PROCESS-ORIENTED SYSTEMS INTEGRATION

Yimin Zhu¹, Raja R.A., Raja R.A. Issa² and Ivan Mutis³,

ABSTRACT

Existing mainstream integration solutions, which are based on sharing common semantic models, are mostly static. The semantic models are designed based on a priori knowledge to the heterogeneity of collaborating systems. However, the problems that those integration solutions intend to resolve are not static, which requires that an integration solution being adaptive to constantly changing information processes.

The adaptiveness of collaborating systems also puts significant constraints on the integration strategy because it is the integration strategy that coordinates and controls the information flow among collaborating systems and the accessibility on which the shared information is based. First, the changes of the composition of supporting information systems throughout the lifecycle of a project such as a construction project will cause the technical infrastructure that supports computer-mediated communication and collaboration to change, which further requires the integration strategy to be able to adjust to different integration situations dynamically. The integration strategy needs to be able to deal with the newly added tools as well as to maintain access to the data in the old format. Secondly, formal business processes such as change of order processes and procurement processes often have significant variations from case to case. Let alone to say many informal business or information processes that team with a typical construction project. The variations also have great impact on the integration solutions because for different cases the information needs are different. Therefore, the changes in technical infrastructure as well as business processes require that the information system of a project be adaptive.

The writing of this paper is motivated by recent studies in computer-mediated information processes and integration strategies. Recently research in areas such as process mining and dynamic enactment of workflow processes makes computer-mediated information processes more adaptive to human-oriented business processes. Meanwhile, the hybrid strategy for systems integration in AEC has been introduced and discussed. The hybrid strategy relies on a community-specific representation that serves as a *de facto* standard for collaborating systems. The community-specific representation is an ontological representation that addresses the access to local data definitions, semantic mappings between local definitions, and other related issues.

To support integration solutions that address the dynamics of a project and evolve as project environment changes, this paper explores theories and methods to establish a link

¹ Professor, Dept. of Construction Management, Florida International University, Miami, FL 33174, U.S.A, Phone (305) 348-3517, zhuy@fiu.edu

Rinker Professor, Rinker School of Building Construction, University of Florida, Gainesville, FL, 32611; Phone (352) 273 1152; raymond-issa@ufl.edu

³ Ph.D. student, Rinker School of Building Construction, University of Florida, Gainesville, FL 32611; Phone (352) 273 1178; imutis@ufl.edu

between processes and the community-specific representation so that the community-specific representation evolves as processes change.

KEY WORDS

Integration, process-oriented, hybrid integration, community-specific conceptual representation

INTRODUCTION

With the advent of the digital information age, the number of information sources and the volumes of data collected and warehoused in enterprises are growing at a phenomenal rate. The data in these information sources are usually heterogeneous and have been used in business management, government administration, scientific and engineering data management, construction management, and many other applications. In business, such data captures information used to identify sales opportunities and quality/cost control, and to improve corporate profitability. In building construction management, such data helps to enhance the ability to share and process information so that the capital project or capital facility operations can timely access accurate and complete construction information.

The need to improve this situation imposes major challenges on the supporting information systems to be 1) *ready* to integrate with heterogeneous data sources; 2) *adaptive* to the dynamic interactions among project participants as well as business process changes; and 3) *sensitive* to the dynamic needs of information ownership, as well as knowledge protection and sharing.

The *readiness* of systems for integration is reflected in two aspects: the feasibility and the practicality of the integration strategy used by the collaborating systems, and the responsiveness of the integrated systems in providing answers to dynamic, even ad-hoc queries. Past experience of using the shared semantics and mapping approaches has raised questions regarding their feasibility and practicality. Existing mainstream integration approaches based on shared semantics have significant limitations as those approaches assume a priori knowledge of heterogeneity during system design time, which many studies in computer science have pointed out to not be practical (Goh et al 1999; Rahm and Bernstein 2001; Mitchell et al 2002; Firat et al 2002; Do et al 2002; Aberber et al 2003; Doan et al 2003). For example, past experience in construction shows that it is difficult for relative static standards to keep up with the complexity and the dynamics of business applications (Zamanian and Pittman 1999; Amor and Faraj 2001; Turk 2001), especially when significant involvement of human beings is present, as is the case in the construction management process. On the other hand, studies have also pointed out that mapping between heterogeneous data sources, which seems practical, is not a simple task, e.g., due to the existence of semantic heterogeneity and the need for bi-directional mappings (Amor and Faraj 2001). Most of those approaches are semi-automatic, which from time to time requires input from domain experts (Do et al 2002; Aberber et al 2003; Doan et al 2003). Unfortunately, there is still a lack of mechanisms to assist domain experts in managing and controlling the semi-automatic process.

In addition, the data and process dynamics of project management often puts project participants in situations requiring them to retrieve information in a dynamic or ad-hoc fashion from various data sources. Both integration strategies discussed above are not responsive enough to such dynamic information needs. Thus, a new integration strategy needs to be developed.

The adaptiveness of collaborating systems also puts significant constraints on the integration strategy because it is the integration strategy that coordinates and controls the information flow among collaborating systems and the accessibility on which the shared information is based. First, the changes in the composition of supporting information systems throughout the lifecycle of a project (a construction project) will cause the technical infrastructure that supports computer-mediated communication and collaboration to change. These changes will further require the integration strategy to be able to dynamically adjust to different integration situations. The integration strategy needs to be able to deal with the newly added tools as well as to maintain access to the data in the old format. Secondly, formal business processes such as change order processes and procurement processes often have significant variations from case to case. In addition, formal business processes are often coupled with informal business or information process such as brown-bag meetings at construction jobsites. The informal processes are often ad-hoc and may change the information requirements of the related formal processes. The dynamics of formal and information processes thus has a great impact on the integration solutions because the information needs for different cases are different. Therefore, the changes in technical infrastructure as well as business processes require that the information system of a project be adaptive.

The sensitivity of the information access is about secure access control. In project management protecting knowledge is as important as sharing it. For example, the knowledge and information that construction companies possess can be roughly classified as public information, project information and company information. Public and project information are shared and accessible by the general public or by authorized persons. Although company information is generally "private", from time to time a construction company will need to share some of it private knowledge or information with the rest of the participants in a construction project, i.e., the owner, the subcontractors or the architects/engineers. Currently, this is done mostly in form of paper-based memos and faxes or phone conversations. Is this the ultimate boundary between the computer world and the human world? Or can information systems be intelligent enough to, for example, detect the needs for information sharing and protection, and to allow users to "flag private" information and knowledge stored in different systems to be accessible to partners based on certain criteria such as roles, tasks, time periods, as well as the information/knowledge, and the owner's confidence and willingness to share? Such sensitivity to information sharing and protection will inevitably tend to enhance the user's confidence in using computer systems for collaborative work.

PROPOSED FRAMEWORK

The primary objective of the proposed research is to streamline project management in a web-based business process through enhanced interoperability among project activities with different representations of resources and services. The problem will be addressed in the context of the building construction application domain, where much standardization has

been established but which still lacks a uniform and systematic way for supporting efficient collaboration among project participants using different standards.

The key to enabling interoperability among various developed standards is a *mapping* process in relating conceptual resource representations or *ontologies*. The relationship between two ontologies can be characterized in their *similarity* (how closely they resemble each other in their representations) and *affinity* (to what degree they are coupled in the context of usage). Both characteristics require syntactic and semantic information from the resources. Statically, this information can be derived from the structure of the representation, vocabulary in the resource description, type and constraints of the resource usage, and other meta-data. Dynamically, the system can derive additional semantic information from the context and usage patterns as observed in the workflow process, or from an expert system or through user intervention. The proposed onto-semantic framework facilitates both static and dynamic semantic mapping.

Figures 1 and 2 show the overall process-centered framework that integrates business processes and information processes with an engine to drive progressive integration of heterogeneous data resources. The framework uses existing Browsers that are the interfaces to the Ontology Editor. The Ontology Editor generates a new ontology and allows the clients to browse through it. The Ontology Repository contains higher-level ontologies, which are specified by Ontology Web Language OWL (http://www.w3c.org/), and also provides functions to retrieve, to update and to validate them. The Ontology Service and Context Service feed semantic information to the Mapping Module, which generates mapped meta data for interoperability in the Workflow Management System. The Ontology Server uses the specifications, provided by the Ontology Repository, to allow the search for specific mappings, requested by the Mapping Module, in order to map two ontology data representations. As was mentioned previously, these requested searches are based on specific classification systems or semantic interpretations included in the ontology. For example, as shown in Figure 2, the Workflow Management System supplies two data representations: Ontology 1 can be the representation of a resource specified in Master-Format (a US building construction standard that describes what components are) while Ontology 2 might represent an equivalent resource in UniFormat (a counterpart ASTM standard that in addition describes where the components are located). However, due to potential incomplete or contradictory information of the ontology data representations and/or the general layer ontology, the mappings might not be successful.

Therefore, as also shown in Figure 1, an additional *Knowledge Base Repository* is necessary to assist the process. The implementation of the *Context Service* will be based on the concept of the *Markov Model Mediator* (MMM). It also interacts with the *Ontology Service* through the *Mapping Module*. The *Mapping Module* in Figure 2 is the central component of the system. It contains a *Semi-Automatic Mapping Engine* that coordinates the requests and responses with the ontology and context servers. It uses a *Wrapper* that serves as a translator or provides mapped meta data for resources used by the *Work Flow Management System*. The *Mapping Engine* also allows eventual human intervention when the *Ontology Server* and *Context Server* do not provide adequate mapping results. The *Engine* generates ontology representation instances or mapped data, expressed in XML format, as a final result of the process. This result is intended for the specific use of the

internal application of the *Workflow Management System* and it will enable readable human interpretation of the XML format that can be expressed in other formats such as HTML.

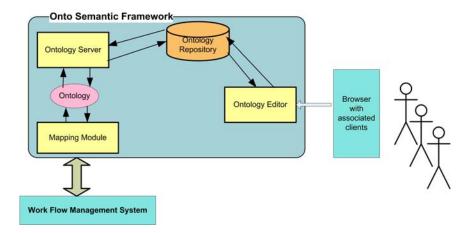


Figure 1: Onto-Semantic Framework.

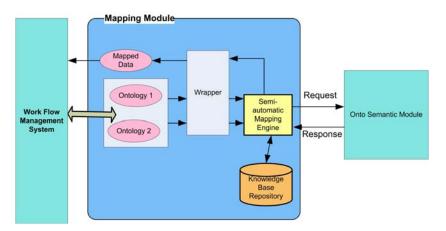


Figure 2: Semantic Mapping Module.

The shared semantic model and general ontology mapping are infeasible to implement. The uniqueness of our research is to make the solution tractable and practical through the use of *a semi-automatic ontosemantic mapping scheme* and *a community-specific conceptual representation* of heterogeneous resources in conjunction with workflow activities to extract dynamic semantic information for ontological integration.

RESEARCH ASSUMPTIONS

This research hypothesizes that effective data, information, and knowledge integration that addresses the dynamics of human-centered project management activities can be achieved through the fusion of *community-specific conceptual representation*, *contextual semantics* and *adaptive workflow models*, which are the three most important concepts of the proposed

scope with respect to the system framework. In addition, from the data point of view, the adaptiveness of workflow models is determined by who has access to what information and at what time. Therefore, modeling state-dependent access control is at the core of adaptive workflows.

COMMUNITY-SPECIFIC CONCEPTUAL REPRESENTATION

To some application scenarios such as project management, the concept of applying community-specific conceptual representations for systems integration seems very attractive, practical and feasible (Zhu and Chen 2004, Zhu et al. 2005) because of the dynamic, transient and fragmented nature of the project management environment.

The conceptual representation, which has the ability to represent a variety of data from heterogeneous data sources and to provide a uniform access mechanism to information stored in heterogeneous data sources, differs from standardization thinking in that it does not require the underlying heterogeneous data sources to share a common standard prior to integration. Furthermore, it takes advantage of "standards" by generating and maintaining a community-specific conceptual representation of the underlying heterogeneous data sources. The conceptual representation thus can support a variety of decision-support and knowledge extraction applications. In addition, it also adapts to changes in the users' information requirements and in the heterogeneous data sources to support location transparency. This approach also clearly differs from the conventional mapping strategies in that it facilitates data integration through a community-specific *de facto* standard.

CONTEXTUAL SEMANTICS

Utilizing the theories and the concepts of context in multi-databases and other related areas, this research will explore the idea of a *dichotomy of application semantics and contextual semantics* (Goh et al. 1999) so that heterogeneous applications can share *both types of semantics for the purpose of integration*. With application semantic models supporting business applications and contextual semantic models describing semantic models for integration, it is reasonable to infer that context models can be more sharable than semantic models.

SECURE AND FLEXIBLE ACCESS CONTROL

Sharing information and data is not unconditional, instead it is determined by the owner of the information or the data in terms of when, who and what to share. The focus of this research thrust is to determine a mechanism for capturing the owner's interests in sharing data and information, as well as other collaborating parties' needs for access; to design a methodology for managing such access control; and to determine the impact of such access control on process integration as well as its conceptual representation. The research task will leverage the techniques developed for role-based access control (RBAC) (Ferrailo et al. 2001 and Sandhu et al. 1996) to model the security aspect of the contextual semantics. In a building construction project, some of the roles played by the participants in the system are that of owner, architect, builder, contractor, subcontractor or accountant. Objects can also be categorized in the context of access control. The association of role and category in lieu of

subject and object matches naturally with the organizational structure and functionalities of the workflow system, and thus greatly reduces the size of the subject/object domain. It is also more dynamic as the role and category can change as the business process evolves. Roles and categories can also be abstracted as resources, making them flexible in the proposed framework for expressing semantics regarding secure information accesses. Access requirements and access patterns will be modeled as static and dynamic contextual semantic information to facilitate secure and flexible access control in the workflow systems.

Based on the goals of this research and the underlying system integration framework, the technical approaches and this research strategy are summarized in the following section.

TECHNICAL APPROACHES

The first task is to develop case studies on the dynamics of project management, including data and process dynamics, so that information models can be developed to represent detailed requirements related to the readiness, the ability to adapt and the sensitivity of the proposed integration solution.

The characteristics of the data and process dynamics in construction project management will be identified. There exists a large body of knowledge in construction project management research that can be utilized for this study, for example,

- Integration the integration of product, cost and schedule information represents a typical challenge that has received significant amount of research attention in the past, e.g., product modeling (Eastman 1999), the relationship between estimating and scheduling (, Hendrickson and Au 1989; Rasdorf and Abudayyeh 1991; Froese and Yu 1999), and related system architectures for integration (Froese and Yu 2000). In addition, the advantages of having such an integrated information base to support project planning, design and construction were also well discussed (Fischer and Kunz 2003).
- Communication Studies (Barrie and Paulson 1992; Thomas et al. 1998; Schaufelberger and Holm 2000) show that project professionals use both formal media and informal media for communication and collaboration. The means of formal communication between project participants during construction is in the form of different kinds of documents such as contracts, drawings, specification, memos, change orders, shop drawings, payroll reports, invoice and so on (either in paper or electronic format), which are produced, stored and used at separate locations (Abou-Aeid et.al. 1995; Ahmad et al. 1995; Edwards et al. 1996; Finch et al. 1996; Thomas et al. 1998; Schaufelberger and Holm 2000). These documents are exchanged and shared during the process of construction. The media of informal communication between project participants are team-building sessions, "brown bag" meetings, disputes and person-to-person exchanges, which are less structured, less formal, and even ad hoc (Thomas et al. 1998). The successful completion of a project thus depends on the effective communication, both formal and informal, between project participants (Finch et al 1996; Thomas et al 1998; Shahld and Froese 1998; Dawood et al. 2002).

• Information Flow - Communication and collaboration between project participants are often process-oriented (Ndekurgi and McCaffer 1988; Scott et al 1995; Abou-Zeid et al. 1995; Bertonlini and Syal 1997; Shahld and Froese 1998), and are sometimes referred to as information processes (Georgakopoulos et al. 1995). The coordination and control of information processes is identified as a key factor in improving communication and collaboration between construction project participants (Dawood et al. 2002).

This existing body of knowledge on the data and process dynamics of construction project management will help us to formulate requirements in performing case studies and in developing information models that reflect the basic characteristics of construction project management.

Through the literature review and the case study development process, patterns representing the data and process dynamics in construction management will be developed. UML tools, such as case tools, static class diagrams, sequence diagrams, and collaboration diagrams etc., will be used to represent the dynamic patterns as information models and access controls.

EXAMPLE OF DYNAMIC SEMANTICS –

Assume that the *Product* entity of "Product" database has another attribute called "budget", which is the lump-sum cost of a product. Such a budget format serves the application of "Product" very well. However, an ad-hoc executive report may require information regarding the budget amount at a certain date, such as the project encumbrance date. The CSM should not only be able to identify the link between the *Activity* entity of "Schedule" database and *Product* entity of "Product" database, but it must also manage additional dynamic semantics that handle the additional relationship between them, i.e., the budget of that activity up to the date of project encumbrance.

In construction management, a lot of information such as situation changes, relationship changes, process changes, etc. are very dynamic and it is not feasible to store them in the static structures. However, such information is very important in project operations. In this study, we will explore the incorporation of some additional constructs into our proposed framework such as representing the dynamic semantics in Resource Definition Framework (RDF) format in a policy database and adding an additional tuple in the MMMs.

This policy database will store the dynamic semantics (including situation, relationship, and process changes, etc.) as policies for the system, where all the policies are represented in ontology such as RDF.

CONCLUSIONS

Answering the above challenges rests upon a system integration strategy that can effectively manage the heterogeneity and fragmentation of shared resources with interoperability and scalability. The challenge is further compounded by the fact that many business processes are highly human-oriented and that consequently human-machine interactions play an important role in the overall business process. The conventional wisdom in addressing interoperability for the automation of a large enterprise business system is to

assume a *common semantic model* shared by all heterogeneous applications. This approach requires a tremendous front-end loaded effort to establish agreements among various project participants using heterogeneous process standards and information resources, and has been proven difficult to achieve. As a result, some business processes rely on the mapping and the translation of various standardizations that were developed independently to achieve interoperability among different standards. Standard translation using *semantic schema/ontology mapping* seems realistic, however, it is not scalable and the task of fully automated mapping is still very difficult if not impossible to achieve since there is often incompleteness in a standard and contradictory semantics between standards.

In order to exploit as much semantics from all the sources (standards, domain-specific ontology repositories, human experts, and contextual information, etc.) involved in a workflow process to facilitate interoperability, the strategy of this research is based on a two-pronged **onto-semantic** framework, which consists of an **ontology service** for semi-automated ontology mapping and a **context service** for capturing contextual semantics without a shared semantic model. These two complementary services form the core of the integrated system, and are based on a *process-centered* workflow model and a *community-specific conceptual representation* of resources in the workflow system. It is chosen since the workflow characteristics in the building construction domain are very representative for our research assumptions.

REFERENCES

- Aberber, K., Cudré-Mauroux, P. and Hauswirth, M. (2003) "The Chatty Web: Element Semantics through Gossiping." The Proceedings of the 20th International World Wide Web Conference, pp. 197 206.
- Abou-Zeid, A., Russell, J. S., Hanna, A.S. and Park, S.C. (1995) "Data Flow Model for Communications Between Project Participants in a Highway Bridge Project." Canadian Journal of Civil Engineering, vol. 22, pp. 1224-1234.
- Ahmad, I.U., Russell, J.S. and Abou-Zeid, A. (1995) "Information Technology (IT) and Integration in the Construction Industry." Construction Management & Economics, vol. 13, No. 2, pp. 163-171.
- Amor, R. and Faraj, I. (2001) "Misconceptions about Integrated Project Database." ITCON. Vol. 6, pp. 57-66.
- Barrie, D. and Paulson, B.C. (1992) Professional Construction Management, McGraw-Hill, New York.
- Dawood, N., Akinsola, A. and Hobbs, B. (2002) "Development of Automated Communication of System for Managing Site Information using Internet Technology", Automation in Construction, Vol. 11, pp. 557-572.
- Do, H., Melnik, S. and Rahm, E. (2002) "Comparison of Schema Matching Evaluations." In Proceedings of the 2nd Int. Workshop on Web Databases (German Informatics Society).
- Doan, A., Madhavan, J., Domingos, P. and Halevy, A. (2003) "Learning to Map between Ontologies on the Semantic Web." The VLDB Journal, Vol. 12, pp. 303-319.
- Eastman, C. (1999) Building Product Models: Computer Environments Supporting Design and Construction. New York: CRC Press.

- Edwards, D.J., Shaw, T. and Holt, G.D. (1996) "Electronic Document Management Systems and the Management of UK Construction Projects." Building Research and Information, vol. 24, No. 5, pp. 287-291.
- Ferrailo, D, Sandhu, R, Gavrila, S., Kuhn, D., and Chandramouli, R. (2001) "Proposed NUST Standard for Role-Based Access Control." ACM Transactions on Information and System Security, Vol. 4, No.3.
- Finch, E.F., Flanagan R. and March, L.E. (1996) "Electronic Document Management in Construction Using Auto-ID." Automation in Construction, vol. 5, pp. 313-321.
- Firat, A., Madnick, S. and Grosof, B.N. (2002) "Knowledge Integration to Overcome Ontological Heterogeneity: Challenges from Financial Information Systems." In Proc. of ICIS-2002.
- Fischer, M. and Kunz, J. (2003) "Impact of Information Technology on Facility Engineering." Journal of Leadership and Management in Engineering, Vol. 3, pp. 100 103.
- Froese, T. (1992) Integrated Computer-Aided Project Management through Standard Object-Oriented Models. Ph.D. Thesis, Stanford University.
- Froese, T. and Paulson, B. (1994) "OPIS, an Object-Model-Based Project Information System." Microcomputers in Civil Engineering, Special Issue on Innovative Computer Technology in Construction, Vol. 9, pp. 13-28.
- Froese, T. and Yu, K. Q., (1999) "Industry Foundation Class Modeling for Estimating and Scheduling." The Proceedings of International Conference of Durability of Building Materials and Components, pp. 2825-2835, Vancouver.
- Froese, T. and Yu, K. (2000) "Architecture Issues For Distributed AEC/FM Systems", The 8th International Conference on Computing in Civil and Building Engineering, Stanford University, California, USA, August 14 17.
- Georgakopoulos, D., Hornick, M. and Sleth, A, (1995) "An Overview of Workflow Management: From Process Modeling to Workflow Automation Infrastructure", Distributed and Parallel Database, Vol. 3, pp119-153.
- Goh, C.H., Bressan, S., Madnick, S. and Siegel, M. (1999) "Context Interchange: New Features and Formalism for the Intelligent Integration of Information." Journal of ACM Transactions on Information System." Vol. 17, No. 3, pp. 270-291.
- Hendrickson, C. and Au, T. (1989) Project Management for Construction. Prentice Hall, Englewood Cliffs, New Jersey.
- Mitchell, B.S., Mancoridis, S. and Traverso, M. (2002) "Reverse Engineering: Search-Based Reserve Engineering." Proceedings of the 14th International Conference on Software Engineering and Knowledge Engineering, pp. 431- 438.
- Ndekurgi, I.E., and McCaffer, R. (1998) "Management Information Flow in Construction Companies." Journal of Construction Management and Economics, vol. 6, pp. 271-294.
- Rahm, E and Bernstein, P. A. (2001) "A Survey of Approaches to Automatic Schema Matching." The VLDB Journal, Vol. 10, pp. 334-350.
- Rasdorf, W. J. and Abudayyeh, O. Y. (1991) "Cost and Schedule Control Integration: Issues and Needs", Journal of Construction Engineering and Management, Vol. 117, No. 3, pp. 486 501.

- Sandhu, R., Coyne, E. J., Feinstein, H. L., and Youman, C. E. (1996) "Role-Based Access Control Models." IEEE Computers, 29(2), pp. 38-47.
- Schaufelberger, J.E. and Holm, L. (1995) Management of Construction Projects, A Contractor's Perspective, Prentice Hall.
- Scott, J.N., Bradley, D.A., Pengelly, M. and Seward, D.W. (1995) "Issues in the Development of Data Flow Simulation of Construction Sites: the Pursuit of Quality." International Journal of Project Management, vol. 13, No. 5, pp. 335-339.
- Shahld, S. and Froese, T. (1998) "Project Management Information Control Systems." Canadian Journal of Civil Engineering, vol. 25, pp. 735-754.
- Thomas, S.R., Tucker, R.L. and Kelly, W.R. (1998) "Critical Communications Variables", Journal of Construction Engineering and Management, Vol. 124, No. 1, pp58-66.
- Turk, Z. (2001) "Phenomenological Foundations of Conceptual Product Modeling in Architecture, Engineering, and Construction." Artificial Intelligence in Engineering, Vol. 15, No. 2, pp. 83-92.
- Zamanian, M.K. and Pittman, J.H. (1999) "A Software Industry Perspective on AEC Information Models for Distributed Collaboration." Automation in Construction, Vol. 8, pp. 237-248.
- Zhu, Y. and Chen, S. (2004) "A Conceptual Framework of Ontology-based Scope Alignment", submitted to the Second LACCEI International Latin American and Caribbean Conference for Engineering and Technology (LACCET'2004), "Challenges and Opportunities for Engineering Education, Research and Development", Miami, Florida, USA, 2-4 June.
- Zhu, Y., Chen, S. and Shyu, M. (2005) "A Hybrid Integration Strategy: Community-Specific, Context-Mediate and Process-Oriented", accepted for publication in the proceedings of the International Conference on Computing in Civil Engineering, Cancun, Mexico, July 11-12.