Zeroing in on the Holy Grail of Meaningful Information Exchange[1]

By Charles Hoffman, CPA (Charles.Hoffman@me.com)

April 26, 2015

Understanding and harnessing the power of machines to successfully perform work for business professionals in our digital age: from the perspective of a business professional trying to fill a business need.

The volume and richness of information the typical business professional has to work with has increased significantly over the past half century and our global interconnectedness will only contribute to even more information with which business professionals will have to interact. We are all awash in a veritable sea of information.

And the sea of information is so vast and deep that unless you have a super-human memory the only solution to this information overload is the same tool that is causing this information overload: computers.

Computers have three fundamental strengths:

1. **Storage**: Computers can store tremendous amounts of information reliably.
2. **Retrieval and processing**: Computers can retrieve and process stored information reliably and efficiently.
3. **Ubiquitous information distribution**: Computers can make information instantly accessible to individuals and more importantly other machine-based processes[2] anywhere on the planet in real time via the internet.

Further, the information stored on computers can be retrieved and then be the basis for certain kinds of automated reasoning that assists business professionals in making decisions. Such a tool can be an invaluable

[1] This document is inspired and highly influenced by the introduction in the paper *Ontology for the Twenty First Century: An Introduction with Recommendations* by Andrew D. Spear, see http://ifomis.uni-saarland.de/bfo/documents/manual.pdf . While that paper uses examples from the scientific community, this resource hopes to communicate the same information in a form understandable to business professionals.
[2] Wired: *The Web is Dead: Long Live the Internet*, http://xbrl.squarespace.com/journal/2010/9/24/wired-the-web-is-dead-long-live-the-internet.html

resource. For example, for business intelligence[3] or more generally "Information at your fingertips," as Bill Gates put this in 1995[4].

But while it is true that computers can put information at your fingertips; computers can also perform work and even change the way you work[5] in very fundamental ways. For example, CAD/CAM (computer aided design, computer aided manufacturing) software[6] changed not only how blueprints were created but rather the entire design supply chain and interaction between draftsmen, architects, engineers, and builders.

In CAD/CAM software architectural objects have relationships to one another and interact with each other intelligently. For example, a window has a relationship to the wall that contains it. If you move or delete the wall, the window reacts accordingly.

In addition, in CAD/CAM software machine-readable architectural objects maintain dynamic links with construction documents and specifications, resulting in more accurate project deliverables. When someone deletes or modifies a door, the door schedule is automatically updated in your local application's database and perhaps even in the database of the door supplier. Spaces and areas are update automatically when the size of a room is changed and calculations such as total square footage are always up to date. That means, say, that the amount of paint necessary to cover a room or an entire building is always updated.

Well organized machine-readable information has other uses as well. Domains of knowledge articulated in machine-readable form can leverage the power of computers to more rigorously communicate that information. For example, ambiguity can be reduced from the US GAAP conceptual framework which is the basis for financial reporting in the U.S.[7] Today, less reliable humans are used to remove ambiguity. Research of a domain of knowledge, such as the FASB Accounting Standards Codification (ASC)[8], can

---

[3] *Business Intelligence Strategies*, Chet Phillips, https://www.youtube.com/watch?v=ArOFlLzblHo
[4] Information at your fingertips video, 1995, Bill Gates, https://www.youtube.com/watch?v=efPwChPPJXI
[5] See *Digital Financial Reporting Will Change Accounting Work Practices*, http://xbrl.squarespace.com/journal/2014/3/20/digital-financial-reporting-will-change-accounting-work-prac.html
[6] See *Understanding the Basic Mechanics of a Financial Report*, http://xbrl.squarespace.com/journal/2015/2/7/understanding-the-basic-mechanics-of-a-financial-report.html
[7] Accountants Understand Utility of Ontology for Reducing Ambiguity Conceptual Framework , http://xbrl.squarespace.com/journal/2015/4/19/accountants-understand-utility-of-ontology-for-reducing-ambi.html
[8] FASB ASC, https://asc.fasb.org/

be made easier and more reliable leveraging machine-readable semantic information. Both text-based search but even more compelling is semantic-oriented search[9].

But, you have to get the computer to perform this work.

Major obstacles to harnessing the power of computers

However, there are a number of major obstacles to harnessing the power of computers to perform work for business professionals within one department, in an organization or across an entire supply chain. These obstacles include:

1. **Business professional idiosyncrasies**: The first obstacle is that different business professionals use different terminologies to refer to exactly the same thing.
2. **Information technology idiosyncrasies**: The second obstacle is that information technology professionals use different technology options[10], techniques[11], and formats[12] to encode information and store exactly the same information.
3. **Inconsistent domain understanding of and technology's limitations in expressing interconnections**: A third obstacle is that information is not just a long list of facts, but rather these facts are logically interconnected and generally used within sets which can be dynamic and used one way by one business professional and some other way by another business professional or by the same business professional at some different point in time. These relations are many times more detailed and complex than the typical computer database can handle.
4. **Computers are dumb beasts**: The forth obstacle is that computers don't understand themselves, the programs they run, or the information that they work with. Computers are dumb beasts. What

[9] *The Future of Search*, see section *Semantics - Giving Search Meaning*, https://www.linkedin.com/pulse/future-search-kurt-cagle

[10] See *Understanding Database and Query Options (Part 2)*, http://xbrl.squarespace.com/journal/2014/4/27/understanding-databasequery-options-part-2.html

[11] See *Understanding Database and Query Options (Part 1)*, http://xbrl.squarespace.com/journal/2014/4/26/understanding-databasequery-options-part-1.html

[12] See *Understanding Syntax*, http://xbrl.squarespace.com/journal/2014/3/30/understanding-syntax.html

> computers do can sometimes seem magical. But in reality, computers are only as smart as the metadata they are given to work with, the programs that humans create, and the data that exists in databases that the computers work with.

Computers do not create the magic. Skilled craftsmen who wield their tools effectively are what create the magic. Commuters simply follow instructions.

If two computers use the same information formats and other technology aspects but use different terminology or different information organization strategies the two computers will find it difficult or even impossible to interoperate. If this is the case, the only way to cross the chasm between these two different computers is with human intervention. Often this involves re-keying information. Saying this another way, in order for two computers to interoperate it is essential that every aspect including terminology, world view, information formats, instructions and so forth necessary to translate from one computer to the second computer must be explicitly provided.

Getting computers to perform work is straightforward science: The only way a meaningful exchange of information can occur is the prior existence and agreement on technical syntax rules, business domain semantics rules, workflow/process rules, and the information with which the computer will be working.

Computers are only able to reason with information that they have explicitly been given[13]. Remember, computers are dumb. Computers are incapable of implying meaning. This means that if the information is vague, inconsistent, logically incoherent, contradictory, ambiguous or in any other way unclear; the computer programmed to reason or use such information will produce either nothing at all or results which are likewise vague, inconsistent, logically incoherent, contradictory, ambiguous, or in some other way unclear.

It really is that straightforward: Nonsense-in-nonsense-out.

Computers cannot check the factual accuracy of information against reality. If the person who put the information into the computer made a mistake or

[13] Closed world assumption, http://en.wikipedia.org/wiki/Closed-world_assumption

intentionally entered the wrong information (i.e. fraud); that is exactly what the computer has to work with.

Finally there is setting the right expectations. Business professionals need to understand what computers can and cannot do. Computers cannot perform magic[14]. Computers fundamentally follow the rules of mathematics which follow the rules of formal logic. It really is that straight forward. Computers cannot effectively work with information such as the following:

- fuzzy expressions[15] - "It **often** rains in autumn."
- non-monotonicity[16] - "Birds fly, penguin is a bird, but penguin does not fly."
- propositional attitudes[17] - "Eve **thinks** that 2 is not a prime number." (It is true that she thinks it, but what she thinks is not true.)
- modal logic[18]
  - possibility and necessity - "It is **possible** that it will rain today."
  - epistemic modalities - "Eve **knows** that 2 is a prime number."
  - temporal logic - "I am **always** hungry."
  - deontic logic - "You **must** do this."

Computers can be provided with instructions in the form of explicit information which helps them mimic or seem to be able decipher such information; but it was really the business professional or information technology professional that created the instructions that made that happen.

At this time in history it is not possible for computers to think like human beings. Could it be possible in principle for computers to reason? Maybe. Artificial intelligence researchers have been working on this task for years but have been here-to-fore unsuccessful. IBM's Watson[19] is not intelligent. It only seems intelligent because the information used by Watson is clear, consistent, logically coherent, and unambiguous.

Overstating what a machine such as a computer can do is not wise. It is also not wise to either misunderstand the capabilities of a computer or to

---

[14] *Limitations of First-order logic expressiveness*, http://dior.ics.muni.cz/~makub/owl/
[15] Fuzzy logic, http://en.wikipedia.org/wiki/Fuzzy_logic
[16] Non-monotonicity, http://en.wikipedia.org/wiki/Non-monotonic_logic
[17] Propositional attitudes, http://en.wikipedia.org/wiki/Propositional_attitude
[18] Model logic, http://en.wikipedia.org/wiki/Modal_logic
[19] *IBM's Watson not as smart as you think*, http://www.computerworld.com/article/2507369/emerging-technology/ibm-s-watson-not-as-smart-as-you-think.html

misinterpret what it takes to make a computer successful in performing the work that computers are capable of performing. Computers can perform specific types of work extremely well. Computers are machines that are very adept at reliably performing repetitive mindless tasks accurately. Even very sophisticated repetitive tasks can be performed by computers.

Computers are tools

In the hands of a skilled craftsman the right tools can produce quality results. In the wrong hands, the same tool might produce poor results.

And so as was pointed out there are a number of problems that need to be worked through in order to get computers to successfully perform the tasks that they are well suited to perform: reliably store a tremendous amount of information and reliably and automatically retrieve and work with that information by anyone from anywhere in real time. The idiosyncrasies of business professionals need to be worked through, the idiosyncrasies of programmers, the idiosyncrasies of database builders and other information technology professionals need to be overcome. Computer languages and programs with sufficient expressive power to handle the richness of business information from complex business structures and transactions, different legal and cultural structures and so forth need to be created and implemented by information technology professionals for business professionals. Substantial care needs to be taken to ensure that the things and relations between the things within a problem domain are clear, logically coherent, consistent, unambiguous, and otherwise well-defined, precise, and accurate to reflect the facts of reality as currently reflected and flexible enough to change as today's dynamic business environment changes.

The answer to all of these challenges, the state-of-the-art solution to the real problems of getting many business systems to successfully interoperate with many other business systems in a distributed environment[20] such that a meaningful exchange of information can occur between business systems is "standards based ontology".

[20] *Understanding Distributed Extensibility*, http://xbrl.squarespace.com/journal/2015/1/7/understanding-distributed-extensibility.html

Such a system must be reliable, repeatable, predictable, safe, cost effective, easy to use, robust, scalable, secure as deemed required, auditable (track provenance) as deemed necessary.

Ontology

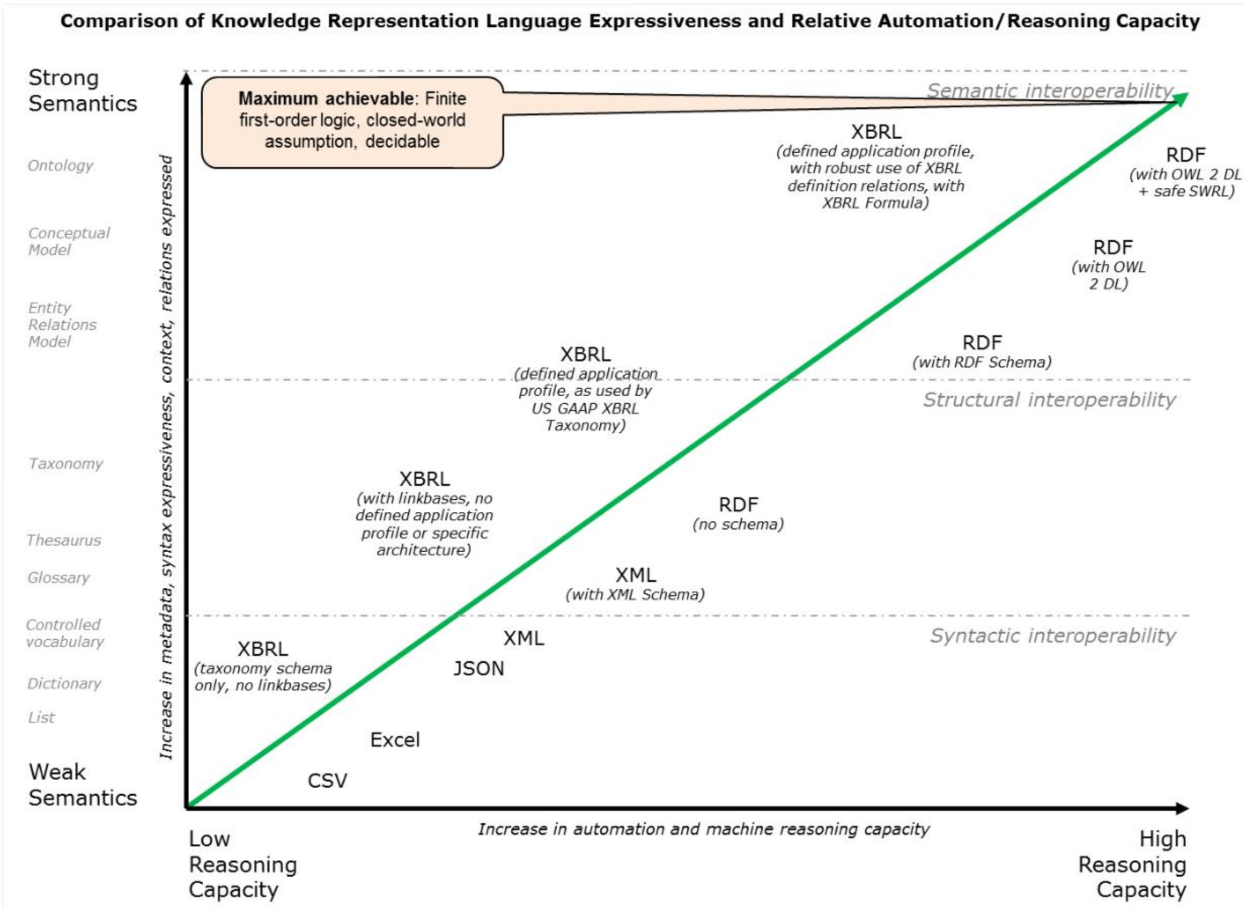
The term ontology has been used in philosophy for thousands of years going back to the father of formal logic, Aristotle[21] (400 B.C.). Ontology is defined as the study of the things and the relations between things that exist in reality. The goal of philosophical ontology is to provide deliberate, clear, coherent and rigorously worked out accounts of the basic structures found in reality.

In more current times, the term ontology has become prominent in the area of computer science and information science. In computer science the term ontology generally refers to the standardization of a terminology framework such that information repositories can be constructed. Ontologies used by philosophers like Aristotle were not machine-readable. Ontologies used by computer are machine-readable.

The problem that ontologies solve is not that of simply coming up with a set of terms such as a dictionary or creating basic relations between terms such as a thesaurus or even more complex relations between terms expressed by a taxonomy. Rather, an ontology defines terms, organizes the terms into categories or classes, and determines as many important relations as practical and necessary between the categories or classes within some business problem domain. Ontologies are the pinnacle of expressiveness.

The diagram below compares the relative reasoning capacity which is achievable give the semantic power or expressiveness of some language. The goal is to maximize the reasoning capacity that can be achieve, or said another way the ability of a computer to automate work.

[21] Aristotle's epistemology, http://en.wikipedia.org/wiki/Aristotle#Aristotle.27s_epistemology



Inspired by similar comparisons from *An Intrepid Guide to Ontologies* http://www.mkbergman.com/date/2007/05/16/ and *Semantics Overview* http://prezi.com/prwsxj8po3ln/semantics-overview/

We are not trying to represent data with ontologies; we are trying to represent information for the purpose of gaining knowledge[22]. Keep in mind that we are consciously using the term information and not data. Don't think "data" when we say "information". This summary helps you to understand the difference:

- **Data**: The basic compound for Intelligence is data -- measures and representations of the world around us, presented as external signals and picked up by various sensory instruments and organs. Simplified: raw facts and numbers.
- **Information**: Information is produced by assigning relevant meaning to data. Simplified: information is data in context.

[22] *Understanding Knowledge Modeling*, http://xbrl.squarespace.com/journal/2014/3/24/understanding-knowledge-modeling.html

- **Knowledge**: Knowledge is the subjective interpretation of information and approach to act upon the information in the mind of the perceiver. Simplified: knowledge is the interpretation of information.
- **Wisdom (or Intelligence)**: Intelligence or wisdom embodies awareness, insight, moral judgments, and principles to construct new knowledge and improve upon existing understanding. Simplified: wisdom is the creation of new knowledge.

Data that is not useable is simply noise. Data without context is not actionable[23]. Information is actionable.

Understanding what ontologies do

Ontologies overcome the four major obstacles of getting a computer system to perform work discussed previously. Remember the goal: reliable, repeatable, predictable, safe, cost effective, easy to use, robust, scalable, secure when necessary, auditable (track provenance) when necessary.

Ontologies both describe the information being worked with and verify information consistency against that description to avoid information quality problems or inconsistencies. Remember: nonsense-in-nonsense-out.

The first two obstacles which related to the problem of business professional idiosyncrasy and technical idiosyncrasy are overcome by using an ontology to standardize terminology. Rather than using arbitrary[24] terminology to express information about some business domain, standard terminology is used. This includes selecting the appropriate important terms and defining the terms in a deliberate, rigorous, clear, logically coherent, consistent, and unambiguous manner. Care is taken to precisely and accurately reflect reality using standard terms.

The third obstacle of expressing the rich logical interconnectedness of facts within and across business systems can be overcome by using general

[23] Understanding the Term Actionable Information, http://xbrl.squarespace.com/journal/2012/1/18/understanding-the-term-actionable-information.html

[24] Understanding the Difference between Standard and Arbitrary, http://xbrl.squarespace.com/journal/2014/8/22/understanding-the-difference-between-standard-and-arbitrary.html

ontological theories disciplined, methodical, and rigorous approach to structuring the relations between terms. Meaning, expressed using machine-readable ontologies, must be exchangeable between business systems not just used within your one system.

Beginning with a rigorous and logically coherent specification of the theoretical information to be implemented makes it possible to address the problems of human idiosyncrasy.

Given the idiosyncratic tendencies of business professionals; interpretations which reflect the arbitrary peculiarities of individuals can sometimes slip in or mistakes can be made when expressing such terminology. Further, parts of our understanding of a business domain can be incorrect and even evolve, improve, or simply change over time.

If different groups of business professional use different terminology for the same concepts and ideas to express the exact same truths about a business domain; those business professionals should be able to inquire as to why these arbitrary terms are used, identify the specific reasoning for this, and specifically identify concepts and ideas which are exactly the same as other concepts and ideas but use different terminology or labels to describe what is in fact exactly the same thing. But to also understand the subtleties and nuances of concepts and ideas which are truly different from other concepts and ideas.

If idiosyncrasies result only in different terms and labels which are used to express the exact same concepts and ideas; then mappings can be created to point out these different terms used to express those same concepts and ideas. Such mappings make dialogue more intelligible and could get groups to accept a single standardized term or set of terminology for the purpose of interacting with common repositories of business information.

If the difference in terminology and expression are rooted in true and real theoretical differences between business professionals, and the different terms express and point out real and important subtleties and nuances between what seemed to be the same terms; then these differences can be made conscious, explicit, clear, and therefore they can discussed, in a rigorous and deliberate fashion because the differences are consciously recognized.

Knowledge engineering

Explaining exactly how to create an ontology is both beyond the scope of this document and not something that the average business professional needs to concern themselves with. There is a significant difference between the skills needed to create an ontology and use an ontology. Most business professionals will use ontologies rather than create ontologies from scratch. Business professionals will highly-likely append ontologies.

Further, there is a significant difference between creating ontologies for any business domain and creating ontologies for one specific business domain. Most business professionals will work within one business domain or perhaps interact with a handful of other business domains.

This is not to say that business professionals have no role or responsibility in creating ontologies. They do have a role. The first role is to understand knowledge engineering[25] enough in order to grasp the moving pieces. The next role is to communicate with information technology professionals to create tools which abstract away as much of the complexity related to creating and using ontologies away so that business professionals can focus on their business domain. Business professional don't need to concern themselves with the details of exactly how everything works, but they need to have some grasp of the big picture and moving pieces.

While complexity can never be eliminated, complexity can be moved. Using the correct software development approaches, complexity can be buried deep within software that exposes simple to understand ideas to business professionals working with the technology in their business domain.

This is much like software developers creating software using higher-level languages and integrated software development environments rather than programming in assembly language using a text editor.

Some business professionals will become skilled knowledge engineers; specialists which help other business professionals create and work with high-quality ontologies to solve specific business domain problems or automate work.

---

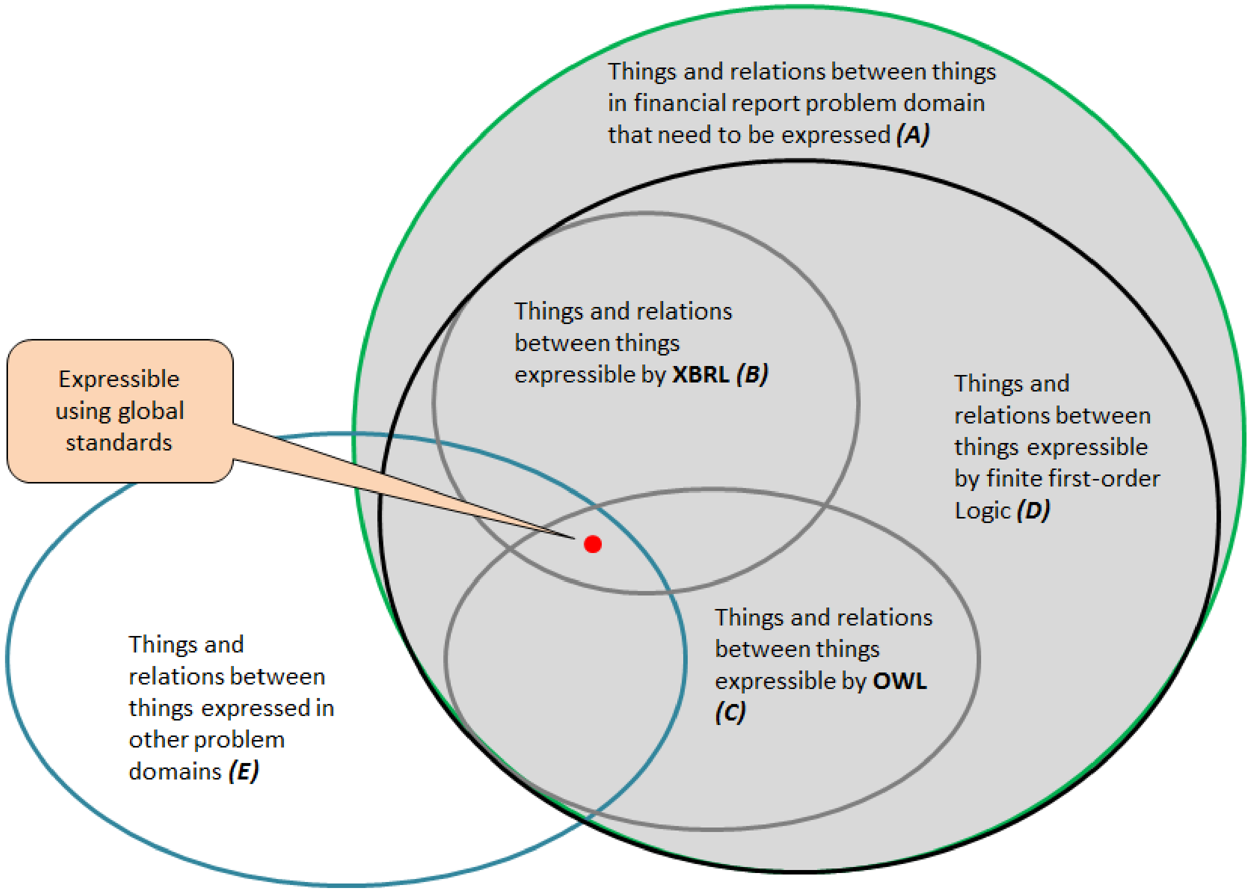
[25] *Knowledge Engineering 101 for Business Professionals*, http://www.xbrlsite.com/2015/Library/IssuesAndConsiderationsInCreatingDigitalFinancialReporting.pdf

The matter of technical syntax

Lastly is the matter of technical syntax. Ultimately, some technology needs to be used to implement a software-based solution. There are two global standard technical syntax options which are very useful in expressing information about a problem domain in the form of an ontology: XBRL[26] and OWL 2 DL[27]. Both global standard technical syntax options have pros and cons. Neither global standard technical syntax option has the full spectrum of expressiveness necessary to articulate what is necessary for most business problem domains.

The graphic below makes this point using the domain of financial reporting which is a business domain with which I am very familiar. I have been trying to figure out how to create an ontology for a financial report[28] for over 5 years. I have used XBRL, OWL, and proprietary approaches. This is what I have discovered:

[26] Extensible Business Reporting Language (XBRL), XBRL International, https://www.xbrl.org/
[27] Web Ontology Language (OWL), W3C, http://www.w3.org/standards/semanticweb/ontology
[28] Financial Report Ontology, http://xbrl.squarespace.com/financial-report-ontology/



This explains the graphic (note that the size of the circles have no meaning):

- **Theoretical goal**: The green circle with the label "(A)" represents the theoretical goal of expressiveness desirable for the business domain of a financial report. It represents every business rule for every relation anyone would every want to express related to a financial report. This is a theoretical goal because it is highly unlikely that this objective will ever be met because of limitations of technology or the ability of the business domain of a financial report to ever discover and express this information.
- **Achievable using technology today**: The black circle with the label "(D)" indicates what is technically possible to implement today given the current state of technology. The best "bucket" that I can use to express the circle is the notion of finite first-order logic. There could be a better bucket and I cannot articulate the boundaries of the bucket, but it seems like the closest correct bucket because it meets two crucial needs and makes one crucial assumption. The assumption

is the closed world assumption[29]. The two needs are the notion of "finite[30]" as contrast to infinite for which systems cannot be built and the notion of "decidability[31]" which eliminates other problems and system blowups.

- **XBRL**: The lighter gray circle with the label "(B)" represents things that are expressible using XBRL currently using global standard approaches. The most important piece of XBRL is what XBRL can do but OWL cannot do. XBRL has two strengths: (1) the ability to articulate information about dimensional relations using the a multidimensional model, (2) the ability to articulate mathematical relations. Clearly dimensional relations and mathematical relations are use cases for financial reporting in particular and business reporting generally. XBRL also has the power to express terms and relations between terms, but in this regard OWL is better equipped than XBRL. However, XBRL does have some expressive power here and it also has the architecture which enables richer expression of relations to be created.
- **OWL**: The other lighter gray circle labeled "(C)" represents things that are expressible using OWL 2 DL and Description Logic SROIQ[32]. OWL 2 DL and Description Logic SROIQ are state-of-the-art technologies which are W3C global standards for expressing ontologies. They meet the two needs of "finite" and "decidability" and also make the closed world assumption. These technologies surpass XBRL's current ability to express relations between terms. However, OWL 2 DL and Description Logic have two major limitations: (1) they do not have a dimensional model and (2) they don't support expressing mathematical relations. There is one additional drawback of OWL which is on the one hand a feature, but on the other hand something that is undesirable. OWL 2 DL is so low-level that it has the flexibility to represent any business domain, scientific domain, or any domain for that matter. But this flexibility comes with a price. The price is that OWL 2 DL is so low-level that it is like working in assembly language and therefore it is extremely difficult for even information technology

[29] For details, see the document *Knowledge Engineering 101 for Business Professionals*
[30] For details, see the document *Knowledge Engineering 101 for Business Professionals*
[31] For details, see the document *Knowledge Engineering 101 for Business Professionals*
[32] See Understanding Description Logic, http://xbrl.squarespace.com/journal/2015/1/8/understanding-the-importance-of-description-logic.html

professionals to make use of, and virtually impossible for business professionals to use.

- **Interoperability with other business domains**: The blue circle labeled "(E)" represents other domains which some business domain must interact and interoperate with. Using my example of a financial report which I am creating, other business domains creating ontologies interact with financial reports. For example, the Financial Institution Business Ontology (FIBO)[33] is likely one of those business domains. FIBO is expressed using OWL 2 DL. Public company financial reports filed with the SEC are XBRL-based. These different technical syntax are not a problem, as long as the business meaning, the semantics, are properly synchronized as pointed out earlier in this document. Other global standard technical syntaxes might be used by other business domains. Certainly proprietary formats will also be used internally by reporting entities.

At this point it is worth refreshing our memories of two things: the goal and how to achieve the goal. The goal is the reliable, repeatable, predictable, safe, cost effective, easy to use, robust, scalable, secure when necessary, auditable (track provenance) when necessary and *meaningful exchange of information between business systems*.

The only way a meaningful exchange of information can be achieved is with the *prior existence and agreement on technical syntax rules, business domain semantics rules, workflow/process rules, and the information with which the computer will be working*.

Both XBRL and OWL 2 DL are global standards. Both will likely progress to be able to serve the needs of business professionals, eventually. But what do we do today? What do we do now? One common denominator for both syntax is finite first-order logic. A partial solution is no solution, it leaves holes which cause problems. These are the complete solution alternatives that I see:

- **XBRL global standard + proprietary**: Using XBRL and supplementing XBRL with proprietary solutions which fill the gap will work. One problem with this is that the OWL-type functionality would need to be recreated in any proprietary solution.

---

[33] Financial Institutions Business Ontology (FIBO), http://www.omg.org/hot-topics/finance.htm

- **OWL 2 DL global standard + proprietary**: Using OWL 2 DL could work, but then you would need to build the multidimensional functionality and the mathematical relations functionality. As I understand it safe SWRL[34] could be used to express mathematical relations. Others say SPIN[35] is a better choice than SWRL. Neither SWRL nor SPIN are W3C recommendations as of yet. The RDF Data Cube Vocabulary[36] could be used to express multidimensional relations. But then, since XBRL is used for actual financial reports, one needs to ultimately serialize information into and likely also read it from XBRL. If this approach is taken, things like open source OWL reasoners[37] can be leveraged.
- **Composite XBRL global standard + OWL 2 DL global standard + proprietary**: It has been suggested before that a composite solution could be built to move things between syntax to leverage both the power of XBRL and the power of OWL 2 DL. That means that any proprietary implementation which fills any gaps that exist would be minimized.
- **XBRL global standard + XBRL-based proprietary**: Another possible solution is to build proprietary, but read the handwriting on the wall and realize where XBRL has to go next and build an XBRL-based proprietary solution. For example, to provide the expressive semantics that OWL provides in XBRL. This can be done using XBRL definition relations[38]. I created arcroles to express all of the types of relations that I see are necessary for what I need to do to make sure a financial report is created correctly. What if someone implemented a semantic reasoner[39] tailored for XBRL-based digital financial reports or other digital business reports?

Software usable by business professionals

As much as possible proprietary solutions should be avoided in favor of a solution which is based on global standard technical syntax. But that still

[34] Semantic Web Rules Language (SWRL), http://www.w3.org/Submission/SWRL/
[35] SPARQL Inferencing Notation (SPIN), http://www.w3.org/Submission/spin-overview/ and http://spinrdf.org/
[36] The RDF Data Cube Vocabulary, http://www.w3.org/TR/2014/REC-vocab-data-cube-20140116/
[37] JAVA-based open source OWL reasoner, http://code.google.com/p/owlreasoner/
[38] State-of-the-Art Use of XBRL Definition Relations to Express Business Rules, http://xbrl.squarespace.com/journal/2015/2/18/state-of-the-art-use-of-xbrl-definition-relations-to-express.html
[39] Semantic Reasoner, http://xbrl.squarespace.com/journal/2013/5/28/semantic-reasoner.html

does not provide business professionals with what they need. Business professionals will want to mainly do things like extend XBRL taxonomies (really they are ontologies). Some business professionals will create big base taxonomies such as the US GAAP XBRL Taxonomy or IFRS XBRL Taxonomy which exist for financial reporting. While those architectures need to be correct and it take more skilled professionals to design an architecture than to simply use an architecture; you will always have professional accountants needing to maintain those taxonomies (ontologies).

But most business professionals will be using taxonomies/ontologies created by other perhaps more highly skilled professionals in the area of knowledge engineering.

Software developers can leverage patterns to make software easier for business professionals to use. Patterns can be combined into composite patterns[40] make working with the technology less like working with low-level assembly language and more like working with Lego blocks.

There are three examples that help you understand what I mean by Lego blocks:

- **Blockly**[41]: This shows the abstract concept of how blocks can be used to work with syntax. Look at the visual, but also note that you can look at the same visual in the JAVA syntax, Python syntax, DART syntax, and XML syntax.
- **Scratch**[42]: This is a tool to help teach elementary school age children about programming. Imagine that the pieces of a financial report or other business report could be put together in this manner.
- **Quatrix**[43]: This is an application which while does not support XBRL, it works very similar to how I would expect an XBRL-based digital financial report creation tool to work.

There are three technically oriented tools that I have worked with to create OWL ontologies:

[40] A Vision for Diagrammatic Ontology Engineering, see Patterns on page 5 and Merging patterns on page 7, http://ceur-ws.org/Vol-1299/paper1.pdf
[41] Blockly, https://blockly-demo.appspot.com/static/demos/code/index.html#5ge5sh
[42] Scratch, created by MIT, watch the video in the upper right hand corner of the web page, https://scratch.mit.edu/
[43] Quantrix videos, watch the Quantrix Key Concepts video, http://www.quantrix.com/en/community/videos/

- **Protégé**[44]: (free download) This is a free software application which is very hard to use to create ontologies. Business professionals would never be able to use this type of tool.
- **Fluent Editor**[45]: (free download) This tool is a little easier to use because the user can simply create an ontology using a controlled natural language. However, you still need to understand how to create a correct ontology. Again, the typical business professional would never be able to effectively use this tool.
- **Top braid composer**[46]: (free download) This is probably the most complex tool that I have used to create ontologies, much too hard for business professionals to relate to.

Imagine the power of the technically oriented tools, an easy to user interface which hides complexity within well designed software which enables business professionals to only create things correctly. Business professionals would work with things that they understand from their business domain and deal with logic. If things act the way they expect, the logic of what they expect and the logic of what the software does are consistent; business professionals could very successfully make use of semantic technologies.

Too much to ask? I don't think so. Creating something that is complex is easy; anyone can do that. Creating something that is simple is hard work. Creating something simple takes thought, creativity, effort, etc. It takes a skilled craftsman. Such a tool will be elegant. Such tools can and I believe will be created.

XBRL International has created the Open Information Model working group[47] to develop a syntax-independent model of a business report. That shows that the understanding that syntax matters less and semantics matters more.

[44] Protégé, http://protege.stanford.edu/

[45] Fluent Editor, http://xbrl.squarespace.com/journal/2015/1/29/fluent-editor-helps-accountants-see-where-financial-reportin.html

[46] Top braid composer, standard edition, http://www.topquadrant.com/tools/modeling-topbraid-composer-standard-edition/

[47] See https://www.xbrl.org/news/open-information-model-call-for-participation/

Holy grail of meaningful information exchange

Ontologies are in essence a very powerful coordination and communications tool. Ontologies describe the things and relations between things and they help to verify the data quality and business logic of information against that description.

Consider this from the perspective of the business professional that must make sure everything works correctly. The business professional needs the best and most complete solution to their problem. They care far less about the technical details of how that solution is provided to them; they are concerned with solving their problem. Business professionals concern themselves with:

- *Complete solutions* are better than *incomplete solutions*
- *Less expensive solutions* are better than *more expensive solutions*
- *Powerful solutions* are better than *simplistic solutions*
- *Easy to maintain solutions* are better than *hard to maintain solutions*
- *Easy to use solutions* are better than *hard to use solutions*
- *Good solution performance* is better than *poor solution performance*
- *More scalable solutions* are better than *less scalable solutions*
- *Standard solutions* are better than *proprietary solutions*

The science part is balancing the concerns and achieving the appropriate equilibrium all things considered. Art is involved to the extent of deciding which of the concerns has priority when a perfect solution to a concern does not exist.

Testing, more testing, and then even more testing answers all questions. One thing that has been very helpful to me is poking and prodding XBRL-based public company financial filings submitted to the U.S. SEC which are all publically available[48].

What exactly is the holy grail? I cannot say for sure, but the moving pieces are revealing themselves. The workflow/process rules remain an outstanding question.

---

[48] Understanding Public Company XBRL-based Financial Report Quality , http://xbrl.squarespace.com/journal/2015/4/7/understanding-public-company-xbrl-based-financial-report-qua.html

Additional reading

If you do want to understand how to create quality ontologies, *Ontology for the Twenty First Century: An Introduction with Recommendations*[49] is one of the best resources in existence of which I am aware.

A book, *Building Ontologies with Basic Formal Ontology*[50], is forthcoming from the same authors.

[49] *Ontology for the Twenty First Century: An Introduction with Recommendations*, http://ifomis.uni-saarland.de/bfo/documents/manual.pdf

[50] See, http://ontology.buffalo.edu/smith/BFO-blurb.pdf