A PRELIMINARY STUDY OF KNOWLEDGE MANAGEMENT IN COLLABORATIVE ARCHITECTURAL DESIGN

JU-HUNG LAN
Department of Architecture
Kao Yuan Institute of Technology
Kaohsiung, Taiwan

Abstract. Collaborative design has been suffered by dealing with huge amount of undesired information today. The variety of information stored in a collaborative design system actually provides a knowledge repository across multiple design domains. This paper presents an information mining approach to capture hidden knowledge within collaborative design information. The discovered knowledge is used to develop an information service mechanism for helping collaborators to access related information during collaborative design processes.

1. Introduction

Collaborative architectural design has been a research topic through the use of information technologies since last decade. The computer-supported cooperated work (CSCW), virtual design studios (VDS) and collaborative design studios (CDS) all explore the capabilities of information technologies from synchronous design and asynchronous design point of views (Chiu et al., 2001; Gross et al., 1999; Kalay, 1999; Kvan, 2000; Lan and Jeng, 2001; Morozumi, 2001). By using a collaborative design system, the collaborators could communicate with each other in a distributed way. They could report design progresses, posting design comments, making design juries and chatting with each other easily. However, what is the next step of collaborative design researches to go?

While the establishment and use of collaborative design systems ease the design process, the collaborators have to deal with vast amounts of information available in the systems and it is very time-consuming for exploring desired information (Chiu and Lan, 2003). What can we do to deal with this kind of information overload problem and really get something

useful from the collaborative design information? In this paper, the research presents an approach to develop an information service mechanism to support collaborative design based on the concepts of knowledge management and information mining.

2. Concepts of Knowledge Management and Information Mining

Researches in knowledge management originated from quality movement and process reengineering. Both two fields emphasize to make product with high efficiency, low cost and high quality based on reorganization viewpoint. The theory of "Spiral of Knowledge" proposed by Nonaka is the starting point of knowledge competing era （Nonaka, 1995）. In 1997, Drucker, the master of management, declared the trend of knowledge works in the future. Also, Gates mentioned the competing age of future enterprises based on knowledge and computer networks （Gates, 1999）. In an era of knowledge economics, the general important issue is how to grasp knowledge and manage it since the ill-defined features of knowledge.

In general, there are two kinds of approaches in knowledge management researches. One is from management of information point of view, and the other is from management of people point of view. The research adopts the former viewpoint to explore the knowledge management issue among collaborative design information. The research assumption is the variety of information stored in a collaborative design system provides a knowledge repository across multiple design domains. Therefore, how to capture knowledge from collaborative design information and manage it is the main task of this research intends to achieve.

The concept of information mining is borrowed from data mining theories and techniques. Data mining is to find previously unknown and potentially useful knowledge from vast amount of data. They are the processes of discovering meaningful patterns through data by using multi-disciplinary technologies （Adriaans and Zantinge, 1996; Han and Kamber, 2000; Mena, 1999）. Based on the assumption mentioned above, this research intends to use the information mining techniques to discover meaningful information service patterns, which means useful knowledge, from collaborative design information. These discovered patterns then could be used to develop an information service mechanism to support collaborative design, which means the purpose of knowledge management in this research.

3. Design Data, Information and Knowledge

The definition about data, information and knowledge is an important issue in the research field of knowledge management. In this paper, design data is defined as a basic data element in a collaborative design system. It could be a single file or a record in database, for instance, a drawing file, a design document or a communication message. The design information is further defined as a set of design data, organizing for communication purpose during design processes. For instance, a set of design documents and drawings are organized for design communication purpose among designers and consultants. The typical design data/information in a collaborative design system includes design documents, drawings, communication messages, charts, 3D models, etc. They are all partial representations related to a specific design project. Typical design data/information in a collaborative design system is shown in Figure 1.

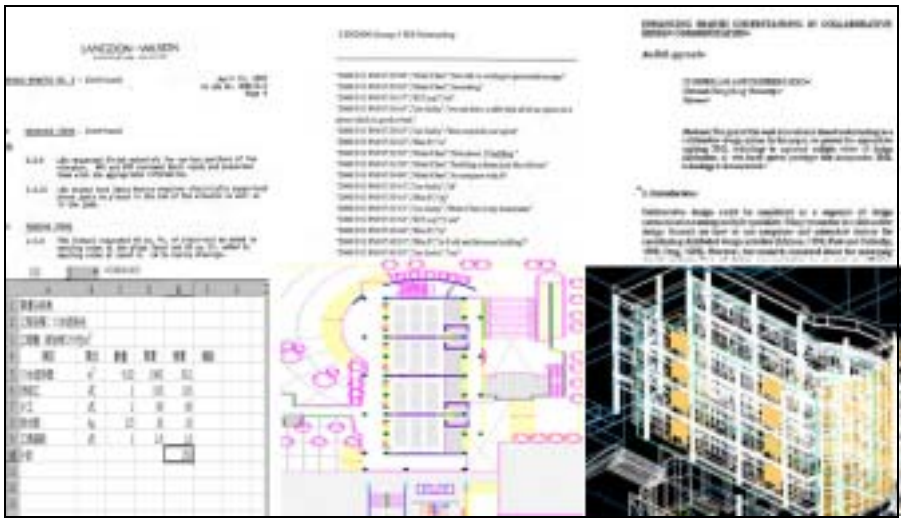

Figure 1. Typical Design Data/Information

As to the design knowledge, how to define it from design information point of view is an important issue in this research. Since design knowledge normally belongs to designers with the essence of ill-defined feature, how to unhide it is critical to successfully represent design knowledge and make it reusable in a collaborative design system. However, the research currently defines design knowledge from information management point of view. Based on the information overload problem mentioned above, this research defines design knowledge as information service patterns regarding with specific design situation. The idea is, in a specific design situation, a useful information service pattern should be able to assist collaborators in finding

most desired information for making situated design decision in a collaborative design system. Therefore, how to discover an information service pattern regarding with specific design situation is the key to capture and manage design knowledge with effectiveness. In this paper, the research proposes a 6W approach incorporating with information mining techniques to achieve the research goal.

4. A 6W Approach

In general, most design documents in Figure 1 are not standardized and without keyword features. Even these can be categorized into some structured formats, but are still difficult for abstracting relations among design data/information. Although general collaborative design systems provide query-based searching mechanisms with Boolean operators to search keywords relating with design documents. However, this kind of searching mechanism has difficulties to deal with vast amount of information （Chiang et al., 2001）. It cannot find such as synonyms, double-entendre, or a word with various meanings. Nor it can provide the searching ability to find associated design information addressing specific design situation.

The variety of information stored in the backend database of a collaborative design system could be classified as structured and ill-structured information. The structured information, for instance meeting briefs or memos, usually provides tree-structure classification functionality of design information in according with specific design information category. The ill-structured information, for instance message board data, usually are sequential records in according with timing of information posted. It is necessary to abstract relations among both structured and ill-structured information for further applications.

When data are organized into design information for communication purposes, a specific design situation related to participants, timing, places, tasks, or rationales occurred. Therefore, a 6W (who, when, where, what, how, and why) approach is proposed for associating related information in this research. For the practical reasons, the collected design information is classified into three levels: project, group, and individual. The information created in the process can automatically be organized into sub-datasets by the 6W-based classification, i.e. who are involved, or when, why, what, where, and how information are occurred. Each participant then can access the information from the individual, group, project or 6W point of views. For facilitating design information services, the correlation of 6W-based information can be developed as a content-awareness mechanism as Figure 2.

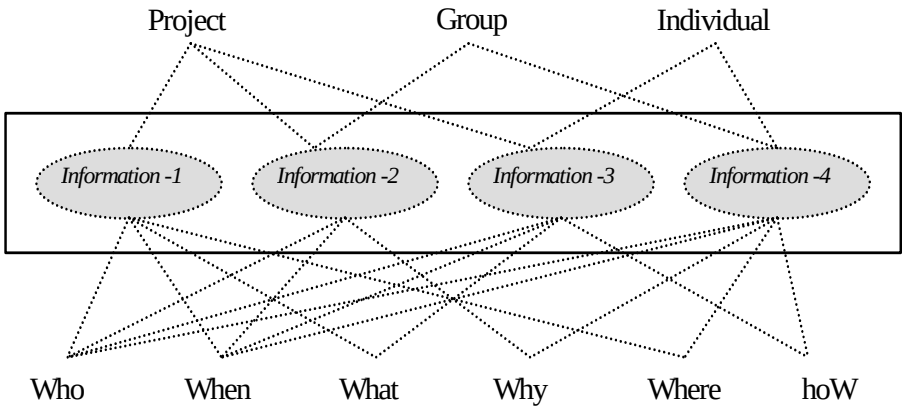

Figure 2. The correlation of 6W-based information

5. Collecting Collaborative Design Information

To support collaborative design activities, this research implemented a prototype of collaborative design system, named as DECADE (a Distributed Environment for Collaborative Architectural DEsign), Figure 3. The functions of DECADE include project management, information browsing and searching, message board, upload/download files, and a viewer for non-text information. While a design project is created in DECADE, this project is assigned to a specific design team. Within DECADE environment, the leader of a design team should set a team-oriented collaborative design process for the team members to follow. During the collaborative design process, each team member could bring up specific design issue for group communication at certain design stage. A whiteboard system mapping with the team-oriented collaborative design process is used to record the information generated by the team members. This research uses DECADE to collect collaborative design information as the platform for mining the information service patterns.

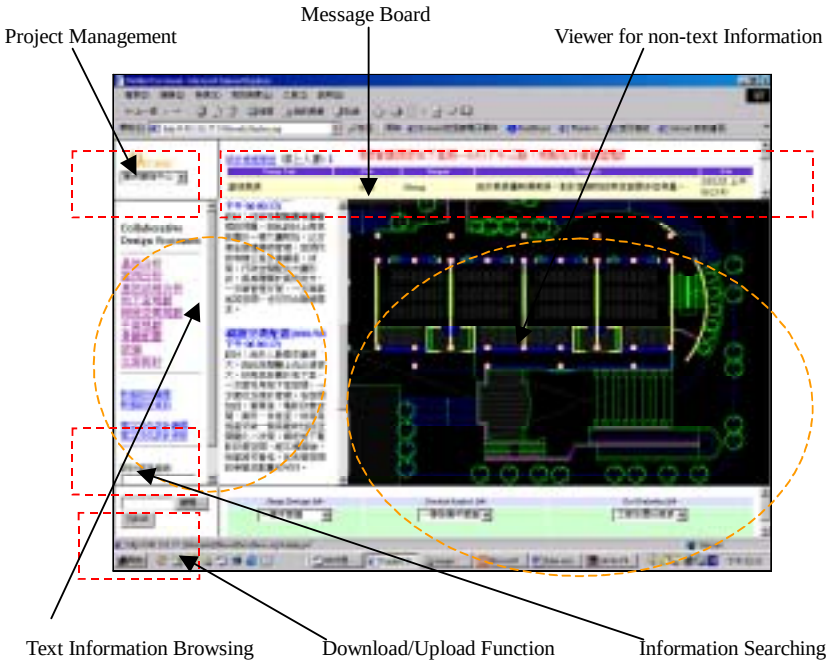

Figure 3. The System Prototype

For the comparison purpose, this research collected collaborative design information by conducting a series of collaborative design projects. These projects are organized in according to the notion of four subgroups, shown in Figure 4. In this research, the main source of design information is collected from collaborative design studios. The students are instructed to generate collaborative design information during collaborative design processes. They are separated as four design teams for doing different design projects. Each design team is further divided into three sub-domains, including architecture design, construction design, and environmental design. The students are advised to specially generate design information from multiple design perspectives. All design teams are instructed to set a team-oriented collaborative design process while conducting their design project. On the other hand, the information collected from design practices, which are not supervised, is only partial design information addressing specific design issue. The research intension is to use information collected from design studio projects as the target data for information mining, and the information collected from design practices could be used to testify and evaluate the mining results.

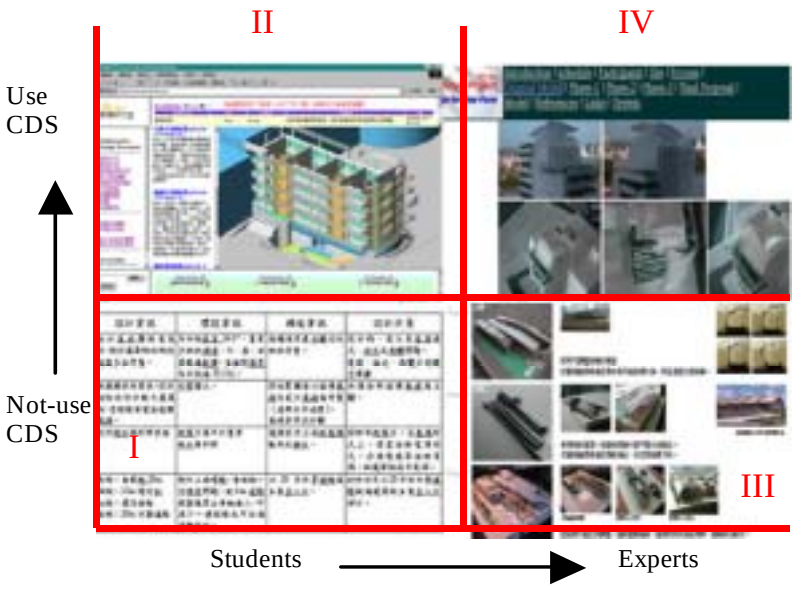

Figure 4. Collaborative Design Projects in this Research

6. Design-issue-oriented Information Framework

While the collaborative design information is ready, the key question is how to discover the hidden information service patterns within it. The research particularly focuses on discovering the classification and association patterns among collaborative design information. Although the 6W features mentioned above already provide information classification and association mechanisms, this research further exploring the hidden information service patterns by information mining techniques.

To do it, the research proposes a design-issue-oriented framework to allocate collected design information for conducting information mining. This framework is a two-dimensional tree-structure for allocating collaborative design information, shown as Figure 5. The x-dimension allocates information addressing main design issues/stages during design processes, while the y-dimension allocates information addressing more detailed sub-design issues. This two-dimensional tree-structure represents the categories of information belonging to, which is design-issue-oriented and serves as the information classification patterns. The attributes setting of each information node in the tree-structure is regarding with 6W features, which serves as the information association patterns. Besides, the information mining techniques are used to discover the hidden information

patterns based on this information framework. These discovered patterns then can be used to enhance the classification and association capabilities among design information.

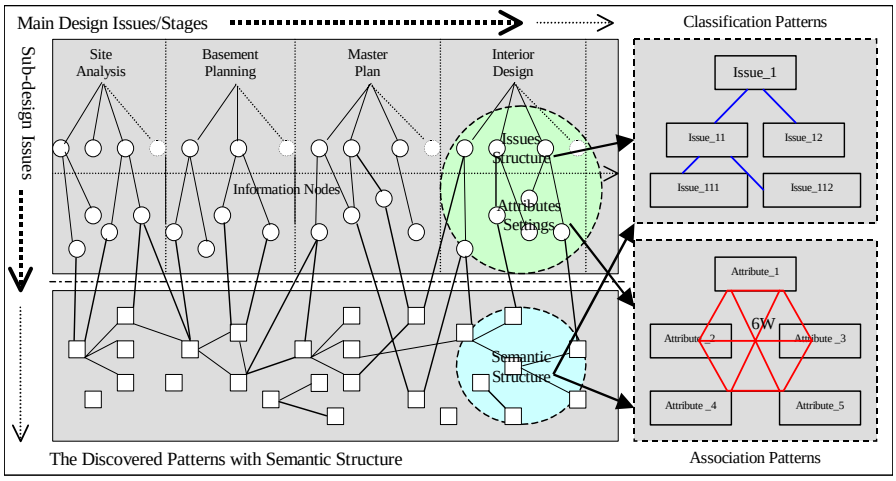

Figure 5. The Framework of Information Mining

7. Design Information Mining

In this preliminary study, the text mining techniques based on information retrieval (IR) theories are used to discover the hidden information patterns. The idea is to find a computational way to represent all collected design information. Then we can calculate the relationships among design information for indexing purposes. In order to represent collected design information, this research adopts a vector space model, which is popular and more effective based on current IR researches. In a vector space model, each design information is represented by a set of design terminologies (T_1, T_2, ⋯, T_n), and each design terminology is assigned a weighting value according to its importance in the design information (W_1, W_2,⋯, W_n). Thus, all design information could be represented by a n-dimensional vector space model. The similarity of target design information (Y) and unknown design information (X) then could be computed by equation (1). As the term-weighting strategy, this research uses a popular *tf-idf* scheme based on term frequency (*tf factor*) and inverse document frequency (*idf factor*) for clustering design information, equation (2)（Baeza-Yates and Ribeiro-Neto, 1960）. The mining processes are introduced as follows.

$$Sim(X,Y) = \cos(X,Y) = \frac{\sum_{k=1}^{n} W_{xk} \bullet W_{yk}}{\sqrt{\sum_{k=1}^{n} W_{xk}^2} \bullet \sqrt{\sum_{k=1}^{n} W_{yk}^2}} \quad (1)$$

$$W_{i,j} = \left(\frac{freq_{i,j}}{\max_l freq_{l,j}}\right) \times \log \frac{N}{n_i} \quad (2)$$

7.1. INFORMATION PRE-PROCESSING

Information pre-processing is to filter out useless data and allocate useful information into the two-dimensional information structure as Figure 5. The task of filtering data is to avoid unsuitable design information collected from design projects, and the task of allocating design information is to enrich the original design information for the purpose of mining meaningful information patterns. The database techniques supporting with domain knowledge are used to filter and enrich the collected design information into an integrated database. The support from domain knowledge is helpful to encode the ill-structured design information such as drawings or charts into text-based information for mining. Furthermore, it is a reliable way to deal with the issues of synonyms, double-entendre or different languages at an initial stage.

7.2. INFORMATION MINING

After establishing the integrated information database, the next step is to do the information mining to find the hidden information patterns. To conduct it, the research starts to discover the design terminologies of collaborative design information. The statistics and database techniques are used to produce a summary analysis of collected information to find frequently used design terminologies. These design terminologies then are used to represent all collected design information by a vector space model mentioned above. Figure 6 shows the design terminologies discovered in collected design information. Currently this research focuses on dealing with Chinese information since the limitation of collected design projects. Table 2 translates partial terminologies regarding with opening design issue into English format for reference convenience.

Figure 6. The Discovered Design Terminologies

Figure 7.The Associating Relationship Among Design Terminologies

Table 2. Design Terminologies Related to Opening Design Issue

Classification Terminology	Association Terminology
Size	Size (Width, Height, Depth)
Style	Style (Type, Material)
Site Analysis	Site Analysis (Orientation, Climate, Sun, View, Noise)
Climate	Climate (Temperature, Humidity, Rain, Wind)
Sunshade	Sunshade (Board, Style, Shading)
Daylight Analysis	Daylight Analysis (Shading, etc.)
Interior Design	Interior Design (Louvers, Lighting, Shading, Furniture, Space
Envelope Design	Layout, Ceiling)
Façade Design	Envelope Design (Façade, Proportion, Aesthetics, Module)
Structural System	Structural System (Column, Structural span)
Mechanical System	Mechanical System (Temperature, Humidity, Air condition)
Regulation …	Regulation (Escape, Retreat) …

Based on discovered design terminologies, the research further studies the associating relationship among these terminologies. The idea is to build item sets of design terminologies for the represented information according to the information framework in Figure 5. These item sets actually correlate with each other by 6W-related or design-issue-oriented structure. The statistics and database techniques, again, are used to discover the association rules among design terminologies. Figure 7 shows the results of associating relationship of discovered design terminologies. Typical association rule with confidence ratio (C) and support (S) ratio in this research is showed as:

{Terminology 1} → *{Terminology 2}, C=0.35, S=0.56* (Agrawal et al., 1993).

7.3. EMPIRICAL STUDY

The research starts information mining from few similar design situations such as opening design, facade design and roof design to associate envelope design information. Figure 8 illustrates partial visualization results of information classification and association in the library design project based on design terminologies and 6W features. For instance, the information in the opening design branch links to the interior design branch by the shading terminology, lighting terminology and 6W attributes.

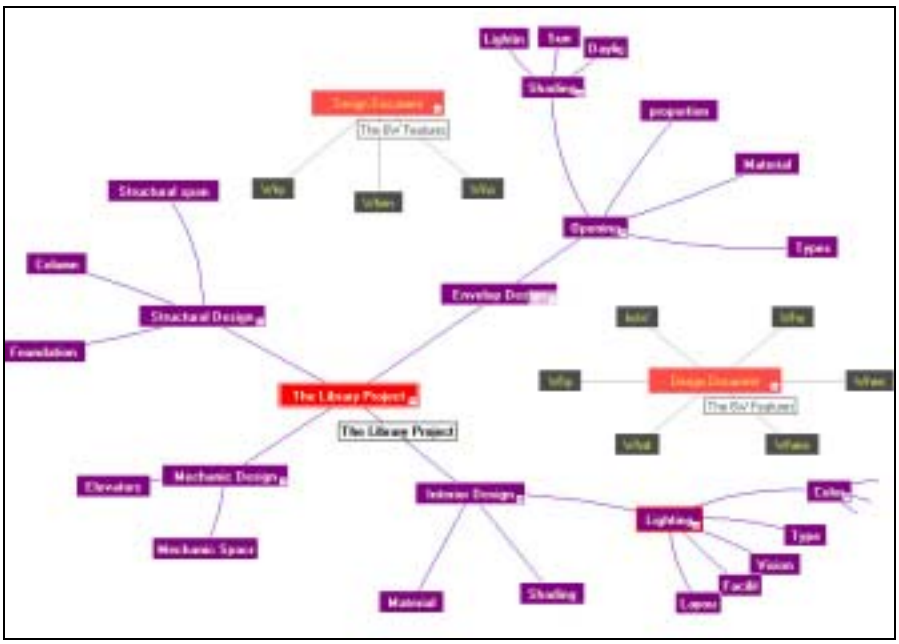

Figure 8. The Visualization of Information Classification and Association

8. Summary and Discussions

In this preliminary study, the research adopts the concepts of knowledge management and information mining to develop an information service mechanism in a collaborative design system. This mechanism provides 6W-related, design-issue-oriented and terminology-based relationships to classify and associate collaborative design information. This information service mechanism shows great potentials to assist collaborators in getting most related information regarding with specific design situation. The

collaborators then could make more effective decisions during collaborative design processes. The study raises some important issues for discussion as follows.

8.1. DISCOVERY OF PATTERNS

In this research, the patterns are considered as various structures to organize information for information service purpose. The information collected in this research is design-issue-oriented with 6W features, which are the keys to discover the classification and association patterns. Besides, the relationship among design terminologies provides a mean to associate information with different ratios. Based on issue detection, 6W features and terminology relationship, the discovered information service patterns could be more effective for supporting CSCW. For further study, the patterns discovered, especially for the design terminology patterns, should consider the issue of generality or specialty since different information resources might lead to different pattern results.

8.2 DESIGN DICTIONARY

The establishment of the design dictionary is subject to domain knowledge, expertise, and applications, while the dictionary is also constrained by how information are generated or prepared. The research built the design terminologies by mining collected design information from typical design projects. However, the dictionary might be restricted because the information collected is based on limited projects. The ideal dictionary could be more comprehensive. In the future, it is necessary to collect more design information from various design projects for building the design dictionary.

8.3 INFORMATION PROVISION

To access the most related information in a collaborative design system, information can be classified and associated by various point of views. How to classify and associate a vast amount of information becomes an important task. The research finds that users often have different levels of needs based on 6W features or issues. Therefore, it is desirable to define different levels of classification and association by users for various purposes, for instance, to choose partial important issues by using first 3W (Who, When, What) for design progress review, and then expanding to another 3W (Why, hoW, Where) for problem definition or decision-making. In a complex project, it is more feasible to use one W (Who) first, followed by design issue consideration, to classify and associate design information. In this manner, the system can facilitate users to handle vast amount of information on hand.

Acknowledgements

This research is supported by National Science Council in Taiwan, NSC92-2626-E-244-001. The author is grateful to this support.

References

Adriaans, P. and Zantinge, 1996, *Data Mining*, Addison Wesley

Agrawal, R., T. Imielinski and A. Swami, 1993, "Mining Association Rules Between Sets of Items in Large Database", *SIGMOD'93*, Washington, D.C., pp. 207-216

Baeza-Yates R. and Ribeiro-Neto B., 1960, *Modern information retrieval*, Addison-Wesley

Chiang, R.H.L, Chua, C.E.H, and Storey, V.C., 2001, "A smart web query method for semantic retrieval of web data", *Data & Knowledge Engineering*, Vol. 38, pp. 63-84

Chiu, M.L., Yamaguchi, S., and Morozumi, M., 2001, "Supporting Collaborative design studios – Scenarios and tools", *CAADRIA'01*, pp. 125-134

Chiu, M.L., Ju-Hung Lan, 2003, "Information and IN-formation - Information mining for supporting collaborative design", *proceedings of eCAADe 2003*, Graz, Austria

Drucker, P., McDermott, R., et al., 2002, *Knowledge Worker's Handbook*, BookZone

Gates, B., 1999, *Business @ the Speed of Thought*, Microsoft Press

Gross, M. D., Do, Y. L., McCall, R. J., et al., 1999, "Collaboration and coordination in architectural design: approaches to computer mediated team work", *Automation in Construction*, Vol.7, Issue 6, No.2, pp.465-473

Han, J. and M. Kamber, 2000, *Data Mining: Concepts and Techniques*, Morgan Kaufmann Publishers

Kalay, Y.E., 1999, "The Future of CAAD: From computer-aided design to computer-aided collaboration", in Eastman, C. (ed), *Proceedings of CAAD Futures'99*, Georgia, USA, 13-30

Kvan, T., 2000, "Collaborative design: what is it?", *Automation in Construction*, Vol. 9, 409-425

Lan, J.H. and Jeng, T., 2001, "Enhancing shared understanding in collaborative design communication: An xml approach", *CAADRIA'01*, pp. 285-290

Mena, J., 1999, *Data Mining Your Websit*e, Digital Press

Morozumi, M., Homma, R., Shounai, Y., et al., 2001, "Web-based collaborative design studio: tools and programs", *CAADRIA'01*, pp. 93-102

Nonaka, I., and Takeuchi, H., 1995, *The Knowledge-Creating Company*, Oxford University Pre