Journal of Information Technology Management

ISSN #1042-1319

A Publication of the Association of Management

STRATEGIES FOR INCORPORATING DATA EXCHANGE STANDARDS IN E-BUSINESS TAXONOMIES

TAMER E EL-DIRABY
UNIVERSITY OF TORONTO, CANADA
tamer@ecf.utoronto.ca

CELSON LIMA
CENTRE SCIENTIFIQUE ET TECHNIQUE DU BATIMENT, FRANCE
Celson.lima@cstb.fr

BRUNO FIES
CENTRE SCIENTIFIQUE ET TECHNIQUE DU BATIMENT, FRANCE
Bruno.Fies@cstb.fr

ABSTRACT

Product models are very suitable for data exchange between software systems. They, however, have less effectiveness when it comes to exchange of information in a semantic web environment. Users need to have the ability to use industry terminology and standard vocabulary. Ontologies have been used by industries to establish effective semantic web environment for e-business and the exchange of information. Mapping product data models and classification systems into ontologies is a meaningful decision. This will allow the use of industry standards and terminology. This paper compares different strategies that have been used for incorporating IFC into semantic web technologies. It also proposes an agent-based distributed architecture for dynamically managing ontology mapping in the construction industry.

Keywords: e-business, Taxonomy, Semantic web

INTRODUCTION

The development of e-business supply chains requires the establishment of shared representation of product names, specifications and supply chain transactions. Effective e-business supply chains transcend the limited view of e-procurement into the realm of virtual collaborative enterprises, where several organizations collaborate on the web on a process level. These organizations assume dynamic roles in different projects [2]. In one project, an organization could be the owner. In a second project, the same organization could be the contractor (manufacturer) and in a third project, it could be a subcontractor or supplier or even a consultant. Such business environment requires developing polymorphic organizations with extremely flexible process structures. On industry level, this requires building e-business taxonomies that cover, not only product data, but, more importantly, the essence of industry knowledge. It is important to develop these taxonomies in a flexible (distributed) manner to allow different organizations to wrap such taxonomies into their proprietary IT systems.

Product models and data exchange standards are shared representation of data structures that are used to facilitate the exchange of product data among several software systems within a domain (or industry). Such data exchange standards are focused on the technical and engineering specifications of products. They are not directly usable for "semantic" exchange of product/process "knowledge" in a web-based supply chain as they are mainly designed to exchange "product data" between "software". Moreover, they lack means to describe supply chain terms. Finally, most of these standards are written in EXPRESS or similar languages, which will normally require mappings to web-friendly languages, for example, XML.

Taxonomies created to support e-business should be closely tied to product models and existing classification systems. System users are very familiar with the structure of legacy classification systems. It would be a tremendous help if new taxonomies use the same names and hierarchy as existing classification systems to ease the transfer to semantic representations. Moreover, these standards link several software systems in order to integrate engineering design. Mapping data exchange standards to e-business taxonomies would assure integration of such "engineering systems" into the e-supply chain and the business aspect of the organization.

This paper presents a comparative analysis of the strategies used to incorporate existing "engineering" data exchange standards into "e-business taxonomies". The construction industry is used as a framework for this analysis. The paper defines the situations where different strategies are applicable and identifies, based on industry needs, the performance criteria that should be satisfied by each transformation strategy. The paper also proposes a framework for a distributed environment for integrating and mapping product data models into ontologies.

THE NEED FOR E-CONSTRUCTION SUPPLY CHAINS

The construction industry has one of the least levels of e-business penetration. The application of e-business concepts has, so far, been limited to the exchange of catalogue data about products and e-bidding. With the ongoing infrastructure privatization and the increasing demands for efficiency, it is expected that more advanced e-business techniques will flourish in the industry. Soon, construction projects will be provisioned through virtual enterprises where several entities will collaborate to develop and operate services for the public simultaneously. For example, a virtual enterprise could be formed between an environmental consulting agency, Department of Transportation and a highway design firm for evaluating the environmental impact of different highway design options. The Department of Transportation will issue a generic file including the scope and objectives of the project. Specific software at the consulting agency will parse this file and develop a set of environmental requirements. The design firm will develop and share design schemes with both entities. The Department of Transportation, through specific web services, will analyze all the options against the set criteria. Various public entities, businesses and the Department of Transportation will then form another dynamic enterprise to evaluate the impact of the highway on local communities. Following that a developer, a bank, the designer, a construction company and the Department of Transportation could form a virtual organization that builds the project. The Department of Transportation and the design firm have to put on too many hats in this scheme where they observe different sets of objectives and balance different players and interests in each case.

As the information culture evolves in the construction industry, specialized companies and brokers could provide "web services" and solutions over the Internet. More advanced than traditional Electronic Data Interchange, web services are poised to become the main approach for facilitating "e-infrastructure". This will require a deliberate level of cooperation and research into the creation of a vibrant information supply chain—which includes:

- Content creation: creation of original multimedia content such as data, photograph, audio and video. These products cover current and future development, productivity data, labour statistics, resource usage information, material pricing, technology briefs, industry best practice, lessons learned, etc.
- Content packaging: modeling, digitization and manipulation of original content into suitable format. i.e. the development of proper means to express construction information in a compatible way (software, components, web services, process modules) that can provide the end-user with a needed service (analysis of one design aspect, communication between different parties, etc.)
- Market creation: developing and managing existing and new markets for information services. i.e. identifying, influencing and satisfying users through awareness and developing end-user oriented software and information commodities.
- Interface and system: establishing a means for the exchange of products among producers and users and between different users through communication protocols, date ex-

change standards, and directories of services, security regulations, etc.

- Customer support: developing interaction systems to solicit and respond to customer needs and requirements both in the software format or the hardware.

The major impact of such a supply chain is the creation of an environment conducive of network-based project development. Virtual dynamic organizations will evolve in the infrastructure development domain providing for more smooth collaboration. This will also require the use of advanced semantic means to standardize industry representation of relevant knowledge through ontologies. Industry-wide taxonomies are being used to provide such needed semantic. For example the bcBuildingDefinitions taxonomy [10] represents a major step forward in the utilization of semantic web technology in the construction industry. The taxonomy included a substantial description of construction product with a simple mapping table to link the bcBuildingDefinitions to IFC.

IFC: INDUSTRY FOUNDATION CLASSES

IFC is the first successful implementation of data exchange standards in the construction industry. IFC is developed under the auspices of the International Alliance for Interoperability (IAI). The standard was created mainly to unify the data structures of CAD software. Later versions of the standards extended this shared platform to scheduling and cost estimating software. Current IAI initiatives are extending IFC into further engineering domains like structural analysis and pre-cast concrete in addition to several related managerial domains such as transactions and cost estimating.

Because of its CAD foundations and its focus on data exchange among software, IFC is fundamentally a description of the data structures/types of basic elements in the engineering design of buildings. For example, a great deal of the standard is dedicated to description of the graphical representation of basic engineering elements(line, curve, space) and their use to describe basic engineering products (column, beam, door window). IFC developed data structures for several theoretical and abstract concepts (containment, space). These concepts are mainly needed to streamline the flow of information between software. Most of these concepts do not capture core industry knowledge [9]. A good percentage of these theoretical concepts are rarely used by the industry during business transactions. In addition, IFC is entirely focused on modeling products, in particular physical products (such as doors, windows, floors, foundations). For example, most of the parameters used in the IFC representation of door are not commonly used by the industry (on a day-by-day basis). These data structure are basically a description of these products engineering details (e.g., measurement, location within a CAD drawing, and geographical details).

COLLABORATIVE SUPPLY CHAINS

Semantic web is the application of knowledge management technologies to support intelligent navigation of the Web. It relies on the use of industry-wide meta-data model (or ontologies) to aid in the classification and recognition of related concepts. One of the main benefits produced by the semantic web is the increased efficiency in the e-business supply chains. Traditionally, supply chains focused on tracking product lists and specifications or order information. The different mismatches between partners information representations has limited the full proliferation of supply chains in many industries. To resolve this, industries have developed product data models to standardize their product specifications. However, the true benefits of e-business supply chains are only realized if all partners collaborate on a process level. In one view, supply chains can be executed at three levels:

- Data exchange level: agreements on consistent product data representations. For example, specification of fire resistance in the Netherlands is by a code such as B30 whereas in the UK it is a time duration such as 30 minutes. A product data model will allow for consistent exchange of this and other specifications in seamless way.
- Information exchange: agreements on consistent order procedures. E-business schemas, such as ebXML, assure the exchange of complex data structures about various business transactions. This allows better tracking of inter-agency orders that goes beyond the specifications level into full information exchange level.
- Process collaboration: agreement between different organizations on integrating their business models and processes to share development knowledge. This level goes beyond the ordering systems into the collaborative development paradigm. This level requires full agreements on the semantics of industry concepts and comprehensive im-

plementation of knowledge management techniques. The corner stone of this level is an agreement on the semantics of concepts in a domain through taxonomies.

Taxonomies vs. Data exchange Standards in the Construction Industry

In the construction industry, IFC has recently gained popularity. Currently, it allows for the exchange of product data between CAD, scheduling and estimation software. Other related domains are currently being integrated into IFC.

Product data models are more suitable for describing engineering data. For example, changes in a concept like fire code specifications for a door lining system can best be introduced by the product data model administrator. These are not normally the domain of an ontology administrator, who is mainly concerned with knowledge representation. If these changes have been introduced into product data models (as normally is the case) first, then a link between an ontology and a product data model will automatically guarantee the inclusion of such new standards into the ontology.

The use of IFC as basis for e-business taxonomies has the following benefits:

- Extending existing integrated engineering platforms into e-supply chains. Such extension will allow the integration of supply chains data and its business aspects into engineering drawings.
- Reducing the overhead associated with any transfer in software systems to adapt to knowledge management concepts.
- Easier introduction of new product model data into ontologies.
- Reduce user confusion and any need for manual transformation.

Requirements for Transformation/Linking Strategies

Any transformation or mapping strategies between existing data structure and e-business supply chains has to achieve the following objectives [6]:

- Coverage: how many concepts are mapped between the two standards?
- Scalability: means to assure that any change in the data structure can be incorporated later into the e-business taxonomy.
- Ease of mapping: minimum time and effort required for mapping the two standards.
- Accessibility/Interoperability: efficient access to the data of one standard from the other on dynamic basis.
- Consistency: Semantic structures should adhere to a given ontology in order to allow for better sharing of knowledge. For example, it should be avoided that people confuse complex instances with attribute types.
- Redundancy: Decentralized knowledge provisioning should be possible. However, when annotators collaborate, it should be possible for them to identify (parts of) sources that have already been annotated and to reuse previously captured knowledge in order to avoid laborious redundant annotations.
- Maintenance: Knowledge markup needs to be maintained. An annotation tool should support the maintenance task.
- Efficiency: The effort for the production of metadata is a large restraining threshold. The more efficiently a tool supports the annotation, the more metadata will produce a user. These requirement stands in relationship with the ease of use. It depends also on the automation of the annotation process, e.g. on the pre-processing of the document.

STATIC MAPPING STRATEGY

Among other things, the e-Construct project [10] developed a web-based system to publish, manage and search manufacturer catalogues. The project defined a taxonomy of major construction physical products (doors, windows, walls, etc.). To incorporate IFC, the project mapped available IFC physical products to their counterparts in the e-Construct taxonomy. By that, anyone who is surfing the e-Construct web site for a door will be able also to search using the word ifcDoor. If a user is searching for say doors, windows, and lamps, the search results, except for the lamps, will be readable by any IFC parser — which will never recognize the concept of lamp as it is not defined in IFC. Please note that e-Construct allows only simple item search (for example, one cannot search for a complete kitchen).

This approach could be explained as using a static mapping table where one-to-one correspondence is established between two concepts in two taxonomies (see Figure 1). In short, the e-Construct approach mapped IFC concepts without any mapping of IFC relationships. Given the fact that e-Construct taxonomy did not include a wide range of relationships, this approach, even though limited, is quite effective in achieving their goals of providing an IFC match to their concepts.

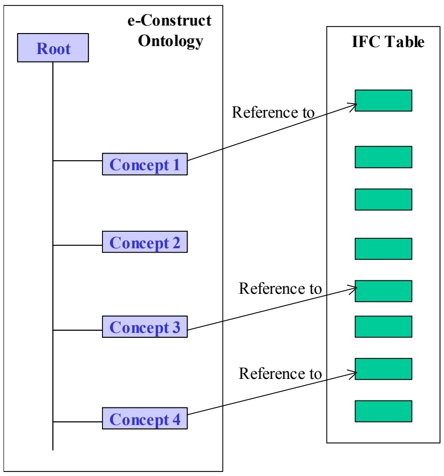


Figure 1: Static Mapping Table Approach

PARALLEL INCLUSION/APPENDING STRATEGY

Growing the e-COGNOS taxonomy through incorporating other taxonomies, was a major objective for the e-COGNOS project. This strategy, basically appends a taxonomical tree to an already existing e-COGNOS concept. This is very beneficial in enriching an upper level taxonomy with concept trees. The strategy include assigning an ID to each concept in the base taxonomy. An agent then appends the concept tree to a concept ID provided by the system administrator. Later in this paper we present an agent that uses a simple Excel sheet to append a concept tree to an existing taxonomy.

INTERPRETATION STRATEGY

ISTforCE project [9] developed a mapping approach that uses an interpreter to develop ontological concepts from IFC concepts. Its knowledge base builds upon the IFC project model and the definition of an Engineering Ontology that enables the translation of 'raw' IFC data to the vocabulary and semantics familiar to the end user. Thus, the end user can retrieve product model information fast and without deeper knowledge of the underlying "technical" data structures. The technique is based on Enterprise Java Beans Technology. A basic engineering ontology layer was created to specify major concepts related to engineering systems, primarily structural design. For example, "Frame", which is a major part of structural engineering vocabulary but is not contained in IFC. An ontology interpreter using specified rules, transfers IFC concepts like "ifcBeam" and "ifcColumn" into the ontological concept "frame". However, this is a very complicated process and will require constant updates to the rules governing the interpreter.

SUPERPOSITION STRATEGY

The e-COGNOS ontology was designed to include terms from IFC, BS6100 and UniClass. All three systems (IFC, BS6100, UniClass) have means to describe the six major domains that have been identified for this ontology. In contrast to BS6100 and Uniclass, IFC is more consistent and closer to an object-oriented approach in developing its terms. It could be immediately seen that IFC has six major categories of concepts: **Project**, **Actor**, **Product**, **Process**, **Resource** and **Relationships**. Relationships in IFC are not semantic. The other five major concepts are fundamental to describing the major elements of construction knowledge. To an extent, these elements tend to describe the physical/obvious elements in construction. The e-COGNOS project added another major concept: **Technical Topics** to cover the softer/logical/scientific issues related to construction. Within the skeleton of these six major domains, IFC, BS6100, UniClass and Talo90 terms will be dropped to enrich the ontology.

In other words, the e-COGNOS approach in incorporating IFC resembles a superposition of ontologies/taxonomies.

1. A quasi IFC ontology (called IFC.daml) was created using DAML (see Figure 2). This taxonomy preserves the relationships between major IFC concepts (an ifcObject is still a child of ifcRoot, an ifcProduct is still a child of ifcObject, etc.).
2. The e-COGNOS ontology was created to reflect, in part, the IFC structure. Not only are the five major domains (**Project**, **Actor, Product, Process, Resource**) common between the two ontologies, but also minor IFC concepts are included (like ifcMaterialLayer and ifcSpace, for example) through the e-COGNOS ontology.
3. The DAML relation "is similar to" was used to match e-COGNOS concepts to IFC.daml

concepts (Root "is similar to" ifcRoot, Actor "is similar to" ifcActor, etc.).

Figure 3 explains the superposition concept. Now if some one searches the e-COGNOS ontology for "kitchen" it will be found under the concept of "Building space" and at the same time ifcSpace. When a user searches for the concept of "engineer" in e-COGNOS, it will be found under the concept of Actor, which is equivalent to ifcActor. Even though kitchen and engineer are not part of IFC, an IFC parser will be able to recognize them in the form of their upper level concepts (in this case, ifcSpace and ifcActor).

The difference between this strategy and the static mapping strategy is that the superposition allows the incorporation of IFC physical product concepts (ifcDoor, ifcWindow, ifcColumn) in addition to IFC logical concepts (ifcSpace, ifcMaterialLayer, etc.) mainly because IFC.daml has captured both IFC entities (be it physical or logical) and their relationships.

The result of this superposition approach is that it will allow almost every item in e-COGNOS ontology to be mapped to an item in the IFC that is either the exact match for it (door matched to ifcDoor) or the best available approximation in lieu of IFC's limited taxonomy (kitchen matched to ifcSpace). More importantly, it will allow any future expansion in IFC (and a lot of them are currently underway) to be dynamically mapped to e-COGNOS ontology through updating IFC.daml and establishing proper "is similar to" relationships to the new added items in IFC and their counterparts in the e-COGNOS ontology.

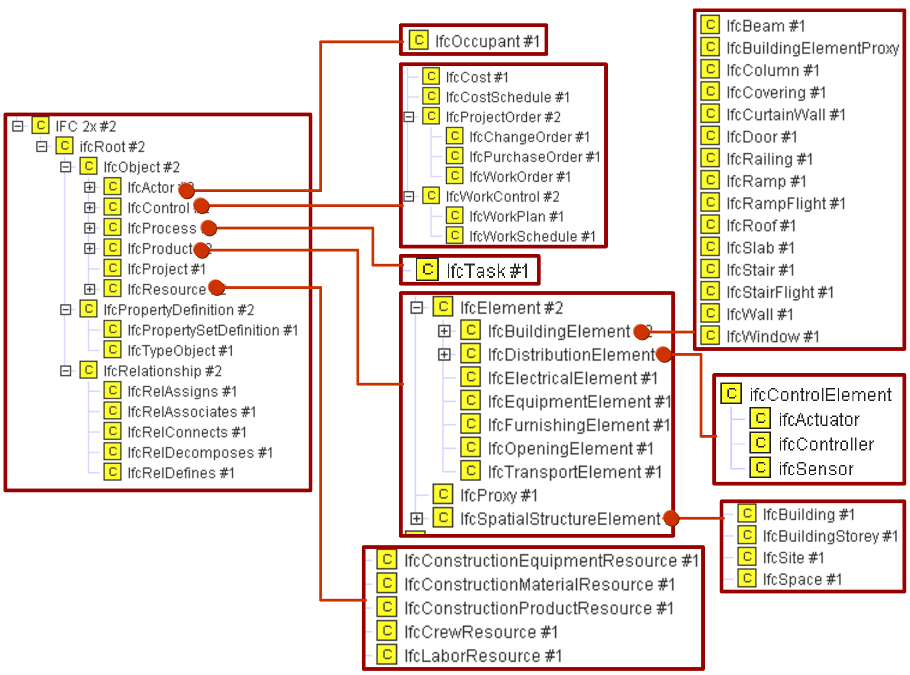


Figure 2: IFC.daml

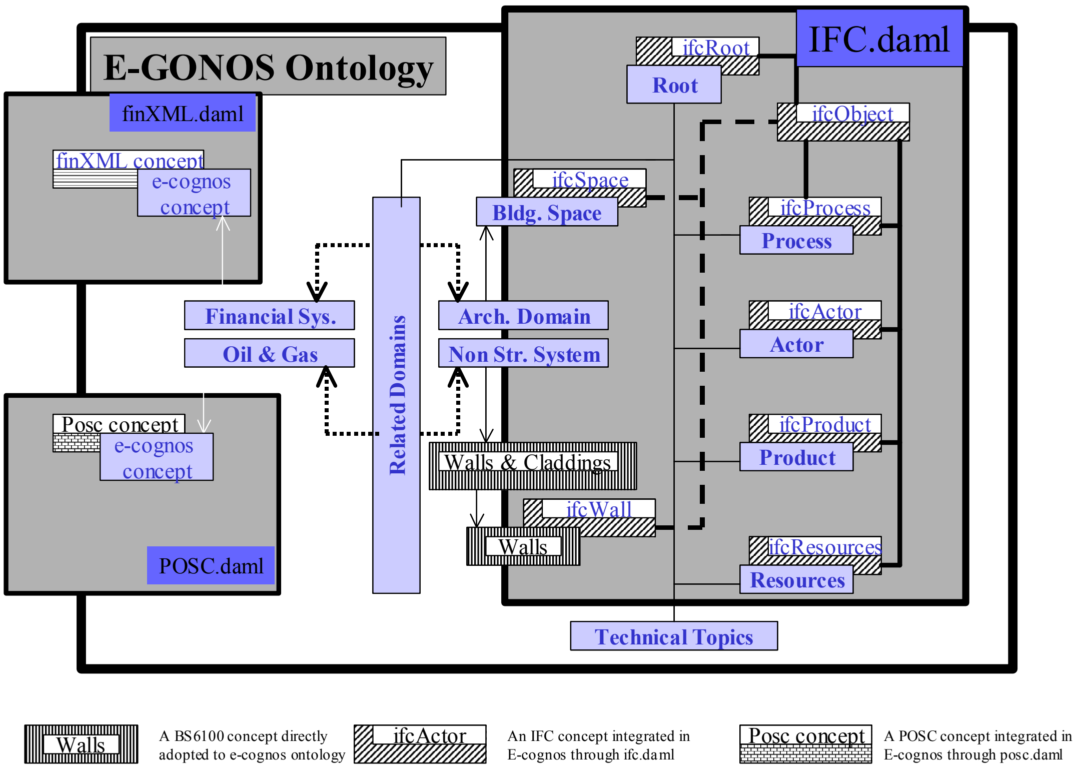


Figure 3: Dynamic Superposition Approach

TYPE-CASTING STRATEGY

In many cases, well established industry classification systems can provide valuable help in developing taxonomy terms. It is advantageous to use these standards in e-business taxonomies as this will introduce already familiar terms to taxonomy users. However, most of these classification systems lack proper object orientation. In this case, researchers have simply subsumed these classification systems into their taxonomies where they fit into the ontological model. For example, the e-COGNOS taxonomy used most of the terms in BS6100 (over 10,000), UniClass and Talo90, which are very well-established national classification systems, by dropping their terms into proper space. For example, Figure 3 shows the inclusion of Walls and Claddings (from BS6100) into the e-COGNOS taxonomy by declaring such term as a product. This is more of type casting already existing systems (that lacks taxonomical hierarchy) into new taxonomical categories. When a user searches e-COGNOS taxonomy for Walls or Claddings, he/she will find them (with the exact same definition and sub elements) but under the product domain.

ANALYSIS OF STRATEGIES

Several industries have faced the issue of cross-standards standardization before. The research team, as part of its work on the e-COGNOS ontology, analyzed the previous strategies. The team also reviewed relevant literature in the field. It was obvious, given the multitude of programming and implementation needs, that no one single approach will be optimum under all conditions. Table 1 shows a simple analysis of the performance of each approach respective to the previously identified criteria. These ratings were developed based on the analysis of the team, reviews of the literature and input from end users. It should be noted that Table 1 was developed for the construction industry. Other industries could have different criteria and/or assessment for the performance of various strategies. Logically speaking, a distributed system that allows combinations of these systems could (eventually) be the optimum solution.

Table 1: Summary of Approach Analysis

Criteria	Mapping Table	Type Casting	Parallel Inclusion	Interpreter	Superposition
Ease of use	Very easy	Difficult	Very easy	Very difficult	Difficult
Efficiency	Low	Moderate	Low	High	Moderate
Convergence	Very low	Low	Adequate	Low	High
Scalability	Very limited	Limited	Unlimited	Very limited	Adequate
Redundancy	Low	Moderate	Very high	Low	Low
Consistency	High	Low	Very low	High	Very High
Navigation accuracy	Low	Moderate	Very low	Moderate	Moderate
Update speed	Very low	Low	Very high	Very Low	Moderate
Execution time	Low	High	Very low	High	Moderate

A PROPOSAL FOR DISTRIBUTED SYSTEM FOR INTEGRATING ONTOLOGIES AND PRODUCT MODELS

A key challenge in building the Semantic Web is streamlining the plethora of ontologies, taxonomies and semantic models that are being developed in almost each domain of knowledge and across these domains (mapping the concepts of one to another). "There is no obvious "oracle" that will make these judgments. For instance, we cannot assume that there is an overarching (possibly "global") ontology that serves as a court of appeals for semantic judgments. Hence the work of ontology reconciliation inevitably involves a human being to do the heavy lifting. The most we can hope for is to provide a formal definition of the problem, and software tools to aid in solving it. The goal of these tools is to develop and maintain ontology transformations. An ontology transformation is a mechanism for translating a set of facts expressed in one ontology (O1) into a set of expressions in another ontology (O2), such that the new set "says the same thing" as the original set [4]."

Several organizations have developed technologies to map, fuse and manage ontology libraries in the same way OMG initiated the IIOP to connect various ORB standards. For example [3]: WebOnto (Knowledge Media Institute, Open University, UK), Ontolingua (Knowledge Systems Laboratory, Stanford University, USA), DAML Ontology library system (DAML, DAPAR, USA), SHOE (University of Maryland, USA), Ontology Server (Vrije Universiteit, Brussels, Belgium), IEEE Standard Upper Ontology (IEEE), OntoServer (AIFB, University of Karlshruhe, Germany) and ONIONS (Biomedical Technologies Institute (ITBM) of the Italian National Research Council (CNR), Italy).

Still, full standardization of data and web services protocols of various industries involved in the infrastructure development may be not achieved. While it is easy to enable wider collaboration through data and web services standardization, static standards are inefficient and will be short lived. It is easy to alienate other participating industries (banking, insurance, Auto, etc.) due to the lack of standardization harmony. This leads to the need to develop dynamic standardization schemes that adapt to change and support inter-standards communications. Essentially, we need to create modular self-describing standardization schemes where the component or web service holds within itself reference to its standards.

Agent technology represents a very suitable technique to handling the mismatches and mappings of ontologies. In a distributed, humanly-supervised environment, several agents could be used to generate sub taxonomies that eventually could be augmented together using the parallel inclusion strategy or more advanced strategies (such as the GLUE system).

Figure 4 shows a proposal for an agent-based distributed system for managing the mapping and integration of ontologies and/or product models. The system is compromised of four main layers. The first layer is dedicated for mapping agents. Each agent can utilize one or more artificial intelligence algorithm to map between ontologies and/or product models. For example, one agent could utilize the GLUE system [4], which applies machine learning techniques, to semi-automatically create mappings between taxonomies based on identifying the most similar concepts in both. Moreover, another agent could use natural language generation systems, such as the IDAS

[13], for generating classifications from text. Semi-structured documents, like dictionaries, could be transformed into rational structures [7]. HTML documents could also be easily transferred to structured text. On the macro level, a system such as Hozo [14] could be used to link and modify several ontologies.

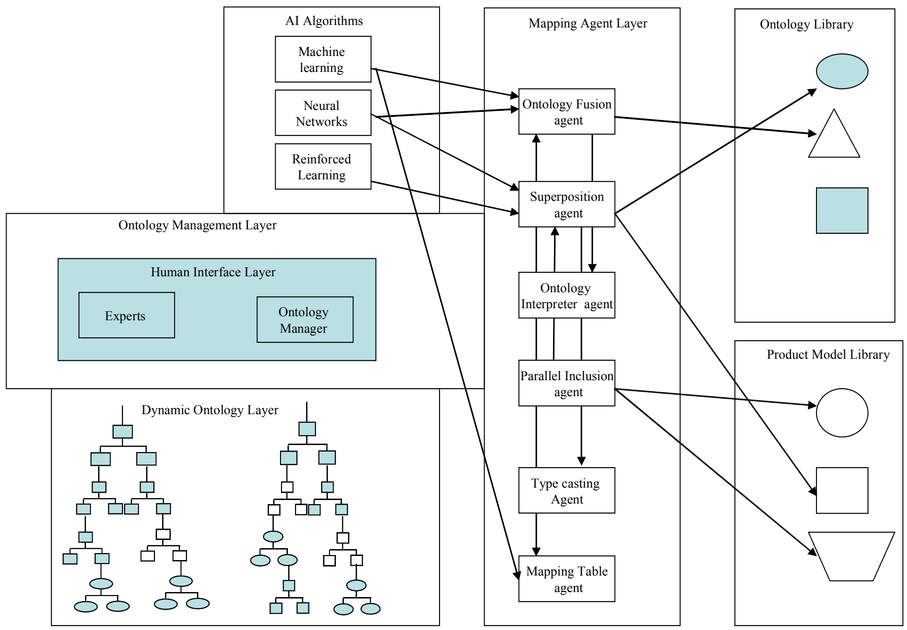


Figure 4: Agent-Based Distributed Architecture for Ontology Management

This will be very helpful for managing large taxonomies as an ontology manager could device sub taxonomies for different departments within one organization. Additionally, several companies could collaborate online using their own taxonomies without the need to change any of their structures as an ontology manager could create a dynamic taxonomy for the specific purposes of their collaboration and then remove this taxonomy from the library when the particular project is finished.

Prototypical Framework

The research team developed a prototypical system to use the gent technology to manage dynamic development of ontologies. The main architecture of the proposed system includes:

1. Establish a domain ontology for an industry: a common ontology that is will be widely used by a sizable number of organizations within an industry. In the construction industry case, the e-COGNOS ontology was used as the domain ontology.
2. Establish enterprise domain ontology: no one ontology will be perfect for every enterprise use. Therefore, each organization should add additional entities and concepts to the e-COGNOS in a manner that suites its specific needs, while preserving the main elements of e-COGNOS ontology. For example, a consulting enterprise could add "services" to its ontology (see Figure 4)
3. Investigate suitable application ontologies: ontology manager should mine the web for relevant domain ontologies.
4. Populate Enterprise domain ontology: apply suitable strategies for incorporating relevant application ontologies into the enterprise domain ontology. The first step would normally be to use the parallel inclusion strategy to populate the enterprise domain ontology by including relevant taxonomies. In this regard, several agents could be used to handle the addition of concepts into the ontology.

5. Establish cross-tree relations: upon incorporating concept taxonomies, another agent could be used to establish relations between concepts on different trees.

The following agents have been created in this research project:

The Parallel Inclusion Agent (PIA):

This agent executes the parallel inclusion strategy. To automate the identification of concepts, this agent mandates the development of a hierarchical concept ID. Each concept should have a concept code. The structure of the concept code indicates the inheritance mechanism. For example, if a “highway product” has an ID of 3, then to specify that a “road” is the first child of this concept, “road” is assigned an ID of 3.1. To indicate that “rural road” is a first child of “road”, “rural road” is assigned an ID of 3.1.1. To facilitate the development of these ID’s, an excel template (Figure 5) is used to get the structure of the taxonomy to be imported. This template is converted into an XML file. A taxonomy generator agent (TGA), then, transfers this table into a taxonomy. The generated taxonomy is appended to the proper place in the domain taxonomy through matching the root concept of the generated taxonomy to an already existing concept in the domain taxonomy. The excel template is structured as follows:

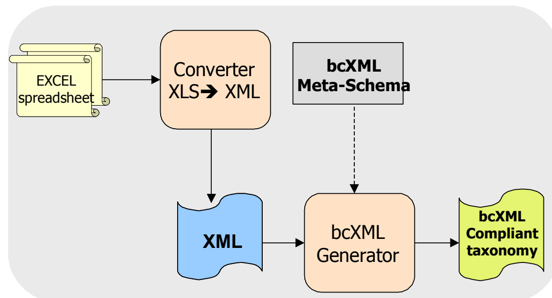
- Provider Information: details about the taxonomy owner;
- Taxonomy General Information: taxonomy-related parameters, namely Native Language, Alternative Languages, Number of Concepts, and Number of Levels;
- For each concept, the following information is required: Concept Code, Level, Concept Name, and Description. Level and Concept Name are mandatory; Concept Code and Concept Description are optional. If the concept is represented in several languages, for each language a par "Concept Name / Concept Description" is required.

Manual Cross Taxonomy Relationship Agent (MRA)

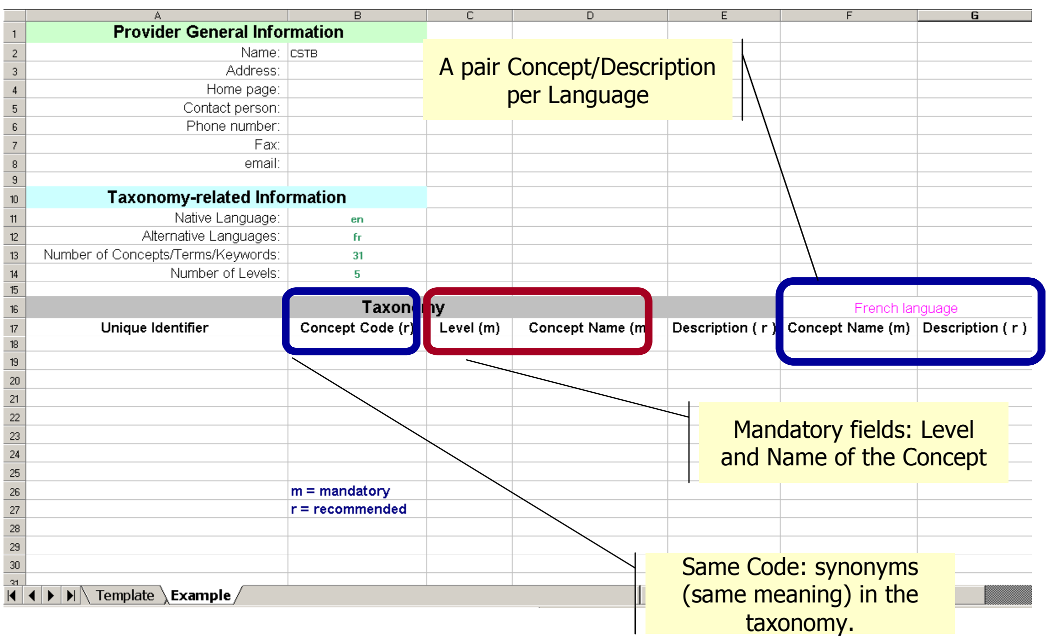
This agent is basically a visual screen that allows a user or an ontology manager to link two concepts.

Automated Cross Taxonomy Relationship Agent (ARA)

This agent uses data mining techniques to find/suggest relationships between closely related/similar concepts.



a: Matching Strategy



b: Excel Template

Figure 5: Parallel Inclusion Agent

Example Implementation;

This example assumes an enterprise interested in privatized highway construction. In this example, we assumed that an enterprise has adopted e-COGNOS ontology as its enterprise domain ontology (without changes). This enterprise is, therefore, interested in a highway taxonomy and a privatized infrastructure taxonomy. The parallel inclusion agent was used to blend concepts from both taxonomies into e-COGNOS taxonomy (see Figure 6). The agent allowed easy and efficient addition of about 2500 concepts to the e-COGNOS taxonomy. For example, financial products (such as bond, loan, fund) were added to e-COGNOS products along with highway products (such as bridge, road, tunnel).

The manual relationship agent was used to establish relationships between concepts of the two taxonomies. For example, the relationship "has" was created between a highway product (on the first taxonomy) and risk (on the second taxonomy).

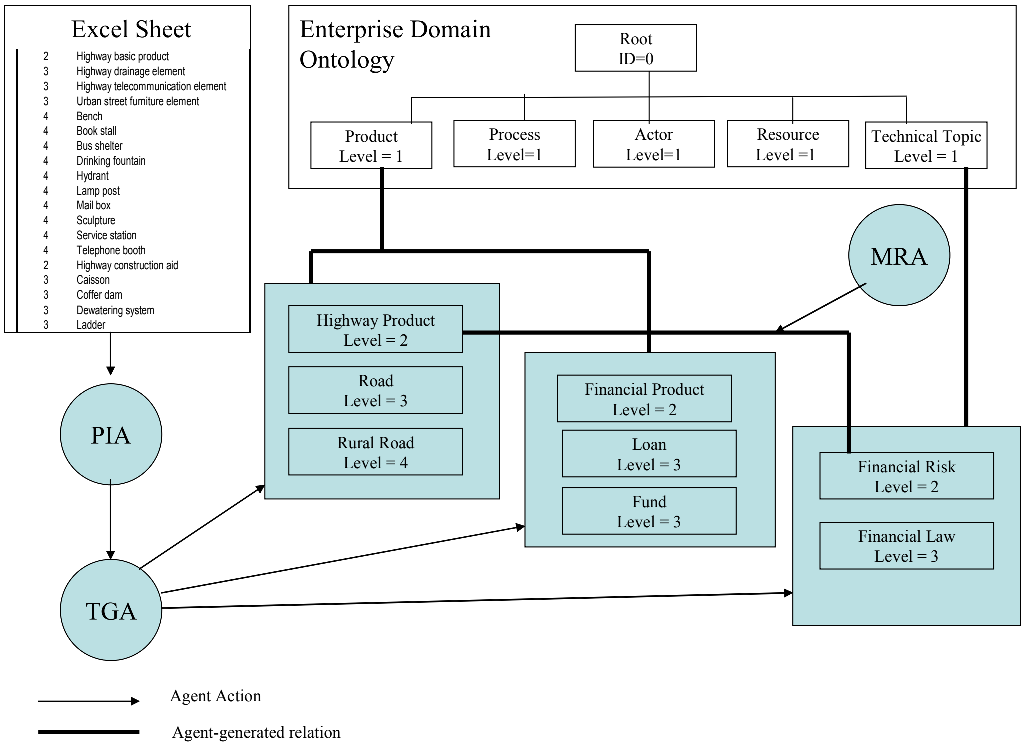


Figure 6: Prototype Implementation

SUMMARY AND CONCLUSIONS

This paper presented a comparative analysis for three strategies to incorporate existing data structures and classification systems into semantic e-supply chain. The strategies included: static mapping, dynamic mapping and ontology fusion. The strategies were analyzed against a set of criteria to test its applicability. Each strategy has been shown to be applicable to one or more situation depending on industry situation, intended uses and extent of use.

REFERENCES

[1] Boer, A., Hoekstra, R., and Winkels, R. "The CLIME Ontology," Second International Workshop on Legal Ontologies (LEGONT), Amsterdam, Netherlands, 2001.

[2] Chapman, L. D., Lathon, R., and Petersen, M. "Demand Activated Manufacturing Architecture: DAMA Model For Collaboration, DAMA-I-1-01 Version 1.0." Sandia National Laboratories, 2000.

[3] Ding, Y. and Fensel, D. "Ontology library systems: the key to successful ontology re-use," *Proceedings of 1st International Semantic Web Working Symposium (SWWS'01)*, 2001.

[4] Doan, A., Madhavan, J., Domingos, P. and Halevy, A. "Learning to Map between Ontologies on the Semantic Web," *Proceedings of the eleventh international conference on World Wide Web*, Honolulu, Hawaii, USA, 2002.

[5] Froese, T. "Future directions for IFC-based interoperability," *Journal of IT in Construction*, Volume 8, 2003.

[6] Handschuh, S., Staab, S., and Maedche, A. "CREAM—Creating relational metadata with a component-based, ontology-driven annotation framework," *ACM K-CAP*, Vancouver, Canada, 2001.

[7] Hearst, M. "Automatic acquisition of hyponyms from large text corpora," *Proceedings of the 14th International Conference on Computational Linguistics*, Nantes, France, 1992.

[8] Kazakov, M., Abdulrab, H., and Babkin, E. "Intelligent integration of distributed components: ontology fusion approach," *Conference on Computational Intelligence for Modelling*, CIMCA, Vienna,

Austria, 2003.

[9] Katranuschkov, P., Gehre, A., and Scherer, R. J. "An engineering ontology framework as advanced user gateway to IFC model data," *Proceedings of the 4th European Conf. on Product and Process Modeling in the Building and Related Industries—eWork and eBusiness in Architecture, Engineering and Construction*. Portoroz, Slovenia, 2002

[10] Lima, C., Stephens, J., and Bohms, M. "The bcXML: Supporting eCommerce and Knowledge Management in the construction industry," *Journal of IT in Construction*, Volume 8, 2003

[11] Maedche, A. and Staab, S."Ontology Learning." Research Center for Information Technologies, 2000.

[12] McDermott, D., Burstein, M. and Smith, D. "Overcoming Ontology Mismatches in Transactions with Self-Describing Service Agents," *International Semantic Web Workshop*, Stanford, CA, 2001.

[13] Reiter, E., and Mellish, C. "Using Classification to Generate Text," *Proceedings of 30th Meeting of the Association of Computational Linguistics (ACL-1992)*, MIT Press, 1992.

[14] Sunagawa, E., Kozaki, K., Kitamura, Y., and Mizoguchi, R. "An Environment for Distributed Ontology Development Based on Dependency Management," *2nd International Semantic Web Conference*, Sanibel Island, FL, 2003.

AUTHORS BIOGRAPHY

Tamer E. El-Diraby is an Assistant Professor and Director of the Centre for Information Systems in Infrastructure & Construction at the University of Toronto, Canada. He holds a PhD in Civil Engineering from the University of Texas at Austin, USA. He is the author of 18 journal articles. His work focuses on the use of information technology and knowledge management tools to optimize the design and construction of civil infrastructure systems.

Celson Lima is a research scientist at the Centre Scientifique et Technique du Bâtiment, France. He holds a PhD in computer science. His work focuses on the use of information technology in construction, including e-business and knowledge management.

Bruno Fies is a research scientist at the Centre Scientifique et Technique du Bâtiment, France. He holds a PhD in computer science. His work focuses on the use of information technology in construction, including virtual organizations and corporate memory.