Third International Conference on Construction in the 21st Century (CITC-III)
"Advancing Engineering, Management and Technology"
15-17, September 2005, Athens

A Hybrid Integration Strategy Using Semantic and Context Models For The AEC Industry

Yimin Zhu, PhD
Assistant Professor, Florida International University, Miami, Florida, U.S.A.

Aarti Pandit, B.E.
Graduate Student, Florida International University, Miami, Florida, U.S.A.

Abstract

In the architecture-engineering-construction (AEC) industry, large-scale capital projects, mainly transportation, energy, commerce and infrastructure projects, display the common characteristic of an extensive utilization of time, capital investment, skilled and unskilled labor, material, equipment and other resources, through the entire project life-cycle. The extensive scope of the project and the complicated structure of its organization make the cost and schedule control and management aspects of these projects extremely complex and critical. The biggest problem, however, is to facilitate dynamic and ad-hoc collaboration between the various project entities to ensure detailed, accurate, updated, timely information transfer to support the decision-making process in addition to other design, planning and construction activities; as a way to overcome problems in communication created by the increasingly fragmented project organization and work structures and the application-specific or domain-specific nature of heterogeneous information systems supporting their business processes.

In this paper, the author discusses systems integration using the common semantic model and mapping strategies as a means of collaboration for decision support, which, although being popular and well-researched, have certain limitations. The paper then suggests hybrid systems integration using context models as a better alternative strategy for systems integration, with a larger spectrum of applications and lesser limitations.

KEYWORDS

Hybrid, Systems, Integration, Semantic, Context

1. INTRODUCTION

In a large-scale capital project, the total project scope is divided into numerous activities or work-packages, the project organization is fragmented into sub-organizations or project entities and the total project cost is split into several cost components, so as to make working with large magnitudes of the project's constituents

and managing the entire project easier and less complicated. Yet, all project entities need to collaborate from the project's conception, through planning and design, as well as for the entire construction phase; as they are working on a single project and the output of their individual efforts is going to be a single product i.e. the constructed facility.
A single piece of information from the entire project information database can be accessed simultaneously and several times by several project entities, who will keep adding to and updating that information piece throughout the project life cycle. To support this collaboration and dynamic information transfer process an integration solution is required. Most constituents of the project's organization require their own information systems to support their business processes and so when these project entities have to collaborate their information systems have to be adaptive enough to support this collaboration. Each information system is more often than not object-oriented and tailor-made to suit that particular business process's and project entity's requirements in order to maximize its efficiency and utility. This further complicates the collaboration process, as now an integration solution has to be provided for heterogeneous information systems with different types of heterogeneities.

This is especially crucial for all decision-making processes wherein decisions are based on information compiled from different sources or information systems. This is also true in the case of collaborative engineering teams, which communicate through electronic, telecommunication and cyber media. In the future too, as projects become larger and the degree of specialization of each task increases, the AEC industry is bound to get more fragmented from geographic, technological and functional points of view, so that the development of an integration solution to enable proper collaboration between the various project entities becomes increasingly critical.

Extensive research has been carried out on the topics of fragmentation and collaboration in the AEC Industry, especially in the case of large-scale capital projects involving many concurrent sub-projects. Systems Integration is the most accepted solution to these problems, for which several alternate approaches and strategies have been researched over the years. Integration, including managerial and technical integration is one of the earliest approaches suggested (Fischer and Kunz 1995). The common semantic model based approach has received the most research attention of all other approaches in the past two decades (e.g. Fischer and Froese 1992, 1996). Schema mapping or schema matching is another well researched integration strategy (Turk 2001), although it requires more expert intervention to identify the mapping/links between the collaborating systems and more machine intelligence to carry out this mapping and is also less stable. Another suggested approach is reverse engineering (O'Brien et al. 2002). This paper however focuses on the hybrid system integration strategy e.g. COIN (Goh et al. 1999) and ECOIN (Firat et al. 2002), a balance between the advantages and disadvantages of the single shared ontology approach (i.e. common semantic model based) and the multiple ontology approach (i.e. mapping based), (Wache et al. 2001).

2. SYSTEMS INTEGRATION USING COMMON SEMANTIC MODELS

Systems Integration using the model based approach is the most popular and well researched strategy of the last two decades. In this approach, collaborating systems share a common semantic model of the product or facility that the project is to construct. Many different types of models have been researched for various application purposes. E.g. COMBINE (Augenbroe 1994), RATAS (Björk 1994), OPIS (Froese and Paulson 1994), COMMIT (Rezgui et al. 1998) and Building Project Model (Luiten at al. 1998)

The semantics of the model and its components are unique, pre-defined and shared by all the collaborating information systems of that project. All systems using the model have to conform to these semantics, making standardization the obvious prerequisite process. Standardization strategies developed so far

include ISO STEP (STandard for the Exchange of Product model data) (ISO 10303-1 1994), IFC (Industry Foundation Classes) (e.g. Froese et al. 1999) and aecXML (aecXML 2001).

Several different software supporting their corresponding project entities, can share this common model in the form of standardized data, process or architectural semantics to enable interoperability or integration between these entities throughout the project life cycle, as shown in Figure 1. This model is more stable than others and the use of machine intelligence is limited to developing and maintaining the shared semantic representation. But, this strategy requires that the heterogeneities among the data sources or collaborating systems be identified and resolved, i.e. the systems be made to conform to the established standards and the common semantic model be established prior to actual integration. However, it has proved to be difficult to establish standards of semantic definitions at an industry level that will be used by all software vendors and the fact that researchers are still looking for alternative solutions to this issue, is proof of the limitations of this approach. Also this strategy can only be used when it is possible for collaborating heterogeneous applications to share a common semantic model, the information databases supporting these applications are stable and the business requirements that these applications have to satisfy can be predicted.

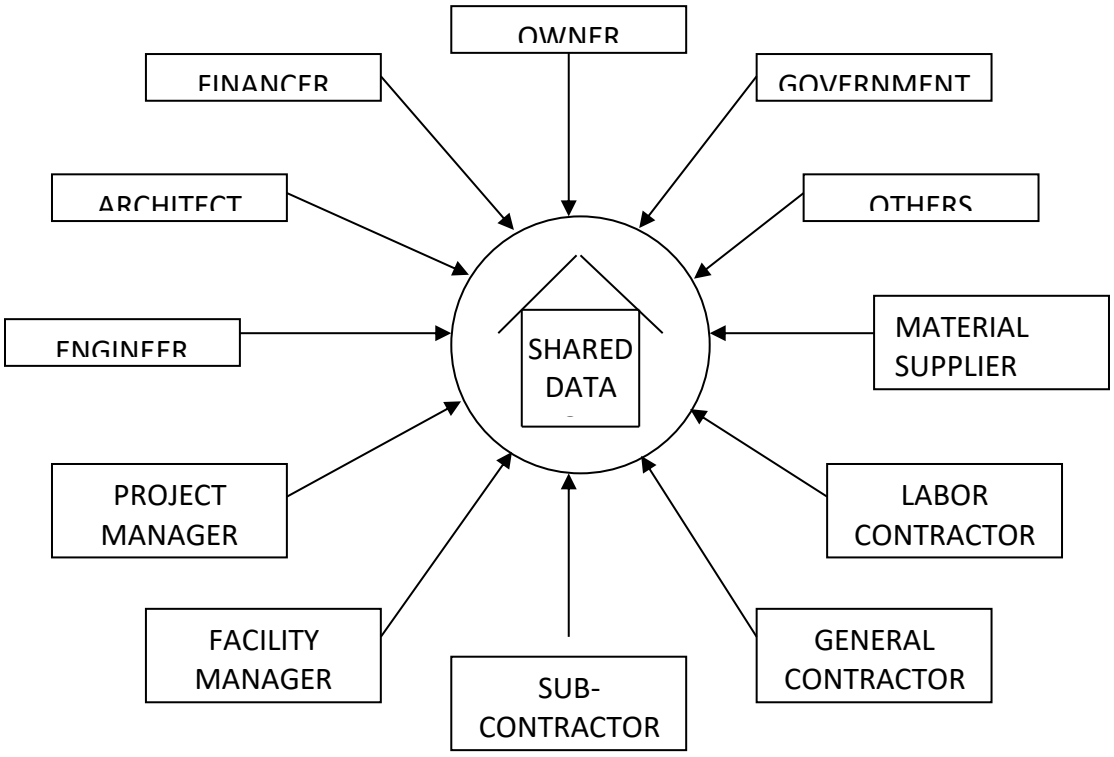

Figure 1: Common/Shared Semantic Model for System Integration

3. CHALLENGES AND MOTIVATION

In a large-scale capital project, which is composed of several concurrent sub-projects and which requires extensive utilization of resources throughout the project life cycle, dynamic and ad-hoc collaboration between the multiple constituent project entities is a common feature. Observations regarding this dynamic and ad-hoc collaboration between these autonomous information systems with different types of

heterogeneities being a common feature of the current and future construction industry have been reported in the FIATECH roadmap documents (http://www.fiatech.com).

The heterogeneous, object-oriented information systems of the AEC industry have to collaborate to satisfy the business requirements of this industry. Hence these systems have to be adaptive and able to establish scalable and robust integration dynamically (Zhu and Chen 2004). This paper stresses that instead of developing a shared semantic model for all collaborating systems prior to integration, it will be more beneficial to establish a community or project specific conceptual representation of and only for the collaborating information systems. Also the establishment of this model is considered as a process rather than a pre-defined static state.

One of the major challenges associated with this integration strategy is to determine the level of semantic details to be included in the common semantic model in order to adequately support the real world AEC domain applications. This paper suggests the adoption of a simpler method based on the dichotomy of semantic models and context models. Semantic models are used to support domain applications, thus ensuring that requirements of the domain's business processes are met with maximum efficiency and utility. Context models are representations of semantic models which facilitate integration.

Since collaboration in the AEC industry is dynamic, system integration is actually an ongoing process which should ideally not be based on the `static' common semantic model. Instead, the model used should be a community specific conceptual representation generated through an incremental or progressive process. The major goal of this paper is to suggest a hybrid integration strategy which is community-specific, context-mediated and process-oriented.

4. A HYBRID INTEGRATION STRATEGY

The hybrid integration strategy is based on three concepts:

- Since collaboration and information transfer mostly occurs at the community or project level, the shared model can be a representation of the concerned community or project information database alone. This representation should not be predetermined or static, but should be generated through an incremental process, based on the business requirements of the specific project, so as to strengthen or improve it.
- Often during the design stages of the information model, various possible scenarios of integration are envisioned and the ability of the model to handle these predicted integration scenarios serves as the basis of its design, all of which has to take into consideration the heterogeneities among the collaborating systems. Instead, the dichotomy of semantic and context models in the context-mediated approach for system integration can be used to provide flexibility to the model's architecture and to reduce the constraints normally placed on the design to facilitate integration.
- Due to the dynamic nature of collaboration in the AEC Industry, for certain types of applications e.g. project management, which requires more human, business or management related information, it may be better to resolve heterogeneity, application or human-related, without using a shared semantic model among the collaboration systems.

Based on these concepts, a hybrid strategy is proposed for system integration, which essentially has three major components:

- The model should be a **community-specific** representation of the collaborating heterogeneous data sources.
- **Context models** are used to describe heterogeneities among the heterogeneous collaborating data sources.

- **A process model** is provided to serve as a connecting point linking business and technical integrations, so that the conceptual models can be improved over time.

Integration using a community-specific conceptual representation, has the ability to represent data from heterogeneous data sources and at the same time, to provide a uniform mechanism to access the information stored in these heterogeneous data sources. Its advantage over standardization thinking is that, it does not require the collaborating heterogeneous data sources to share a common standard in order to be integrated, yet it generates and maintains a community-specific conceptual representation of the collaborating heterogeneous data sources.

There are essentially two types of conceptual representations:

- a local Community-Specific Model (CSM) – Each data source is provided with its own individual local CSM, which captures the semantic relationships among the objects within that data source., and
- An integrated Community-Specific Model (CSM) – An integrated CSM is then constructed to integrate the local CSMs, ensuring that the semantic relationships among the objects in each local CSM are captured and retained in this integrated CSM.

As long as each data model of a data source provides the defining objects in that data source along with their corresponding sets of features or attributes, the semantic relationships between them can be maintained in the integrated CSM without mapping between schemas. Hence, the autonomy of the schemas as well as of the databases is preserved as much as possible i.e. this CSM supports schema and database independence. As a result, the CSM can support a variety of decision supporting and knowledge extracting applications. It also adapts to changes from users' information requirements and the heterogeneous data sources to support location transparency. This clearly differs from the mapping strategy, in that it facilitates data integration between collaborating heterogeneous systems through a community-specific de facto standard, without carrying out mapping between the systems.

Integration studies in the AEC domain place more focus on sharing of data and process semantics, as is reflected in current literature on computer integrated construction. Recent studies have tried to utilize the concept of context in the AEC Industry for various purposes such as a high-level context model which a workflow task is associated with (Zhu and Augenbroe 2004). On the other hand, this concept is well studied and established in many areas in the field of computer science e.g. NLP (Natural Language Processing) and database integration, although its meaning varies with its area of application.

Context is defined as "meta-data related to data's meaning, properties (source, quality, precision, etc.) or organization" (Sciore et al. 1994); "the key component in capturing the semantics related to an object's definition and its relationship to other objects" (Kashyap and Sheth, 1996); or "additional constraints or assertions with reference to the domain model" (Goh et al. 1999). The main purpose of context information is to serve as a means for a system to correctly understand and interpret the data of another system during their integration or collaboration.

Utilizing theories and concepts of context in multi-databases, this paper explores the idea of a dichotomy of semantics and context (Goh et al. 1999), such that collaborating heterogeneous applications can share semantics as well as context. With semantic models supporting business applications and context models describing semantic models for the purpose of integration, the resulting models, mainly the context model will be more sharable than the semantic model alone. Also, sharing at the terminological level will prove to be more beneficial than sharing the entire context model. (Kashyap and Sheth 1996)

In the system architecture for achieving Hybrid Integration, there are two major components related to context (Figure 2), which also make the dichotomy of semantic models and context models obvious from an architectural point of view. The `Context Knowledge Base' component is where the knowledge of the

heterogeneities among collaborating systems or databases, generated from expert input or from machine learning processes, is stored in the from of computer interpretation representations. It is critical that the generation of context knowledge is managed and controlled, for which a solid software infrastructure is required. The 'Context Services' component is a software module that can provide context information required at the time of integration, e.g. retrieving data from heterogeneous data sources.

The hybrid integration strategy is thus based on a conceptual representation which is generated dynamically an improved over time so as to satisfy new business requirements. Integration is process-oriented, in that the integration process becomes a series of connecting points for business and technical integration. To limit the scope of the study, this paper will only focus on the context and hypothesizes that if context information is defined and available independent to semantic information, systems integration can be realized with a result similar to mainstream integration strategies.

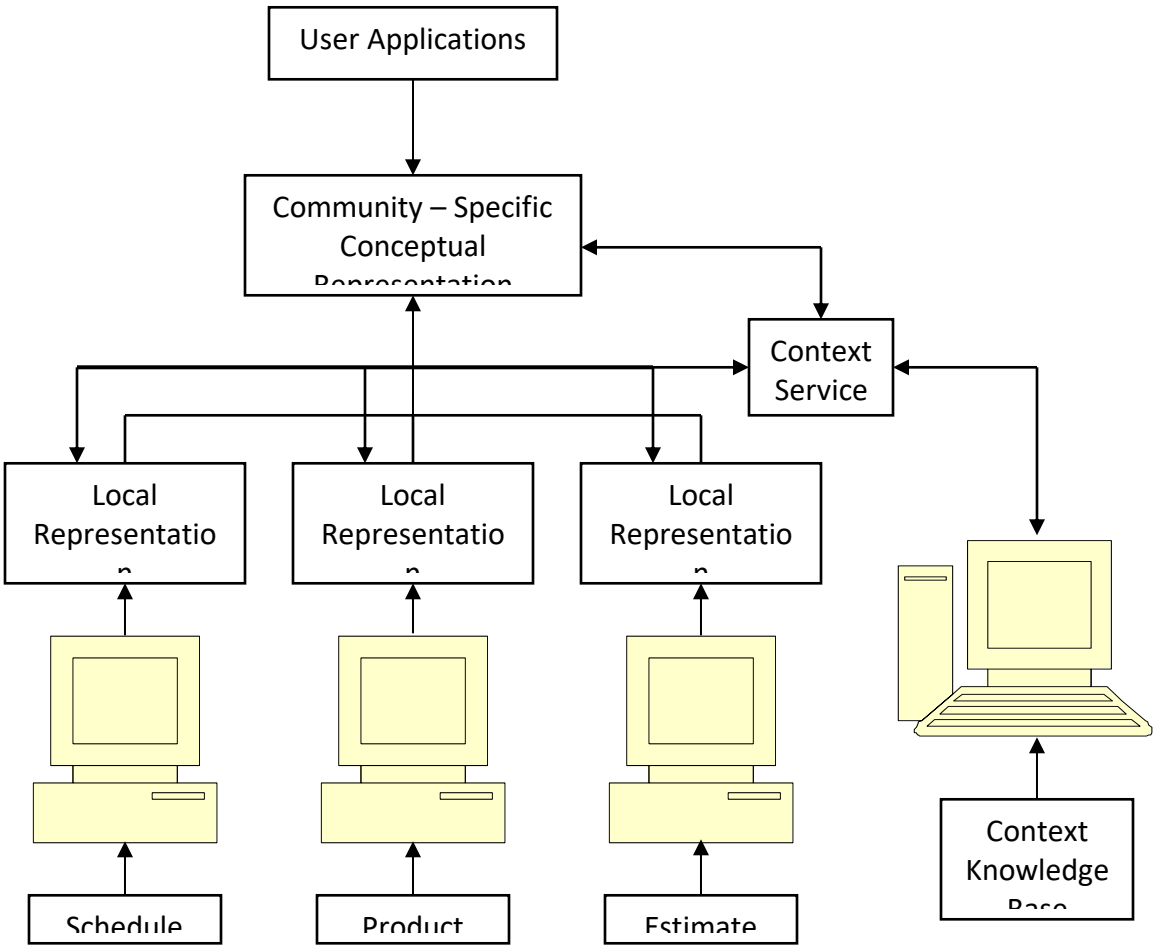

Figure 2: A High-Level System Architecture for Achieving Hybrid Integration

5. CONCLUSIONS

In the AEC Industry, for decision-making and during the entire project life cycle, several organization working on the same project need to collaborate, requiring the heterogeneous applications supporting their business processes to be integrated. Several integration strategies have been researched till date, the most well–researched of which is the strategy of sharing a common semantic model among the collaborating systems, even more popular than the schema mapping or schema matching strategy. But in cases where it is not possible to identify heterogeneities among the collaborating systems before integration occurs, these systems are unable to share a common semantic model.

This paper proposes a community-specific, process-oriented and context-mediated hybrid integration strategy as an alternative strategy for system integration, which will be a solution to the above mentioned problems and therefore will be more sharable than other proposed strategies. In this strategy, a semantic model is used to support business applications and a context model to describe the semantic model for the purpose of integration. However, certain aspects of this strategy, such as the application of context in the AEC Industry, need to be researched further, before the full potential of this strategy can be realized and utilized.

REFERENCES

aecXML. (2001). "aecXML Framework", aecXML/IAI domain committee, http://www.iai-na.org/aecXML/docements.php

Augenbroe, G. (1994). "Integrated Building Design Systems in Context of Product Data Technology." Journal of Computing in Civil Engineering. Vol. 8, No. 4, pp. 420-435.

Björk, B.C. (1994). "RATAS Project – Developing an Infrastructure for Computer-Integrated Construction." Journal of Computing in Civil Engineering, Vol. 8, No. 4, pp. 401-419.

FIATECH website. "Life Cycle Data Management and Information Integration Tactical Plan", http://www.fiatech.org/projects/roadmap/integrateddata.html

Fischer, M. and Froese, T. (1992). "Integration through Standard Project Models." CIB 78 Workshop, Montreal, Canada.

Fischer, M. and Froese, T. (1996). "Examples and Characteristics of Shared Project Models." Journal of Computing in Civil Engineering, vol. 10, No. 3, pp. 147-182.

Fischer, M. and Kunz, J. (1995) "The Circle: Architecture for Integrating Software". Journal of Computing in Civil Engineering, vol.9, No.2, pp.122-133

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.

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.

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.

ISO 10303–1. (1994). Part 1: Overview and Fundamental Principles. International Standard, ISO. Geneva.

Kashyap, V. and Sheth, A. (1996). "Semantic and Schematic Similarities between Database Objects: A Context-Based Approach." VLDB Journal, Vol. 5, No. 4, pp.276-304.

Luiten, G., Tolman, F.P. and Fischer, M. (1998). "Process-modeling in AEC to integrate design and construction". Computers in Industry, Vol.35, No.1, pp. 13-29.

O'Brien, W. J. et al. (2002). "SEEK: Accomplishing Enterprise Information Integration across Heterogeneous Sources." ITCon, Vol. 7, pp. 101-123.

Rezgui, Y., Cooper, G., and Brandon, P. (1998). "Information Management in a Collaborative Multi-actor Environment: The COMMIT approach." J. of Computing in Civil Engineering, Vol. 12, No. 3, pp.136-144.

Sciore, E, Siegel, M and Rosenthal, A. (1994). "Using Semantic Values to Facilitate Interoperability among Heterogeneous Information Systems." ACM Transactions on Database Systems, pp. 254-290.

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.

Wache, H. et al. (2001). "Ontology-based integration of information – a survey of existing approaches." In Stuckenschmidt, H., ed., IJCAI-01 Workshop: Ontologies and Information Sharing, pp. 108-117.

Zhu, Y. and Chen, S. (2004) "A Conceptual Framework of Ontology-based Scope Alignment", the Proceedings of 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.

Zhu, Y. and Augenbroe, G. (2004). "A Reference Model for Supporting Inter-Organizational Information Process Integration", under review by Automation in Construction.