www.itcon.org - Journal of Information Technology in Construction - ISSN 1874-4753

VALIDATING ONTOLOGIES IN INFORMATICS SYSTEMS: APPROACHES AND LESSONS LEARNED FOR AEC

REVISED: July 2012
PUBLISHED: October 2014 at http://www.itcon.org/2014/28
EDITOR: Rezgui Y.

Tamer E. El-Diraby, Associate Professor,
University of Toronto;
tamer@ecf.utoronto.ca, www.i2c.ca

SUMMARY: *In their pursuit to represent a human-savvy machine interpretable model of knowledge, informatics ontologies span three dimensions: philosophy, artificial intelligence, and linguistics. This poses several challenges to ontology validation. Within the scope of knowledge models, four types of validity are relevant: statistical, construct, internal and external. Based on benchmarking some tools and best practices from other domains, a map is proposed to link specify a set of tools to support the handling of four validity types (statistical/conclusion, internal, construct, and external) in each of the three dimensions. The map advocates a debate-based approach in validating the philosophical dimension to allow for innovation and discovery; use of competency questions and automated reasoning tools for the artificial intelligence dimension; and experimenting with lexical analysis tools (especially web contents) for the linguistic dimension. A set of best practices are proposed based on benchmarking other domains. These include falsifying the conceptual frameworks of research methodologies, scope management, iterative development, adequate involvement of experts, and peer review.*

KEYWORDS: *Informatics, Ontology, Research Validation, Knowledge Management.*

REFERENCE: *Tamer E. El-Diraby (2014). Validating ontologies in informatics systems: approaches and lessons learned for AEC, Journal of Information Technology in Construction (ITcon), Vol. 19, pg. 474-493, http://www.itcon.org/2014/28*

1. INTRODUCTION

Informatics (at least as defined here) is different from AI-based research. AI-based knowledge management research attempts to use operations research methods to reason about decision. What is the equation or model that can be used to make a decision? How can we mimic the cognitive decision making process in a human mind? In contrast, Informatics systems emerged as means to utilize the best in AI and IT to support human-oriented exchange of "knowledge". Consequently, informatics systems aim, fundamentally, at representing knowledge in the first place. Such representation can be used then for communication purposes or to support an AI-based reasoning tool. In most cases, this is done through an ontology. The obvious difference here is that representing knowledge is closer to the core mission of informatics systems, while such representation is secondary in AI-based systems. Of course, such delineation is not universally agreed upon. Some researchers do not distinguish between AI-based research and informatics research.

Informatics ontologies are at the intersection of three major domains of discourse: epistemology (the theoretical assumptions about knowledge held by the ontology), artificial intelligence (the formal/codified aspect of knowledge), and linguistic (the lexical representativeness of ontology term). This very feature has major implications on the validation approach that should be used to evaluate informatics ontologies. On one hand, informatics ontologies can be seen as AI (artificial intelligence) models, which are meant to provide consistent and rigorous means for reasoning. This means that typical AI validation tools are applicable to ontologies. On the other hand, and true to their genesis in metaphysics, informatics ontologies are based on a conceptual/philosophical view of the knowledge within a domain. Directly or indirectly, ontologies have a major interest in the philosophy of knowledge—granting merit for the applicability of philosophical (more accurately, epistemological) debates in their evaluation. Lately, a third dimension has been added to this dichotomy: semantic communication. The rich semantics of ontologies makes them a major tool in the social semantic web revolution. This means that typical semantic, communication, and linguistic evaluation techniques are also applicable to the validation of informatics ontologies.

This paper presents the results of a research work that aimed to synthesize validity assessment efforts in ontology development and means to benchmark some of the best practices to the benefit of research into AEC ontologies. The fundamental thesis of this paper is that validating informatics ontologies has to adopt a mix of tools to address these three dimensions: a (rather relaxed) debate-based approach in discussing the epistemology of the ontological model; a strict (technical-based) approach when it comes to the artificial intelligence dimension, and a (highly suggested) well-informed analysis of its lexical aspects.

2. SCOPE: THE NEED FOR BENCHMARKING

An informatics ontology is a formal description a domain of knowledge. On its face, it encompasses three fundamental components: taxonomy of concepts (domain vocabulary), a set of relationships (to link and triangulate domain terms), and axioms (a set of rules to describe and limit the behaviour of concepts). Many tools for coding these three components have been developed over the years. The IDEF5 method provides a modeling notation and specification approach to assist in creating, modifying, and maintaining ontologies. The Web Ontology Language (OWL) is the de facto standard in coding ontologies in machine interpretable manner.

The above, rather mechanical, view of ontology is very suitable to IT-based semantic systems. However, at its core, an ontology is essentially a philosophical claim regarding the theory of knowledge in a domain—and this is where the true challenge of ontology development and validation exists.

Almost all work on ontology in AEC informatics can be classified as AI-inspired and mostly interested in the mechanical aspect of ontology development (the face value of ontology). The work by Ugwu et al. (2002) and Tesfagaber et al. (2003) represent some of the early work on ontology and its potential in AEC informatics systems. The majority of subsequent work reported the use of ontologies to support "product modeling and information management" as part of a larger IT system—for example, representing space and their related construction methods (Akinci et al. 2002); development of service infrastructure for ubiquitous computing (Liu et al. 2003); supporting cost calculations (Staub-French et al. 2003); infrastructure products (Osman and El-Diraby 2006); hazard analysis (Wang and Boukamp 2009); managing context-sensitive information (Wang et al. 2010). Lately, some work in AEC informatics started to examine the role of ontology in supporting humancentred "knowledge management systems" (see for example, Lima et al. 2005; Rezgui 2006; Zhang and El-Diraby 2009; Wender and Hübler 2009, El-Gohary and El-Diraby 2010b). In parallel, some researchers developed work that was focused/dedicated to stand-alone ontology building (see for example, El-Diraby et al. 2005; El-Diraby and Kashif 2005; Kim and Grobler 2007; Aksamija and Grobler 2007; Lin et al. 2009; El-Gohary and El-Diraby 2010a). Some work is emerging on AEC-specific methods for ontology merger (El-Gohary and El-Diraby 2011).

The methodology of ontology development did not receive extensive discussions in most of the work above. This is typical and is expected in the early stages of the use of a new tool in research. To this end, this paper aims at providing a roadmap for ontology validation in AEC that is based on benchmarking some best practices form other domains.

The paper starts with review of some relevant concepts in philosophy and theories of knowledge along with a review of validity definitions (especially in social sciences). This lays the foundations to discuss three major dimensions in ontology development (epistemology, AI and linguistics), including analysis of some main approaches for ontology development methodologies. Next an analysis section scopes the challenges posed by the philosophical and linguistic dimensions and provides arguments about the need for road-mapping ontology development and validation in AEC. The paper then summarizes some of the lessons learned in research methodology that could help in enhancing the validity of ontology. Based on that, a map is then proposed to link the three dimensions of ontology with four major types of validity.

3. BACKGROUND: THEORIES OF KNOWLEDGE

Philosophy has two main branches: metaphysics and epistemology. Metaphysics relates to the description of the "known" through an ontology—a formal model of our "knowledge" about a domain. Such ontologies are developed based on a set of assumptions about the world and through a "knowing" process. An ontology is the formalization of concepts, their relationships, and axioms in a certain domain. Epistemology is the branch of philosophy that studies the nature and theory of knowledge. It deals with the process and means of "knowing", including its assumptions, sources, essence, and acquisition. Epistemology can also include the study of justified belief (a core issue in knowledge): how to justify claims and/or beliefs to be true or false.

Over the years, many theories were developed about the structure and evolution of knowledge and the role of beliefs and justification in reaching the truth. These include the following:

- Foundationalism: According to foundationalism, our justified beliefs are structured like a building (or a tree): they are divided into a foundation and a superstructure. Beliefs belonging to the foundation are *basic*. Beliefs belonging to the superstructure are *non-basic* and receive justification from the justified beliefs in the foundation (Fumerton 2009).
- Coherentism: coherentism relaxes the strict structure of foundationalism by viewing knowledge as a loosely structured *web* of truths. A belief is justified if it fits (coheres) with related, nearby beliefs in the web of knowledge. Coherentists, then, deny the tree-like structure of beliefs, starting with the denial of basic beliefs. Some philosophers such as Quine attempted to balance the two schools by adopting a star-like view, where beliefs at the core of the star are more important and basic than those at the outer end of the star (Quine and Ullian 1978).
- Reliabilism: reliabilism shifts the attention from the structure of knowledge into critiquing the processes of reaching such knowledge. A belief is justified only if it is produced by processes that typically yield a sufficiently high ratio of true to false beliefs (Hacking 1981).
- Fallibilism and falsifability: According to Popper (2002), every belief is contingent. No number of positive outcomes of testing can confirm a scientific theory. On the other hand, a single counterexample (or failure in a single test) is enough to fault the theory. A theory should be considered scientific if and only if it is falsifiable! The aim is not to fault theories. Rather, it aims at constantly testing viable theories to uncover its limitations, hence enhancing its scope, reliability and advance the search for newer theories to address the shortcoming of the one under testing.

4. VALIDITY

Typically, discussions about validity in construction IT research relates to verification and/or validation of computational models. Verification seeks to answer the question: is the model being built right? In other words, does the (coding of the) computational model realize the intended objectives? This starts, normally, with checking the development process, identifying errors, and procedural testing and debugging of propositions and codes (especially for software). This can be done manually or in an automated way. Validation aims to answer a more serious question: is the right model being built? Typically, validation refers to the representativeness and applicability of the model. Does it represent or match reality? How applicable it is to various situations. Many statistical (and operations research) methods, including expert interviews and actual implementation are used to validate the relevance and applicability of these models.

However, it is almost unanimously agreed that no computational model is ever fully verified or validated (akin to Popper's views). In fact, Gödel proved that nothing could ever be proved (Hofstadter 1980)! Or, more formally, that no theory could be both complete and self contained. This is more so the case in models that deal with softer issues such as human decision making and rationale, information processing, knowledge flows, and human behaviour and selections such as motivation and trust. Validation in this arena is reduced to a matter of establishing credibility in the model.

4.1 Validity in social sciences

Validity spans a variety of types--for example, content validity refers to the soundness of the contents of a model. This is typically contrasted with face validity, which refers to the degree to which a test or a theory (at the outset) gives the impression of measuring what it is intend to measure (without analysis of its contents) to an observer (typically, not an expert). Criterion-related validity compares the new model against established models or observations. It includes two sub-types. Predictive validity examines the degree of correlation between the outcomes/results or predictions of a model with what it actually is meant to measure (Brown 1996). Concurrent validity tests the relationship between measures made with existing tests. The existing test is thus the criterion (i.e. a new measure of a phenomena should cohere with existing measures/tests).

In social sciences where subjectivity is a major concern, researchers have used statistical methods (especially surveys) to introduce more objectivity to the validation exercise. One of the most important validity types in this regard is construct validity which refers to whether a test or an assessment tool is actually measuring what it is supposed to measure. In other words, the experimental demonstration that a test is measuring the phenomenon/criteria it claims to be measuring.

There is quite an overlap between the various definitions of validity types. In many cases, the difference is fundamentally a naming issue. For example, the definition of content validity overlaps clearly with what some scholars call "internal" validity; the scope of predictive and concurrent validity is similar to what some scholars call "external" validity.

Cook and Campbell (1979) presented a widely approved classification of the term validity (Black 1993; Clader et al. 1982): "the best available approximation to the truth or falsity of a proposition". They divided validity into four major types: statistical, construct, internal and external. Many researchers use the term conclusion validity instead of statistical validity. It refers to the degree to which conclusions we reach are reasonable. The use of the term "conclusion validity" is meant to emphasise that testing validity of outcomes is equally necessary to qualitative research and to avoid the typical overuse of statistical tests to prove a theory.

Cook and Campbell's definition of construct validity is similar to the above. However, Trochim (2006) argues that construct validity should be seen as independent from the "testing" realm –especially the statistical testing. In essence, it should measure the validity, relevance and representativeness of the claimed constructs (akin to content validity). Are the parameters/ concepts being claimed by a model (its constructs) the most basic or refer to the most possible basic terms/concepts. Can other (especially, simpler) constructs be used.

The proposed roadmap adopts Cook and Campbell categorization (with some deviation on the scope of construct and statistical validity which will be shown below). This does not discard other types or definitions of validity but was meant to assure consistency. The four types are defined as follows:

4.1.1 Construct validity (Validity of the claims)

Constructs are the elements of a model. They represent collections of behaviours that are associated in a meaningful way to create an image or an idea invented for a research purpose. In other words, models are composed of constructs: parameters, concepts, assertions, assumptions. Normally, a set of operationalizations (testing tools) are deployed to study the relevance of the model constructs. Do they make sense? Do they cohere to or explain reality? Why? Construct validity is the degree to which inferences can be made from operationalizations in a study to the constructs on which those operationalizations are based. In addition to this definition that comes from the social surveying and quasi-experimentation realm, we have to test the relevance and validity of the constructs themselves too. Theoretical as they may be, these constructs have to be contrasted to reality. To establish construct validity we must first provide evidence that data supports the theoretical structure. We must show that the theory has some correspondence with reality. Construct validity encompass three sub-types. Convergent validity is assured if we can show that measures that should be related are in reality related. In contrast, discriminant validity is concerned with showing that measures that should not be related are in reality not related. Finally, according to Cronbach and Meehl (1955), we have to establish also nomological validity: a lawful/formal network that includes the theoretical framework that needs to be measured/evaluated, an empirical framework for how we are going to measure it, and specification of the linkages among and between these two frameworks.

4.1.2 Internal validity (validity of relations/structure)

While construct validity is interested in the relevance and representativeness of the proposed parameters, this type is closer to testing their interrelationships. This type addresses the quality of the inner structure of a model or a theory. How cohesive is the theory? How accurate are the interrelationships between variables? In particular, when asserting causality, is one variable causing the other, or does a third (or a multitude of) variable(s) causes both variables?

4.1.3 External validity (scope of applicability)

This type is related to the scope and repeated applicability of the whole model or theory. While construct validity relates to “matching” or cohering to reality, External validity is related to generalization. How general is the model? What limitations are applicable to the model? Under which conditions, can the proposed model be applied reliability. Typically, statistical methods are used to test the applicability of a model to its claimed domains.

4.1.4 Statistical/conclusion validity (soundness of statistical methods, if used)

Given that Cook and Campbell work was related to quasi-experimentation, which is essentially statistics-intensive, statistical analysis is a major type in their taxonomy of validity. In ontology validation, statistical validity is only relevant if statistics is used. While the above three validity types relate to the model (its parameters, relationships and applicability), this type merely relates to the validation process itself. When sampling and other statistical approaches are used, the validity of the research work is contingent on the soundness of statistical methods used. Data-intensive samples can be used as sources for extracting/finding concepts or variables. They can also be used to observe/detect co-variations between variables. In other cases, statistical tests are needed to test hypothesis. In all cases, when statistics is used to validate part of a model (or a stage of research), typical statistical analysis best practices have to be used, including the relevance, quality, suitability, and quantity of samples, and the relevance and accuracy of the statistical tests.

5. BUILDING ONTOLOGIES: THE PHILOSOPHICAL DIMENSION

One can observe three dimensions for ontology development and use: philosophical (epistemological), AI, and linguistics. This section and the next two discuss some of the issues and current practices related to these dimensions (see also Guarino 1997 for an interesting discussion). These will be contrasted later with the four types of validity above.

When it comes to the validity of an ontology, three epistemological issues are of major relevance:

- Conceptualization, abstraction and reductionism: the fundamental contribution of the philosophical dimension of an informatics ontology is its conceptualization of the domain. Conceptualization includes

abstracting some aspects of a domain knowledge to reach for its underlying philosophy or essence. This should be encouraged. Some could confuse abstraction with reductionism, which reduce the multitude of concepts or issues within a domain of knowledge into a limited set of constructs to the exclusion of other worthy one. This should be avoided. In short, an ontology should expose, develop and use the basic fundamental "constructs" in a domain based on a genuine, thoughtful conceptualization of the essence of the knowledge in the domain. On the one hand, listing terms and parameters (even if they work fine in an AI equation/algorithm) is not the aim. Rather, it is the qualified discovery of universals (and their quality, representativeness, interrelationships, and their match to reality) that matters. Basic Formal Ontology (BFO) is a typical example in this regard. BFO adheres to an extremely well-structured and strict philosophical/logical rules about categorisations that assure high level of consistency. On the other hand, strict enforcement of a (limited) set of rules or universals for all situations (as is being attempted by BFO) could be counterproductive and could lead to reduced representativeness of the terms used (reductionism) (Dumontier and Hoehndorf 2010).
- Positivism, logicalism, and coherence: while an ontology has to be logical and should attempt to use formal logic in its structure (especially when it comes to axioms), it should avoid strict adherence to logical positivism. Logical positivism is the strong flavour of positivism, which asserts a universal and complete view of sciences. In one angle, positivism is related to the dream of an integrated theory. On another angle, positivism was (to an extent) a result of the euphoria of mechanicalism that emerged in the Newtonian era of physics. Both angles are highly debateable in today's world (if existent). In the domain of AEC, which has major relationships with management and social sciences such foundationalism is not achievable. While ontologies have to adhere to the rules of logical consistency in their internal structure, they should only be required to cohere to nearby formal knowledge in the domain. Of course, when it comes to very established well-structured domain (such as structural analysis) the coherence of a new ontology (in this specific domain) has to be almost at the level of foundationalims. i.e. it has to have almost a 100% coherence and, in this case, logical positivism is justified.
- Phenomenology, constructivism, and fallibilism: while in traditional philosophical debates it is acceptable that personal introspection leads to the formation of a theory (as could be the case in a pure application of phenomenology), this is not acceptable in the AEC domain. AEC knowledge modeling should be seen and practiced as an endeavour in social and formal sciences. A mix of scientific phenomenology and constructivism must be adopted. In all cases, falibilism is a virtue that must be embraced in such a dynamic and subjective domain of knowledge.

6. BUILDING ONTOLOGIES: THE AI DIMENSION

The methodological debates related to the AI and formal parts of ontology development and testing are more straightforward and crisp. Grüninger and Fox (1995) led the way (at least in Engineering ontologies) with the development of TOVE (TOronto Virtual Enterprise Ontology). In their view, an ontology is an AI model. To validate an ontology, you have to follow typical modeling procedures: develop a set of competency questions and then test if the ontology (typically laden with formal logical notations) satisfy/adhere to these questions. Their approach follows the following steps:

- Capture a motivating scenario: these are story problems, which serve also to provide a set of intuitively possible solutions to the scenario problems.
- Develop a set of informal competency questions (CQ): The questions serve as constraints on what the ontology can be, rather than determining a particular design.
- Collect/specify terms: using the informal CQ as basis, develop and then formalize concepts that relate to/answer/clarify these questions. These concepts have to be presented in a formal logical format (such as first order logic).
- Formalization and test CQ: review the informal CQ and use typical logic validation methods to test the adherence of axioms to the CQ.

The main features of TOVE are its high level of formality (very fitting for reasoning), modularity, and consistency. However, while TOVE's limited scope and focus on reasoning have saved it from the universality trap, it full adoption of AI tools has resulted in weak representational aspects that make TOVE fitting mainly to interoperable and formal domain representation to "computers".

Uschold and King (1995) used a methodology akin to typical software development processes. It starts with identifying a purpose, then iterative development of ontology terms and structure. This is followed by coding into a formal language. Their evaluation system is also akin to software verification methodologies (see also Uschold and Gruninger 1996).

In the same direction, Gomez Perez and her colleagues at the Laboratory of Artificial Intelligence at the Polytechnic University of Madrid developed an approach (METHONTOLOGY) which embraces reliabilism and balances the modeling/ AI orientation of Grüninger and Fox and the software/IT approach of Ushold and King. This approach is porcessual (focuses on the integrity of the development process) and emphasis a life cycle approach to ontology development. It includes the following major steps (see Fernández-López et al. 1997; Gómez-Pérez 2001; Fernández-López and Gómez-Pérez 2002):

- Project Management Activities: planning, control and quality assurance;
- Development-Oriented Activities: Specification (define the scope and intended user); conceptualization (investigate the philosophical constructs of knowledge); formalization (build a formal logical structure of the main constructs); and implementation (coding in a computational language).
- Support Activities: parallel to the above, the following activities have to be conducted: knowledge acquisition, evaluation, integration, documentation, and configuration management. Evaluation makes a technical judgment of the ontologies, their associated software environments during each phase and between phases of their life cycle. Integration is required when building a new ontology and reusing other ontologies that are already available. Configuration Management records all the versions of the documentation, software and ontology code to control the changes.

Brank et al. (2005) classified (AI-based) ontology verification approaches into the following categories:

- “Association/comparison to a “golden standard”, which could be an upper ontology or a theory of knowledge—for example, the work of Maedche and Staab (2002);
- Testing through using the ontology in an application and evaluating the results (e.g. Porzel and Malaka 2004);
- Comparisons with collection of documents (e.g. Brewster et al. 2004);
- Human assessment (e.g. Lozano-Tello and Gómez-Pérez 2004).”

DOLCE (Descriptive Ontology for Linguistic and Cognitive Engineering) represents the school of thought that attempted to mix AI-approach with methods inspired by software and computer science practices. DOLCE is an ontology of particulars that attempted to adopt pragmatic philosophical aspects along with practical terms to describe domain particulars using a flexible set of universals.

In the same line, many AI (artificial intelligence) researchers attempted to enhance the above computational practices by blending elements of philosophy and knowledge modeling into the validation mechanisms (please see: Bench-Capon 1990; Gruber 1995; Chandrasekaran and Josephson 1997). After reviewing these and other research in AI and ontology evaluation, Visser and Bench-Capon (1997) developed a comprehensive analysis of the issue of ontology effectiveness and representation. They concluded that the following criteria should be considered during ontology evaluation:

- Epistemological adequacy: The degree to which the ontology resembles the cognitive sentence:
  - Epistemological clarity: Do all concepts and relations in the ontology have a unequivocal meaning?
  - Epistemological intuitiveness: Do the ontological concepts and relations provide a vocabulary that matches the intuition of the experts in the domain?
  - Epistemological relevance: Are all concepts and relations in the ontology relevant for the domain?
  - Epistemological completeness: Does the ontology cover all relevant concepts? Are there any entities that cannot be modeled n the ontology?
  - Discriminative power: Does the ontology have enough discriminative power in that it provides distinctions at a sufficiently high granularity level (viz. sufficient detail)?
- Operationality: Effort required to implement ontological concepts and relations in a representational language:

  - Encoding bias: Does the ontology rely on symbol-level choices? An encoding bias results when a representation choice is made purely for the convenience of notation or implementation.
  - Coherence: Is the ontology coherently defined in that it is internally consistent?
  - compositionality: Does the ontology provide a basis for (computational) representation, and is this representation computationally adequate?
- Reusability: The degree in which the ontology can be reused to conceptualize new concepts:
  - Task-and-method reusability: is the ontology dependent on certain (types of) tasks and methods?
  - Domain reusability: Is the ontology dependent on certain (types of) sub-domains?

7. BUILDING ONTOLOGIES: THE LEXICAL DIMENSION

The knowledge acquisition part of ontology building is the main bottleneck in the whole development process and also is a major threat to validity (Spyns and Reinberger 2005). Computational linguistics approaches provide a major help in information and concept extraction, pattern detection and analysis, and relationship semantics. These tools are typically applied to text corpus and databases. Lately, however, the web is becoming a major source for raw data. Its less formal structure is, however, challenging to the typical formal tools of computational linguistics and lexical analysis. One of the major tools in this regard is Folksonomies: ad hoc taxonomy-like vocabularies used in social web contexts. Originally defined by Wal (2010) as a combination of "folk" and "taxonomy", it arises from tagging and annotation of web resources, done by its users, and without the constraints of a predefined taxonomy.

Many researchers have found this to be an interesting starting point for finding not only the relevant concepts for a domain, but also their structure and in many cases, the purpose of the whole ontological exercise. This has given rise to the use of tools (mostly bottom-up) such as text mining, peer review systems, communication pattern analysis, machine learning, and linguistic and lexical analysis. However, the most important impact of this shift has been an increasing interest in the study and modeling of relationship.

Maedche and Staab (2000) developed a methodology and an ontology management environment (OntoEdit) that integrates text mining and other lexical features in ontology development. It uses a semi-automated approach to enrich existing ontologies from text sources. DOGMA (Developing Ontology-Guided Mediation for Agents) is another ontology engineering approach that integrates Object Role Modelling (ORM) method with lexical analysis tools. It aims at 1) reducing natural language texts into a set of basic beliefs formulated in binary elementary sentences, and 2) discovering a meta-lexicon through meaning disambiguation (Reinberger and Spyns 2005).

However, a major problem for this approach (which hinders its internal and construct validity) is the quality and formalizations of relationships and axioms generated through these techniques—especially in the unsupervised approach (where no human expert is involved). The supervised approach, which requires human intervention in text preparation, presents a suitable solution. This usually starts with compilation/adaptation of a semantic dictionary (base terms), defining and labelling seed terms (for text extraction) and adopting/coding grammar rules for the domain. For the actual extraction of terms, distributional analysis, statistical analysis, machine learning techniques, neural networks can be used. To discover semantic relationships between concepts, one can rely on already established semantic networks, co-occurrence patterns, machine readable dictionaries, association patterns or combinations of all these (Auger and Barriere 2008; Reinberger and Spyns 2005) .

8. ANALYSIS: THE NEED FOR RETHINKING THE VALIDITY EXERCISE

So far, all research in AEC ontology has focused on the AI dimension. It is argued that we need to expand our research horizon into the other two dimensions as well. To be relevant, ontology research in AEC has to address the need to enrich its contents with linguistic elements given the ongoing interest in linguistics and human communication as a result of the increasing prevalence of social web and systems. Consequently, it is further argued, that this requires ontology research to 1) embrace more rigorous methodological approaches with particular focus on the various validity types, and 2) seriously consider the epistemological foundations of any ontology. The following section provides a rationale and some analysis of these arguments.

8.1 Computational model vs domain theory

The typical computational approach for validation is no longer sufficient for informatics ontologies. The progress of social semantic web means that ontology use is now more related to human communication rather than just a computational artefact. In the pure AI approach (for example, TOVE), ontologies are just an AI model that formalizes concepts for computational use and interoperability. In contrast, in the social semantic systems arena, Gruber's famous definition of ontology should/can (now) be further enriched/modified as followed: informatics ontologies are practical "theories" of domain usable knowledge. The difference between computational models and domain theories is quite fundamental. Typically, a computational model (AI approach) focuses on the formality of knowledge representation and its ability to be tested using computational means (such as reasoners). Theories tend to emphasis the epistemological aspects and the essence of domain knowledge beyond computational issues. BFO represents a case in point in this regard. It has strict adherence to (realist) epistemological definitions of (limited) universals. This could make some concepts too metaphysical for practical use in communication or reasoning exercises. Overemphasizing epistemological rules can lead to excessive abstraction (some can argue, reductionism) that is suitable for a philosophical ontology in contrast to an informatics ontology. DOLCE is closer to achieving the role of ontology as defined above given that it addresses and uses practical terms. This is mainly because of two reasons: the adoption of a pragmatic epistemology in DOLCE and its consideration of linguistic aspects (albeit limited).

In short, not considering the epistemological foundations of an informatics ontology reduce it to a computational model. Overemphasising epistemology could cast the informatics ontology into a typical philosophical ontology (with all the associated vagueness). It seems that focusing on communication and usage by users, and adhering to practical aspects is what distinguishes informatics ontologies from its two ancestors (computational models and philosophical ontologies).

8.2 Theory and theory building

defining informatics ontologies as theories (not just computational models) and emphasising its epistemological aspects raises the interest in the process of ontology building itself (or what is called theories-in-use). In practical sciences, we (first) experience, observe and then try to understand how a phenomenon presents itself and works through a set of ideas that constitute an informed knowledge framework or what can be called theories-in-use. "They are put into practice or use precisely because they help us to understand, explain, anticipate, know, and act in the world in better and more informed ways, and to better ends and outcomes (Argyris and Schon 1996)." Theories-in-practice can be transferred into theories by one of two means a leap of faith or through inquiry and discovery (Lynham 2002a). She further explains that "Theory building can be described as "the purposeful process or recurring cycle by which coherent descriptions, explanations, and representations of observed or experienced phenomena are generated, verified, and refined (Lynham 2000b)."

Good theory building should result in two kinds of knowledge: outcome knowledge, usually in the form of explanative and predictive knowledge, and process knowledge, for example, in the form of increased understanding of how something works and what it means (Dubin 1976). Good theory and theory building should also reflect two important qualities: rigor and relevance (Marsick 1990a), or what are also termed validity and utility (Van de Ven 1989). Theory building achieves these two desired knowledge outputs and empirical qualities by use of what is called "the logic-in-use" and the "reconstructed logic". By following a logical cognitive style in the development and application of the theory and by explicitly reconstructing, or making explicit the "logic-in-use" one gains more than just developing a model or an ontology, but also enriches the wisdom and practices of doing so (Kapaln 1964).

Scoping ontology beyond a mere computational model poses the following fundamental challenges to any validation exercise (which has to be carefully managed if AEC ontologies are to be relevant and practical theories of the domain):

8.3 Artificiality

Ontologies are artificial/theoretical models of knowledge. In fact, a firm outcome in this exercise is not possible, because the "dedicated sceptic cannot be proven wrong by logical arguments (Starmer 1999)". For more on this, see Schram (2005).

8.4 Complexity

The current state of advancement in scientific theories and their ever-increasing connectivity along with the desire for interoperability across domains, make the scope of most reasonable ontologies very complex. The Duhem-Quine thesis shows that one can never test a theory (or hypothesis) in isolation. One is always simultaneously testing various auxiliary hypotheses. Attempts to limit the scope of an ontology may not be the solution as their negative impact (such as imputation and reduced domain coverage or lack of rigour) may outweigh the positive impacts of such action (such as higher levels of validity and/or simplicity).

8.5 Inter-tensions between validity objectives

The complex and artificial nature of ontology development and usage create an ever exiting tension between internal and external validity. An internally valid design will yield results that are robust and replicable. External validity refers to the possibility of generalizing the conclusions to situations that prompted the research. Where internal validity often requires abstraction and simplification to make the ontology more tractable, these concessions are made at the cost of decreasing external validity (for more on this, see Loewenstein 1999).

8.6 Subjective dynamism

The above challenges could have limited impacts on validity had the target domain been more quantitative. The domain of modeling in ontology is knowledge. Unlike data and information (which are easily defined and codified), knowledge is a more fuzzy concept. Knowledge is "a fluid mix of framed experience, values, contextual information, and expert insight that provides a framework for evaluating and incorporating new experiences and information. It originates and is applied in the minds of the knowers (Davenport and Prusak 1998)". More importantly, this fluid mix changes dynamically.

9. SOME RELEVANT LESSONS LEARNED

To meet the epistemological and linguistic challenges of ontology building in future AEC research we need to benchmark best practices in other domains. Luckily, we can draw on the expertise of other domains in this regard—particularly, social and behavioural sciences, management sciences, and economics. The following summarizes some of these:

9.1 Falsify your methodology (expose your intermediate theory)

If a researcher examines the composition of his/her research conceptual framework, he/she may discover not only the limitations of the methodology but also the hidden theory leading him/her in the conceptualization of the model being developed. Critiquing this underlying theory and experimenting with variations of it, is one of the most effective tools to discover a theory about the truth (see also Maxwell 1996; Johnson 1994).

Along with Kaplan and Lynham arguments, Shields and Tajalli (2006) suggested the following micro-conceptual frameworks as building blocks for the "theory as a tool" approach:

- Formalizing Working hypothesis (relevance: during the exploration phase of research): These are loosely formatted provisional, hypothesis meant to support advancing the investigation and the discovery of other critical facts. They should be carefully articulated at the start and questioned throughout the development exercise.
- Development of categories (relevance: during the description phase): clustering concepts into categories and/or universals.
- Practical Ideal type (relevance: gauging): developing criteria to judge the effectiveness of inquiry and then to collect evidence to contrast the reality of the program against the criteria.
- Model of Operations Research (relevance: decision making): how to use the resulting model in decision making.
- Formal hypothesis (relevance: explanation and prediction): this micro-conceptual framework coincides with the "hypothetico-deductive" method of inquiry (for more on this, see Kaplan 1964).

9.2 Triangulation

The systematic mixing and integration of a variety of tools can help in assuring validity as they may corroborate each other. Miles and Huberman (1994) distinguished five kinds of triangulation in qualitative research:

- Triangulation by data source (data collected from different persons, or at different times);
- Triangulation by data type (e.g., combining quantitative and qualitative data);
- Triangulation by method (observation, interviews, documents, etc.);
- Triangulation by researcher; and
- Triangulation by theory (using different theories to explain results);

9.3 Scoping

Researchers should resist at all costs the idea that an ontology is universal. Informatics ontologies are developed for a specific domain of knowledge. In fact, researchers should strive for a minimum ontological commitment (especially if they are not developing an upper ontology). However, the famous minimalist approach (suggested by Gruber 1995) should be dedicated only to ontology scope (do not model outside your target domain). Minimalism is not acceptable within the ontology itself in particular to construct and internal validity. Close attention to problem definition and requirements analysis procedures are keys for developing a clear and coherent scope. Competency questions have also a great role in making sure that the scope of the ontology is clear. The rigorous application of competency questions provides means to assure comprehensive coverage of concepts once the scope has been defined.

9.4 Iterative development

Applied theory-building research spirals (circulate and deepen) between two major phases theorizing to practice and practice to theorizing. Each of these phases strengthen our knowledge frameworks and produces in-process outputs that guide the applied theory-building research and, ultimately, result in a trustworthy, rigorous, and relevant theory for improved action. “An essential output from the theorizing component of theory building is a coherent and informed theoretical framework, which encapsulates and “contains” the explanation of the phenomenon, issue, or problem that is the focus of the theory (Lynham 2002a)

Like any other model, the minute an ontology is developed it helps enhance our knowledge of the domain including our assumptions. Consequently, it requires some modification. Development steps and basic assumptions has to be revisited and calibrated frequently in an iterative manner before a stable and adequate ontology is reached. Repeated categorization and questioning the micro-conceptual framework of the ontology are of particular importance here. Emphasis on life cycle management (as suggested by Gómez-Pérez et al. 2007) and the triangulation of research methodology provide very helpful means in this regard.

9.5 Expert and users involvement

An ontology is basically a formal model for a domain and its use (by industry). To assure theoretical consistency, during the different phases of ontology development, experts and users should be involved in evaluating the work through workshops. (El-Diraby and Gill 2006), used the following structure for expert interviews

- Navigation/Traversing: The interviewees were asked to locate 25 concepts in the taxonomy, and assess how easy it was to locate them.
- Categorization and abstraction Consensus: Interviewees were asked to indicate whether they agreed or disagreed with the categorization of the 25 concepts (above) as suggested in the taxonomy. Further, interviewees were asked to categorize 25 additional new concepts, which were intentionally removed from the ontology. Experts were asked to rate the ease of this categorization.
- Representation: Now that the interviewees became aware of the contents of the taxonomy, they were asked if any major concepts are missing and/or should be added to the taxonomy.
- Quality of the ontological model: Experts were asked about the overall validity of the taxonomy; if the taxonomy adequately covered the domain well, and whether the taxonomy was biased to one sub-domain.
- Usability: assess the potential benefits of the use of the taxonomy.

9.6 Application development

Developing a software based on the ontology puts the usability and its coherence to the test. It can also help in assessing the lexical relevance of the terms especially if the software uses social web tools (see for example, Lima et al. 2005).

9.7 Benchmarking and Inter-disciplinary peer review

Comparing the ontological model to ontologies in other domains provides significant insights. Similarly, input from peers from other domains can be very enlightening.

10. A ROADMAP FOR ONTOLOGY VALIDATION

The following is a suggested map for ontology validation in future AEC informatics domain. It is based on the assumption that philosophy and linguistics play a major role in informatics as we have to understand in a human (not artificial) way the meaning of concepts, their relationships and their frames of references (context). This in contrast to the current general trend (that follows the AI-based approach), which emphasises reasoning, computaionality, axiomization and support for mathematically-based decision making. It should be noted here that the prevailing epistemology (of the author) is contemporary pragmatism, which, acceptably, "taints" the proposed map.

The proposed roadmap is shown in Table 1. As appropriate, validation should consider the fundamental four aspects of validity: statistical/conclusion, construct, internal and external. In each case, threats to each type of validity should be identified and means for overcoming/addressing these threats should be devised. Orthogonal to the above four types, validity can be conducted at three dimensions: philosophical, AI, and lexical. Given the current state of ontology development and usage in the AEC industry, validation should, at minimum, address the AI dimension of ontology. At least in the view of this author, the philosophical dimension of the map below is more important (albeit harder) and should be seriously considered. Finally, the communication/lexical aspect of ontology development merits substantial consideration not only because of the increasing role of social and semantic web, but more importantly, because of the sheer importance of linguistics and communication in a subjective domain of knowledge such as AEC.

Table 1 shows a matrix of the proposed validation model: the vertical dimension lists the four types of validity along with a set of typical threats and a sample of tools (breakers) that can be used to address them. The horizontal dimension shows the three dimensions (philosophical, AI and lexical). Each cell is further divided into two sub-cells. The top sub-cell indicates the relative importance of the validity type in realizing the philosophical, AI and linguistic dimension. The lower sub-cell shows some of the validation tools that can be used in this regard.

10.1 Threats and their breaker tools

The following represent a short synthesis of generic threats and threat breakers in each of these types:

1) Statistical validity: threats to this type relate to typical problems of statistical analysis such as bad or unrepresentative samples, limited sample size, missing data, miss representing statistical results (see Grimes and Schulz, 2002 for an interesting discussion).
2) Construct validity: threats to this type relate to the adequacy of analysis and problem modeling such as personal bias, reductionism and over simplification. Model imputation and lack of clarity are also two main concerns (for more on this, see Petter et al. 2007; Dascal 1992).
3) Internal validity: threats to this type are specific to the internal structure of a model, in particular its relationships. This includes over generalization of the relationships, context-insensitivity in the relationships, and disregard to dynamism in defining the relationships.
4) External validity: threats to this type are similar to the internal validity but relate to the applicability of the whole ontology—over generalization of possible model use scenarios, context-insensitivity (ambiguity) in relation to scope of applicability, and disregard to dynamic changes in problem scope (see Guala 2002)..

Table 1: Validity Roadmap

		Objective	Philosophical Dimension Knowledge discovery abstraction, and representation.	AI Dimension Reasoning, Usability in software	Linguistic Dimension Human communication
Validity	statistical	Threats: representativeness, relevance, quality and coverage of sources, limited sample size, missing data, miss representing statistical results	Not applicable	Not Important	Fundamental
		Sample breakers: diversify, prune sources, define scope		Interviews, scoping and domain analysis, literature reviews	Text Mining, natural language processing, folxonomies (as source of terms)
	Construct	Threats: personal bias, reductionism and over simplification. Model imputation and lack of clarity are also two main concerns	Fundamental	Relevant	Not important
		Sample Breakers: investigator triangulation, theory triangulation	Logical debates on quality of ontological models, formal logic, benchmarking.	Benchmarking, comparative analysis, iterative development and testing, using existing upper ontologies	folxonomies (as benchmarks for terms), comparative linguistic analysis, machine learning and (communication) pattern analysis
	Internal	over generalization of the relationships, context-insensitivity in the relationships, and disregard to dynamism in defining the relationships	Important	Fundamental	Important
		Sample Breakers: negative case sampling, pattern matching	Logical debates, use cases (scenarios), depth and consistency of axioms.	Formal Logic, competency questions, error checking and reasoners	Text parsing, lexical analysis, computational linguistics (relationship/semantic analysis).
	External	over generalization of possible model use scenarios, context-insensitivity (ambiguity) in relation to scope of applicability, and disregard to dynamic changes in problem scope.	Relevant	Important	Relevant
		Sample Breakers: test triangulation.	Academic Peer review, usage levels	Interviews, software testing	Peer review (social), usage levels, interviews

10.2 The Philosophical dimension

The main objective of this dimension is a pursuit of deeper and innovative analysis of the foundations of AEC knowledge and its fundamental constructs and their relationships. On one hand, this dimension should push the limits of our exploration and challenge our existing beliefs. On the other hand, work on this dimension should benchmark the rich literature on philosophy of science and epistemology. In short, we need to inter-breed as we push the envelope. Luckily, validation tools in this domain tend to be relaxed, subjective and less formal and less conclusive than their counterparts. Consequently, they are conducive for innovation and out-of-the-box thinking. Some of these tools are listed below:

1) Statistical validity: statistical validity should not be a major concern in analyzing validity at this dimension. This dimension is about idea generation and challenging or discovering beliefs. The statistical relevance of these ideas is far less important than their logical relevance.
2) Construct validity: in the view of the author, this is the most important validity type in this domain. The philosophical validity and relevance of an ontology is not in its enumeration or listing of concepts or even axioms. It is in its claims on the fundamental constructs of these concepts and axioms. What is the theory of knowledge being claimed? What are the components of this theory? Which "universals" are being claimed? How they behave and how do they interact? Why? What epistemology is being adopted? What propositions are being claimed? How does time and context impact these propositions?

   To scope this cell (construct validity of the philosophical dimension) is not an easy task. However, one telling example should serve as a fundamental guide to AEC practitioners in this regard: the debate/comparison between DOLCE and BFO. Both are upper ontologies but both have completely different epistemological views and, consequently, constructs.

   Some of the relevant (generic) tools in this cell include: logical debates about the quality and relevance of the proposed constructs and how to define them and how to develop their transient behaviour over time and context; benchmarking existing ontologies (whether in informatics or pure philosophy); and formal logical testing/analysis (if the ontology is being written in a formal logic format).

3) Internal validity: this cell is quite important as it relates to the soundness of the ontological model. Fundamentally, there is a need to "stress" the interrelationships between the constructs and "test" the resilience of their "structure". Do they follow logically? Can anomalies exit? When (under which conditions)? Why (which assumption, conditions or axioms lead to that)?

   Some of the (generic) tools in this regard include: logical debates; and utilization of real or hypothetical use cases (scenarios) to test the consistency of the axioms and concept categorization. Analysis of the depth and consistency of axioms in this regard is very crucial. While definitional axioms (axioms that just define a concept or some of its attributes) are very important to the automated semantic analysis (by computers), axioms that are developed to describe and/or limit the "behaviour" of a concept or a group of concepts are far more important. They should be the basis of assuring the internal validity of the philosophical dimension.

4) External validity: this type is relevant and applicable to the philosophical dimension albeit being very low in the importance scale. The rationale here is that irrespective of the applicability of an ontology, we should allow free (yet vetted) thoughts to surface and be discussed. Brainstorming and discovery (indeed, challenging existing beliefs) is far more important, in this dimension, to the real applicability (in the world of pragmatics) of the proposed ontology.

   When it comes to testing the external validity of the philosophical dimension of an informatics ontology (especially if the author claims its immediate applicability to current/future contexts), the following tools are of use: peer review (by academics and industry experts), and levels and extent to which others use/support the main claims of the "ontological model" and its findings about universals.

10.3 The AI dimension

This part is by far the most developed and most formal dimension in ontology studies. While the fundamental (at least philosophical) objective of ontology is to, just, represent knowledge, AI-inspired ontologies aim at understanding human cognition and representing means to mimic their reasoning in order to optimize or make a decision. Consequently, this dimensions normally requires that the ontology be coded into a formal logic format and/or formal languages-- starting from UML in the notational aspect, to OWL in the documentation/representation aspect, to one of the rule-checking tools (such as RACER) in the axiomization aspect, to general programming languages (such as Java) in the application aspect.

1) Statistical validity: depending on the AI method/approach being used, statistical validity could be important or less important. In all cases, it is relevant. If the approach being used is to synthesis existing

data, then the statistical relevance and representativeness of these target (sample) data is important. If the approach is pure modeling based on experience and introspection within a specific target (as is the case of Grüninger and Fox approach), then statistical validity (of sources) is not that important. In their view, an ontology is just a model. It should adhere to and be tested against competency questions. Consequently, in this inside-out approach, there is a limited need to scan sources before modeling the domain at hand. This "modeling" approach is somehow prevalent (not necessarily unanimous) amongst researchers who work on AI-inspired ontologies.

Some of the relevant validation tools here include: critical analysis of literature, formal requirements analysis procedures, expert interviews and detailed definition and management of scope and limitations.

2) Construct validity: the typical lack of interest in pure knowledge presentation (and focus on mimicking human reasoning) reduces the interest and criticality of construct validity in typical AI-inspired ontologies. Researchers tend to be very pragmatic and practical in this regard. There is no pressing desire for abstraction and discovery of core constructs. One of the most important tools that are applicable here is using existing upper ontologies (which assures that the constructs of the new ontology are soundly benchmarked after the upper ontology). Other relevant tools include: benchmarking other models, comparative analysis, and iterative development of the new ontology.
3) Internal validity: given the reasoning aim of AI-inspired ontologies, assuring internal validity and soundness of ontology logical structure are of paramount importance. Progressive consistency between axioms is at the core of this cell in the map. There should be no conflict between them and all combinations of axioms have to yield their intended result rigorously. Fortunately, because of the long history of AI systems and the precedence of the AI school in informatics ontologies, a set of mature and formal tools are available in this regard—for example, formal logic (particularly first order logic), competency questions, automated reasoners (such as RACER), debuggers and error checkers.
4) External validity: in a typical AI application, external validity is rather very important. However, it is contended here that this is not the case in AI-inspired ontologies. One of the main challenges to ontology external validation is its name, which is typically associated with universality. Contrary to that, and in agreement with Fox and Grunninger approach, an ontology is just a model of knowledge in a domain. In so far as it means addressing relevance and applicability to a target domain, external validity is required for ontologies. However, proving universality should be avoided at all costs. Some of the relevant tools in this raged include: interviews (not necessarily extensive-as universality is not a requirements), using the ontology in software systems and testing the usability and value of ontology-supported software (see for example, Lima et al, 2005).

10.4 Lexical/communication dimension

While this dimension has flourished in other domains (especially those with extensive formal database and textual croups such as the Biomedical domain), the AEC industry may not be fully ready for its use given the historical lack of formalized vocabulary and the lack of extensive textual databases. However, this dimension is very promising and more attention should be directed to it given its potential large impact on external validity (see for example, May and Lausen 2004).

1) Statistical validity: the representativeness and quality of source data that is used in lexical, semantic or other computational linguistic exercises are of fundamental importance to the validity of any work in this regard. Some of the available tools in this regard include: text and data mining, natural language processing and the use of Folksonomies as a source of terms.
2) Construct validity: while construct validity is generally advocated by this map, practically it is not of a major concern in this dimension. The reason is that this whole dimension (lexical and communication aspects of an ontology) is seen as a domain of pragmatics. In other words, this dimension is not the place to start or establish the basic constructs of the terms people use. To the contrary, we should be concerned with and using tools for discovering what language and terms people use, hence the paramount importance of statistical validity in the area of lexical validation of ontologies. To suggest a strong importance to construct validity within this dimension may be beyond the capabilities of existing tools that we have at the moment. Some of these tools include: Folksonomies (as benchmarks for

terms), comparative linguistic analysis, machine learning and (communication) pattern analysis, semantic matching (see for example, Giunchiglia and Shvaiko 2004; Reinberger et al. 2003).

3) Internal validity: internal validity in this dimension refers to the consistency of linguistic nature (aspects) of the terms selected. Investigation of the internal validity at the lexical level in any ontology can be the best means to discover and refine the relationships used in the ontology (see Wong et al., 2009) and, hence, add required depth to the semantic and triangulation of concepts. Some of the relevant tools in this regard include: Text parsing, lexical analysis, computational linguistics (relationship/semantic analysis). See for example, Budanitsky and Hirst (2000).
4) External validity: if we agree that an ontology is needed for pure knowledge representation, reasoning and to support human communication, then external validity at the lexical level is of great importance. Upon usage, people should find the terms of the ontology relevant and meaningful. However, we should avoid the quest for universality. Some of the relevant tools in this regard include: Peer review by practitioners about the usability proposed terms, analysis of usage levels and any usage problems/issues, and formal interviews of users (again, not necessarily and extensive number of these as universality is not a requirement).

11. SUMMARY

Informatics deals with knowledge management (not just information technology), which, by its very nature, is social, cognitive and tied to linguistics and human communication. In fact (beyond the mechanics, notational, and coding aspects), at the core of ontology research methodology is the theorization and the sound creation and testing of conceptual frameworks (Steiner 1988). Conceptual frameworks are maps that link and give coherence to all aspects/ stages of a research endeavour (e.g., problem definition, purpose, literature review, methodology, data collection and analysis).

It is further argued that the use of linguistic and epistemic analysis are not only methodological issues (needed to enrich ontologies), but also a necessity to handle the inexorable rise in published material (scientific, business and social ones). This is why text mining and linguistic approaches have flourished in the legal and medical domains. The medical domain in particular has had significant success in using these approaches mainly due to three factors: the stability of medical terms (many technical terms have stabilized over centuries if not millennia), the extensive use of formal database (which makes lexical analysis more accurate), and the sheer volume of publications in the domain (Rebholz-Schuhmann et al 2005). Similarly, in the legal domain, the extensive use of documents as the fundamental communication means was very conducive to the use of text mining and linguistic analysis tools. The significant role of logical analysis (of cases) has also enhanced the interest and use of epistemic analysis. In contrast, AEC information is fundamentally presented graphically (in drawings) and the industry (especially the construction sector) intensely depends on oral communication (Ballan and El-Diraby 2011). Given the recent history of ontology development in AEC, it is possible that such tools will receive further attention from researchers in the field.

The fundamental keys to achieve validity in such a challenging environment are the integration of formal and rigorous epistemic inquiry (not just the mechanical aspects) of ontology development; and a sound self-questioning research methodology. The proposed map tried to address that by integrating four types of validity (statistical/conclusion, internal, construct, external and along with three dimensions of inquiry (philosophical, AI and linguistic). These four types of validity were selected because they present a cohesive and consistent definition of validity. This, of course, does not preclude other types of validity that can be the topic of future work. Along with these dimensions, a set of threats, threat breakers and tools are presented to support building a good theory while developing an ontology.

12. ACKNOWLEDGEMENT

This paper was possible because of a probing question from Prof. Lucio Soibelman and input from my former students Hesham Osman and Nora El-Gohary.

13. REFERENCES

Akinci, B., Fischer, M., Kunz, J., and Levitt, R. (2002). “Representing Work Spaces Generically in Construction Method Models.” J. Constr. Engrg. and Mgmt., ASCE, 128, 296.

Aksamija, A., and Grobler, F. (2007). “Architectural Ontology: Development of Machine-Readable Representations for Building Design Drivers.” Conf. on Computing in Civil Engineering, ASCE.

Argyris, C., and Schon, D. A. (1996). “Organizational learning II. Theory, method, and practice.” Addison-Wesley, New York, NY.

Auger, A., and Barriere, C. (2008). “Pattern-based approaches to semantic relations extraction.” John Benjamins Publishing Co., Philadelphia, PA.

Ballan, S., El-Diraby, T. E. (2011) A Value Map for Communication Systems in Construction, ITcon Vol. 16, pg. 745-760, http://www.itcon.org/2011/44

Bench-Capon, T. J. M. (1990). “Knowledge Representation; An Approach to Artificial Intelligence.” APIC series, No. 32, Academic Press, London, United Kingdom.

Black, T. R. (1993). “Evaluating social science research: an introduction.” Sage Publ. Inc. Thousand Oaks, CA.

Brank, J., Grobelnik, M., and Mladenić, D. (2005). “A survey of ontology evaluation techniques.” Conf. on Data Mining and Data Warehouses, Ljubljana, Slovenia.

Brewster, C., Harith, A., Srinandan, D., Yoric, W. (2004). “Data driven ontology evaluation.” Int. Conf. on Language Resources and Evaluation, Lisbon, Portugal.

Brown, J. D. (1996). “Testing in language programs.” Prentice Hall Regents, Upper Saddle River, NJ.

Budanitsky, A., and Hirst, G. (2000). “Semantic distance in WordNet: An experimental, application-oriented evaluation of five measures.” Workshop on WordNet and Other Lexical Resources, Pittsburgh, PA.

Calder, B. J., Phillips, L. W., and Tybout, A. M. (1982). “ Beyond external validity.” J. of Consumer Research, 10 (1), 112-114.

Chandrasekaran, B., and Josephson, J.R. (1997). “The Ontology of Tasks and Methods.” Working Notes of the AAAI Spring Symposium on Ontological Engineering, Stanford Univ., CA.

Cook, T. D., and Campbell, D. T. (1979). “Quasi-experimentation - design and analysis issues for field settings.” Rand McNally College Publ. Inc., Chicago, IL.

Cronbach, L. J., and Meehl, P. E. (1955). “Construct validity in psychological tests.” Psychological Bulletin, 52 (1), 281-302.

Dascal, M. (1992). “Why Does Language Matter to Artificial Intelligence?” Minds and Machines, 2 (2), 145-174.

Davenport, T., & Prusak, L. (1998). “Working Knowledge.” Harvard Business School Press, Boston, MA

Dubin, R. (1976). “Theory building in applied areas.” Handbook of industrial and organizational psychology, Dunnette M. D., ed., Rand McNally, Chicago, IL., 17-39

Dubin, R. (1978). “Theory building (Rev. ed.).” Free Press, New York, NY.

Dumontier, M. and Hoehndorf, R. (2010) “Realism for Scientific Ontologies”, 6th International Conference on Formal Ontology in Information Systems, Toronto, Canada.

El-Diraby, T. E., and Kashif, K. (2005). "A distributed ontology architecture for knowledge management in highway construction," J. Constr. Engrg. and Mgmt., ASCE, 131 (5), 591-603.

El-Diraby, T. E., Lima, C., and Fies, B. (2005). “A Domain Taxonomy for Construction Concepts: Toward a Formal Ontology for Construction Knowledge.” J. Comp. in Civ. Engrg., ASCE, 19 (3), 394-406.

El-Diraby, T. E., and Gill S. (2006). "A taxonomy for construction terms in privatized infrastructure finance: toward a semantic cross-industry exchange of knowledge,” J. of Construction Management & Economics, 24(2), 271-285.

El-Diraby, T. E., Lima, C., and Fies, B. (2006). “Strategies for incorporating data exchange standards and classification systems in e-business taxonomies,” J. of Information Technology Management, XVI (4), 26-38.

El-Gohary, N., and El-Diraby, T. E. (2010a). “A Domain Ontology for Processes in Infrastructure and Construction.” J. Constr. Engrg. and Mgmt., ASCE, 136 (7), 730-744.

El-Gohary, N., and El-Diraby, T. E. (2010b). “A Dynamic Knowledge-Based Process Integration Portal for Collaborative Construction” J. Constr. Engrg. and Mgmt., ASCE, 136 (3), 316.

El-Gohary, N., and El-Diraby, T. E. (2011). “Merging Architectural, Engineering and Construction (AEC) Ontologies.” J. Comp. in Civ. Engrg., ASCE (In Press).

Fernández-López, M., Gómez-Pérez, A, and Juristo N. (1997). “METHONTOLOGY: From ontological art to ontological engineering.” Spring Symposium on Ontological Engineering of AAAI, Stanford, CA.

Fernández-López, M. and Gómez-Pérez, A. (2002). “Overview and analysis of methodologies for building ontologies.” The Knowledge Engineering Review, 17(2), 129-156.

Fumerton, R. (2009). "Foundationalist Theories of Epistemic Justification." The Stanford Encyclopaedia of Philosophy (Winter 2009 Edition), Edward N. Zalta E. N., ed., (May 2, 2009).

Giunchiglia, F., and Shvaiko, P. (2004). “Semantic Matching.” The Knowledge Engineering Review, 18(3), 265-280.

Gómez-Pérez, A. (2001). “Evaluation of ontologies.” Int. J. of Intelligent Systems, 16(3), 391-409.

Gómez-Pérez, A., Fernández-López, M., and Corcho, O. (2007). “Ontological engineering: what are ontologies and how can we build them?” Semantic Web Services: Theory, Tools and Applications, Cardoso, J., ed, IGI Global, Hershey, PA.

Grimes, D. A., and Schulz, K. F. (2002). “Bias and causal association in observational research.” Epidemiology series, the Lancet, Vol. 359.

Gruber, T. (1995). “Toward principles for the design of ontologies used for knowledge sharing.” Int. J. of Human Computer Studies, 43(5-6), 907-928.

Grüninger, M. and Fox, M. S. (1995). “Methodology for the design and evaluation of ontologies.” IJCAI95 Workshop on Basic Ontological Issues in Knowledge Sharing, Montreal, Canada.

Guala, F. (2002). “On the Scope of Experiments in Economics: Comments on Siakantaris.” Cambridge J. of Economics, 26 (1), 261-267.

Guarino, N. (1997). “Understanding, building, and using ontologies—A commentary to Using explicit ontologies in KBS development by van Heijst, Schreiber, and Wielinga.” Int. J. Computer-Human Studies, 46(2–3), 293–310.

Hacking, I. (ed.) (1981). “Scientific revolutions.” Oxford Univ. Press, New York.

Hofstadter, D. R. (1980). “Gödel, Escher, Bach: An eternal golden braid.” Vintage Books/ Random House, New York, NY.

Johnson, R. B. (1994). “Qualitative research in education.” SRATE Journal, 4 (1), 3-7.

Kaplan, A. (1964). “The Conduct of Inquiry: Methodology for Behavioural Science.” Chandler Publishing Co., Scranton, PA.

Kim, H., and Grobler, F. (2007). “Building Ontology to Support Reasoning in Early Design.” Conf. on Computing in Civil Engineering, ASCE.

Lima, C., El Diraby, T. E., and Stephens, J. (2005). “Ontology-based optimization of knowledge management in e-construction.” J. of IT in Construction (ITCON), 10, 305-327

Liu, D., Cheng, J., Law, K. H., Wiederhold, G., and Sriram, R. D. (2003). “Engineering Information Service Infrastructure for Ubiquitous Computing.” J. Comp. in Civ. Engrg., ASCE 17, 219.

Loewenstein, G., (1999). “Experimental Economics from the Vantage-point of Behavioural Economics.” The Economic Journal, 109, F25-F34.

Lozano-Tello, A., and Gómez-Pérez, A. (2004). “Ontometric: A method to choose the appropriate ontology.” J. Database Management, 15(2), 1–18.

Lynham, S. A. (2000). “Theory building in the human resource development profession.” Human Resource Development Quarterly, 11(2), 159-178.

Lynham, S. (2002a). “The general method of theory-building research in applied disciplines.” Advances in Developing Human Resources, 4, 221-242.

Lynham, S. (2002b). “Quantitative Research and Theory Building: Dubin's Method.” Advances in Developing Human Resources, 4, 242-261.

Maedche, A., and Staab, S. (2000). “Mining Ontologies from Text.” 12th European Workshop on Knowledge Acquisition, Modeling and Management, Juan-les-Pins, France.

Maedche, A., Staab, S. (2002). “Measuring similarity between ontologies.” European Conference on Knowledge Acquisition and Management- EKAW-2002, Madrid, Spain.

Marsick, V. J. (1990a). “Altering the paradigm for theory building and research in human resource development.” Human Resource Development Quarterly, 1(1), 5-24.

Maxwell, J. A. (1996). “Qualitative research design.” Sage, Newbury Park, CA.

May, W., and Lausen, G. (2004) “A uniform framework for integration of information from the web.” Information Systems, 29(1), 59-91.

Miles, M. B. & Huberman, A. M. (1994). “Qualitative Data Analysis.” Sage, Thousand Oaks, CA:

Osman, H., and El-Diraby, T. E. (2006). “Ontological Modeling of Infrastructure Products and Related Concepts.” Transportation Research Record *No. 1984, Highway Facility Design*, TRB.

Petter, S., Straub, D., and Rai, A. (2007). “Specifying formative constructs in information systems research.” MIS Quarterly, 31 (4), 623-656.

Popper, K. (2002). “The logic of scientific discovery” (J. F. K. Popper & L. Freed, Trans.). Routledge Classics, New York, NY.

Porzel, R., Malaka, R. (2004). “A task-based approach for ontology evaluation.” ECAI 2004 Workshop Ontology Learning and Population.

Quine, W.V., and Ullian, J.S. (1978). “The Web of Belief, 2nd edition.” Random House, New York, NY.

Rebholz-Schuhmann, D., Kirsch, H., and Couto, F. (2005). “Facts from Text—Is Text Mining Ready to Deliver?” PLoS Biol, 3(2), e65.

Reinberger, M.-L. & Spyns, P. (2005). “Unsupervised text mining for the learning of DOGMA-inspired ontologies.” Ontology Learning from Text: Methods, Applications and Evaluation, P. Buitelaar, P. Cimiano & B. Magnini (eds), IOS Press, Amsterdam, the Netherlands

Reinberger, M.-L., Spyns, P., Daelemans, W., and Meersman, R. (2003). "Mining for Lexons: Applying Unsupervised Learning Methods to Create Ontology Bases." CoopIS, DOA, and ODBASE, Springer-Verlag.

Rezgui, Y. (2006). “Ontology-Centered Knowledge Management Using Information Retrieval Techniques.” J. Comp. in Civ. Engrg., ASCE, 20, 261

Schram, A. (2005). “Artificiality: the tension between internal and external validity in economic experiments.” CREED, University of Amsterdam, the Netherlands.

Shields, P. M., and Tajalli, H. (2006). "Intermediate Theory: The Missing Link to Successful Student Scholarship." J. of Public Affairs Education, 12(3).

Spyns, P. and Reinberger, M-L. (2005). “Lexically Evaluating Ontology Triples Generated Automatically from Texts.” Second European Semantic Web Conf.

Starmer, C., 1999: “Experimental Economics: Hard Science or Wasteful Tinkering?” The Economic Journal, 109, F5-F15.

Staub-French, S., Fischer, M., Kunz, J., and Paulson, B. (2003). “An Ontology for Relating Features with Activities to Calculate Costs.” J. Comp. in Civ. Engrg., ASCE, 17, 243.

Steiner, E. (1988). “Methodology of theory building.” Educology Research Associates, Sydney, Australia.

Tesfagaber, G., Bouchlaghem, N. M., Baldwin, A. N., Cutting-Decelle, A. F. and Khalfan M. A. (2003). “Ontology Modeling of IDEF3 Process Description for AEC Applications Interoperability.” Construction Research Congress, ASCE.

Trochim, William M. (2010). “The Research Methods Knowledge Base, 2nd Edition.” <http://www.socialresearchmethods.net/kb/> (June 20, 2010).

Ugwu, O. O., Arcisweski, T. and Anumba, C. J. (2002). “Teaching Agents How to Solve Design Problems: A Mixed Initiative Learning Strategy.” Information Technology, ASCE.

Uschold, M. and King, M. (1995) “Towards a methodology for building ontologies.” IJCAI'95 Workshop on Basic Ontological Issues in Knowledge Sharing, Montreal, Canada, 6.1-6.10.

Uschold, M. and Gruninger, M. (1996). "Ontologies: Principles, Methods, and Applications." Knowledge Eng. Rev., 11 (2), 93–155.

Van deVen, A. H. (1989). “Nothing is quite so practical as a good theory.” Academy of Management Review, 14(4), 486-489.

Visser, P. R. S., and Bench-Capon, T. J. M. (1997). “A Comparison of Four Ontologies for the Design of Legal Knowledge Systems.” Artificial Intelligence and Law, 6 (1), 27-57.

Wal T. V. (2010). “Folksonomy definition and wikipedia.” http://www.vanderwal.net/random/entrysel.php?blog=1750 (viewed 10/4/2010)

Wang, H. H., and Boukamp, F. (2009). “Ontology-Based Job Hazard Analysis Support.” Conf. on Computing in Civil Engineering, ASCE

Wang, H. H., Boukamp, F., and Elghamrawy, T. (2010). “An Ontology-Based Approach to Context Representation and Reasoning for Managing Context-Sensitive Construction Information.” J. Comp. in Civ. Engrg., ASCE.

Wender, K. and Hübler, R. (2009). “Enabling More Human-Oriented Exploration of Complex Building Information Models.” J. Comp. in Civ. Engrg., ASCE, 23, 84

Wong, W., Liu, W., and Bennamoun, M. (2009). “Acquiring Semantic relations using Web for constructing lightweight ontologies.” 13th Pacific-Asia Conf. on Knowledge Discovery and data mining (PAKDD), Bangkok, Thailand.

Zhang, J. and El-Diraby T. E. (2009). “SSWP: A Social Semantic Web Portal for Effective Communication in Construction.” Journal of Computers, 4 (4).