The Open Ontology Repository Initiative: Requirements and Research Challenges

Kenneth Baclawski[*] kenb@ccs.neu.edu

Abstract

Very large data sets are increasingly common both in science and industry. However, incorporating multiple data types from multiple sources to solve major problems is a significant interoperability challenge. Furthermore, documents and other artifacts created in the past can be as important as recently created data sets, but interoperability with such legacy data can also be difficult. The problem with such data sets is not only the differences in recording media and formats but also the enormous changes in terminology over time. Data sets run the risk of rapid obsolescence as the meaning and formats of the data fields are forgotten or no longer available.

Semantic technologies based on logic, databases and the Semantic Web can address the problem of meaningful access to and integration of data both today and over decades and centuries. This paper discusses an initiative to develop and deploy a new federated interoperability infrastructure called the Open Ontology Repository (**OOR**). The OOR is intended to support the full data management lifecycle for the communities that it serves. The OOR grew out of the Ontolog community that has existed for over 6 years and continues to grow in both size and diversity. An initial OOR server based on BioPortal has been deployed, and further development will emphasize technological solutions that build on existing ontology repositories as well as proven architectures and standards. Nevertheless, many research challenges remain to achieve the requirements that have been identified. This paper reports on the research challenges of the OOR initiative and the requirements that gave rise to them. Ultimately, it is hoped that the OOR initiative will result in the deployment of a robust, federated knowledge repository that can collectively correct for multiple points of failure and can foster collaborative stewardship of knowledge and metadata.

Keywords: ontology registries, ontology repositories, federated repositories, metadata repositories, interoperability infrastructures

1 Introduction

Many communities and companies are producing very large data sets that are increasingly complex and diverse. These data sets are very well suited for particular narrowly-defined, discipline- and sector-specific purposes. In principle, these data sets could be used for solving more broadly-defined problems. However, incorporating multiple data types from multiple sources to solve these problems remains a significant challenge.

While the data sets that are currently being developed typically engender the greatest level of enthusiasm by the communities that are creating them, data sets created in the past can have equal

[*]College of Computer and Information Science, Northeastern University, Boston, MA 02115

importance for related communities. The problem is not just the differences in formats but also the enormous changes in terminology over time. Current data sets run the risk of rapid obsolescence as the meaning of the data fields is forgotten even by the individuals who introduced them.

To ensure that data remains usable, it is important that it be annotated with metadata. As the structure and processing of data has become more complex, metadata has also more complex, and the the role of metadata has become increasingly important. There is a critical need for a robust, scalable infrastructure for metadata. This infrastructure facilitates data interoperability, sharing, search, retrieval, and long-term preservation.

Scientific research communities exemplify the problems of increasingly large and complex data sets. The importance of metadata for scientific research communities was recognized in a recent joint report of the National Academies of Science and Engineering, and the Institute of Medicine[14] concerned with data usability and integrity, "Enhancing the Integrity, Accessibility, and Stewardship of Research Data in the Digital Age." This report states, "Metadata are a critical part of the context needed to assess the integrity of data and use data accurately." The report then emphasizes this point when it states, "... and it is generally impossible to judge the integrity of processed data without access to the metadata documenting how they were processed." Moreover, the report states that the role of metadata has become increasingly important over time, "As digital data has become more important in a variety of disciplines and fields, the scope and value of metadata have grown, leading to the development of metadata standards. Metadata standards are an agreed set of terminologies, definitions and values to be provided for data in a given field or community."

Given the recognition of the importance of metadata by the National Academies, the OOR has the potential for having a major impact on the conduct of modern data-intensive scientific research. However, simply creating the infrastructure and making it available will not automatically ensure that it is utilized in spite of the mandates in the National Academies report cited above. One can also say the same about metadata and ontologies in industry and government. While ontology search engines, registries and repositories have been developed, and they have had a significant impact on some communities such as many in biomedicine, they have yet to achieve general acceptance by most of the communities that could benefit from the infrastructure.

In a recent meeting, the state of the art of ontology repositories was evaluated and the research challenges were identified[2]. The OOR initiative is currently focusing its attention on challenges that are the most likely to increase the level of adoption of the ontologies in general, and the OOR in particular, by data-intensive and knowledge-intensive communities. They are also the challenges that have not yet been adequately addressed by existing ontology repository projects. Some of the most important of these challenges are federation, modularization and education which will be discussed in more detail below.

The National Academies report[14] raises several concerns about data. One of these is a concern with the long-term preservation of data, "... our focus in this report is on a specific aspect of utility that we refer to as data stewardship – the long-term preservation of data so as to ensure their continued value, sometimes for unanticipated uses. Stewardship goes beyond simply making data accessible. It implies preserving data and metadata so that they can be used by researchers in the same field and in fields other than that of the data's creators." One consequence of this concern is the need to maintain all versions of metadata standards so that preserved data will still be meaningful as standards evolve. Another consequence is the need to map older metadata to newer metadata so that data from different time periods can be compared.

Another concern raised in the National Academies report is the need for incentives to ensure that researchers make their data accessible. As the report makes clear, researchers must also provide metadata, and research communities are strongly encouraged to institute ways to reward researchers who do this. While the OOR will certainly not solve the problem of ensuring that researchers release their data; nevertheless, a common, easily used infrastructure for metadata registration and storage would reduce some of the barriers for researchers and research communities to provide standard metadata annotations for their data.

2 Requirements

At the Ontology Summit 2008[21] and in subsequent virtual meetings of the OOR community, a series of requirements was identified and refined. These requirements are now discussed.

2.1 Support the full data preservation and access lifecycle

To truly support discovery, innovation and learning well into the future the OOR collaboration will manage the full metadata life cycle by providing an architecture and an infrastructure that supports the creation, sharing, searching, and management of ontologies, including database and XML Schema data definitions. Complementary goals include fostering the ontology community, the identification and promotion of best practices, and the provision of services relevant to ontologies. Examples of anticipated services include automated semantic interpretation of content expressed in knowledge representation languages, the creation and maintenance of mappings among disparate ontologies and content, and inference over this content. The OOR will support a broad range of semantic services and applications of interest to enterprises and communities.

The OOR will develop efficient logic programming-based reasoning methods that amalgamate Semantic Web-based ontologies and rules with extended Prolog and Answer Set Programming, to be used for reasoning over the ontologies, instances, and rules of the repository. [25, 24, 30] The OOR will design and implement service-oriented architectures and services, including automated and semi-automated service orchestration and parallel optimization to support the repository. [12, 23, 31]

The following are the core requirements for the OOR[21]:

1. The repository architecture shall be scalable.
2. The architecture shall be optimized for sharing, collaboration and reuse.
3. The repository shall be capable of supporting ontologies in multiple formats and levels of formalism.
4. The repository architecture shall support distributed repositories.
5. The repository architecture shall support explicit machine usable/accessible formal semantics for the meta-model of the repository.
6. The repository shall provide a mechanism to address intellectual property and related legal issues/problems.
7. The repository architecture shall include a core set of services, such as support for adding, searching and mapping across ontologies and data related to the stored ontologies.
8. The repository architecture shall support additional services both directly within the province of the repository and as external services.
9. The repository should support all phases of the ontology lifecycle.

2.1.1 Data deposition/acquisition/ingest

The OOR initiative is developing requisite ontology-based architectures, including ontology lifecycle management, theories and implementations of ontology modularity, upper and middle ontologies, and methods for automatically and semi-automatically aligning and mapping ontologies. Logical relationships between ontologies will be supported within the repository, including mutual consistency, extension, entailment, semantic mappings, intelligent search, and decision support. [3, 15, 19, 26, 28, 27]

The OOR will support internal and external services and applications including: ontology creation tools, ontology editors, ontology differencing tools, ontology modularization tools (clustering, etc.), ontology export, ontology visualization (e.g., graph visualization), version management and access control. While the emphasis is on the metadata level, it is awkward to exclude all instance data. Depending on the ontology, instance data will also be included. For knowledge-rich domains, ontologies can include all of the data as well as the metadata. For other domains, the data will be managed by special-purpose applications, and the ontology will be only part of the database, playing the roles that are most appropriate, such as encoding access policies and procedures, enabling discovery and interoperability, helping to ensure integrity, and assuring that data remains accessible and understandable over time periods of decades or more.

2.1.2 Data curation and metadata management

The OOR distinguishes between gatekeeping and quality control. Gatekeeping criteria are a set of minimal requirements that any ontology within the OOR has to meet. The latter are intended to enable the users of the OOR to quickly find ontologies that fit their needs; the criteria are not supposed to ensure the quality of the ontologies. The metadata in the OOR must meet the following criteria: (1) The metadata are submitted in a publicly described language and format; (2) The metadata are read accessible; (3) The metadata are expressed in a formal language with a well-defined syntax; (4) The authors of the metadata provide the required provenance and other annotations; (5) The metadata have a clearly specified and clearly delineated scope; (6) Successive versions of the metadata are clearly identified; (7) The metadata is organized into ontologies that are appropriately named; It is especially important that the ontologies specify the process that is employed to create and maintain them.

Ontologies may be regarded as engineering artifacts. Consequently, there must be support for versioning and configuration management. The Ontology Metadata Vocabulary (OMV), Dublin Core, ISO 11179, ISO 19763, and other existing approaches to provenance and versioning support will be used as the basis for annotating ontologies in the OOR. An empirical approach will be used to identify and evaluate repository metadata. Proposals for repository metadata already exist, and they will be evaluated using use-case scenarios. These scenarios both motivate the use of the repository metadata and help establish best practices.

Metadata occurs on many levels. While it is not always possible or necessary to distinguish these levels, it is expected that the metadata in the repository will usually be organized according to Figure 1. In this figure data is from 3 different domains. In the first, data is in the form of image data, in the second data, is sensor data, and in the third, data is a knowledge base. The dashed line separates the data being managed by the OOR and data managed separately from the OOR. Both image data and sensor data requires domain-specific storage, retrieval and processing software and hardware so it is managed separately. Knowledge bases, on the other hand, are well suited to storage, retrieval and processing by the OOR. Each data set is annotated with metadata that is specific to the data set. This metadata includes provenance information as well as any structural and processing specifications that are specific to the data set. Each domain has its own ontology that specifies those aspects of the metadata annotations that are not covered by the OOR ontology. The domain-specific ontologies and data set-specific metadata are annotated as artifacts in the repository. This is called repository metadata in the figure. The semantics of the repository metadata is specified by the OOR ontology.

The organization shown in Figure 1 is well suited for data sharing and interoperability within one domain, but not for cross-domain interoperability. Achieving the latter requires that the domain ontologies be related to each other. Although the figure does not explicitly show this, the various ontologies will be related to each other in a web of interconnections. Ontology matching and mediation is an active area of research. The OOR will support the specification of relationships between ontologies, both at the level of whole ontologies (e.g., ontology importing) and at lower levels of granularity (e.g., declarating that two resources in different ontologies are the same). The OOR will also provide APIs that allow one to plug in ontology matching modules.

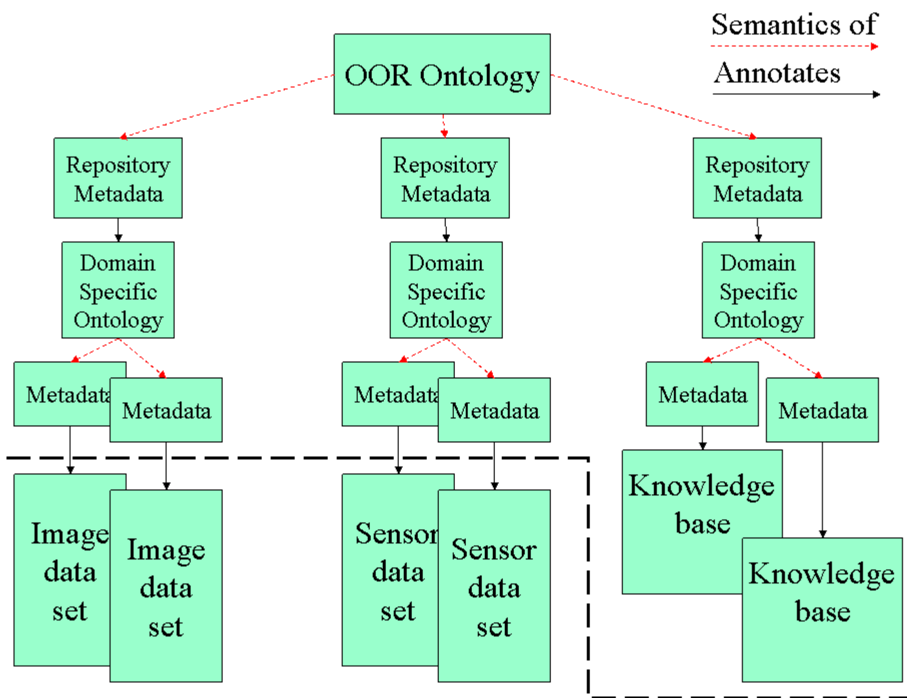

Figure 1: Metadata Levels in the OOR

2.1.3 Data protection

A repository requires mechanisms for effective management. The understanding is that as the repository and its infrastructure evolve, more management support mechanisms will be included. The core mechanisms to be provided in the first version include enforcement of policies for access, submission, governance, change management, and control over user and administrative access. Subsequent versions of the OOR will provide capabilities to: create usage reports, validate syntax, check logical consistency and automatically categorize a submission.

2.1.4 Data discovery, access, use, and dissemination

Achieving these goals will help reduce semantic ambiguity whenever and wherever information is shared, thereby allowing information to be located, searched, categorized, and exchanged with a more precise expression of its content and meaning. The artifacts of the repository will provide a semantic grounding for diverse formats and domains, ranging from the conceptual domains and specific disciplines of communities to technical schemas such as WSDL, UDDI, RSS, and XML schemas, and of course expressed in standard ontology languages such as RDF, OWL, Common Logic, and others. Perhaps most importantly, the repository will enable wide-scale knowledge re-use and reduce the need to re-invent the wheel when defining concepts and relationships that are already understood.

To facilitate knowledge discovery the repository shall provide metadata capabilities to support search capabilities, governance process, and management. The repository will support discovery by at least the following criteria: domain, author/creator/source, version, language, terminology and controlled vocabularies, quality, mapping, and inference. The interfaces for discovery will be suitable for both specialist and non-specialist users, using GUIs, web services, and language-based APIs.

2.1.5 Data interoperability, standards, and integration

To support the sharing and reuse of ontologies within the repository the OOR will store both ontologies and metadata for ontologies. The metadata will allow users to:

- determine whether an ontology is suitable for a user purpose;
- capture the design rationales that underlie the ontology;
- find information about author, author credentials, and source of ontology reference material;
- retrieve ontologies for use in domain applications;
- retrieve ontologies to be integrated with other ontologies;
- retrieve ontologies that will be extended to create new ontologies;
- determine whether or not an ontology can be integrated with given ontologies;
- determine whether a set of ontologies retrieved from the repository can be used together; and
- determine whether an ontology in the repository can be partially shared.

There will be policies for creation and modification of metadata and documentation of ontologies and the management of the persistence and sustainability of ontologies.

Users (including end-users, ontology and repository developers, subject matter experts and stakeholders) should participate in the collaborative ontology development life cycle and in decisions regarding what metadata are suitable for ontologies in the repository.

The metadata will include both logical metadata (logical properties of the ontology independent of any implementation or engineering artifact) and engineering metadata (properties of the ontology considered as an engineering artifact).

2.1.6 Data evaluation, analysis, and visualization

It is not sufficient for the OOR just to store ontologies; it also needs to enable the evaluation of the ontologies within it. The OOR will offer functionalities like those on social networking sites which would allow users to comment on ontologies and rank them. Further, the OOR will enable selective views of the repository using tags provided by subcommunities that characterize ontologies with respect to their chosen criteria. For example, such a view might select ontologies for specific fields of research or industries, or ontologies satisfying specific quality criteria or levels of organizational approval.

The OOR will develop methods, practices, services, and artifacts to support automated and human reviewed evaluation and comparison of ontologies stored in the repository. [9, 17, 16]

2.2 Research Challenges

Ontology engineering is a rapidly developing research area. There are both computational and storage challenges as the size of a knowledge base grows, since the computational complexity of logical inference is much greater than the complexity of traditional database query processing. Ontology repositories constitute a subfield within ontology engineering, and has its own community and conferences. In a recent meeting, the state of the art of ontology repositories was evaluated and research challenges were identified[2].

2.3 Inadequacies of Current Ontology Repositories

Existing ontology repositories tend to have many of the same features, such as registration, submission and upload; browsing and search; description and documentation; metrics and statistics. There are many missing features, but the most noticeable is the lack of structure among the ontologies.[1] The ontologies in the repository are treated as indepedent entities. Another missing feature is the lack of sufficient metadata annotating the ontologies, although some ontology repositories have started to grapple with this problem, such as BioPortal[13] and Cupboard[5, 4].

While ontologies are claimed to be a mechanism for interoperability and communication between data sources, ontologies themselves are nearly always built in isolation. There is no common representation of metadata annotations of ontologies and no common ways to identify versions of ontologies. The various ontology repositories use a variety of techniques and do not enforce any standard conventions.

2.4 Representation Languages

The number of existing representation languages for metadata and ontologies is enormous and growing. One can use text, XML, frames, graphs, OWL DL, HOL, UML and SQL to name just a few. Languages for representing relationships between ontologies are also emerging. The development of the Common Logic standard (ISO/IEC 24707:2007) is especially important because it can act as a common language for most of the existing ontology representation languages. The challenge to the ontology community to develop tools for CL. The particular challenge for ontology repositories is to Construct a repository of first-order ontologies that will serve as a testbed for ontology evaluation and integration techniques, and that can support the design, evaluation, and application of ontologies in first-order logic. The Common Logic Ontology Repository (COLORE) being developed by Michael Gruninger is a prominent example of such an effort.[6]

2.5 Policies and Best Practices

Experience has shown that the following to be the factors that are well correlated with reuse:[7]

1. Small development teams with larger user communities

2. Commitment to users and to continuous improvement
3. Publication of maintenance policies, URI naming conventions and policies, useful documentation

Therefore, it is important to have well-specified policies for vocabulary management, metadata, and provenance specification to enable trust. It is also critical to have a commitment to forming, accommodating, serving, and working with a community of users. This emphasizes the importance of outreach and education, including the identification and promotion of best practices. Since the OOR will be an open, federated architecture and infrastructure, it is intended to be utilized by communities to host their own ontologies as well as allowing the communities to adapt previously established ontologies for their own purposes. The challenge is to develop appropriate educational mechanisms for a very diverse set of communities. An important example is the Ontolog Forum which has been engaging in educational and outreach activities for 6 years, reaching over 40 distinct communities. Examples include communities in bioinformatics, national command and control, and intelligence. [11, 10, 8, 20, 18, 22, 26, 29, 30, 31]

Ultimately the OOR initiative should provide:

- Guidance for data providers, metadata providers, and ontology providers;
- Organized references on all facets of metadata needs and solutions;
- Services targeting semantic interoperability in the respective and related domains, including vocabulary lists, ontology repository and associated services, and vocabulary creation and maintenance tools, services, and guidance;
- Community collaboration environment (shared files, email archives, and web pages, either public or secure);
- Access to work in progress on metadata tasks and projects.

In addition, the OOR collaboration can provide purposeful capabilities:

- Targeted identification and evaluation of resources (vocabularies, standards, tools, services);
- Identification and engagement of key community participants (projects or individuals) in metadata initiatives;
- Training and workshops in metadata technologies and techniques, particularly dealing with semantic tools and services, including vocabulary and ontology development, metadata standards and their application, as well as metadata-enlightened architectural development and analysis;
- Community environment(s) to advance particular topics or discussions.

3 Acknowledgements

This report is the result of contributions by a large number of individuals who have been involved in the Ontology Summits, the Ontolog and OOR meetings, and particularly the following individuals: Leo Obrst, Mark Musen, Mike Dean, Peter Yim, Michael Gruninger, Elisa Kendall, Bart Gajderowicz, Mathieu d'Aquin, Todd Schneider, Steve Ray and Katherine Goodier.

References

[1] C. Allocca, M. d'Aquin, and E. Moaa. DOOR: Towards a formalization of ontology relations. In *International conference on knowledge engineering and ontology development*, 2009.

[2] K. Baclawski. Ontology repository research issues workshop, August 2009. Founder, convener and chair.

[3] M. Daconta, L. Obrst, and K. Smith. *The Semantic Web: The Future of XML, Web Services, and Knowledge Management*. John Wiley, June 2003.

[4] M. d'Aquin, J. Euzenat, C. LeDuc, and H. Lewen. Sharing and reusing aligned ontologies with cupboard. In *International Conference on Knowledge Capture*, 2009. Demo.

[5] M. d'Aquin and H. Lewen. Cupboard: A place to expose your ontologies to applications and the community. In *European Semantic Web Conference*, 2009. Demo.

[6] M. Gruninger. COLORE: Common Logic Ontology Repository, August 6 2009. http://ontolog.cim3.net/file/work/OOR-Ontolog-Panel/2009-08-06_Ontology-%?Repository-Research-Issues/Colore--MichaelGruninger_20090806.pdf.

[7] E. Kendall. OOR for public sector use: Metadata questions, August 6 2009. http://ontolog.cim3.net/file/work/OOR-Ontolog-Panel/2009-08-06_Ontology-%?Repository-Research-Issues/PublicSectorUse--ElisaKendall_20090806.pdf.

[8] J. Luciano. A pathway to pathways. In *Genome Technology*, pages 15–16, June 2006.

[9] J. Luciano and L. Obrst. Ontology evaluation methods and metrics, 2008. Proposed MITRE internal research.

[10] J. Luciano and R. Stevens. e-Science and biological pathway semantics. *BMC Bioinformatics*, 8(Suppl 3), 2007.

[11] J. Luciano and R. Stevens. OWL: PAX of mind or the AX? Experiences of using OWL in the development of BioPAX. In *OWL: Experiences and Directions*, Gaithersburg, MD, USA, April 1-2 2008.

[12] D. McCandless, L. Obrst, and S. Hawthorne. Dynamic web service assembly using OWL. In *W3C Workshop on Semantic Web in Energy Industries; Part I: Oil and Gas*, Houston, TX, Dec. 9-10 2008.

[13] M. Musen, N. Shah, N. Noy, B. Dai, M. Dorf, N. Griffith, JD Buntrock, C. Jonquet, MJ Montegut, and DL Rubin. BioPortal: Ontologies and data resources with the click of a mouse. *AMIA Annual Symposium*, 2008.

[14] Enhancing the integrity, accessibility, and stewardship of research data in the digital age, July 22 2009.

[15] L. Obrst. Ontological architectures. In M. Healy, A. Kameas, and R. Poli, editors, *TAO Theory and Applications of Ontology, Volume 2: The Information-science Stance*. Springer, 2008.

[16] L. Obrst, W. Ceusters, I. Mani, S. Ray, and B. Smith. The evaluation of ontologies: Toward improved semantic interoperability. In C. Baker and K.-H. Cheung, editors, *Semantic Web: Revolutionizing Knowledge Discovery in the Life Sciences*. Springer, 2007.

[17] L. Obrst, T. Hughes, and S. Ray. Prospects and possibilities for ontology evaluation: The view from the national center for ontological research (ncor). In *Workshop on Evaluation of Ontologies for the Web (EON2006)*, Edinburgh, UK, May 22 2006.

[18] L. Obrst, T. Janssen, and W. Ceusters, editors. *Ontologies for Intelligence Analysis*. IOS Press, 2008. forthcoming.

[19] L. Obrst, H. Liu, and R. Wray. Ontologies for corporate web applications. *Artificial Intelligence Magazine, special issue on Ontologies*, pages 49–62, Fall 2003. http://portal.acm.org/citation.cfm?id=958676.

[20] L. Obrst, D. McCandless, S. Stoutenburg, K. Fox, D. Nichols, M. Prausa, and R. Sward. Evolving use of distributed semantics to achieve net-centricity. In *Regarding the "Intelligence" in Distributed Intelligent Systems*, Arlington, VA, Nov. 8-11 2007.

[21] April 2008. http://ontolog.cim3.net/cgi-bin/wiki.pl?OntologySummit2008.

[22] M. Parmelee, D. Nichols, and L. Obrst. A net-centric metadata framework for service oriented environments. *International Journal of Metadata, Semantics and Ontologies (IJMSO)*, 2008. forthcoming.

[23] M. Prausa. *Parallelized Dynamic Service Orchestration*. PhD thesis, Colorado Technical University, September 2008. Obrst: PhD advisor.

[24] K. Samuel and L. Obrst. Answer set programming: Final report on a comparison between ASP and Prolog for Semantic Web ontology and rule reasoning. Technical report, MITRE, October 2007. Result of MITRE Sponsored Research Innovation Grant experiment.

[25] K. Samuel, L. Obrst, S. Stoutenberg, K. Fox, P. Franklin, A. Johnson, K. Laskey, D. Nichols, S. Lopez, and J. Peterson. Applying prolog to semantic web ontologies and rules: Moving toward description logic programs. *The Journal of the Theory and Practice of Logic Programming (TPLP)*, 8(03):301–322, May 2008. http://journals.cambridge.org/action/displayAbstract?fromPage=online&aid=1853440.

[26] S. Semy, M. Pulvermacher, and L. Obrst. Toward the use of an upper ontology for U.S. government and U.S. military domains: An evaluation. Technical Report MTR 04B0000063, MITRE, November 2005. http://www.mitre.org/work/tech_papers/tech_papers_05/04_1175/index.html.

[27] S. Stoutenburg. *Refining Ontology Alignment: New Methods for Relationship Acquisition and Advanced Web Applications*. PhD thesis, U. of CO, Colorado Springs, CO, 2009. Obrst: PhD committee member.

[28] S. Stoutenburg, J. Gray, and J. Kalita. Refining ontology mapping techniques: Acquiring properties beyond similarity and equivalence. In *ODBASE*, 2008. submitted.

[29] S. Stoutenburg, L. Obrst, D. McCandless, D. Nichols, P. Franklin, M. Prausa, and R. Sward. Ontologies for rapid integration of heterogeneous data for command, control, and intelligence. In *Ontologies for the Intelligence Community Conference*, Columbia, MD, Nov. 28-30 2007.

[30] S. Stoutenburg, L. Obrst, D. Nichols, P. Franklin, K. Samuel, and M. Prausa. Ontologies and rules for rapid enterprise integration and event aggregation. In *Vocabularies, Ontologies and Rules for the Enterprise (VORTE 07), EDOC 2007*, Annapolis, MD, Oct. 15-19 2007.

[31] S. Stoutenburg, L. Obrst, D. Nichols, K. Samuel, and P. Franklin. Applying semantic rules to achieve dynamic service oriented architectures. In *RuleML 2006: Rules and Rule Markup Languages for the Semantic Web*, volume 4294, pages 581–590, Heidelberg, November 10-11 2006. Springer-Verlag. http://www.mitre.org/work/tech_papers/tech_papers_06/06_0904/index.html.