez-Ruiz, E., Matentzoglou, N., Parsia, B., and Steigmiller, A., editors (2014). *ORE 2014: OWL Reasoner Evaluation Workshop*. Number 1207 in CEUR Workshop Proceedings. CEUR-WS.org, Aachen, Germany.
- Belleau, F., Nolin, M., Tourigny, N., Rigault, P., and Morissette, J. (2008). Bio2RDF: Towards a mashup to build bioinformatics knowledge systems. *Journal of Biomedical Informatics*, **41**(5), 706–716.
- Bodenreider, O. (2004). The Unified Medical Language System (UMLS): integrating biomedical terminology. *Nucleic Acids Res*, **32**(Database issue), D267–D270.
- Cote, R., Jones, P., Apweiler, R., and Hermjakob, H. (2006). The ontology lookup service, a lightweight cross-platform tool for controlled vocabulary queries. *BMC Bioinformatics*, **7**(1), 97+.
- Del Vescovo, C., Gessler, D. D., Klinov, P., Parsia, B., Sattler, U., Schneider, T., and Winget, A. (2011). Decomposition and modular structure of biportal ontologies. In *The Semantic Web–ISWC 2011*, pages 130–145. Springer.
- Gonçalves, R. S., Parsia, B., and Sattler, U. (2011). Analysing multiple versions of an ontology: A study of the nci thesaurus. In *24th International Workshop on Description Logics*, page 147. Citeseer.
- Grau, B., Horrocks, I., Motik, B., Parsia, B., Patelschneider, P., and Sattler, U. (2008). OWL 2: The next step for OWL. *Web Semantics: Science, Services and Agents on the World Wide Web*, **6**(4), 309–322.
- Hanna, J., Joseph, E., Brochhausen, M., and Hogan, W. (2013). Building a drug ontology based on rxnorm and other sources. *Journal of Biomedical Semantics*, **4**(1), 44.
- Hoehndorf, R., Dumontier, M., Oellrich, A., Wimalaratne, S., Rebholz-Schuhmann, D., Schofield, P., and Gkoutos, G. V. (2011). A common layer of interoperability for biomedical ontologies based on OWL EL. *Bioinformatics*, **27**(7), 1001–1008.
- Hoehndorf, R., Slater, L., Schofield, P. N., and Gkoutos, G. V. (2015). Aber-owl: a framework for ontology-based data access in biology. *BMC Bioinformatics*.
- Horridge, M., Drummond, N., Goodwin, J., Rector, A., Stevens, R., and Wang, H. (2006). The Manchester OWL Syntax. *Proc. of the 2006 OWL Experiences and Directions Workshop (OWL-ED2006)*.
- Horridge, M., Bechhofer, S., and Noppens, O. (2007). Igniting the OWL 1.1 touch paper: The OWL API. In *Proceedings of OWLED 2007: Third International Workshop on OWL Experiences and Directions*.
- Horrocks, I., Sattler, U., and Tobies, S. (2000). Practical reasoning for very expressive description logics. *Logic Journal of the IGPL*, **8**(3), 239–264.
- Jupp, S., Malone, J., Bolleman, J., Brandizi, M., Davies, M., Garcia, L., Gaulton, A., Gehant, S., Laibe, C., Redaschi, N., Wimalaratne, S. M., Martin, M., Le Novre, N., Parkinson, H., Birney, E., and Jenkinson, A. M. (2014). The EBI RDF platform: linked open data for the life sciences. *Bioinformatics*, **30**(9), 1338–1339.
- Kazakov, Y., Krötzsch, M., and Simancík, F. (2011). Unchain my \mathcal{EL} reasoner. In *Proceedings of the 23rd International Workshop on Description Logics (DL'10)*, CEUR Workshop Proceedings. CEUR-WS.org.
- Kazakov, Y., Krötzsch, M., and Simancík, F. (2014). The incredible elk. *Journal of Automated Reasoning*, **53**(1), 1–61.
- Manola, F. and Miller, E., editors (2004). *RDF Primer*. W3C Recommendation. World Wide Web Consortium.
- Matentzoglou, N., Bail, S., and Parsia, B. (2013). A corpus of owl dl ontologies. *Description Logics*, **1014**, 829–841.
- Motik, B., Grau, B. C., Horrocks, I., Wu, Z., Fokoue, A., and Lutz, C. (2009). Owl 2 web ontology language: Profiles. Recommendation, World Wide Web Consortium (W3C).
- Noy, N. F., Shah, N. H., Whetzel, P. L., Dai, B., Dorf, M., Griffith, N., Jonquet, C., Rubin, D. L., Storey, M.-A. A., Chute, C. G., and Musen, M. A. (2009). Biportal: ontologies and integrated data resources at the click of a mouse. *Nucleic acids research*, **37**(Web Server issue), W170–173.
- Patel-Schneider, P. F., Hayes, P., and Horrocks, I. (2004). Owl web ontology language semantics and abstract syntax section 5. rdf-compatible model-theoretic semantics. Technical report, W3C.
- Salvadores, M., Horridge, M., Alexander, P. R., Ferguson, R. W., Musen, M. A., and Noy, N. F. (2012). Using sparql to query biportal ontologies and metadata. In *The Semantic Web–ISWC 2012*, pages 180–195. Springer.
- Salvadores, M., Alexander, P. R., Musen, M. A., and Noy, N. F. (2013). Biportal as a dataset of linked biomedical ontologies and terminologies in rdf. *Semantic web*, **4**(3), 277–284.
- Sazonau, V., Sattler, U., and Brown, G. (2013). Predicting performance of owl reasoners: Locally or globally? Technical report, Technical report, School of Computer Science, University of Manchester.
- Seaborne, A. and Prud'hommeaux, E. (2008). SPARQL query language for RDF. W3C recommendation, W3C. <http://www.w3.org/TR/2008/REC-rdf-sparql-query-20080115/>.
- Sioutos, N., de Coronado, S., Haber, M. W., Hartel, F. W., Shaiu, W.-L., and Wright, L. W. (2007). Nci thesaurus: a semantic model integrating cancer-related clinical and molecular information. *Journal of biomedical informatics*, **40**(1), 30–43.
- Steigmiller, A., Liebig, T., and Glimm, B. (2014). Konclude: System description. *Web Semantics: Science, Services and Agents on the World Wide Web*, **27**(1).
- The Uniprot Consortium (2007). The universal protein resource (uniprot). *Nucleic Acids Res*, **35**(Database issue).
- Tobies, S. (2000). The complexity of reasoning with cardinality restrictions and nominals in expressive description logics. *J. Artif. Int. Res.*, **12**(1), 199–217.
- Tudose, I., Hastings, J., Muthukrishnan, V., Owen, G., Turner, S., Dekker, A., Kale, N., Ennis, M., and Steinbeck, C. (2013). Ontoquery: easy-to-use web-based owl querying. *Bioinformatics*, **29**(22), 2955–2957.
- Williams, A. J., Harland, L., Groth, P., Pettifer, S., Chichester, C., Willighagen, E. L., Evelo, C. T., Blomberg, N., Ecker, G., Goble, C., and Mons, B. (2012). Open phacts: semantic interoperability for drug discovery. *Drug Discovery Today*, **17**(2122), 1188 – 1198.
- Xiang, Z., Mungall, C. J., Ruttenberg, A., and He, Y. (2011). Ontobee: A linked data server and browser for ontology terms. In *Proceedings of International Conference on Biomedical Ontology*, pages 279–281.