Reuse of Semantics in Business Applications

Michael G. Bennett

Hypercube Ltd. 89 Worship Street, London EC2A 2BF, United Kingdom
mbennett@hypercube.co.uk

Abstract. This paper sets out an exploration of the considerations for the reuse of ontologies in the creation of an industry-wide business ontology. The paper introduces the Financial Industry Business Ontology (FIBO), a finance industry initiative and describes a number of ways in which meaning has been reused from available ontological resources. Some common themes are identified for reuse of semantics and the socializing of business meanings from different sources within an organization or industry standard ontology. Examples are given of ways in which different ontologies have been referenced by or incorporated into the FIBO models. Business applications range across a wide range of subject matter and arguably require a more well-grounded approach to meaning in the ontology than would be appropriate for stand-alone applications. Some pitfalls are identified in the re-use of ontologies which have been created with different purposes in mind.

1. Introduction

Research carried out by the Enterprise Data Management (EDM) Council [1] identified an awareness at board level for the need for consistent and standardized terms, definitions and relationships across the enterprise, and a growing recognition of enterprise-wide data management as a business issue not an IT issue.

The financial crisis of 2008 exposed a number of weaknesses in the way that data is managed in financial institutions. Firms which were exposed to the failure of Lehman Brothers and other distressed or failing institutions found themselves with all the data they needed to calculate their exposures and yet took days or even weeks to turn that information into actionable knowledge.

The lesson from this is that the focus should not be on data but on the business concepts that those data represent. This requires a holistic approach to business meaning.

In response to the crisis, the Enterprise Data Management Council commissioned the creation of a common industry semantic model, or ontology, called the Financial Industry Business Ontology (FIBO). Here, the word “ontology” is used in the sense of “a specification of a conceptualization” [13] with the specification being expressed in the Web ontology Language (OWL) [2] and the conceptualization being that of business subject matter (securities, business entities and the like).

An additional motivation for the use of OWL in FIBO is that it enables applications to take advantage of the formal representation of business concepts to draw inferences from available data. These kinds of application are gaining traction because of the opportunities they present for analyzing information available within the organization.

Another motivation is that the use of formal first order logic in OWL makes it possible to frame business rules, which require higher orders of logic to express them, using the representations in the ontology for the basic concepts on which those rules operate.

In the development of FIBO the Council is committed to the reuse of existing ontology resources where possible. There is a set of defined treatments for ways in which ontologies and other sources of semantics may be referenced, but there is a need for a clearly defined method for assessing the suitability of available ontologies for re-use or reference within the FIBO ecosystem.

In this paper we consider some cases in which publicly available semantic resources are identified and re-used for FIBO, and then explore what are the evaluation requirements for such resources, with reference to the available literature and recent observations. Addressing these questions is part of the evolving methodological framework for the management of the FIBO ecosystem itself.

2. Semantics Re-use in the Financial Industry Business Ontology

The Financial Industry Business Ontology (FIBO) is an industry collaborative ontology for use across financial industry firms, created by the EDM Council [1]. The first formally published FIBO specification is FIBO Foundations, published through the Object Management Group (OMG) [8] which has an annex (Annex B) detailing the required treatments for re-use of semantics from other sources. Additional OMG FIBO specifications include Business Entities and Indices and Indicators

FIBO comes in two parts: the published OMG specifications such as FIBO Foundations, and a set of models maintained by the EDM Council itself comprising the overall ontology of the financial domain, also referred to as a conceptual ontology.

The ontology models include a set of top level elements known as partitions. These are so called because they allow for the partitioning of the ontology content at a very abstract level, for example to distinguish between concepts which have differing temporality. These together constitute the upper ontology of FIBO.

One way in which these partitions are used is to disambiguate between a thing in itself, a thing in a role, and the context in which a thing in a role is defined. These are known as independent, relative and mediating things after Sowa [9].

In a typical application the role which something plays is implicit from the context of the application. In the ontology we are able to make this contextual information explicit by the use of these partitions. So for example a database may use the word “customer” without needing to distinguish between those assertions which are always true of a person or company, and those concepts which relate to its role as a customer.

While some data sources may make these kinds of contextual nuance clear, others may not. As an example, a data source about wine would typically use the word "varietal" as a general term for a kind of grape, and this label makes clear that the grape is being described specifically in the context of wine-making, that is as a "thing in a role", the role being its role in the creation of the wine. Other data sources and applications may make this less explicit, and so some skill and awareness of the nature of meaning is required to dispose otherwise similar terms under the framework provided by the upper ontology partitions.

Similar considerations apply in the treatments of temporality for different concepts such as events and activities, where a lattice partitions of continuant versus occurrent things is provided in the FIBO upper ontology.

2.1 Methodological Requirements

In addition to the formally defined technical treatments for the re-use of ontologies, FIBO needs a well-defined method for assessing the suitability of ontologies for re-use or reference within the FIBO ecosystem.

The treatments defined in FIBO [8, Annex B] are: incorporation by reference (OWL Import); use of a snapshot of an ontology at a given point in time, and use of a snapshot of a sub-set of the terms in an ontology. The use of a sub-set of a given ontology is recommended when it is desirable to avoid importing assertions which are not relevant to the FIBO ontologies.

Meaning does not automatically follow from the use of Semantic Web syntax. The ability of such applications to draw inferences provides some confidence that the concepts in the ontology are potentially meaningful, but is not itself the source of meaning. Meaning itself requires a more sophisticated approach. For this we believe it is necessary to apply knowledge representation methods.

A standards-based ontology like FIBO would ideally use concepts drawn from the appropriate communities of practice, framed within an overarching ontology framework.

This leads to an interesting distinction when considering the re-usability of a given ontology: the best source of knowledge about the business itself, may or may not be the best-formed ontology of that subject matter from a technical point of view. Converseley, ontologies which are well-designed for an individual application may contain semantic inaccuracies, such as the use of "country" where "jurisdiction" is really meant. The presence of such inaccuracies in an imported or referenced ontology would degrade the accuracy of semantic querying or reasoning applications which are derived from the overall FIBO ontology.

2.2 Transaction Semantics Alignment in FIBO

The FIBO partitions described above were used in the alignment of concepts for transactions, using the REA (Resource, Events, Agents) ontology for transactions [10]. We were able to frame the REA concepts in relation to double entry bookkeeping concepts as used in the XBRL reporting standard [11]. This is described in [12].

In this activity, some transaction concepts needed to be made more abstract so as to re-use them elsewhere. For example, when framing the concept of a commitment, REA defined commitment specifically in the transaction context whereas there will be other kinds of commitment in the enterprise as a whole. Therefore some REA concepts were used as the basis for more general concepts, with refinements of these being defined to correspond to the original transaction-specific concepts.

Meanwhile the REA concept of “event” was seen to differ significantly from the FIBO event concept, the latter being similar to the event ontology design patterns used elsewhere. In REA what was labeled as an event corresponded to what FIBO would call an “activity”, and this was framed with reference to the continuant versus occurrent partition in the lattice pattern.

Another innovation in this alignment activity was the extension of the “relative thing” partition to define that which is an aspect of some thing. This was used to take the definition of either of side of a transaction in REA and re-frame this from the perspective of one or other party to the transaction, as the “aspect” of each transaction side from the perspective of that party. For example a transaction event which is a kind of payment has both a payer and a payee. Similarly a delivery event has both a deliverer and a recipient of goods or services. The view of these events from the perspective of the different parties to the transaction is what is reflected in ledger accounts and reported in financial reports.

In this way the FIBO partitions were able to form a bridge between the view of transactions “in the round” as provided in REA, and the view of transactions as seen by individual participants in those transactions and as reported in accounting standards. This results in a model which can be used to create detailed semantic representations of derivatives trades, securities transactions and so on, while also being able to represent the positions of exposures of an individual financial institution.

This is one example of how the FIBO upper ontology partitions could be used to integrate semantics from different sources. In considering additional ontologies for re-use we therefore have two things to think about: how to assess their suitability for re-use in FIBO, and how to potentially redispose the incorporated terms with reference to the FIBO upper ontology partitions and to existing high level abstractions which are already present in the FIBO conceptual ontology.

3. Insights from Other Sources

One of the most widely cited referenced for ontology evaluation is he “OntoClean” method described in Gomez-Perez (2001) [14]. This deals with aspects of an ontology which is to be used in a reasoning based application.

The need for evaluation metrics for the re-use of ontologies for common meaning is addressed in a further paper by Gomez-Perez et al [20].

Pinto and Martins [17] describe a generalized approach to evaluating ontologies for possible re-use and integration, which includes pointers that may be developed further to meet the needs described in this paper.

Bontas, Mochol and Tolksdorf [8] analyze the challenges related to the reuse process with reference to examples.

Simperl [19] argues for the need for a context- and task-sensitive treatment of ontologies, and identifies reuse processes which could profit from such an approach. She argues for the need for ontology reuse methodologies which optimally exploit human and computational intelligence to operationalize those reuse processes.

3.1 The Annual Ontology Summit

The Ontology Summit is an annual event put on by the ontology community.

The Ontology Summit of 2013 was titled "Ontology Evaluation Across the Ontology Development Lifecycle" [21] and was an exploration of ontology evaluation from a number of different perspectives, along with a number of practical events called "hackathons" in which different tools and techniques for ontology evaluation were trialed.

A key take-away from the 2013 Summit was the need to deal with semantic issues in addition to syntactical or performance considerations. Various Summit presentations highlighted the need for distinct but overlapping treatments for the evaluation of ontologies to be used in individual applications, and ontologies to provide a source of reference to business meaning, for the benefit both of conventional and semantic technology applications.

Since the OWL language lends itself well to both applications, ontologies which are to be re-used for an industry standard in the business space must be evaluated as to their suitability to this purpose.

Another key finding of the 2013 Ontology Summit was the need to adequately deal with possible differences between the underlying theoretical assumptions behind different ontologies. A detailed treatment of these issues was described both by Smith [15] and by Partridge [16]. For example an ontology which is based or grounded in a four dimensional view of the world might not be usable by reference or import from an ontology or set of ontologies which take a more conventional three dimensional view of the world. This is similar to the way in which a model which uses Newtonian calculations may not be compatible with a model of the same subject matter which uses quantum theory.

The 2013 Ontology Summit also included a set of hackathon activities aimed at understanding the available tools and techniques for ontology evaluation. One of these was organized around the FIBO standard, using the OQuaRE [22] and OOPS! [23] ontology evaluation frameworks, integrating these around the requirements for assessment of ontologies within FIBO itself, as well as looking at the OntoQA toolset [24] for evaluating knowledge bases used in tests. The lessons from this hackathon could in principle also be applied to the evaluation of ontologies for reference or inclusion within the FIBO framework. The output of the hackathon was a table of standard software metrics, adapted to ontologies and cross-referenced to individual assessment tools from OQuaRE and OOPS! These include measures of ontology quality as well as metrics of interest about an ontology such as the clustering of concepts and the depth of the subsumption hierarchy.

In order to take the lessons from the Ontology Summit 2013 FIBO hackathon and apply these to external ontologies, we would need to identify what are the desired metrics for such ontologies and then apply or adapt the tools to carry out those measurements on candidate ontologies for re-use.

Another hackathon at the 2013 Ontology Summit explored a practical application of the GOEF Methodology [25], a method for ontology evaluation which focuses on the original intended use case of an ontology. The techniques described in this work can also be applied directly to the FIBO requirements for ontology re-use. In the GOEF methodology, the use case is split into functional objectives, design objectives and semantic components.

The 2014 Ontology Summit was titled “Big Data and Semantic Web Meet Applied Ontology” and one stream of work this year explored considerations in sharing and reusing semantic content. Again there were lessons in ontology evaluation, in particular the evaluation of the suitability of ontologies for specific intended uses. Findings included the need to understand the intended use case for ontologies which are being evaluated, modularity considerations, the use of standard ontology design patterns and the value of well-annotated ontologies [26].

The kind of semantic content covered by these explorations extended to formal and informal ontologies, vocabularies and other formats in which meaning is captured. Ontologies modeled in OWL represent a significant proportion of these resources, however there are many communities of practice which use other means, such as extensions of the Unified Modeling Language or other less formal notations, as well as business-facing vocabularies and community-specific languages. The communique published as an outcome the summit includes discussion of some of these issues [4].

3.2 Lessons from the Risk Reuse Hackathon at Ontology Summit 2014

The 2014 Ontology Summit again included a set of “hackathon” activities [5]. The hackathon on ontology reuse [6] focused on risk since this would require integration of concepts across a range of concerns namely impacts, goals, events, and so on.

This hackathon used an ontology design pattern for events which was presented at the Summit [7] alongside a trajectory ontology [27]. The latter was used as the basis for a new ontology describing journeys in order to describe their risks.

A number of re-use techniques were demonstrated in this hackathon: abstraction from available ontologies; extension of patterns, and direct modeling from available data.

The concepts modeled for the travel risk hackathon could in principle be extended further to accommodate logistical concepts and cashflow representations for securities payment structures, as needed for financial risk applications.

There was an interesting observation about the considerations for re-usability of ontologies, as demonstrated by the Trajectories ontology. This ontology has a number of properties with no declared domain or range, these being applied via restrictions.

The logic behind not declaring domains or ranges for properties is that when adding a new property to an ontology, the developer should avoid applying the domain and range for which the property was originally conceived, since it is likely that they may need to re-use that property in another or broader context. At the same time, there is no business case for analyzing what are all the possible domains and ranges for a given property if one is building a single application, and even if these abstractions were identified, it is unlikely that the ontology for a stand-alone application would have those abstractions in place. For this reason, a recognized “best practice” has arisen whereby designers of application ontologies define most properties without a domain and range, and then apply these to classes via restrictions.

However, those properties have no machine readable distinctions between them. While appropriate for a stand-alone implementation, this is not appropriate for a system-wide ontology or for an industry standard ontology such as FIBO. What is best practice for application ontology development, is the opposite of what is appropriate for an industry common language or standard.

This practice will affect whether or how ontologies may be re-used within an ontology which is built to define a common industry language. Many of the ontologies which we would identify and want to re-use will have been designed in this way. Prevalence of this approach to ontology design may also mean that ontologies which are intended for use as industry standards may also in some cases have properties which are underspecified in this way.

Therefore when re-using an ontology in which the properties do not have domain and / or range specified, properties derived from that ontology should be given appropriate domains and ranges in the target model, these being the most abstract classes of thing to which the property may apply or to which it may refer. This will be the case for most ontologies which have been built for a stand-alone application or for a limited number of use cases.

4. Conclusions and Recommendations

The potential re-use of ontologies involves a number of different factors. When evaluating different ontologies for potential re-use some of these factors may need to be balanced against others. For example one ontology may represent an authoritative source for the meanings of some concepts while another ontology may cover the same terms in a more logically complete and consistent way. Some ontologies may have been optimized for use in semantic technology applications while others may have been created with a view to providing common meaning across an industry or an enterprise. Some application ontologies may contain simplifications or short-cuts whereby similar but distinct concepts are conflated, such as countries and jurisdictions.

Reusing semantics published by others means one is able to make use of the knowledge of standards bodies and other communities of practice who understand the concepts in ways which may not be apparent to a non-expert in that field. However, many of the semantic models created by such communities are not made with re-use in mind, and so any methodology for re-use must take account of this. Some communities of practice may not have used OWL or other first order logic in framing their concepts. Some may not use any formal logical notation. The re-use of the concepts in each case requires detailed analysis of the definitions and in literature provided by the originators of the model.

When reusing third party semantics a key requirement is that the re-using party is able to identify the business application context for which the ontology was created. Metadata within such an application is a help in this, but may not exist in some cases. Once the intended meanings are clear these would need to be framed with reference to the upper ontology partitions of the overall ontology, so that similar but distinct concepts can be contained within the same overall model and the relationships between different nuances of meaning are clearly identified.

An interesting question to address with an ontology which is to be re-used, is the domains and ranges of properties. Some ontologies which would be considered suitable for re-use may have properties with no domain or range; these would need to be added using classes which are available in the overall set of ontologies, in order for those properties to be considered meaningful.

Another important consideration is the use of ontology design patterns in ontologies which are to be re-used or referenced. These patterns may also correspond to semantic abstractions of concepts, though some patterns provide commonality in different ways.

It is also important to recognize that ontologies may be framed under different theories about the world. Not all ontologies can be used or referenced directly if these are based in different theories. However, a well-constructed upper ontology may provide the means for at least some seemingly incompatible ontologies to be integrated.

The intended use of the ontology or set of ontologies which is intended to make use of these resources also affects the evaluation requirements, so that for example the considerations when re-using an ontology for a reasoning or semantic querying application may be different to the considerations for re-using ontologies within an industry standard framework such as FIBO.

Creating a business-wide or industry-wide ontology requires some treatment of formal semantics, and this is more than a matter of using a syntax such as RDF or OWL. Where an ontology is intended to provide an industry common “language” to address problems of data standardization, transparency, reporting or risk management, such as in the financial services industry, that ontology should follow established knowledge representation principles to an extent which may not be important for stand-alone ontology applications. Concepts in this kind of ontology must be semantically grounded, and where possible such grounding would take the form of semantically primitive concepts, that is classes and properties which represent the simplest kind of thing in given set of types, for example the simplest thing which is a contract, a transaction, a commitment and so on. Wherever possible, common concepts should be derived from suitable industry communities of practice if the semantics of the model are to be widely reusable.

One requirement for common meaning is the means to unify the different theories that may have been applied in different ontologies which one wants to re-use. A coherent system of semantics helps in structuring the model and gives the business the confidence that the semantics in the overall model can be referenced and mapped to existing data schemas and message models.

There is a wealth of research and information covering ontology evaluation generally, and the principles explored in those resources must themselves be understood and applied selectively according to the intended requirements of a project. The evaluation criteria to be used for an industry standard ontology like FIBO will be very different to the criteria that would be appropriate for integrating a small set of ontologies for a single application.

It would be of value to a number of different industries if there were some kind of cross-industry consensus on potentially re-usable semantic resources. This would be particularly valuable in the case of cross-industry concepts such as business entities, contractual and transaction concepts, mereology, units of measure and the like.

Defining common meaning would not be achieved by creating industry vertical ontologies because concepts in reality do not always respect those boundaries. For example some concepts which may represent specialist knowledge in one industry are widely extended in others. An example of this is contracts, which form the basis of financial securities, insurance products and a host of other industry vertical concepts.

References

1. EDM Council: www.edmcouncil.org
2. OWL 2 Web Ontology Language Quick Reference Guide (Second Edition), W3C Recommendation 11 December 2012. At http://www.w3.org/TR/2012/REC-owl2-quick-reference-20121211/
3. Ontology Summit 2014: http://ontolog.cim3.net/cgi-bin/wiki.pl?OntologySummit2014
4. Ontology Summit Communiqué: http://ontolog.cim3.net/cgi-bin/wiki.pl?OntologySummit2014_Communique
5. Ontology Summit 2014 Hackathon page: http://ontolog.cim3.net/cgi-bin/wiki.pl?OntologySummit2014_Hackathon
6. Berg-Cross, G., Bennett, M.: Reuse Hackathon Summary. At http://ontolog.cim3.net/file/work/OntologySummit2014/2014-04-28_29_OntologySummit2014_Symposium/Hackathon-Reuse_OntologySummit2014_Symposium-Report--MikeBennett-GaryBergCross_20140429.pdf
7. Event ontology. Available at: http://purl.org/NET/c4dm/event.owl
8. FIBO Foundations v1.0 Beta1 specification. Available from www.omg.org
9. Sowa, J.F.: Knowledge Representation, Logical, Philosophical and Computational Foundations. Brooks/Cole, Pacific Grove, California (2000)
10. ISO/IEC 15944-4:2007, Information technology — Business Operational View — Part 4: Business transaction scenarios — Accounting and economic ontology. ISO/IEC (2007)
11. eXtensible Business Reporting Language, `www.xbrl.org`

12. Bennett, M.: Adopting and Extending REA Terms in the Financial Industry Business Ontology: A Case Study. VMBO presented paper, Feb 2014
13. Gruber, T.R.: A translation approach to portable ontologies. Knowledge Acquisition, 5(2):199-220, 1993
14. Gómez-Pérez, A.: Evaluation of ontologies. International Journal of Intelligent Systems, Special Issue: Verification and Validation Issues in Databases, Knowledge-Based Systems, and Ontologies, Volume 16, Issue 3, pages 391–409, March 2001
15. Smith, B.: Ontological realism as a strategy for integrating ontologies. Ontology Summit, 2013.
16. Partridge, C.: Ontology Architecture: Top Ontology Architecture. Ontology Summit 2013.
17. Pinto, Martins: A methodology for ontology integration. At K-CAP'01
18. Bontas, E. P., Mochol, M., Tolksdorf, R.: Case studies on ontology reuse. In Proceedings of the 5th International Conference on Knowledge Management IKNOW05.
19. Simperl, E. P. B.: Reusing ontologies on the Semantic Web: A feasibility study. Data Knowl. Eng., 68, 905-925, (2009).
20. Gómez-Pérez, A., Suárez-Figueroa, M-C., Villazón, B.: Ontological Engineering. Introduction to the Semantic Web Tutorial, ISWC 2008
21. Ontology Summit 2013: Ontology Evaluation across the Ontology Development Lifecycle. At: http://ontolog.cim3.net/cgi-bin/wiki.pl?OntologySummit2013
22. Duque-Ramos, A., Fernández-Breis, J.T., Stevens, R., Aussenac-Gilles, N.: OQUARE: A SQuaRE-based Quality Evaluation Framework for Ontologies. Ontology Summit 2013.
23. Poveda-Villalón, M, Suárez-Figueroa, M.C., Gómez-Pérez, A.: Did you validate your ontology? OOPS! Ontology Engineering Group, Departamento de Inteligencia Artificial, Facultad de Informática, Universidad Politécnica de Madrid. Available at: http://2012.eswc-conferences.org/sites/default/files/eswc2012_submission_322.pdf
24. Tartir, S., Budak Arpinar, I., Moore, M., Sheth, A.P., Aleman-meza, B.: OntoQA: Metric-based ontology quality analysis. IEEE Workshop on Knowledge Acquisition from Distributed, Autonomous, Semantically Heterogeneous Data and Knowledge Sources (2005).
25. Luciano, J., Obrst, L., Stoutenberg, S., Cohen, K., Stanford, J.: Ontology Evaluation: Methods and Metrics. Available at: http://www.slideshare.net/joanneluciano/luciano-pr-08849ontologyevaluationmethodsmetrics-8294436
26. Westerinen, A.: Reuse of Content from ISO 15926 and FIBO. Ontology Summit 2014.
27. Hu, Y., Janowicz, K., Carral, D., Scheider, S., Kuhn, W., Berg-Cross, G., Hitzler, P., Dean, M., Kolas, D.: A Geo-Ontology Design Pattern for Semantic Trajectories. Spatial Information Theory, Lecture Notes in Computer Science Volume 8116, 2013, pp 438-456.