Chapter 12: Ontologies, Semantic Technologies, and Intelligence, Looking Toward the Future

Leo OBRST[a,1], Werner CEUSTERS[b], T erry JANSSEN[c]
[a] *The MITRE Corporation, USA*
[b]*State University of New York at Buffalo,* [c]*Lockheed Martin Corporation*

Abstract: This chapter looks at the intersection of intelligence and ontologies and semantic technologies, and tries to characterize the impact of these in the future. It provides a view into some emerging technologies such as query languages and rule standards for the Semantic Web. It also provides some guidance from a different domain, the biomedical domain, and tries to show that realist ontologies, ontologies based on common real world characterizations, have an effective impact on applications in those domains. Finally, it looks at the potential impact of these technologies on intelligence collection and analysis in the future, and makes some predictions.

Keywords: Ontology, information-sharing, intelligence community, semantic technologies, healthcare.

Introduction

In this chapter we look ahead: at ontology and semantic technologies and standards that are emerging, at the prospective evolution of the intelligence community, at where things could go in the future. The future direction and its success depend on many factors, including the commitment to embracing these technologies and the quickness and sophistication of their adoption. To assist our consideration of issues in technology adoption that could affect the intelligence community, we look at a test case, that of the adoption of these technologies by the healthcare community, and its prospective lessons for the intelligence community. Finally, we describe our projections and hopes.

---

[1] Corresponding Author: Leo Obrst, The MITRE Corporation, 7515 Colshire Drive, McLean, VA 22102-7508, USA; E-mail: lobrst@mitre.org.

1. Emerging Ontology and Semantic Technologies and Standards

We've largely focused on Semantic Web technologies in this book. Why? Because Semantic Web technologies represent n emerging set of global standards that are commercially rooted and also driven by a standards process that tends to be shorter in lifespan than older standards processes. This typically shorter, differently regimented standards process, usually enacted under the World Wide Web Consortium (W3C), does not guarantee better standards, but standards that typically are more immediately adapted to multiple communities of the Internet, e.g., researchers, Web and service developers, database practictioners, digital librarians, ontologists, etc. This is not to diminish the value of ISO standards and, in particular, ISO Common Logic, which is a valuable standard for representing very expressive logical ontologies.

1.1. Complexity of Applications and Costs

However, a general maelstrom of activity and acclamation does not guarantee the success of the technologies touted. The value of the work the technologies accomplish, and in fact, the greater value and lesser cost of the work they accomplish even as prorated over time, must be demonstrated. Value, potential value, cost over time – all of these must be estimated. However, as is usual with technological evolution, there is a spectrum or continuum behind the potential adoption of technologies, because there is a spectrum behind the expressivity of the models and the complexity of the potential applications that those models can provide, as Figure 1 depicts [1].

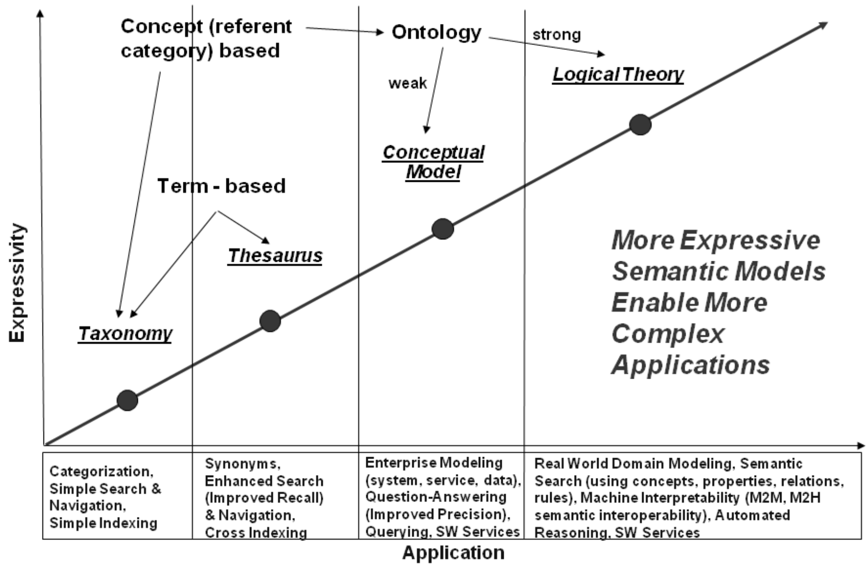

Figure 1. More Expressive Semantic Models Enable More Complex Applications

In Figure 1 shows that as the expressiveness of the semantic model increases, so does the possibility of solving more complex problems. Note that we distinguish *term* and *concept* here, where their definitions are the following (from [1]).

Terms (terminology) are natural language words or phrases that act as indices to the underlying meaning, i.e., the concept (or composition of concepts). The term is syntax (e.g., a string) that stands in for or is used to indicate the semantics (meaning).

A *concept (a universal category for referents)* is a unit of semantics (meaning) in the mental or knowledge representation model. In an ontology, a concept is the primary knowledge construct, typically a class, relation, property, or attribute, generally associated with or characterized by logical rules. In an ontology, these classes, relations, properties are called concepts because it is intended that they correspond to the mental concepts that human beings have when they understand a particular body of knowledge (subject matter area or domain) but at the philosophical universal level, i.e., as kinds of entities. In general, a concept can be considered a placeholder for a category (way of characterizing) of specific real world referents. From a realist perspective, as will be discussed in the next section, these concepts as placeholders are dispensable.

For simple applications, controlled vocabularies, terminologies, and classificational systems, usually structured in topic taxonomies or thesauri, are sufficient. For more complex applications that require precise semantics, more expressive models, i.e., ontologies, are required.

Costs of development and maintenance of models have to be tied to use cases and requirements, initially as they exist but also as they evolve over time. A larger cost initially that engenders an ascending benefit over time may be preferable to a much lower initial cost that generates a plateau or even descending slope of accrued benefit, sometimes within a short period of time. Figure 2 [1] notionally depicts this tradeoff between the cost and complexity of the semantic model developed and the prorated value over time of the benefit of using such a model, including reduced maintenance costs.

In recent years, there have also emerged better models to estimate the cost of developing ontologies, such as ONTOCOM [2, 3, 4] which also include estimation software in the form of detailed spreadsheets.[2]

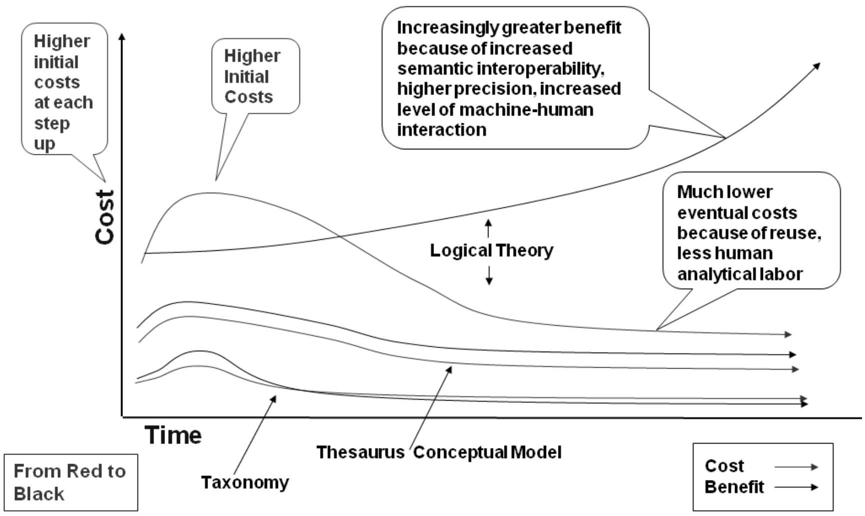

Figure 2. Approximate Cost/Benefit of Moving up the Ontology Spectrum: From Simpler Taxonomies to Ontologies

[2] ONTOCOM tools. http://ontocom.sti-innsbruck.at/tools.htm.

1.2. Emerging Technologies for Ontologies and the Semantic Web

Among the technologies focused on the Semantic Web, in recent years, a number stand out as potentially very useful, for many kinds of applications, but especially for intelligence analysis. These technologies include query languages, repository structures, and rules for rule-based reasoning and interchange. There are also many more kinds of inference engines, from description logic based classificational reasoners to first-order logic and logic programming based reasoners, and many new end-user Semantic Web applications. This section discusses the SPARQL Protocol and RDF Query Language (SPARQL) [5], triple-stores, the Rule Interchange Format (RIF) [6], and a range of inference engines. These technologies built on the established and more mature Semantic Web languages of the Resource Description Framework (RDF), the Resource Description Framework Schema (RDFS), and the Web Ontology Language OWL, the latter two of which are ontology description languages and the first is a graph-structured instance language. These languages became W3C standards in 2004. OWL 2, however, is a relatively new proposed standard, and increases the expressivity of OWL [7] by providing more datatype support, support for declarations and annotations on ontologies, and "syntactic sugar" for more succinctly and easily defining certain constructs in OWL.

The SPARQL query language [5] is a standard graph-based query language defined by the W3C to work RDF triple stores (i.e., n-tuple stores) which are graph-structured, potentially exist anywhere on the Internet or within an Intranet, and are exposed as so-called SPARQL endpoints. A SPARQL endpoint is a way of indexing a triple store, typically by providing an International Resource Identifier (IRI), so that it is known to a query engine. Currently, SPARQL is defined only over RDF, but many Semantic Web inference engines have extended SPARQL support to include the ontology languages RDFS and OWL. These engines extend the SPARQL support by enabling ontology reasoning methods over the queries, in addition to strict retrieval of graph-based instances.

Some representative triple stores are: OWLIM [8], Garlik 4Store [9], AllegroGraph [10], Jena [11], Sesame [12], Oracle 11g [13], Mulgara [14], and OpenLink Virtuoso [15]. Some triple stores advertise high storage sizes and various other high access, query, and load rates, with high-end triple stores reporting the ability to store billions of triples. But these claims are not yet independently confirmed. Also, many of these triple stores also support inference engines or link to existing inference engines, both those based on description logics (OWL is mostly a specific kind of description logic) and those based on more general logic and rules. Description logic reasoners include Pellet [16], RacerPro [17], FaCT/FaCT++ [18], etc. More general logic and rule reasoners include Jena, KAON2 [19], SILK [20], and various logic programming (Prolog) engines such as SWI-Prolog [21], Ciao Prolog [22], XSB Prolog [23], HighFleet's (formerly Ontology Works) High Performance Knowledge Server [24], Cyc [25], and TopQuadrant's TopBraid, [26] etc.

Rules are IF/THEN constructs that specify constraints on classes, relations, and properties (see [27] for more discussion on rules and related Semantic Web notions, from which this section is adapted). They thereby constrain how new classes, relations, and properties are defined, prevent contradictory information from being added to a knowledge base, and enable discovery of new information without explicitly asserting the information. Examples of common rules are 1) a rule that prevents a "child" from being its own "parent", and 2) a rule that says a "parent" of a "parent" that has a “child” is a “grandparent.” Rules are very closely associated with proof, i.e., rules require a proof mechanism to realize their value. In fact, a class of rules, inference rules, are directly associated with proof insofar as those inference rules license valid deductions as steps in an automated proof.

The de facto standard Semantic Web Rule Language (SWRL) is an example of a language for expressing rules and is based on OWL [28]. There is an emerging W3C standard rule language based on XML syntax called the RIF [6], which will probably supersede SWRL.

Because RIF tries to accommodate many different kinds of rule engines and existing deployed and used implementations, RIF is structured into multiple versions, called dialects or profiles, including the following: *Core:* the fundamental RIF language and a common subset of most rule engines (providing a basic Datalog, where Datalog is a simplified logic programming language); *BLD (Basic Logic Dialect):* this adds to Core, by providing logic functions, equality in the then-part, and named arguments (providing a basic Horn Logic, which is the foundation of Prolog, the primary logic programming language) ; and *PRD (Production Rules Dialect):* this adds a notion of forward-chaining rules, where a rule fires and then performs some action, such as adding more information to the store or retracting some information (providing an expert system-like capability).

2. A Prospective Lesson for Intelligence: Realist Ontologies in Healthcare

In the domain of healthcare information technology (HIT) it has been commonly accepted for some years now that both the development and use of clinical terminology should be supported by formal methods. Although this is a thesis that we strongly support, we wish no less strongly to insist that formal methods alone are not enough. The use of a Description Logic-based system appears, for example, not to have provided any guarantee for the absence of errors in SNOMED-CT [29], one of the most popular formal biomedical terminologies today.

With the extremely positive response to the creation of the Open Biomedical Ontologies (OBO) Foundry [30] it became clear that a role had to be played by *realist ontology* in making better biomedical terminologies. Realist ontology helped in detecting errors and in ensuring intuitive principles for the creation and maintenance of systems of a sort that can help to prevent errors in the future. More importantly still, however, it helps in ensuring that terminologies are compatible with each other. Note that we say ‘*realist ontology*’, in order to distinguish ontology in our understanding from the various related things [31] which go by this term in contexts such as formal knowledge representation. It is a realist conception of ontology which underlies statements such as:

> *The UMLS is an extensive source of biomedical concepts. It also provides a large number of inter-concept relationships and qualifies for a source of semantic spaces in the biomedical domain. However, the organization of knowledge in the UMLS is not principled nor consistent enough for it to qualify as an ontology of the biomedical domain [32]*

In the tradition of analytical philosophy, ontology is understood by the OBO Foundry community not as a software implementation or as a controlled vocabulary, but rather as '*the science of what is, of the kinds and structures of objects, properties, events, processes and relations in every area of reality*' [33]. Ontology as it concerns us here is a theory of those higher-level categories which structure the biomedical domain, the representation of which needs to be both unified and fully coherent – and as closely allied as possible to the representations used by clinicians in formulating patient data – if terminologies and coding systems are to have the requisite degree and type of interoperability. Ontology in this realist sense has successfully been used as a method to find inconsistencies in terminologies and clinical knowledge representations [34] such as the Gene Ontology [35] or the UMLS Semantic Network [36]. The method has also proved useful in drawing attention to certain problematic features of the HL7 RIM [37, 38, 39].

One of the major insights brought about by realist ontology in the healthcare domain is that biomedical terminologies can only be compared amongst each other, or used without loss of information within an electronic healthcare record (EHCR) system, if they share a common framework of top-level ontological categories [40]. Often one talks in this connection merely of a shared or common *semantics*, meaning thereby the sort of regimentation that can be ensured through the use of enabling technologies such as RDF(S) [41] and OWL [42] that currently enjoy a wide interest through their association with the Semantic Web project, not to forget systems such as Protégé that are able to cope with them in a user-friendly way [43]. On closer inspection, however, one discovers that the 'semantics' which comes with languages like RDF(S) and OWL is restricted to that sort of specification of meaning that can be effected using the formal technique of mathematical *model theory*, which is to say that meanings are specified by associating with the terms and sentences of a language certain abstract set-theoretic structures, taking Alfred Tarski's 'semantic' definition of truth for artificial languages as paradigm [44]. But model theory is metaphysically and ontologically almost completely neutral. Merely to formulate statements in a language such as OWL is far from building an ontology in the sense of ontology that is employed by analytical philosophers, and neither would translating a terminology into OWL turn it into an ontology. Such translation would indeed allow consistent reasoning about the 'world' – but only in the model-theoretic sense of 'world' that signifies not the flesh-and-blood reality with which biomedicine is concerned, but rather merely only some highly simplified set-theoretic surrogate. The task of ensuring that the latter somehow corresponds in broad terms to the real world of what happens and is the case, was in the semantics biomedical literature almost never addressed. Now it has become clear that the whole detour via semantic models is in fact superfluous: the job of ontology is not the construction of simplified models; rather, a biomedical ontology should directly correspond to reality itself in a manner that maximizes descriptive adequacy within the constraints of formal rigour and computational usefulness.

Applying realist ontology to terminologies and EHCR architectures means in the first place applying it to those entities in reality to which these artifacts of the human intellect refer, such as concrete patients, diseases and therapies. We do this to serve at least one important goal, namely making terminologies coherent, both internally as well as in their relation to the EHCRs in or for which they are used.

Already a very superficial analysis of a coding system such as the International Classification of Diseases [45] reveals that this system is not in fact a classification of *diseases* as entities in reality. Rather it is a classification of *statements about disease* *phenomena which a physician might attribute to a patient*. As an example, the ICD-10 class *B83.9: Helminthiasis, unspecified* does not refer (for example) to a disease caused by a worm belonging to the species *unspecified* which would be some sub-species of Acanthocephalia or Metastrongylia. Rather, it refers to a statement (perhaps appearing in some patient record) made by a physician who for whatever reason did not specify the actual type of *Helminth* the patient was suffering from. Neither OWL nor reasoners using models expressed in OWL would complain about making the class *B83.9: Helminthiasis, unspecified* a subclass of *B83: Other helminthiasis*; from the point of view of a coherent ontology, however, such a view is nonsense: it rests precisely on a confusion between ontology and epistemology [46].

A similar confusion can be found in EHCR architectures, model specifications, message specifications or data types for EHCR systems. References to a patient's gender/sex are a typical example. Some specifications refer to it as "administrative sex" (leaving it to the reader of the specification to determine what this might actually mean). The possible specifications of *administrative sex* are then *female*, *male*, *unknown*, or *changed*. *Unknown*, here, does not refer to a new and special type of gender (reflecting some novel scientific discovery); rather it refers to the fact that the actual gender is not documented in the record.

An interpretation along these lines does not work in every case, however. Consider those specifications which refer explicitly to "clinical observations", as is the case for Corbamed-COAS ("Clinical Observations Access Server'), which consists of:

> *any information that has been captured about a single patient's medical/physical state and relevant context information. This [information] may be derived by instruments such as in the case of images, vital signs, and lab results or it may be derived by a health professional via direct examination of the patient and transcribed [sic]. This term applies to information that has been captured whether or not it has been reviewed by an appropriate authority to confirm its applicability to the patient record. [47]*

When in a EHCR system that claims to follow the COAS specifications the specification "*unknown*" would be registered for gender, then that specification has to be interpreted that an observation has been made with respect to the patient's gender, and that as a result of that, an unknown kind of gender has been observed. Of course, that is not supposed to be the idea.

European and international efforts towards standardization of biomedical terminology and electronic healthcare records were focused over the last 15 years primarily on syntax. Semantic standardization was restricted to terminological issues around the semantic triangle paradigm [48] on the one hand and to issues pertaining to knowledge representation (and resting primarily on the application of set-theoretic model theory) on the other hand. Moves in these directions are in indeed required, and the results obtained thus far are of value both for the advance of science and for the concrete use of healthcare telematic applications. We can safely say that the syntactical issues are now resolved and also that the problems relating to biomedical terminology (polysemy, synonymy, cross-mapping of terminologies, …) are well understood – at least in the community of specialized researchers. Now, however, it is time to solve these problems by using the theories and tools that have been developed so far, and that have been tested under laboratory conditions. This means using the right sort of ontology, i.e. an ontology that is able explicitly and unambiguously to relate coding systems, biomedical terminologies and electronic health care records (including their architecture) to the real world.

To do this properly will require a huge effort, since the relevant existing standards need to be reviewed by experts who are familiar with the appropriate sort of ontological thinking (and this will require some effort in training and education). Even before that stage is reached, however, there is the problem of making all constituent parties – including patients (or at least the organizations that stand up for them), healthcare providers, system developers and decision makers – aware of how deep-seated the existing problems are. Having been overwhelmed by the exaggerated claims on behalf of XML and similar silver bullets of recent years, that would solve everything, they must be informed about the fact that XML alone isn't a silver bullet. And for sure, we must also be careful in not giving realist ontology a similar silver bullet status.

The message of realist ontology is that, while there are various different views of the world, this world itself is one and unique. It is our belief that it is only through that world that the various different views can be compared and made compatible. To allow clinical data registered in electronic patient records by means of coding (and/or classification) systems to be used for further automated processing, it should be crystal clear whether entities in the coding system refer to diseases or rather to statements made about diseases, or to procedures and observations, rather than statements about procedures or observations. As such, coding systems used in or for electronic healthcare records should be given a precise and formal semantics that is coherent with the semantics of the record as well as with the real world parts that are described by them.

3. Intelligence, Ontologies, and Semantic Technologies

The previous section underscored that realist ontologies are important, and they are important for intelligence collection and analysis. Realist ontologies are based on common understandings of the real world, and try to avoid conceptualist pitfalls (where concepts are introduced without direct origin in real world objects, relations, and properties) and epistemological, belief-based, or evidential (we use these terms synonymously) observational knowledge. The latter knowledge or approximations of knowledge are extremely important and are the basis of intelligence analysis and collection, but they largely address *instance* knowledge of the real world, i.e., individuals or particulars, about whom there may be many sources of data, much of which are contradictory. Why? Because the data being received from human and machine sensors are uncertain, error-prone, and often subject to noise, misinterpretation, and deception. It is very important to capture this incoming data, present it to intelligence analysts, and attempt to characterize it according to realist ontologies, but those realist ontologies describe the best knowledge of the real world we have, and the best knowledge of the general properties of that incoming data. They do not presume to be able to adjudicate which belief or observation is actually correct. That's what an intelligence analyst does, when he/she stitches together the evidence, generates hypotheses, and then either confirms or assigns a value to those hypotheses according to some strength of belief or evidence.

Ontology addresses the real entities, relations, and properties of the world; epistemology is about the perceived and belief-attributed entities, relations, and properties of the world, empirical evidence gleaned that will be described or characterized by ontology (see [49, 50], from which this is adapted, for further discussion of the differences between ontology and epistemology). Epistemology is employed in the use and qualification of data and actual data as stored in databases or tagged or indexed in documents. If ontology states that human beings have exactly one birth date, the data about a specific person is epistemological: in a given set of databases the person instance named John Smith (we assume we can uniquely characterize this instance, but we may not) may have two or more attributed birth-dates, not one of which are known to be true. Ontology tells us that everything that lives has only one birth-date. Epistemology helps us understand how we can address which one of seven birth-dates is possibly the most accurate, i.e., true. Without ontology, there is no firm basis for epistemology.

Epistemological concerns often distort and push off needed ontological distinctions. Why? Because analysts of information often believe that all is hypothesis and argumentation. They really don't understand the ontological part, i.e., that their knowledge is really based on firm stuff: a human being only has one birth date and one death date, though the evidence for that is multivarious, uncertain, and needs to be hypothesized about like the empirical, epistemological notion it is. Often also the charge that knowledge is just too "dynamic" is unjustified. Instance knowledge is very dynamic, i.e., the particular people, places, things, events we are interested in change all the time. My location is different from minute to minute. My activities change every minute, i.e., the events I participate in are new events that occur in time as time moves forward. I cut my hair or dye it. I marry, have children, divorce, move to another city, change jobs, go back to school, start a new hobby, make new friends, lose money in a new investment, watch and like different television programs, books, music, I eat different food and like different food. I think new thoughts and act on them. But the knowledge behind those instances largely remains the same. I am still a human. Families are still families. Organizations, work relationships, friendships, jobs, locations, kinds of events and activities, interests, etc., are the same. Occasionally this generic, ontological knowledge changes. For example, perhaps I join a new kind of organization where I pay the organization to work there. If this would occur (perhaps it's unlikely), then my ontology about organizations would have to change, to reflect this new real world situation. Perhaps in the future, a collection of men and women can combine to provide genetic material to create a child – in which case, the ontological notion of parent will have to change. The notion of what a parent and a child is, is ontological; which people are the parents of which child is at least partially epistemological: we need evidence, but it is based on our ontological knowledge.

Ontologies and semantic technologies are important and will be increasingly important for the intelligence community in the coming years. We have focused this book primarily on ontologies, representing the high end of semantic models, but semantic technologies more generally include a range of semantic models, some of which such as taxonomies, thesauri, and conceptual models are less expressive than ontologies, but useful for particular kinds of applications. Although predictions are notoriously problematic and often overcome by unanticipated events, we think we can make a number of predictions that will become true over the next ten to twenty years:

- The intelligence community will increasingly use semantic technologies in two forms: in the form of vocabularies that enable diverse sub-communities to use their own terms (words and phrases) to express their knowledge and queries, and in the form of ontologies that represent and model the meanings of those vocabularies, so that common information can be shared despite terminological differences among communities. The community of interest (COI) paradigm embraces these notions, and top-down and mid-level vocabularies such as UCore and Command and Control Common Core begin to address the vocabulary (syntactic) side, though are not yet sufficiently focused on the ontology (semantic) side. Largely this is because practictioners are focused purely on XML technologies, and are knowledgeable primarily about database technologies, i.e., structural methods and local semantics only. This will change over time, is changing now.

- Ontologies and semantic technologies will increasingly provide a basis for intelligence collection and analysis to go beyond local and structural data models, which are the standard currently for the structured data of the database community, and beyond the primarily statistical models, which are the standard currently for the unstructured data of the natural language and information retrieval communities. Semantic analysis and interoperability will be seen to provide more capability over structural and statistical methods and models for sounder and more effective intelligence capture and analysis. For example, greater precision for responses to queries depends on better semantic representation of the data. As is already apparent, analysts and other kinds of users of documents and information, do not have the time to keep performing syntactic, free text searches ala Google – because they do not have the time to read or even skim the documents returned in the result sets of queries, to see whether those documents are really relevant to their queries. How can an analyst tell whether the real answer or best data exists in document 10,000, since he/she will never get to that document?
- Organizations will change to accommodate ontologies and semantic technologies. The primary issue with technological change is not technological, but sociological. The people and the organizations must change, for better technological methods to be employed to solve existing problems. Institutionalization of change is very hard. The intelligence community, like governmental and even commercial organizations in general, can accommodate technological change, but their sociological milieus and organizational structures in general cannot. Software acquisition processes are monolithic and even in the era of service oriented architecture (SOA), which tries to decompose the older systems and systems of systems into service atoms and molecules, organizational and process change really depends on heroes, i.e., managers and directors who are technologically aware and advocates for change, but who are in oversight and guidance positions for only two years. When they leave after two years, any progress they may have initiated and supported evaporates, and the institution is again left bereft. Ontologies and semantic technologies cannot solve sociological and organizational problems.
- Unfortunate events will propel change. This is what we all fear. Correlations will not be made, since the data stores are huge, the sources are immense, noise is rampant, collection and analysis resources are insufficient, and there are no overriding descriptions, models, rules, procedures, processes, nor organizational and sociological support that will enable evidence to be stitched

together, described under common ontological, semantic, and epistemological characterizations, and acted upon in time to prevent bad events.

4. Cautious Optimism

We remain cautiously optimistic about change for the intelligence community, and the prospects for ontologies and semantic technologies to propel those changes. If any technology can be seen to enable a revolutionary leap forward for intelligence collection and analysis, if not for information technology in general, it is that of ontologies and semantic technologies. The authors of these chapters and the editors of this book are primarily technologists, and so, by predisposition, optimists about the use of technology to effectively achieve information-technological goals. But we are also realists, as our predisposition to realist ontologies indicates, and pragmatists: we are interested in these technologies, yes, but to serve a purpose, specifically to increase the effectiveness of intelligence collection and analysis. We are interested in technologies serving a purpose, and from our perspective, the best service for ontologies and semantic technologies is to enable the intelligence community, as it is for other scientific communities, to characterize the real world and thereby find out the truth and the probabilities that surround that truth, and so prevent, correct, and adjust to events that threaten nations and peoples. We wish us all sensible heads, sound technologies, stout hearts, and good luck.

References

[1] Obrst, Leo. 2009. Ontological Architectures. Chapter 2 in Part One: Ontology as Technology in the book: *TAO – Theory and Applications of Ontology, Volume 2: The Information-science Stance*, Michael Healy, Achilles Kameas, Roberto Poli, eds. Forthcoming, Springer.

[2] Simperl, Elena Paslaru Bontas; Christoph Tempich; York Sure. 2006. ONTOCOM: A Cost Estimation Model for Ontology Engineering. Proceedings of the International Semantic Web Conference ISWC 2006. http://ontocom.ag-nbi.de/.

[3] Elena, Simperl; Igor O. Popov; Tobias Bürger. 2009. ONTOCOM Revisited: Towards Accurate Cost Predictions for Ontology Development Projects. In: Proceedings of the European Semantic Web Conference 2009 (ESWC '09), Heraklion, Greece, May 20 - Jun, 04, 2009.

[4] Imtiaz, Ali; Tobias Bürger; Igor O. Popov; Elena Simperl. 2009. Framework for Value Prediction of Knowledge-based Applications. 1st Workshop on the Economics of Knowledge-based Technologies (ECONOM 2009) in conjunction with 12th International Conference on Business Information Systems (BIS 2009) April 27, 28 or 29, 2009, Poznan, Poland.

[5] World Wide Web Consortium (W3C), SPARQL Query Language for RDF, http://www.w3.org/TR/rdf-sparql-query/, (2008).

[6] World Wide Web Consortium (W3C), Rule Interchange Format (RIF), http://www.w3.org/2005/rules/wiki/RIF_Working_Group, (2009).

[7] World Wide Web Consortium (W3C), OWL 2.0, http://www.w3.org/TR/2009/PR-owl2-new-features-20090922, (2009).

[8] Ontotext, OWLIM Semantic Repository. http://www.ontotext.com/owlim/, (2009).

[9] Garlik 4Store. http://4store.org/.

[10] Franz Inc., AllegroGraph RDFStore™, http://www.franz.com/agraph/allegrograph/, (2009).

[11] Hewlett Packard (HP) Labs, Jena, http://www.hpl.hp.com/semweb/, (2009).

[12] Aduna (OpenRDF community), Sesame, http://www.openrdf.org/.

[13] Oracle Database 11g. http://www.oracle.com/technology/tech/semantic_technologies/index.html.

[14] Mulgara Community, Mulgara Semantic Store, http://www.mulgara.org/, (2009).

[15] OpenLink Software, Virtuoso Universal Server Platform, http://virtuoso.openlinksw.com/ , (2009).

[16] Pellet. http://clarkparsia.com/pellet/.
[17] Renamed ABox and Concept Expression Reasoner (RACER) Pro. http://www.racer-systems.com/products/racerpro/index.phtml.
[18] FaCT++. http://owl.cs.manchester.ac.uk/fact++/.
[19] KAON2. http://kaon2.semanticweb.org/.
[20] Semantic Inferencing on Large Knowledge (SILK). http://silk.semwebcentral.org/.
[21] SWI-Prolog. http://www.swi-prolog.org/.
[22] Ciao Prolog. http://clip.dia.fi.upm.es/Software/Ciao/.
[23] XSB Prolog. http://xsb.sourceforge.net/.
[24] HighFleet's High Performance Knowledge Server. http://www.highfleet.com/.
[25] Cycorp's Cyc. http://www.cyc.com/,
[26] TopQuadrant's TopBraid. http://www.topquadrant.com/.
[27] Parmelee, Mary; Leo Obrst. 2010. Technologies for Metadata, Vocabulary, and Ontology Storage. Chapter in: Miguel-Angel Sicilia, ed. Handbook of Metadata, Semantics and Ontologies. World Scientific Publishing Co. Forthcoming.
[28] Horrocks, Ian, Peter F. Patel-Schneider, Harold Boley, Said Tabet, Benjamin Grosof, Mike Dean, SWRL: A Semantic Web Rule Language Combining OWL and RuleML, W3C Member Submission. http://www.w3.org/Submission/SWRL/, (2004).
[29] Ceusters W, Smith B, Kumar A, Dhaen C. Mistakes in Medical Ontologies: Where Do They Come From and How Can They Be Detected? in Pisanelli DM (ed) "Ontologies in Medicine. Proceedings of the Workshop on Medical Ontologies, Rome October 2003" IOS Press, Studies in Health Technology and Informatics, vol 102, 2004.
[30] Smith B, Ashburner M, Rosse C, Bard J, Bug W, Ceusters W, Goldberg LJ, Eilbeck K, Ireland A,Mungall CJ, the OBI Consortium, Leontis N, Rocca-Serra P, Ruttenberg A, Sansone SA, Shah N, Whetzel PL, Lewis S. The OBO Foundry: Coordinated Evolution of Ontologies to Support Biomedical Data Integration. Nature Biotechnology 2007;25:1251-1255.
[31] Guarino, N; P. Giaretta, "Ontologies and Knowledge Bases: Towards a Terminological Clarification". In Towards Very Large Knowledge Bases: Knowledge Building and Knowledge Sharing, N. Mars (ed.), pp 25-32. IOS Press, Amsterdam, 1995.
[32] Bodenreider O. Medical Ontology Research: A Report to the Board of Scientific Counselors of the Lister Hill National Center for Biomedical Communications. May 17, 2001 (http://etbsun2.nlm.nih.gov:8000/pubs/pdf/2001-MOR-BoSC.pdf)
[33] Smith, B. Ontology, in L. Floridi (ed.), Blackwell Guide to the Philosophy of Computing and Information, Oxford: Blackwell, 2003, 155–166
[34] Smith B, Ceusters W. Towards Industrial-Strength Philosophy; How Analytical Ontology Can Help Medical Informatics. Interdisciplinary Science Reviews, 2003, vol 28, no 2, 106-111.
[35] Smith, Barry, Jacob Köhler, Anand Kumar: On the Application of Formal Principles to Life Science Data: a Case Study in the Gene Ontology. In: Erhard Rahm (Ed.): Data Integration in the Life Sciences, First International Workshop, DILS 2004, Leipzig, Germany, March 25-26, 2004, Proceedings. Lecture Notes in Computer Science 2994, Springer 2004, 79-94.
[36] Schulze-Kremer S, Smith B, Kumar A. Revising the UMLS Semantic Network Medinfo 2004. Proc. Medinfo 2004.
[37] Lowell V: Actions in Health Care Organizations: An Ontological Analysis in: Proceedings of MedInfo 2004, San Francisco.
[38] Lowell V, Smith B: Speech Acts and Medical Records: The Ontological Nexus. in: EuroMISE 2004, Prague.
[39] Smith B, Ceusters W. HL7 RIM: An Incoherent Standard, Stud Health Technol Inform. 2006;124:133-138. (Presented at MIE2006)
[40] Smith B, Ceusters W. An Ontology-Based Methodology for the Migration of Biomedical Terminologies to Electronic Health Records. AMIA 2005, October 22-26, Washington DC;:669-673.
[41] RDF Semantics. W3C Recommendation 10 February 2004 (http://www.w3.org/TR/rdf-mt/)
[42] OWL Web Ontology Language Semantics and Abstract Syntax. W3C Recommendation 10 February 2004 (http://www.w3.org/TR/owl-semantics/)
[43] Protégé OWL plug-in (http://protege.stanford.edu/plugins/owl/)
[44] Model Theory. Stanford Encyclopedia of Philosophy (http://plato.stanford.edu/entries/model-theory/)
[45] World Health Organisation. ICD-10 - The International Statistical Classification of Diseases and Related Health Problems, tenth revision (http://www.who.int/whosis/icd10/).
[46] Bodenreider O, Smith B, Burgun A. The Ontology-Epistemology Divide: Case Study in Medical Terminology. Submitted to the Third International Conference on Formal Ontology (FOIS) 2004.

[47] 3M, Care Data Systems, Inc., CareFlow/Net, Inc., HBO & Company, Philips Medical Systems, Protocol Systems, Inc. CORBAMED-COAS: clinical observations access server specification. Version 1, April 2001. (http://www.medcom.dk/picnic/deliverables/01-04-06%20coas%20specs.pdf.
[48] Ogden, C.K., & Richards, I.A. (1927). Meaning of meaning. New York: Harcourt, Brace & Company.
[49] Obrst, Leo. 2010. Ontological Architectures. Chapter 2 in Part One: Ontology as Technology in the book: TAO – Theory and Applications of Ontology, Volume 2: The Information-science Stance, Michael Healy, Achilles Kameas, Roberto Poli, eds. Forthcoming, Springer.
[50] Poli, Roberto; Leo Obrst. 2010. The Interplay Between Ontology as Categorial Analysis and Ontology as Technology. Chapter 9 in Part One: Ontology as Technology in the book: TAO – Theory and Applications of Ontology, Volume 2: The Information-science Stance, Michael Healy, Achilles Kameas, Roberto Poli, eds. Forthcoming, Springer.