

EXPLORING ONTOLOGY ADOPTION AND DIFFUSION IN THE CONSTRUCTION VIRTUAL ENTERPRISE

Yacine Rezgui

Informatics Research Institute, University of Salford, Salford, UK.

y.rezgui@salford.ac.uk

The construction sector has adopted for decades the modus operandi of the virtual enterprise. In this context, a large number of semantic resources have been developed to support seamless collaboration between individuals and teams on projects. However, these tend to be (a) discipline oriented and geared towards servicing specific user community needs or applications; and (b) incomplete in terms of concept, and relationships between concepts covered and conveyed. Therefore, none of these (a) enable effective knowledge management across projects and organizations, (b) are accepted across disciplines, and (c) provide a true conceptualization of the construction domain. The paper argues the case for the development of knowledge-rich ontologies, and explores the factors that may hinder or promote their wide adoption and diffusion.

1. INTRODUCTION

Construction is a knowledge intensive industry characterized by its unique work settings and virtual enterprise like modus operandi [1]. Buildings have long been designed and constructed by non co-located teams of separate firms, with various levels of IT maturity and capability, which come together for a specific project and may never work together again. Moreover, the Construction sector is fragmented and the major consequence is the difficulty to communicate effectively and efficiently among partners during a building project or between clients and suppliers of construction products. Several initiatives led by standardization and / or industry consortia have developed data / product models aimed at facilitating data and information exchange between software applications. These efforts include STEP [2] and the Industry Foundation Classes (IFCs). Several other initiatives at a national and European level have developed dictionaries, thesauri, and several linguistic resources focused on Construction terms to facilitate communication and improve understanding between the various stakeholders operating on a project or across the product supply chain. However, these initiatives tend to be country specific and not adapted to the multi-national nature of the sector. Also, given the vast scope of Construction, these semantic resources tend to be specialized for dedicated applications or engineering functions, e.g. product libraries and HVAC (Heating, Ventilation and Air Conditioning), respectively. Some recent research has started highlighting the need for an ontology in the sector [3], while others have already started referring to the IFCs as being an ontology or suggesting that they be extended to become an ontology [4].

The aim of the paper is to explore adoption and diffusion factors of a domain ontology. As such, the core research question is: *What are the factors that may promote or hinder the adoption of a construction ontology?* First, a critical discussion on product data technology is provided arguing the case for knowledge-rich ontology. The requirements for such ontology are then provided supported by an illustration of the eCognos ontology [3]. The paper then discusses the validation process of the eCognos ontology, and answers the main research question. Finally the paper provides concluding remarks and future directions for research.

2. THE NEED FOR RICHER CONCEPTUALISATIONS THROUGH ONTOLOGY

The progress made so far in arriving at the Building Information Model (BIM) concept and its associated tools is undoubtedly a sizeable step forward in the management, communication and leveraging of construction project information. Both the BIM models used by the commercial vendors and the international standards developed for construction such as STEP, IFC and CIS/2 do however still exhibit a number of shortcomings identified in [1; 5; 6] related to the degree of expressiveness, lifecycle support, and comprehensiveness of these product models. Other problems that render data level integration in the STEP or IFC mould less effective than might otherwise be the case. To understand this position it is necessary to consider the way in which data integration mandates a considerable degree of work up-front. This is required in order to agree upon standards, construct a schema for integration, adapt applications to the standards etc. all before any benefits are realized. These issues become more onerous the larger the scope of agreement one is trying to achieve (inter-organizational, national, international etc.). Finally, for large international standards efforts, agility is something of a problem. Once the standard is agreed, changing it can take a considerable amount of time, which in an age of rapidly evolving business needs can turn a formerly helpful system into a hindrance.

This is where an ontology can leverage many of the above problems. Various definitions of what forms an ontology have been formulated and have evolved over time. A good description of these can be found in [7]. From the author's perspective, the best definition that captures the essence of an ontology is the one given by Gruber [8]: "an ontology is a formal, explicit specification of a shared conceptualization". As elaborated in [9]: "Conceptualization refers to an abstract model of some phenomenon in the world which identifies the relevant concepts of that phenomenon. Explicit means that the types of concepts used and the constraints on their use are explicitly defined. Formal refers to the fact that the ontology should be machine processable".

The use of an ontology or multiple ontologies of the construction domain could act as a semantic abstraction layer above current standards and models, which would help to resolve some of the issues outlined above and further integrate project data in a more intelligent fashion. For example, an ontology with appropriate mappings into the underlying data models could be used to provide a more intuitive view of project data for any given actor based on their particular disciplinary concepts and terminology. That same ontology could also provide the view for an actor from a

different discipline, based on the relationships explicated within the ontology itself providing links to the appropriate terminology for the same data items. This type of 'translation' function becomes more compelling when used to view initial project briefs or client constraints and later when viewing the rationale for changes as it helps all actors to understand the reasoning involved in a language they can comprehend easily. With respect to unstructured project information, the use of ontologies in tandem with other techniques drawn from information retrieval/extraction could be used to automatically infer links between the structured and unstructured information and indeed between items of unstructured information, based on the links defined in the ontology. These links lend a greater degree of context to each item relative to the project as a whole. Benefits may also be derived from uncovering previously unseen linkages between various elements of project data using such analysis methods. The e-Cognos project for example, developed and used a construction oriented ontology to augment the services that it offered as part of the collaborative knowledge management environment also developed on the project [1,4]. In the following section, we introduce the requirements of such ontology and provide an illustration of a potential ontology for the sector.

3. REQUIREMENTS FOR ONTOLOGY DEVELOPMENT IN CONSTRUCTION

A critical analysis of the structure and overall practices in the industry [1] as well as available semantic resources, ranging from taxonomies to thesauri, have helped formulate a set of requirements that ought to be addressed in order to maximize the chances of a wide adoption of any ontology project in the construction sector. These requirements are listed below:

- The ontology should not be developed from scratch but should make as much use as possible of established and recognized semantic resources in the domain.
- The ontology should be built collaboratively in a multi-user environment: the construction sector involves several disciplines and communities of practice that use their own jargon and have specialized information needs.
- There is a need to ensure total lifecycle support, as the information produced by one actor within one discipline should be able to be used by others working in related disciplines.
- The ontology must be developed incrementally involving the end-users. This is important given the multi-disciplinary and multi-project nature of the industry, and the fact that each project is a one-off prototype.
- The ontology should be flexible and comprehensive enough to accommodate different business scenarios used across projects and disciplines.
- The ontology should be user friendly, i.e., easy to use and providing a conceptualization of the discipline / domain being represented that embeds the technical jargon used in the sector.
- The ontology should be a living system and should allow for future expansion.

Given the following factors [1]: (a) the fragmented and discipline-oriented nature of the construction sector; (b) the various interpretations that exist of common concepts by different communities of practice (disciplines); (c) the plethora of semantic resources that exist within each discipline (none of which have reached a consensual agreement); (d) the lifecycle dimension of a construction project with information being produced and updated at different stages of the design and build process with a strong information sharing requirement across organizations and lifecycle stages; a suitable ontology development methodology should accommodate the fact that the ontology should be specific enough to be accepted by practitioners within their own discipline, while providing a generic dimension that would promote communication and knowledge sharing amongst these communities.

4. THE ECOGNOS ONTOLOGY

Given the above requirements, an ontology is developed, referred to as eCognos [3]. The eCognos ontology is structured into a set of discrete, core and discipline-oriented, sub-ontologies. Each sub-ontology features a high cohesion between its internal concepts while ensuring a high degree of interoperability between them. These are organized into a layered architecture (three layers) with, at a high level of abstraction, the core ontology that holds a common conceptualization of the whole construction domain enabled by a set of inter-related generic core concepts forming the seeds of the ontology. These generic concepts enable interoperability between specialized discipline-oriented modules defined at a lower level of abstraction. This middle layer of the architecture provides discipline-oriented conceptualizations of the construction domain. Concepts from these sub-ontologies are linked with the core concepts by generalization / specialization (commonly known as IS-A) relationships. The third and lowest level of the architecture represents all semantic resources currently available, which constitute potential candidates for inclusion into eCognos either at the core or discipline level.

There are a large variety of available semantic resources that can form the basis for building the eCognos core ontology. These range from classification systems to taxonomies. The latter do deserve particular attention as argued in [10]. One of the principal roles of taxonomies is to facilitate human understanding, impart structure on an ontology, and promote tenable integration. Furthermore, properly structured taxonomies: (a) help bring substantial order to elements of a model; (b) are particularly useful in presenting limited views of a model for human interpretation; and, (c) play a critical role in reuse and integration tasks. Improperly structured taxonomies have the opposite effect, making models confusing and difficult to reuse or reintegrate [10]. IFCs, being more recent and also the closest taxonomy currently in use in the sector, are therefore the preferred candidate semantic resource that can provide the skeleton on which such a core ontology can be built.

A particular approach is adopted for building and / or expanding the discipline-oriented sub-ontologies. This involves selecting and making use of a large documentary corpus used in the discipline and ideally produced by the end-users. The sub-ontologies are then expanded and built from index terms extracted from commonly used documents using information retrieval techniques [3].

5. TECHNICAL EVALUATION OF THE ONTOLOGY

Two main metrics are used to evaluate the eCognos ontology:

- (a) performance evaluation in terms of response time delivered by the use of the ontology; and,
- (b) the retrieval performance in terms of relevance of the retrieved document set through the use of ontology.

A sample document set comprising about 12,000 documents is used. This was gathered from several recently completed projects. 15 queries have been formulated by discipline experts involved in the research. These were based on their own information needs experiences on projects, traditionally performed manually or using ad-hoc search facilities.

The experts were asked to identify manually the document set matching each formulated query. While the identification process was achieved in a matter of days, this did reflect (timescale wise) the current information and document search practices in the construction industry. These identified relevant document sets have then been used as a basis to compute the average Recall (the fraction of the retrieved relevant documents) and Precision measures (the fraction of the retrieved documents which is relevant). These have been quantified in two scenarios: (a) through the simple use of index terms; and, (b) through the use of the eCognos ontology.

In the first scenario, a full-text summarization of the documents and queries was performed, while in the second scenario the summarization relied exclusively on concepts of the ontology. The discipline experts that generated the queries in the first place and identified the relevant document sets have been involved in the relevance assessment work.

In terms of response time, the difference between the two approaches (with and without the use of ontology) was comparable. The difference was in the retrieval performance provided by the ontology. For further details please see (Rezgui 2007). The main results of the evaluation are summarized below.

Despite reporting missing concepts, the search functionality with ontology outperformed the traditional full-text search approach as more relevant document subsets have been retrieved. The knowledge representation technique, based on ontology, provides a more accurate user and machine interpretable summarization of documents, as illustrated by the field trial results with significant improvements in the Precision and recall factors.

The use of integrated services articulated around a common in-house ontology promotes the wide adoption of common standards and the sharing of a common understanding of terms and concepts. However, subtle variations of the semantics of certain terms may exist across companies. This suggests that discipline ontologies in the construction sector may need further refinement to be adapted to the norms and values of an organization. Some business processes involve the use of concepts from more than one discipline ontology. While this is supported by the eCognos methodology through the generic core ontology, it triggers another issue in relation to supporting dynamic views or perspectives that involve concepts drawn from more than one discipline ontology.

6. ECOGNOS ONTOLOGY ADOPTION AND DIFFUSION

A number of studies have been reported in the literature describing various theories and models related to information technology adoption, diffusion, and innovation into the workplace, as reported in [11]. Some of these theories describe transition processes and mechanisms, including Rogers's stage model of innovations in organizations [12]; whereas others define causality among factors to predict successful transition of a technology, including Davis's Technology Adoption Model (TAM), [13]. TAM argues that end-user acceptance and use of information systems innovations is influenced by their beliefs regarding the technology. In particular, it proposes that perceived usefulness and perceived ease of use influence the use of information systems innovations and that this effect is mediated through behavioural intentions to use (Davis, 1989). The model highlights the critical role of extrinsic motivation and, in particular, expectations of task-related performance gains in end-users' adoption and use of Information Systems innovations (Davis, 1989). TAM is used to discuss the adoption and diffusion of the eCognos ontology. It has been applied in the context of the validation process described in the previous section, whereby a set of KM services articulated around the eCognos ontology have been developed, using the Web services model, tested and validated in a real organizational setting [14].

6.1. The Case Study

The eCognos ontology and overall solution has been introduced within OTH, a French construction engineering company, as an enhanced service of the documentation department. The latter is, by nature, a knowledge centre for the whole organization. Its mission is to collect, organize, classify and disseminate knowledge throughout the group on regulatory, standardisation, technical and commercial issues. Its manager has specific skills in information indexing, searching and retrieval and more generally knowledge management. She has been assigned the role of eCognos manager. The field trial is used to evaluate the integration capacity of the system and the benefits that can be derived from integrating within it various knowledge sources already available and used throughout OTH. In the context of the field trials, deployment of the system has been limited to a consistent and knowledge-friendly group of users drawn from 3 OTH subsidiaries. Five users per subsidiary have been retained supported by the manager of the documentation department. It was decided to base the trial on the "electrical design" domain. An OTH-specific ontology has been developed. It contains general concepts relevant to OTH at a high level, and more detailed concepts in the field of electrical design at lower levels. It includes over 830 ontological concepts. The 3 subsidiaries have been granted access to the deployed eCognos system through an existing Virtual Private Network (VPN).

Three scenarios have been developed simulating (a) the process of setting-up a specialised corporate memory, (b) locating a knowledge asset, and (c) creating value by exploiting available knowledge.

6.2. Perceived Usefulness of the eCognos Ontology

Overall, the trials confirmed the positive role that the use of an ontology can have in addressing the knowledge management needs of employees. Also, it was found that developing and using an in-house ontology gave the users a strong feeling of acceptance of the system. In particular, the use of the knowledge extractor service [14] was highly valued as the mapping of index terms with ontological concepts enables intelligent, automatic indexing (the documentation department no longer need to index manually their documents). The search service was enhanced through the use of the Ontology and provided substantial improvements to the existing system. These have in particular highlighted the advantages that can be gained from an organizational as well as a technical level by providing an integrated and unique access point to all OTH applications. Moreover, by integrating legacy systems within eCognos as well as having access to more advanced functionality, users felt that they could continue using their accustomed tools to manage their projects and at the same time share their knowledge with colleagues in a transparent, integrated and effective way. KM becomes part of the day-to-day work and takes place in the background without intrusion to the practice. Application integration through the system did highlight interesting potentials for improved process integration. The perceived usefulness of the eCognos ontology has been reported on several fronts:

- **Organizational support:** eCognos is perceived as having the potential to leverage intellectual capital both inside as well as outside the organization. It also contributes to lowering project costs through the use of an integrated mechanism for managing centralised or distributed services. Also, the semantics of Construction terms and concepts hold different meaning for different users depending on their work discipline as well as their education background. The use of integrated services articulated around a common in-house ontology promotes the wide adoption of common standards and sharing of common understanding of terms and concepts.
- **Team support:** collaboration is perceived now as possible across distance and time zones as users can now easily publish documents within and across the organization while gaining access to others' experiences. Physical location of people and knowledge sources become irrelevant. Teams on projects have now the potential to better combine distributed competences via eCognos.
- **User support:** ICT systems are generally perceived by users to be invasive and individuals need constantly to adapt to continuous introduction of new emerging technologies in the workplace. The integration of new functionality through the ontology-based web services model becomes transparent to users who do no longer have to use multiple windows, requiring each its own authentication, to conduct a specific task. Activities and business processes encapsulate relevant knowledge sources accessible through a role / right-based approach across the organization and projects. Access to useful and timely information becomes possible regardless of its location, native format or system. The search facility is perceived as a strong advantage, and the information that is pushed to users is filtered and personalised thus avoiding information overload.

6.3. Ease of Use of the eCognos Ontology

In terms of ease of use of the eCognos ontology, it is worth noting that users were not exposed directly to the ontology as they were interacting with it transparently through the invocation of services that relied on the ontology. But still, the trials provided means to assess the level of familiarity and comprehension of the ontological concepts. In that respect the end-users highlighted two important aspects of the ontology:

- The modular structure of the ontology: The ontology takes into account the fragmented nature of the construction sector and its organization into established disciplines. It, therefore, mirrors the discipline-oriented nature of the industry.
- The support for the multiple interpretations of concepts across disciplines: the construction disciplines have their own norms and values, reflected in the development of dedicated semantic resources. The proposed ontology architecture is modular, with sub-ontologies dedicated to each established discipline, federated by a core ontology defined at a higher level comprising generic concepts applicable across sectors and enabling reconciliation of different terminologies used across disciplines.

Also, it is worth noting that the Ontology has been developed with the objective of facilitating its comprehension and use by end-users. In that respect, the following is worth highlighting:

- The collaborative nature of the ontology development process: It supports parallel development of the various ontology modules while promoting their integration through the core ontology.
- The iterative nature of the ontology development process: given the large-scale dimension of the construction industry and its project-oriented nature, an iterative process that proceeds by refining and extending the ontology over time helps converge towards a complete and true conceptualization of the domain.
- The ontology development approach: this is semi-automated and relies on discipline-oriented documentary corpuses to identify concepts and relationships using tf-idf and metric clusters techniques [3], which are then validated by human experts.

7. CONCLUSION

The paper presented an ontology initiative to support the knowledge management needs of practitioners in the construction sector, including support for seamless collaboration. The ontology developed to date is far from being complete, and will probably never be, as an ontology should be viewed as a living system. Similar efforts, such as the ISO STEP project and its application to various industry sectors (including manufacturing), have taken almost a decade or longer to come to fruition. The eCognos ontology is by no means different. The issue of the existence of a unique ontology for an entire sector remains open. This suggests that while the

eCognos Core Ontology forms a robust basis for interoperability across the discipline-oriented ontologies, the latter will need adaptation and refining when deployed into an organization and used on projects. Another issue that was raised is that related to the adoption of user specific views or perspectives on the global ontology. In fact, in many instances, some actors might be required as part of their job to deal with more than one discipline ontology to conduct a task. This necessitates some flexible mechanisms that can enable the rapid combination of two or more discipline ontologies into a single view / perspective. This constitutes one of the author's future research objectives.

Also, TAM, the model used to assess the adoption and diffusion of the ontology – with its original emphasis on the design of system characteristics – does not account for social influence in the adoption and utilization of new information systems [13]. Furthermore, It has been reported that when TAM is applied to collaborative systems, it is often observed that the belief structures (perceived ease of-use and perceived usefulness) are not stable, and that the influence of these belief structures act in combination with the effect of social influences to determine the use of the system [15, 16, 17]. The author encourages researchers to explore further models to investigate the adoption and diffusion factors of any ontology project.

8. REFERENCES

- [1] Rezgui, Y. and Zarli, A. (2006) Paving the way to digital construction: a strategic roadmap, *Journal of Construction Engineering and Management* (Journal of the American Society of Civil Engineers), Volume 132, Issue 12, pp. 767-776.
- [2] ISO 10303-1:1994. Industrial automation systems and integration – product data representation and exchange – Part 1: Overview and fundamental principles. International Standards Organization. TC 184/SC 4, 1994. <http://www.iso.ch/cate/d20579.html>.
- [3] Rezgui, Y. (2007) Text Based Domain Ontology Building Using tf-idf and Metric Clusters techniques, *Knowledge Engineering Review* (Cambridge Press), Volume 22, Issue 4, pp. 379-403.
- [4] Lima C, Al Diraby T, Fies B, Zarli A, Ferneley E. The E-COGNOS Project: current status and future directions of an ontology-enabled IT solution infrastructure supporting knowledge management in construction. In: Molenaar Keith R, Chinowsky Paul S, editors. *Winds of change: integration and innovation of construction*. Construction research congress 2003, Honolulu, Hawaii, USA, 19– 21 March 2003, 2004.
- [5] Eastman CM, Sacks R, Lee G. Strategies for realizing the benefits of 3D integrated modeling of buildings for the AEC industry. In: *Proceedings of ISARC-19th international symposium on automation and robotics in construction*, Washington, DC, 2002. p. 9–14.
- [6] Boddy, S, Rezgui, Y., Cooper, G., Wetherill, M, (2007), *Computer Integrated Construction: A review and Proposals for Future Directions*, *Advances in Engineering Software* (Elsevier), Volume 38, Issue 10.
- [7] Corcho, O, Fernando-Lopez, M, Gomez-Perez, A, 2003. "Methodologies, tools and languages for building ontologies. Where is their meeting point?" *Data and Knowledge Engineering*, 46 pp 41-64.
- [8] Gruber, T., 1994. "Towards principles for the design of ontologies used for knowledge sharing" *International Journal of Human Computer Studies* 43(5/6) pp 907-928.
- [9] Studer, R., Benjamins, V., and Fensel, D, 1998. "Knowledge engineering: Principles and methods" *IEEE Transactions on Data and Knowledge Engineering* 25 pp.161-197.
- [10] Welty, C, and Guarino, N, 2001. "Supporting ontological analysis of taxonomic relationships" *Data and Knowledge Engineering* 39 (1) pp. 51–74.
- [11] Zmud, R.W., 2000. *Framing the Domains of IT Management: Projecting the Future Through the Past*. Pinnaflex, Cincinnati.
- [12] Rogers, E.M., 1995. *Diffusion of Innovations*, fourth ed. Free Press, New York.

- [13] Davis, F.D., 1993. User acceptance of information technology: system characteristics, user perceptions and behavioural impacts. *International Journal of Man-Machine Studies* 18 (3), 475–487.
- [14] Rezgui, Y. (2006) Ontology Driven Knowledge Management Using Information Retrieval Techniques, *Computing in Civil Engineering (Journal of the American Society of Civil Engineers)*, Volume 20, Issue 3, pp. 261-270.
- [15] Easley, R.F., Devaraj, S., Crant, J.M., 2003. Relating collaborative technology use to teamwork quality and performance: an empirical analysis. *Journal of Management of Information Systems* 19 (4), 247–268.
- [16] Rezgui, Y. Role-Based Service-Oriented Implementation of a Virtual Enterprise: A Case Study in the Construction Sector, *Computers in Industry (Elsevier)*, 58(1), pp 74-86, 2007.
- [17] Rezgui, Y. Exploring Virtual Team-Working Effectiveness in the Construction Sector, *Interacting with Computers (Elsevier)*, 19(2), pp 96-112 , 2007.