The EDM Council Semantics Repository - Considerations in Ontology Alignment

Michael G Bennett

EDM Council Inc.

mbennett@edmcouncil.org

Abstract. This paper describes the EDM Council Semantics Repository and in particular the steps taken to align the content of financial industry semantics with common cross industry terms. The long term goal of the Repository is to be able to align with and reuse semantics from industry led initiatives across a range of industries. This paper describes the initial approach to alignment, the lessons learnt and the plans for more formal alignment in the future.

1 Introduction

The Enterprise Data Management Council [1] is a not-for-profit industry association in the financial services sector, based in Washington DC. Members include leading banks, brokerages, data providers and data utility vendors. The Semantics Repository [2] is a two year project sponsored by the Council which has recently been elevated to Beta status.

In the course of creating the Semantics Repository we identified the need to incorporate terms from a wide range of industry domains outside of the financial services industry. Initially we created the necessary terms as a separate upper ontology in order to make reference to these, but in the longer term we would prefer to align more formally with the semantics of recognized industry bodies. This paper looks at a number of considerations in implementing this.

2 The Semantics Repository

The EDM Council's Semantics Repository is an industry collaborative project to define common terms and definitions for financial securities terms at a business level. This is achieved using semantic technology in order to provide a technology neutral view of terms and definitions. This is described in more detail in [3].

The repository is built using a conventional computer engineering tool rather than semantic technology tools. The output of the Repository is displayed on a website, with diagrams and tabular information generated from the tool. This website is used to communicate the terms and definitions to the industry and to elicit feedback. Regular web-based review sessions are held using the underlying repository tool to add new terms and to review and sign off on existing terms and definitions. The underlying philosophy of the Council is that terms and definitions are not held out as being complete and correct until industry subject matter experts have agreed on them.

2.1 Project Scope

The EDM Council Semantics Repository covers the following scope:

- Traded financial instruments (bonds, shares, traded derivatives etc.)
- Over the counter derivatives
- Market indices, rates and indicators
- Collective investment vehicles (funds)
- Business entities and legal entities

For the above classes the following types of term are represented:

- Static or reference data terms
- Date and time dependent terms (market data, pricing, analytics etc.)

Also to be covered in future releases are corporate actions and events, securities transaction processing terms and loan terms.

2.2 Output Formats

The requirements for the Repository are met by the following output formats:

- Tabular and spreadsheet report output, capturing the terms, definitions, synonyms and property details
- Simple "boxes and lines" diagrams, in which there are no language artifacts that would be unfamiliar to business experts.

2.3 Underlying Tool configuration

The output formats are generated from a UML modeling tool called Enterprise Architect, from Sparx Systems [4].

The available choices were to extend UML in our own way or to adopt a standard set of UML extensions for OWL and adapt these as necessary for business reviewability. In line with the EDM Council's commitment to adhere to standards wherever possible, we based the work on an early draft of the Ontology Definition Metamodel (ODM) standard [5], published by the Object Management Group [6]. This was adapted to provide the user diagram formats required.

2.4 Theory of Meaning

The underlying theory is based on set theory semantics, as generally used in OWL modeling. This is described to business subject matter experts without reference to OWL construct names, as follows:

- Everything is a kind of Thing
  - This is formalized by having everything in a taxonomy with "Thing" at the top
- Every Thing has facts which distinguish from other Things
  - Simple facts: numbers, text, dates etc.
  - Relationship Facts: relate one kind of thing to another by way of a relationship, e.g. Share confers ownership of Equity
- Each kind of Thing is defined as a set theory construct - if something has those facts true of it, then it is a member of that class of things.
- Sets which may not have common members are shown as mutually exclusive
- Some sets are logical unions of two or more sets

In this way we are able to explain the models to business domain experts without reference to OWL or other semantic web terminology.

Meaningful representations of each kind of thing are thereby defined by way of facts which disambiguate that thing from all other things in the model. In addition, formal human readable definitions are given, and these are reviewed and agreed upon by industry subject matter experts.

3 Modeling Philosophy

3.1 Taxonomy Structure

There was a choice to be made in terms of whether the concepts in the financial instruments model were directly modeled as sub-types of the universal class "Thing", or were descended from more primitive kinds of thing.

Given that new financial instruments are being invented all the time, it was considered important that we relate each class to the simplest kind of thing that it is. This would enable new concepts to be defined in the future with reference to these primitive concepts and without the brittleness often associated with conventional data models.

This need for an "atomic" set of concepts from which to derive terms, led us to recognize the need for common primitive terms which are not themselves part of the financial securities domain. For example all securities are contracts which is a legal concept; debt securities are defined in terms of a relationship to the real thing called "Debt" which is defined in accountancy literature, and so on.

Similarly, there are facts about many securities which relate to geographical concepts, for example bonds may be issued by sovereign nations and by municipalities. Similarly concepts to do with time, mathematical formulae, transactions and the like, would all need to be specialized to define terms in securities models.

3.2 Upper Ontology Requirements

We needed to support the simplest primitive types of thing for the full range of concepts in the financial securities problem space, both for common ancestry of securities, transactions etc. themselves, and for common ancestry of the concepts which are the ranges of the various object properties. For example an equity or share instrument gives the holder some share of the equity in a company, and that fact necessarily refers to the concept of "equity" as defined in the accountancy literature. This will be a specialization for the class of equity represented by that share (ordinary equity, senior equity etc.) which in turn is a specialization of the common concept of equity itself, as defined in the accounts equation which relates assets, liabilities and equity.

In modeling the full range of concepts needed for financial instruments, both for securities classes and the ranges of object properties, we identified the following subject areas in which semantically primitive concepts needed to be defined:

- Dates and times - including complex specifications of dates;
- Countries, cities etc - as issuers of sovereign debt, municipal bonds etc.;
- Law - securities are contracts, with contractual terms, jurisdiction etc.;
- Mathematical constructs - for example formulae for determining variable interest;
- Financial Accounting terms - interest accrual, equity, debt, payments, money;
- General business terms - transactions, securities trading, parties etc.;
- Legal and Business Entities - legal persons, companies, individuals, trusts, governments;
- Information - identifiers, classification schemes; market data, prices and so on;
- Events - including payments, default (non-payment) events, corporate actions etc.;
- Activities - terms about processes such as the securities issuance process;
- Risk - different types of risk and the factors that go into analyzing these.

The decision was made to put together models of the semantic primitive concepts in these domains, and to identify and start to re-use suitable material from the relevant industry bodies as soon as such material could be identified. Some of this material would need to be reverse engineered from formats that do not embody common logic, for example from data models.

3.3 Upper Ontology Lattice

Having determined that the upper levels of the ontology should be common semantic primitives like contracts, geographical entities and so on, we then looked at whether to define each primitive concept directly as a sub type of "Thing", or to partition the top of the model.

A good partitioning structure can be found the "Knowledge Representation" lattice of theories in the book of the same name by John F Sowa [7], known as the KR Lattice. The top layer of the KR Lattice was adopted as the framework within which to define the primitive concepts we needed.

The partitions defined in the KR Lattice are fairly fundamental in nature, and can therefore be explained coherently to business domain experts. These are:

- Independent/Relative/Mediating;
- Continuant/Occurrent;
- Concrete/Abstract.

These are explained as follows:

Independent/Relative/Mediating: this distinguishes between a thing in its own right, a thing defined in the context of some role it plays, and the context in which that relative thing is defined. For example a business entity is an independent thing, whereas "Issuer" or "Underwriter" are parties, a kind of relative thing. A mediating thing is the context in which some relative thing is defined, such as securities issuance or underwriting.

Continuant/Occurrent: this distinguishes whether something has some ongoing existence across a period of time or is defined with reference to some point in time.

Concrete/Abstract: whether something is a real material thing or a non material construct. Note that many non physical things such as dematerialized securities are regarded as concrete here. The partition of "Abstract Thing" is reserved for things which by their nature are necessarily abstract, such as goals.

We also added concepts for parts and sets, and for time constructs.

The model was structured with these lattice categories as sub-classes of "Thing". No axioms are given for these, other than the identity relation between Relative Thing and Independent Thing, the context relationship between Mediating Thing and Relative Thing, and the whole / part relations. We did not use logical axioms to formally define the partitions.

All other terms are defined in relation to the concepts in this lattice. In principle, each of the concepts at the top of the taxonomy of terms, such as contracts, goals, actors and so on, is defined as having one parent from each of the above three partitions. In practice the concrete class was not always applied, but is implied.

3.4 Archetypes

We recognized that for each simplest or atomic concept there were necessary facts or axioms which define what that thing is, and that are either inherited or specialized for all sub-classes of that thing. These were formally defined as "archetypes" and given their own unique appearance in the modeling framework, using UML stereotype indications.

The archetypes all belong in the upper ontology. These are generally terms from outside of the financial securities universe and therefore would correspond to concepts in standards from other industries. Examples include Contract, Equity, Party, Transaction and so on. Where possible, we defined the necessary facts about each archetype with reference to existing standards or ontologies.

4 Upper Ontology and Standards

4.1 Existing Standards Commitment

Our commitment to standards leads us to want to reference and reuse existing standards wherever possible. In addition, we took the view that the provenance of semantics should not lie with self-proclaimed ontologists but with the industry bodies responsible for the relevant terms. This means that while we needed to create upper ontology material from which to derive the terms in the financial instruments model, we would ideally want to replace these with standards material at the earliest practical opportunity.

4.2 Initial Approach

In order to be able to populate the Semantics Repository with terms for securities and other financial instruments, we decided to create the required upper ontology first, and seek to align with existing standards later.

In doing this, we identified a small number of sets of terms which we were able to use right away. These were included in the main body of the upper ontology sections for the time being. These were treated as follows:

Time: there is a W3C Time ontology [8] which covers the most fundamental concepts of time, such as instant and interval, along with a number of terms that were not relevant such as calendar scheduling. This was used as the basis for a much broader Time upper ontology section. Additional requirements for time terms in financial securities were identified by inspection of the derivatives messaging standard FpML [9], most of which are also relevant for non derivatives time terms such as debt interest payment schedules.

The required time terms included periods of time specified in days, offsets specified either in business days or calendar days, and so on. This required that we segregate terms which are real, "calendar" time, from terms which specify a formula or other arrangement for working out some unknown future date. Another common feature of financial securities time terms is the specification of date roll rules, such that when a scheduled payment falls due on a non working day it is rolled forward to the next working day, and when a payment which is rolled forward lands in the following month it is rolled back to an earlier working day. These terms all take the form of specifications of times and not times themselves. Most of these terms are also denominated in days rather than in hours, minutes and seconds.

In the initial versions of the Repository the above time terms were specified in relation to the core W3C Time Ontology terms by including those terms as the core of the Time section. Additional higher level terms were also needed above and alongside the top of the W3C Time Ontology, for example for specified time and for time specifications.

Financial (Accounting): The terms which form the ranges of object properties of bonds and equities include basic accounting concepts such as debt and equity. These have well defined relations in the accounting literature, such as the well known "Accounts Equation" [10] that relates assets, liabilities, equity and additional paid in capital.

The eXtensible Business Reporting Language (XBRL) standard [11] defines terms in this domain. The standard itself does not provide the vocabulary but provides a method for creating vocabularies (called "taxonomies" in XBRL) for various financial reporting jurisdictions.

Meanwhile the XBRL standard has the potential to create meaningful relationships among terms. It was not practical to interrogate those relationships in the original and see if they represented formal logical axioms. Therefore we identified the basic accounting concepts with reference to XBRL and defined the relationships in accordance with the accounting literature.

Geographical: We did not identify a suitable geographical ontology at the start of the work, but later we came across the UN-FAO ontology [12]. On the basis that this is the most authoritative body for country information, we felt that this should be treated as the source for countries semantics.

In a later draft of the Geographical upper ontology we therefore included some terms from this ontology. The scope of the FAO Ontology does not include intra-country information such as states / provinces, counties, municipalities, cities and the like. Also the core concept in the FAO Ontology is Territory not Country, so we needed to identify whether the scope of this term is identical to the scope of "Country" as used in the Repository. This corresponds to the meaning used by the standard ISO 3166 which allocates codes to countries. If not, there is a non equivalent relation to be identified between our Country term and the FAO Territory term.

Another interesting issue was that the FAO Ontology was not written with the assumption that it would be used in other business contexts, so there may be some interesting challenges in integrating this material. It was noted for example that some aspects of the ontology were addressed using enumerated data ranges, which would be modeled as classes in our ontology.

Transactions: One ontology we used was the Resource Event Agent (REA) Ontology [13]. This ontology provides an academically well grounded account of the terms needed to model transactions. We did not include this in the earliest drafts of the Semantics Repository, but added REA content when we started work on over the counter derivatives, since these are fundamentally based on transactions.

We were able to create derivative transaction models, both for simple derivatives and for swaps, based on this high level model. That is, the fundamental concepts in the REA Ontology were able to form the basis of a "grammar" of primitive concepts and relationships between these, that could be specialized and could hold true for all the transactions we needed to model. There were also some terms in the REA Ontology which we did not use.

Events: Early on in the modeling process we made contact with the owners of the DOI Indecs standard for digital rights management [14]. This is a diagrammatic standard which has at its core the idea of a "context", this being what was considered necessary to give terms some meaning. On analyzing this, it was clear that what Indecs defines as a "Context" for digital rights management is something which has a time and a place. Something which has a time and a place is also known as an Event. We therefore created a model of Event which was fundamentally based on the Indecs model but used the word Event instead of Context. This was then further extended to provide the concepts of Activity, and then of processes and process steps as kinds of activity. In this way we were able to define all the building blocks of process models, in a semantic format based on logical axioms.

4.3 Lessons Learnt

Based on the above initial experience in incorporating known standards material, we have identified the following possible issues that may have a bearing on formally importing and aligning ontology material:

- Not all standards are in a semantic format - many will be vocabularies in which meanings are asserted only, or are grounded in human readable definitions rather than in any formal logic;
- Some standards (e.g. XBRL) are defined in a format which does potentially support semantics, but the internal logic of the terms and relationships is not amenable to inspection and validation as being meaningful
- It is possible to create something in OWL and call it an ontology, which is in fact a logical data model expressed in OWL syntax.
- Some ontologies may be created for specific purposes and were not written with a view to being reused or referenced in business conceptual ontologies

- Many ontologies are grounded in different theories of meaning or have different underlying philosophies which are incompatible with the approach taken with the Semantics Repository. For example some ontologies make strong claims about three versus four dimensional theories, or claim to be extensional rather than intensional.

5 Alignment Considerations

To adequately deal with the above issues, we need some formal means of dealing with the nature of ontologies and other vocabulary resources, such that it is possible both to import and to reference them in the Semantics Repository. This section looks at some of the potential issues and possible solutions.

5.1 Non Semantic Standards

Many standards bodies have vocabularies of terms and definitions which are not stated within a formal ontology. Typically, a group of industry participants get together and define a message schema with terms and definitions, for machine to machine data interchange. These include meaningful business terms, but generally these become designs which implement a vocabulary rather than being the business vocabulary themselves. For example, common data structures may be combined into common message elements, without regard for their semantics. This is good design, but good design leads to weak semantics and it is not practical to overload a message design with business semantics.

Some standards take a positive step towards resolving this by creating a separate, logical model of the message content. Again, a logical model is design, not business semantics, however the meanings of terms are more easily discerned in a logical model and a greater number of terms in the model may have clearly defined semantics. Terms which have only one intended meaning may therefore have one unambiguous definition, so that it becomes a simple matter to axiomatize that definition in a formal logical notation such as OWL.

5.2 Application-Specific Ontologies

The fundamental premise of the EDM Council Semantics Repository is that this provides a business conceptual model of industry terms and definitions, which is distinct both from a logical and a physical model of technology designs such as database schemes or message models. This is framed in terms of standard engineering design methodology, which segregates logical, physical and conceptual models. An example of this is the Zachman Framework [15]. There are a number of well documented requirements in the engineering literature, which a conceptual model must satisfy. For example it must be formal, it must be independent of any technical implementation, and it must be understood and owned by the business. In other words, part of the art of creating an ontology as a business conceptual model, is the art of not designing something. The end result must be owned by the business alone.

However this is not the only way in which semantic technology is used. A great number of semantics applications are written with a particular technical goal in mind, such as the ability to efficiently interrogate a knowledge base using semantic querying. These technical requirements include considerations such as decidability, efficiency and so on. As such, these are designs, and not business conceptual models, even though they may be in OWL.

It should therefore be anticipated that some OWL based ontologies have been created with technical considerations in mind. However, unlike logical model designs, models which are in OWL are specified in formal logic. The expressive power of that logic may have been explicitly limited in order to meet the design requirements of the application, but the business concepts will be specified using formal axioms which define meaningful set theoretical constructs. This means that some imported ontologies may not have the level of detail which we would use to fully define all the necessary facts about a given kind of thing, but the facts which they do specify will be meaningful.

5.3 OWL Models

It is possible to take a logical data model and import it into OWL. The result is a logical data model in the OWL syntax. The use of OWL syntax does not make it an ontology. By extension therefore, it is possible to find an OWL model described as an ontology, which is in fact not an ontology at all.

Similarly, is it possible to find an ontology model built along similar lines to the "Pizza" ontology tutorial example [16], in which quite low-level terms are shown as direct sub-types of the universal class "Thing", without regard to possible intermediate terms. This particular ontology is intended as a tutorial exercise only, but it is possible for modelers to have replicated this approach in order to satisfy a one-off requirement.

These and other considerations raise the question of what makes something an ontology, and what makes it a good ontology for reuse. As a minimum the set theoretical semantics of OWL [17] should have been adhered to. That is, each class in the model represents a set theory construct which corresponds to some real set of things in the problem domain. There may also be appropriate choices of modeling pattern. An example is the common use of enumerated data ranges where in fact what is enumerated should not be data literals but types of thing. This is a common challenge when reverse engineering data models, but is also seen in some OWL models.

These considerations require some formal approach in aligning the concepts in an imperfect or incomplete ontology, with the formal logic required to axiomatize the terms we would need to refer to.

5.4 Differing Philosophies

A number of commentators describe conflicting theories of reality, such as three dimensional versus four dimensional modeling approaches. Trying to refer to a three dimensional ontology from within a four dimensional one, or vice versa, would appear to be problematic.

In fact, it should be possible to resolve these issues by not being dogmatic. For example, in a three dimensional ontology a "Continuant" would be regarded as something with three dimensional extent, which continues over time, as distinct from an "Occurrent" which takes place at some point in time. A four dimensional ontologist would argue that in their way of modeling the world, there is no such thing as a continuant and that everything has four dimensional extent or volume, with the time dimension being treated exactly the same as the three spatial dimensions.

This may be less of a problem than it seems. The problems of integrating different ontologies is dealt with in detail in John F Sowa's recent paper on sharing and integrating ontologies [18], which makes specific reference to the 3D versus 4D problem. The point made by Sowa is that the underlying theories may be significant within a given model but these models cannot contradict observed reality, which therefore defines a common boundary at which such claimed inconsistencies cannot exist.

If a definition of "Continuant" may be framed which is independent of 3D and 4D notations, then that single meaning for Continuant subsumes both models. If "Continuant" is one of the top level partitions in the ontology framework, then the semantics of this term may be grounded in written definition rather than in logical axioms, and the individual model versions of the term have their meanings defined in relation to this term and without reference to one another.

This is the approach we have taken in the Semantics Repository. The lattice partitions are defined textually but are not axiomatized in formal logic, so in effect the meanings of these terms are not semantically grounded but the meanings of other terms are grounded in relation to these.

In general, different models may be expressed in different kinds of formal logic but it should be possible to relate these to a common logical framework. For any given pair of ontologies, it must be possible to identify a common ontology which subsumes both and which contains the common concepts from which both are specialized. For more on this, see also the note on the Ontolog Forum from John Sowa [19].

5.5 Comparing Lattices

In general, any model expressed in some variant of logic will contain a lattice or hierarchy of terms which go from the most general (the universal set of which everything is a member) to the most specific (the empty subset of which nothing is a member). Adding axioms to classes makes them more specialized, so for example adding the axiom "has backbone" to a sub-class of "Animal" defines that sub-set of animals which we call vertebrates. Removing axioms from classes makes them more general. The class with no axioms is the Universal class.

A third possibility is when a class in one lattice or hierarchy has no more and no fewer axioms than a class in another lattice, and is in fact a representation of the same real world thing. This means that the two terms may be related to one another simply by re-labeling - the second item is simply a relabeled version of the first [18].

In the Semantics Repository we have tried to use terms directly if they are the same thing rather than relabeling and having two classes with the same meaning in different namespaces, but this will not always be possible. For example, two ontologies which we may want to refer to, could both have a term for the same meaningful thing, such as country or legal entity.

Where this situation arises, we may decide to make use of the Simple Knowledge Organization System (SKOS) standard [20], which provides the terms with which to define kinds of relationship between concepts in different knowledge resources, including ontologies.

5.6 Top Level Lattices

Meanwhile the existing material in the Semantics Repository is already defined such that every top level primitive concept is itself descended from terms in each of the three lattice partitions as described earlier. This includes terms which we wish to replace with terms derived from authoritative standards bodies. Meanwhile, we have also defined these terms as archetypes, an ontology decoration which is not supported in other modeling formats and therefore won't be present in material from other ontologies.

This means that two approaches are available to us:

- Define parallel sets of terms, with the ones in the external ontologies being referred to via some equivalence relationship (for example using SKOS terms) from existing terms in the upper ontology;
- Incorporating copies of the external ontologies directly, but adding missing axioms, archetype tags.

The second approach is preferred as it results in a single coherent model, however we need to make clear which axioms belong in the original model. For this second approach, we would also need to remove the "Thing" at the top of each external ontology copy, and replace this with relations to the lattice concepts used in the model. This is similar to the SKOS "Top level concept" except that there will be more than one.

6 The Future

6.1 Overview

The EDM Council Semantics Repository has been built from a pragmatic, business point of view, and is a repository of business terms and definitions not a platform for semantic applications. In order to support the formal definitions that we needed to create for financial securities and related terms, we needed to define upper ontology material, but we intend to replace as much of this material as possible with terms derived from the appropriate industry standards bodies, assuming that these can be defined in logical, semantic terms.

6.2 Ontology Registration

The Council intends to register the Semantics Repository with the Open Ontology Repository (OOR) project [21]. However, the Semantics Repository is maintained in a UML file format using OWL constructs, and not natively in an OWL or Common Logic format. There is work ongoing to provide an OWL version of the Repository but this will not become the main point of control of repository content. Therefore it is hoped that at some point the OOR project will provide the capability to register the Repository in its native format as a binary UML file.

Also unlike many ontologies, the use of and derivation of terms from common, non financial industry terms is essential to the very nature of the Repository. There are almost no terms which exist in their most primitive form as financial industry terms. This means that the question of sharing and integrating ontology material is central to the nature of the Semantics Repository itself.

6.3 Additional Ontology Features

In the course of developing the Semantics Repository we have also identified a number of additional ontology features which are not intended to be covered in the OWL syntax but which are essential to the way terms are managed in this repository, such as synonym. Some of these terms can be dealt with by the SKOS standard, as can terms relating non matching relationships between internal and external ontology terms.

We have also identified future extensions such as classification facets (identifying what facts about a set of sub classes was used to define that set), and the indication of exhaustive sets of sub classes. For these we would hope to see some emerging consensus whereby the terms that we all wish to add to ontologies are mechanized in a consistent way, for example as a standard set of OWL annotation properties.

One feature is the identification of simple types of concept as "Archetypes". There is no reason why this should not be adopted, but it is not something other ontologies have attempted to date. Most of our archetypes are terms in external ontologies since this is where the most primitive form of each meaningful concept is to be found.

We therefore need to be able to refer to concepts in ontologies developed by different industry groups, and enhance these with our own meta-ontology terms, including identifying the archetypes and ensuring that all the necessary axioms are in place.

6.4 Ontology Analysis

Another thing which we hope to see is some common consensus within the ontology community about what constitutes a well formed ontology. That is, it should be possible to identify for any given ontology how well the standard set theory semantics and formal logic have been applied, and provide workarounds for ontologies which have shortcomings in this regard. It should also be clear whether ontologies conform to a common model theory in terms of their correspondence to items in the problem domain.

It would also be beneficial to be able to formally analyze and identify the ways in which different theories have been applied within external ontologies, such as three versus four dimensional modeling. This may make it easier to relate the terms in these ontologies, to the common lattice terms.

6.5 Ontology Import or Cross-reference

We need to explicitly refer to external, industry-owned ontologies directly from within the EDM Council Semantics Repository so that the Repository itself provides a framework in which the meanings of the financial securities terms are fully defined.

External ontologies by definition will have a universal class ("Thing" or equivalent) at the top of their hierarchies. In order to relate ontologies to one another, and the financial industry terms to these, these top level terms would have to be disregarded locally within the Repository, and replaced with parent relationships to the various partitions in the KR-derived lattice of terms. So for example a given ontology for countries might define Territory as a child of "Thing", whereas we would necessarily redefine this as a child of Independent Thing, of Continuant Thing and of Concrete Thing.

6.6 Common Lattice

It would be beneficial to arrive at an industry consensus lattice. This could be used to provide axiomatically neutral high level theories which any reasonably well formed external ontology can be defined in relation to, thereby allowing us to refer to the terms from those ontologies as we now refer to the temporary terms for those subject areas. This is an area which would benefit from further work and research in the near future.

Much of the required notation is available in the SKOS standard [20] and we are looking at how to enhance our existing model using SKOS and adding other meta-level terms to relate any available ontology to a set of lattice terms. In addition, we would like to work with other ontology curators to arrive at a mutually useful set of lattice terms that can be used across the full range of industrial ontologies.

It should also be possible to apply this analysis to any ontology that is defined in any suitable logic syntax, not just OWL. This would include any variant of Common Logic, as well as other semantically capable syntaxes such as XBRL. There would need to be a commonly accepted theory against which non CL based syntaxes can be evaluated.

7 Conclusions

There are a number of things that we have needed to do at the EDM Council in order to be able to provide a meaningful semantic model of terms and definitions for our membership base of financial firms, data providers, technology providers and the like. We have done this in such a way that terms are unambiguously defined in relation to the simplest possible concepts from which they are derived, and this has been necessary in order to ensure that future innovations within the industry can be supported without the brittleness that is common in data models. This is also good practice since it means that where simpler terms are used in a wider range of industries, we are not defining localized equivalent terms that duplicate well attested industry terms elsewhere.

This opens the way to a high degree of interoperability across a range of industries that have no obvious boundary between them, principally investment, commerce, banking, accounting, insurance, risk management and so on. This interoperability can be capitalized on by all industry participants if we are able to come up with consensual approaches to ontology development. This would include consensus on ontology good practice, on meta-terms about ontology elements, and on how different theories may be described in relation to neutral, higher level lattice terms. Where industry bodies do not currently have logic based ontologies of their terms, there would be benefit in encouraging them to use the same approaches we have used, but this will only sometimes be possible.

It is to be hoped that the current ontology repository initiatives such as the Open Ontology Repository initiative will themselves support some consensus building on these aspects of ontology management. In the meantime, our strategy is not to go too far on our own in developing these features, but rather retain the flexibility to adapt to any consensus as it emerges from within the ontology community. In this way we would be in a position to participate in a single, coherent framework of meaningful terms, which will be of benefit to our members.

An immediate benefit of this is that those who are developing linked data resources, semantic applications, open data repositories and so on, will have a sound and well maintained base of ontologies to which they can refer, so that linked data knowledge bases can have a complete and consistent semantic grounding of terms.

References

1. The Enterprise Data Management Council, www.edmcouncil.org
2. EDM Council Semantics Repository, www.hypercube.co.uk/edmcouncil
3. Bennett, M.: The EDM Council Semantics Repository Case Study, Semantic Technology 2010, http://www.edmcouncil.org/PDFs/20100623.EDMC.SemTech2010.pdf
4. Enterprise Architect, http://www.sparxsystems.com.au/products/ea/index.html
5. Ontology definition Metamodel. The Object Management Group. OMG Ref ptc/2007-09-09
6. The Object Management Group, www.omg.org
7. Sowa, J. F.: Knowledge Representation: Logical, Philosophical and Computational Foundations. Brooks/Cole (2000)
8. W3C Time Ontology, http://www.w3.org/TR/owl-time/
9. The Financial products Markup Language, www.fpml.org
10. Internet Center for Management and Business Administration: The Accounting Equation, http://www.quickmba.com/accounting/fin/equation/
11. eXtensible Business Reporting Language, www.xbrl.org
12. Kim, S., Sucasas, M. I., Caracciolo, C., Viollier, V., Keizer, J.: Integrating country-based heterogeneous data at the United Nations: FAO's geopolitical ontology and services. Semantic Universe (2009)
13. McCarthy, W. E.: The REA Accounting Model: A Generalized Framework for Accounting Systems in a Shared Data Environment. The Accounting Review Vol. LVII, No. 3 (July 1982) pp. 554-78.
14. Rust, G., Bide, M.: The <indecs> metadata framework: Principles, model and data dictionary. Document reference WP1a-006-2.0. Indecs Framework Ltd. (2000)
15. Zachman, J. A.: A framework for information systems architecture. IBM SYSTEMS JOURNAL, VOL 26. NO 3,. 1987. Actual framework available for free on registration at http://www.zachmaninternational.com/index.php/the-zachman-framework
16. Drummond, N., Horridge, M., Stevens, R., Wroe, C., Sampaio, S.: Pizza Ontology (tutorial ontology) http://www.co-ode.org/ontologies/pizza/
17. McGuinness, D. L., van Harmelen, F.: OWL Web Ontology Language Overview - W3C Recommendation (10 Feb 2004)
18. Sowa, J., Integrating Semantic Systems, Semantic Technology conference 2010, http://www.jfsowa.com/talks/iss.pdf
19. Message on Ontolog Forum by Jon F Sowa, http://ontolog.cim3.net/forum/ontolog-forum/2010-08/msg00131.html
20. W3C Recommendation 18 August 2009, Simple Knowledge Organization System: http://www.w3.org/TR/skos-reference/
21. Open Ontology Repository Initiative: http://ontolog.cim3.net/cgi-bin/wiki.pl?OpenOntologyRepository