A Case for the Ontological Expression of E-Business Standards

Prepared for EIDX Quarterly Meeting
Semantic Harmonization Session
Menlo Park, CA
December 1, 2004

Prepared by

Kurt Conrad
President, The Sagebrush Group
conrad@SagebrushGroup.com

Brian (Bo) Newman
Founder, KM-Forum
Bo.Newman@KM-Forum.org

Bob Smith, Ph.D.
Professor Emeritus, CSU
Bob@1TallTrees.com

1.0 Introduction

This document is based on work done by members of the Ontolog Forum (http://ontolog.cim3.net) in conjunction with the CCT-Representation project [CCT-Rep], a project to develop a formal ontology based on the ebXML Core Component Types.

One of the primary goals of the CCT-Representation project is to explore the rationale for using ontological modeling and representation languages in the development and expression of E-Business data standards.

This white paper presents an emerging realization of the potential value of ontological engineering when applied to the standards development process. It was developed through a series of ongoing conversations amongst project participants over the course of many months. The authors consolidated and extended their ideas to create the business case that you see presented, here. Please note that while this work is based on the input of many, it is not necessarily the argument that any of them, individually or collectively, would make.

In identifying potential benefits, we considered a number of different approaches for augmenting current standards development processes.

The first approach is the default condition of developing data standards without the incorporation of any visible ontological engineering. This approach served as the basis for identifying the general advantages that ontological engineering could be expected to offer.

The second approach envisions the standards development process progressing largely unchanged, with ontological engineering methods applied after the specifications are largely stable. This reflects the approach used in the CCT-Representation ontology project, where the team is attempting to develop a normative KIF (Knowledge Interchange Format) representation of the already-existing ebXML Core Component Types.

The third approach involves applying ontological engineering in a more systematic fashion, shifting the methodological focus from data modeling to conceptual modeling. While this approach would have the greatest disruptive impact on the committee's work processes, it is also expected to have the greatest comparative advantage.

Implicit in both of the ontological engineering approaches which were considered is the use of a standardized upper ontology. The CCT-Representation project is using SUMO (the Suggested Upper Merged Ontology) [SUMO], which is owned by the IEEE [SUO] and available for public use. Many of the use, reuse, integration, and semantic harmonization benefits which are presented in this white paper result from the use and extension a common upper ontology [Niles].

2.0 Expected Qualitative Impacts

A basic lifecycle model was used to inventory the potential benefits of applying ontological engineering methodologies during the development of standards:

Development and Maintenance
Review and Approval
Publishing
Decision to Adopt
Implementation

In addition, we also identified a number of benefits that are "transcendent" and impact multiple portions of the lifecycle.

2.1 Development and Maintenance Impacts

Standards development is hard work. Most standards bodies work harder than they have to. Not only is getting people to agree often difficult, but getting them to use the same terms and mean the same thing when they use them has been the downfall of many a fine standards body. From an ontological perspective, the basis of any robust data standard is a shared conceptualization: agreement on an underlying conceptual model which is reflected in the resulting specifications.

Most of us are familiar with situations where the committee could not identify the core, underlying issues and/or reach needed agreements, resulting in standards that contained "land mines" that undermined the quality of the standard.

It is expected that the use of ontological engineering approaches would lower overall cost and time to deliver standards. Ontological engineering can illuminate ontological gaps and misalignments among committee members and other interests. This should enable them to accelerate agreement on key terms to more efficiently communicate, in general, and identify and resolve the truly fundamental issues associated with the targeted standard.

Ontological engineering can also be expected to improve the stability of the resulting standards through a number of mechanisms.

> First, better identification and resolution of critical issues can be expected to improve the stability of the resulting standard, directly.
>
> Second, an improved ability to balance competing interests and reduce the tendency to couch implicit policies, hidden agendas, and unaligned operational models in inherently ambiguous language will improve the quality and transparency of committee’s decision making process.
>
> Third, the development of ontological formalizations will allow the semantics of the specification to be expressed in a more complete fashion, further reducing ambiguity and improving knowledge transfer from the committee to downstream communities.

Expressing specifications as formalized ontologies will also enable the standards bodies to make use of more advanced tools that can automatically perform various quality and constancy checks, create conforming alternate renderings, reduce drudgery, and further improve speed and quality [Denno].

As standards are subject to a wide variety of potentially disruptive interpretations and change vectors, the introduction of ontological engineering can be expected to help, not only with the initial development of the standard, but also in its downstream maintenance.

Tool and technology changes can introduce new policy objectives, semantic requirements, and emergent opportunities. It will be much easier to evaluate an ontologically-complete specification for potential impacts, identify the root causes of semantic conflict [Pollock], and determine which, if any, changes to the standard are advisable.

Just as technology changes are potentially disruptive, changes in the natural ontologies that comprise the operational context for standards can also be expected to drive misalignments that necessitate revisions. New business models, theoretical advancements, economic trends, and social and political change – both planned and unplanned – can dramatically change the operational context and, thus, relevancy of a standard. For example, privacy rights have taken on whole new meanings after 9/11.

This is not to say that these benefits can be realized without additional resources. Clearly, the addition of ontological engineering tasks to a standards development effort will drive the explicit awareness of a series of knowledge gaps that will need to be filled either with training or the addition of skilled personnel to the effort.

Critical to the understanding of these knowledge gaps is the distinction between data modeling and conceptual modeling. The risk of only relying on data modeling to create the standards in question is that it is a technique that may be unable to illuminate the critical issues that should be addressed to create a complete, well-articulated, and stable standard. Conceptual modeling, in contrast, is a different activity that requires a different set of skills. While the addition of those skills can be expected to increase initial development costs, they can also be expected to reduce overall lifecycle costs for a standard, especially if they can be amortized across multiple standards development efforts.

The importance of conceptual modeling also points to a significant difference between the two implementation approaches which were considered. An approach that introduces ontological engineering late in the process – after the standard has "stabilized" – to express a data model more formally is either unlikely to benefit from the additional power of conceptual modeling, or is almost certain to identify conceptual modeling issues that destabilize the data models which were produced [McGuinness and Guarino].

2.2 Review and Approval Impacts

Just as the introduction of ontological engineering practices to standards development can be expected to increase the efficiency of the development process, the improved clarity of the resulting specifications can be expected to simplify review and approval.

The same automated tools that check consistency, verify quality, and provide alternate renderings for the development community will save time and effort for reviewers, as well.

But perhaps more importantly, the resulting specifications should be easier for people to read and understand, because they are based on higher quality conceptual models that have fewer semantic gaps and breakdowns and can, therefore, be expressed using clearer and less complex natural language descriptions. It can also be expected that the supporting documentation will better reflect social consensus and will be less likely to suffer from divergent communication styles which can be traced to multiple authors working from independent, unaligned conceptualizations.

2.3 Publishing Impacts

A viable standard has to be read and understood by a global community. The development of formal logical models dramatically simplifies a wide array of localization issues for the standards committee, multinational users, and transnational supply chains.

Where multiple implementations are anticipated to rely on different expressions of a given specification (relational database schema, XML schema, ASN, etc.) implementation-specific expressions can be easily derived from the formalized ontology, retaining as much expressivity as the targeted language will allow, and ensuring consistency and validity of those constructs which have be translated.

2.4 Impacts on Decisions to Adopt

Adoption of standards is a form of reuse. All too often, reuse is hampered by an inability to efficiently understand the artifact at a level of detail sufficient to comfortably envision its behavioral implications. The explicit articulation of a standard is often matched by an even greater body of implicit and tacit knowledge which was developed by the standards body, much of which deals directly with the decision points, tradeoffs, and optimizations around a particular theory of operations. that are so critical to the decision to adopt.

Faced with what appears to be a daunting learning exercise (with an uncertain outcome), many choose, instead, to develop their own approach. For those who commit to applying a standard, the knowledge gaps can be staggering. In many cases consultants, companies, and even entire industries have emerged to fill these knowledge gaps.

As has already been established, an ontology-based standard is likely to be more complete, have greater internal consistency, and be more explicitly expressed. All of these factors speed the learning curve for decision makers. Furthermore, the articulation of semantics as specific axiomatic properties can be expected to make it much easier to understand and project the behavioral implications of specification. By exposing the underlying theoretical foundation of the standard, potential users will have a much easier time determining where their own operational context matches and/or diverges from the design baseline.

By providing “full disclosure” to the “buyer” of a standard, the decision to adopt becomes a much easier task. Assuming that the standard has widespread applicability, this should result in much more rapid adoption curves, with a corresponding multiplier effects on the development of associated tools, service providers, and realized business value.

2.5 Implementation Impacts

Consistent with the general benefit of improved knowledge transfer, ontology-based standards will make it easier to “educate” those individuals who are tasked with implementing the standard. As was described earlier, the ability to automatically derive implementation-specific subsets of the standard can also be expected to reduce the time and cost to implement.

Data migration can also be a significant implementation issue. The richer and more exact semantic specifications that an ontology-based standard would provide can be expected to make it much easier to relate existing data sets and associated schemas to the new data model to determine specific translation and conversion requirements. Likewise, the existence of more complete target specifications are likely to significantly reduce data conversion costs.

More importantly, the addition of formal ontological components to automated systems can be expected to dramatically reduce the cost of future migrations and implementations. The use of ontologies represents the next step in loose coupling: a long-term trend to abstract out critical business rules and support logic in a way which enables the use of more generic (and hence stable) computational engines. Given that the ontologically-expressed logic can be revised without destabilizing the code base, applications can be “dynamic”, both in terms of handling greater complexity and context-sensitivity and because the maintenance process is simplified.

By making the integration process more predicable and reliable, it is even conceivable that systems will become self-describing and self-integrating [Ray].

2.6 Transcendent Benefits

Many of the potential advantages of ontological engineering in standards development are not specific to a given set of interests but impact the entire standards lifecycle, thus transcending individual lifecycle phases.

The use of an upper ontology in standards development can be expected to have significant benefits to all who interact with the standards. An upper ontology is a set of ontological formalizations that are general in nature and have widespread applicability. They are typically extended with a set of application/domain-specific formalizations (mid-level and domain ontologies).

An upper ontology provides an initial set of conceptual models, which – through their reuse – can dramatically simplify the complex task of shifting from a data-modeling methodologies to conceptual-modeling methodologies. In addition, as an upper ontology represents a consensus around these core conceptual models, reuse of an upper ontology reduces the potentially disruptive impacts of individual conceptualizations on the conceptual-modeling process.

The reliance on an upper ontology as a conceptual anchor point can be expected to have benefits not only within the development process but on through implementation, where the need to extend a standard would be more robust and consistent if done using the same upper ontology and formalization techniques used to develop the standard in the first place.

Perhaps most significant over time, to the extent that the same upper ontology is used across versions of a standard and across different standards, integration and interoperability should be much easier, potentially providing “out of the box” semantic harmonization.

Semantic conflicts among related standards are all too common. To avoid a combinatorial explosion of mapping and integration points associated with harmonizing multiple specifications, many are currently seeking to utilize a hub and spoke approach to semantic harmonization, where a single standard is chosen for the hub and all other specifications are mapped to it [Korsgaard]. This raises an important set of questions. Which specification should serve as the hub? What are the requirements to be an effective hub standard?

As the old adage goes: You can’t manage what you can’t measure. Semantic conflicts are often the consequence of implicit and tacit semantics. As such, it is not unreasonable to expect the explicit formalization of semantics to be a necessary precursor to effective semantic harmonization and that standards that result from a data-modeling approach – no matter how rigorous – will be difficult, if not impossible, to harmonize.

Likewise, any specification which seeks to serve as a hub for semantic harmonization should be expected to be delivered in a form where the formalized semantics are explicitly expressed, typically as logical axioms. In addition, to be a viable hub, a standard needs to be based on conceptual models that are internally consistent and complete enough in scope to enable sufficient mapping to the targeted "spoke" specifications to meet functional requirements.

Fundamentally, this points to the need to integrate and harmonize various data and conceptual models throughout the standards lifecycle. Just as standards bodies are often formed with multiple draft proposals, each with their own underlying conceptual framework, so too does the need to relate and incorporate multiple conceptual models exist after the standard is published (e.g., as part of an implementation, harmonization, or maintenance effort). Fundamentally, the development and application of standards can be viewed as an series of ongoing exercises in ontological alignment.

3.0 Conclusions

- Standards development is hard work. Most standards bodies work harder than they have to.
- Standards-setting bodies are susceptible to ontological gaps and breakdowns that hamper progress and threaten both the expressiveness and semantic stability of the resulting specifications.
- Ontologically formalized standards should be easier to adopt and provide numerous migration, integration, and interoperability advantages.
- A systematic approach that incorporates conceptual modeling, ontological engineering, and the use of a standardized upper ontology in the development of standards will yield the greatest benefits.
- Adopting an ontological engineering approach will create a variety of knowledge gaps which will need to be addressed, but should improve the flow of knowledge both within the standards committee and to downstream communities.
- Businesses and other communities can be expected to enjoy standards that are more stable, easier and less expensive to adopt, and provide more rapid returns on investment.

Ontological engineering is being applied in the design and development of computing applications with increasing frequency [Obrst]. So too, we expect that these ontological engineering approaches are sufficiently compelling that they should be adopted by the committees that design the standards that these computing systems conform to.

The potential benefits which have been outlined in this white paper need to be tested in actual standards development efforts. We solicit committees to contact the Ontolog Forum to identify opportunities for incorporating ontological management and engineering methods in their standards development projects.

4.0 Contributors

The following Ontolog Members have made significant contributions to the articulation of this business case:

Mike Brenner
Pat Cassidy
Peter Denno
Ed Dodds
Robert Garigue
Brand Niemann
Monica Martin
Leo Obrst
Adam Pease
Steve Ray
Evan Wallace
Peter Yim

In addition to the members listed above, the CCT-Representation project has also benefitted from the contributions of:

Nenad Ivezic
Bill McCarthy
Tim McGrath
Garret Minakawa
Sue Probert
Alan Stitzer
Susan Turnbull
Holger Knublauch

5.0 References

[CCT-Rep] Ontolog Forum, “CCT-Representation Project Home Page”, http://ontolog.cim3.net/cgi-bin/wiki.pl?CctRepresentation

[Denno] Peter Denno, “Exploratory Development of a UBL Validation Tool”, http://ontolog.cim3.net/file/work/CCT-Representation/wip/UBLTool-NIST.ppt

[Guarino] Nicola Guarino and Chris Welty, “Conceptual Modeling and Ontological Analysis”, http://www.cs.vassar.edu/faculty/welty/aaai-2000/sld001.htm

[Korsgaard] Stig Korsgaard, Alan Stitzer, and Sue Probert, “TBG17 Core Component Harmonisation Presentation”, http://ontolog.cim3.net/file/work/CCT-Representation/Core_Components_Harmonisation--TBG17_20040310.ppt

[McGuinness] Deborah L. McGuinness, “Conceptual Modeling for Distributed Ontology Environments”, http://www.ontology.org/main/papers/iccs-dlm.html

[Niles] Ian Niles and Adam Pease, “Towards a Standard Upper Ontology”, http://home.earthlink.net/~adampease/professional/FOIS.pdf

[Obrst] Leo Obrst, “Ontologies for Semantically Interoperable Systems”, http://ontolog.cim3.net/file/resource/tutorial/OntologiesForSemanticallyInteroperableSystems-ONTOLOG--LeoObrst_20040115b.htm

[Pollock] Jeff Pollock, “Semantic Conflicts Chart” and “Semantic Conflict Solution Patterns Chapter”, http://jtpollock.us/semanticconflicts/

[Ray] Steve Ray, “Tackling the Semantic Interoperability of Modern Manufacturing Systems”, http://colab.cim3.net/file/work/SICoP/Semantic_Technology_Conference_2004_0908/Ray_20040909.ppt

[SUMO] Adam Pease, “Ontology Portal: The Suggested Upper Merged Ontology Web Site”, http://www.ontologyportal.org

[SUO] IEEE P1600.1 – Standard Upper Ontology Working Group, “Home Page”, http://suo.ieee.org/