An Ontology-driven Approach for Monitoring Collaborative Design Knowledge

Y-C. Lai & M. Carlsen
Dept of Building Technology & Structural Engineering, Aalborg University, Aalborg, Denmark

Proceedings of the 5th European Conference on Product and Process Modelling in the Building and Construction Industry - ECPPM2004, 8-10 September 2004, Istanbul, Turkey. Leiden, A.A. Balkema Publishers, 2004. (pp. 311-317)

ABSTRACT: The use of meeting minutes has been confined as summaries that are recorded corresponding to the discussion content, and sometimes as agenda for the next meeting. The conventional approach implemented to arrange the summaries in tree-structure format results in design rationales and decision intents are implicitly contained in the written plain-text. The conventional meeting minutes are also incapable of integrating pieces of design information. This paper hypothesizes that semantically structured meeting minutes may serve as dynamic medium to record key design information by presenting the design intents explicitly. This paper describes a prototype system that is developed based on semantic web technologies and ontological approach. The prototype support design progress meeting by generating dynamic records with respect to the content of discussion.

1 INTRODUCTION

It has been aware by the building industry that an ergonomic fit project information management system is one of the fundamental needs at the project outset so that knowledge corresponds to the activities conducted throughout the project life can be managed effectively. The project information management system has evolved from the conventional paper-based mechanism to the nowadays digital based mechanism. However, the implementation of the digital based project information system does not respond perfectly to the demands of the building industry yet indicated weaknesses identified from the current web technologies. This paper presents a hypothetical knowledge management system supported by the semantic web in which ontological approach is applied for knowledge modeling purposes. The outlines of this paper is divided into two parts where the current practice of project information management will be analysed followed by a comprehensive description of the hypothetical system that is to improve the pitfalls identified from the current practice.

1.1 *Background*

Design process of the building industry was characterized as a sequential conversion flow that transforms information from technical standards, legislations and other design specialties into solutions and product specifications. The conventional design process as such is typically disciplinary orientated, which means that different team actors concern mainly about their respective interests and knowledge to formulate technical solutions corresponding to the design requirements under their disciplinary specialties. Reworks in designs are the consequences of this fragmented design process in order to maintain the coherence among the numerous decisions made throughout particularly the ambiguous and frequently changed briefing and early design process (Mesquita, 2002). Close collaboration amongst the multidisciplinary team actors has therefore become a necessity from the outset of a project.

2 COLLABORATION THROUGH KNOWLEDGE MANAGEMENT

The characteristic of building project is unique, which is to involve multi-participants from different business nature but are required to collaborate closely throughout the project life. Reuse and share of knowledge are underlying activities that have been conducted without much awareness of project team actors as a mechanism to improve collaboration amongst them. This is primarily due to not so many of us are aware that information that can lead to effective action can also be defined as explicit knowledge (Davies, Fensel & Van Harmelen, 2003) while both the tacit and explicit knowledge have been recognized as the important strategic resource of an organization (Nonaka & Takeuchi, 1995). It

has been the custom for building industry to organize the vast amount of information, particularly those generated during the early design phase, with particular mechanisms in a way that future reuse and retrieval are possible. Organizing the mass quantity of information has however been realized as an uneasy task (Fruchter, 2002).

2.1 *The Attempts*

The basic need of an ergonomic fit project information management system is to be able to support the traditional project management tasks of planning, monitoring, reporting and control of baseline scope, cost, time and quality (Archer et al. 1997). Such a system is also expected capable with the mechanisms of trend forecasting and change control, and also be able to manage documents in a manner that would track issues, provide fast retrieval of relevant documents and support the time limited process for the resolution of disputes (Archer et al. 1997). Several attempts have been conducted including the concept of project web, which tends to apply the fast developing information communication technologies (ICT) to manage the existing information base more effectively. A comprehensive discussion in regard to project web associated with the technology behind will be given in the following section of the paper.

3 METADATA AND PROJECT WEB

3.1 *What is Metadata*

In brief, metadata can be defined as data about data, and obviously itself is also data. Metadata can be embedded in the document that it describes/represents, or exist separately from that document. Metadata can be used to describe any object in the universe. Generally, metadata about an object is structured to provide a description. The structure is common for all instances of the same type of object. A very typical example for this is the library card system where each library card contains description of a book such as title, author, keywords, publishing date, and so forth (NAEH, 2004b). Metadata used in the library card system is to facilitate the library user for books searching by considering a book as an object, which has a number of properties that can be represented with descriptive metadata. With this respect, it has been obvious that the application of metadata has its long history in the aspect of information management.

3.2 *Project Web*

Project web is a project-level mechanism, which is to function as a centralised repository for project team actors to exchange project information during the phases of planning, design, construction and facility management (FM). Digitalised information, such as design drawings, progress reports and meeting minutes are available in this information container.

3.3 *Use of metadata in the building industry*

Integrating information, as a mechanism to improve the efficiency of information monitoring, becomes a crucial task for the A/E/C professionals. The mass amount of information produced at the project outset may have a big variety of formats, including the well structured data stored in database management system (DBMS), the semi-structured HTML and/or XML files, and also the weakly structured texts/graphics/multimedia files (Maher & Simoff, 1998). Both technical and managerial approaches have been investigated for the purpose of improving information integration (Fisher & Kunz, 1995). These approaches involved the use of a centralized project model that adopts data standards range from the previous ISO-STEPS to the recent IFC, in which structured data integration was primarily of concern. Apart from that, some recent researches started to take unstructured data integration into consideration. Amongst these efforts were design tool that could capture, share and reuse project information (Fruchter, 2002), approach that could extract concepts from textual design documentation, the use of arbitrarily metadata that could markup documentation (Brüggemann et al., 2000), as well as the use of controlled vocabularies that could integrate heterogeneous data representations (Kosovac et al., 2000). Apparently, metadata has been one of the approaches under investigation by the A/E/C professionals tended to improve data integration.

3.4 *The current practice*

The use of project web by the A/E/C professionals becomes more and more widespread. However, the conscious use of metadata in the construction industry is seldom. The use is often limited to unconscious use of more or less occasional metadata elements in desultory situations. Therefore the potential benefits of using metadata are lowered.

3.5 *Case studies*

Several semi-structured interviews were conducted by the author of the paper to investigate project webs operated for the building industry. During the interviews, limitations of project webs with respect to their respective efficiency in information dissemination were delineated by the interviewees, and were described hereafter. Information was first categorized based on some sort of relations, and then was archived under different electronic file folders corresponding to the information categories. Semi- and

unstructured information such as briefing notes, design rationale, and e-mail messages, was usually not stored in the project web. In general, e-mail messages were collected in another project-level digital information source while the paper-based information was kept in company-level paper-based archives such as file cabinets. Drawings were generated at every stage with respect to the change of design, but only the final version was uploaded to the project web. In brief, such descriptions reflected the implication of fragmentary communication and information flow in project web. As a consequence, project web turned out to be highly dependent on inefficient human efforts in processing such as searching, browsing and extracting the stored information.

Comments about the contribution of project web in the regular basis design progress meeting were also given by interviewees. It was commented that, meeting participants (i.e. project team actors) required spending rather long time to read a specific piece of information in order to comprehend its context during discussion in meeting. It was also observed that meeting participants faced the difficulty to find a specific piece of document during the meeting, particularly when the need of that document arose at random. On the other hand, the person who was responsible for making meeting minutes tempted to capture the discussion content on papers. The captured information would then be transformed to digital format with word processor tool after the completion of meeting before it could be uploaded to project web, which resulted in repetition of workload. The conventional notes-taking approach also structured the meeting summaries in tree-structure format with design rationales and decision intent implicitly contained in written plain text. The implicit design rationales and decision intents could only be interpretable rapidly by those who attended the meeting and actively joined the discussion. For those who did not sit in on the meeting but were interested in following the design progress, extra time needed to be spent to collate and review the series of meeting minutes that were stitched with time element. The conventional meeting minutes were also incapable of integrating pieces of design information that had been produced throughout the early design process. This increased the time needed to review the stitches of meeting minutes in particular when the necessity to gather the relevant but scattered design information arose.

3.6 *The analysis of questionnaire survey*

Questionnaire Agency for Enterprise and Housing to study a blend of eight Danish and international project webs to what extent metadata was applied in their respective systems (NAEH, 2004a). The purpose of the study was to understand the current situation of metadata implementation in order to look into potential improvements when necessary. The analysis with respect to the result of the survey study indicated that there was between the systems a big diversity if users were obliged to use metadata to connect to files. Only in half of the systems a search on metadata could be done otherwise they were just indicated when the user browsed the files. For instance it was only possible to make a direct search on metadata in a very few systems. User reactions therefore claimed that the retrieval of documents often took place in a list named e.g. ‘New Documents’ or ‘Since last Time’. This list was sorted by date and contained therefore documents concerning quite different subjects. The list was emptied after each visit and the documents were automatically stored in folders that correspond to categories of the different subjects. This user pattern illustrated that either the users were not aware of the functionalities or more likely that document retrieval that functioned based on metadata were not well-developed.

The connection of metadata to files was either done automatically by the system or manually by the user. The automatically connected elements were often such as Date, User ID and File Name (NAEH, 2004b). Only a few elements such as Document Number and Revision were mandatory while the rest were optional. The optional elements were sometimes used with different approaches or just ignored by the system user.

In most cases, the metadata elements did not follow a common standard. The elements were defined on the basis of the provider’s preferences and not on actual user needs or user patterns. The standards that were used to define the metadata elements were either out of date or not commonly used and therefore useless (NAEH, 2004b).

For the further work with metadata in project webs systems with respect to these shortcomings, it would be naturally to follow a standard. ISO/IEC 82045-5 ‘Document management - Part 5: Application of metadata for the construction and facility management sector’ are a standard for the utilization of metadata in the construction industry and provides 4 metadata sets each of which directed to specific phases of the construction process (ISO/IEC CD 82045-5). The standard is under development.

For some project team actors such a standard would make some limitations in their work procedure and requires modifications of this, but in accordance with common understanding ‘a poor standard is better than no standard’. With the standard proposed by ISO/IEC 82045-5, it would be possible to raise the use of project web systems from merely document containers to more intelligent document management systems. By following a standard the users would be presented with the same user interface independent of which project web system the actual construction project are using. The transfer of

data from one system to another would then be less problematic.

4 THE HYPOTHETICAL SYSTEM

4.1 *The concept*

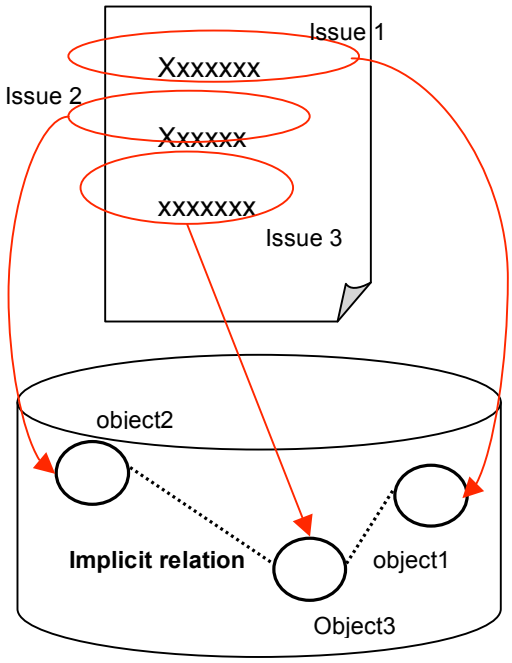
With reference to the comments given by interviewees and the analyses of questionnaires survey, a semantic web based knowledge management system is developed to improve the management of project information which is a crucial means to enhance collaboration amongst project actors (Lai et al, 2002; Lai et al., 2003). The system is primarily devised to integrate pieces of information generated at the iterative early design stage in order to provide decision making support in a multi-actors environment where information is archived in heterogeneous sources. Pieces of information in this case are hypothesized as information chunks each of which can be represented as an object while document is represented as the container for the information objects (see Figure 1). Annotate information chuck with metadata is a method to integrate information in a way to make the relationships between the different pieces of information explicit.

Progress meeting is one of the important collaboration activities since the project outset. Notes-taking has been a common practice to record discussion content of a meeting. This prototype system is thus developed to support progress meeting, and meeting minutes is chosen as the medium for integrating information. The prototype may be an alternative method of capturing discussion content of design progress meeting rather than the conventional notes-taking approach. In light of the limited use of conventional meeting minutes, the prototype system is also devised to provide fast and precise semantic search, and to capture the intent and rationale behind decisions made during the early design process. The prototype system is envisioned to fulfil the following tasks:

- To integrate information that is distributed in heterogeneous sources without using one central repository to reduce repetition of workload.
- To capture and store discussion content wherein design rationale and decision intent are intrinsically encompassed.
- To organise the captured information in a way that is both human and machine readable.
- To contextualise the captured information in representation that may improve the human's efficiency to interpret its implicit meaning.

4.2 *The tools used*

The semantic web (Bernes-Lee, 2001) technologies were chosen as the core of the prototype system. In order to fulfil its tasks, the prototype system was built based on an underlying ontology model so that the discussion content is organized in a semantic-based network. Resource



Description Framework and its Schema (RDF(S)), the de facto standards proposed by the industry group W3C(W3C, 2002) are adopted to develop the ontology model of the prototype system mainly due to the availability of several open source RDF(S) tools.

Protégé 2.01, an open source ontology editor is used to develop the form-based user interface of the prototype system. This user interface is to facilitate the system user establishing RDF data file based on the lightweight ontology model, which is written in RDFS. Sesame 1.0 (detail see http://sesame.aidministration.nl), an open source RDF(S) based repository and querying facility is used as the development base of the prototype system. RQL, query language used in Sesame, is also implemented in this prototype system as the means of accessing information in RDF(S) (detail see Broekstra & Kampman, 2001).

4.3 *The Implementation*

The underlying ontology model of the prototype system consists of few modular components, which respectively is ontology, as illustrated in Figure 2. Each of this ontology describes another aspect of interest, for instance the "team-profile ontology" describes the profile of the design team. The modular characteristic with respect to the ontologies network as such is to streamline the prototype's flexibility for future expansion. Each modular component within the ontologies network is accessible through uniquely specified URI (Uniform Resource Identi-

fier). This modular characteristic permits the scattered information including the existent data and the respective ontologies not to be collected under one central repository.

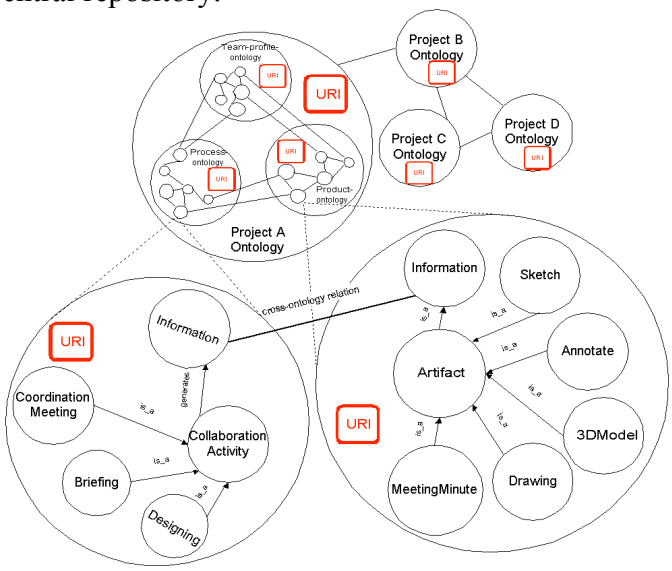


Figure 2. The modular characteristic of ontology network

Form-based user interface is chosen as the mechanism to annotate the content of the meeting minutes with a set of metadata that is pre-defined in the ontology model, including Infoblock Author (denotes the person who raised the discussion issue), Discussion Date, Text (denotes the discussion content), Title (denotes the title of the discussion issue) and so forth. With this set of metadata, the content of the meeting minutes could be semantically structured, and become readable to the machine and easily interpretable to the human. As shown in Figure 3, for instance: <has_action_taker>Joe Young</has_action_taker> illustrates that "Joe Young would take the action" on "something" that was discussed in the meeting. The form filling user interface is chosen because form-filling has been a familiar activity for most computer users. The filled-up form represents a dynamic meeting minute with all of the annotated information populates in the RDF(S) based repository. The annotated information can be stored separately from its corresponding ontology. Queries can be established by system user to initiate semantic search (for detail please see Lai et al, 2003). The searched result, which is a list of URIs, is accessible by the system user provided that all of the relevant repositories are connected to the Internet.

5 DISCUSSION

The prototype is to test if ontology-driven approach is contributable in collaborative design knowledge management. At this stage, the reasoning structure of the prototype system is being evaluated by populating instances to the underlying ontology model via the form-filling user interface. The interim findings has identified that the current metadata initiatives are insufficient for the prototype to fulfil all of its tasks. These initiatives focus on the encoding of primary content attributes of resources (e.g. documents, datasets, etc), such as author, date, location ID, and so forth, with the purpose merely to improve information retrieval and interoperability. In order to fulfil all of its tasks, the prototype system is devised one step forward to take the challenge claimed by Goel (Goel, 1995), which is to provide possible means for analyzing the contents from group discussions so that the idea flow can be traced.

Ideas with respect to various issues were generated, shared, and discussed during design progress meeting. These ideas comprise newly defined or existing design problems, propositions to solutions, as well as the solutions themselves. The relationship between these ideas was implicitly written as plain-text messages in the conventional meeting minutes. The implicit relationship between these ideas can be made explicit by contextualizing them with reference to the propositions given by Shum et al. (2002) so that the flow of ideas can be traced:

1 The intellectual lineage of ideas, for instance, where has this idea come from, is this idea a problem or proposition, has this problem been solved, are there any precedent cases?
2 The impact of ideas, for instance, what was the impact of this proposition to its problem and to other proposition?
3 Inconsistencies, for example, did the solution gain unanimous agreement from the project team, what was the reason given as opposition?

Contextualizing information in such a way is similar to overlaying interpretation of contents explicitly based on the semantic network derived from the underlying ontology model. As shown in Figure 4, the contextual map, which is part of the prototype system devised for this purpose, allows system users to model the semantic relationships between information graphically by binding the different sets of annotated ideas with context dependent relations (or properties as defined in RDF(S)). For detailed explanation, please see Lai (2004). With reference to Figure 4, <*Agreement*> and <*Problem*> were two of the examples of metadata used in the prototype system to annotate the content of information, and *solved_by* was the example of relations used to disclose the semantic relationships between information. Briefly, disclosing the semantic relationships between information graphically as illustrated in Figure 4 may reduce the time users will spend to digest the non-relevant information and therefore enable the users to manage information of interest more efficiently.

At this stage, metadata used in the prototype was arbitrarily defined based on its meaning in natural language. Vocabularies were chosen on the basis of their expressive semantic in describing collaborative design process. Use of arbitrary metadata is a pitfall that may hinder effective information interoperability as already be identified in the project web system (see Section 3.6)

6 CONCLUSION & FUTURE WORK

Design rationale and decision intent are intrinsically embodied in the discussion content of the design progress meeting. Discussion content has been conventionally captured in meeting minutes simply as a piece of plain-text document that is circulated amongst the project stakeholders. Sources of design information that is referred to during discussion are usually specified in this plain-text record. By making use of the technologies of Semantic Web and ontologies, the conventional meeting minutes is envisioned upgradeable to a dynamic and semantically structured medium. The implication is that this medium may handle the mass quantity of design information effectively by eliminating repetition of workload for uploading purposes. The dynamic meeting minutes may also allow the design intents be explicitly presented instead of implicitly described in plain-text records. This envisioned system will be a medium for project team actors to manipulate (store, index and retrieve) knowledge as well as the corresponding meta knowledge effectively.

Further research will examine the possibility to incorporate the standards for metadata proposed by the ISO/IEC 82045-5 in the ontology modeling process. This consideration may avoid repeating the same imperfection of project web while offering a coordinated strategy for better mapping of metadata between different knowledge management systems. A promising standard may adapt the prototype knowledge management system not only sufficient to serve the early design stage but also the whole building life cycle.

REFERENCES

Archer G.O., Futcher K. & McMahon M.A. 1997. IT support for construction re-engineering. Robin Drogemuller (ed). CIB Proceedings. Publication 208. JamesCook University of North Queensland.

Bernes-Lee, T., Handler, J., Lassila, O. 2001. The Semantic Web A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. Scientific American. 284(5): 34-43.

Broekstra, J. & Kampman, A. 2001. Query Language Definition, On-To-Knowledge Project Deliverable 9. Project EU-IST On-To-Knowledge IST-1999-101.

Brüggemann, B. M., Holz, K. -P., and Molkenthin, F. 2000. Semantic documentation in engineering. Proceedings. of the ICCCBE-VIII, Palo Alto, CA: 828-835.

Davis, J., Fensel, D. and Van Harmelen, F. (eds). 2003, Towards the semantic web: ontology-driven knowledge management, John Wiley & Son Ltd.

Fischer, M., Kunz, J. 1995 The circle: architecture for integrating software. Journal of Computing in Civil Engineering, 9(2): 122-133.

Fruchter, R. 2002. Metaphors for knowledge capture, sharing and reuse. in Turk & Scherer (eds.). Proc of eWork and eBusiness in Architecture, Engineering and Construction, ECPPM Conference 2002: 17-26.

Goel, V.: 1995, Sketches of Thought, Cambridge, MA: MIT Press.

Kosovac, B., Froese, T., Vanier, D. 2000. Integrating heterogeneous data representations in model-based AEC/FM systems. Proceedings of CIT 2000. Reykjavik. Iceland. 1: 556-566.

Lai, Y-C, Carlsen, M., Christiansson, P., Svidt, K. 2003. Semantic-web supported knowledge management system: An approach to enhance collaborative building design. in Proc of ASCE Nashville 4th Joint Symposium on IT in Civil Engineering, Nashville: 15-16 November.

Lai, Y-C. 2004. Contribution of semantic web to collaborative design. In Proc of CAADRIA2004. Seoul: 91-106.

Maher, M.L., Simoff, S.J. 1998 Ontology-based multimedia data mining for design information retrieval. in Proc. of Computing in Civil Engineering, ASCE: 212-223.

Mesquita, M.J.M., Fabricio, M.M., Melhado, S.B. 2002. E Concurrent engineering in construction: studies of brief-design integration. Proc IGLC-10. Gramado Brazil.

NAEH (The Danish National Agency for Enterprise and Housing). 2004a., Analysis B. Projektwebkonsortiet: Bygherrekrav-projektweb Delrapport B-Analyse af Projektweb. Project Web Konsortium. Copenhagen.

NAEH (The Danish National Agency for Enterprise and Housing). 2004b, Analysis C. Projektwebkonsortiet: Bygherrekrav-projektweb Delrapport C-Datanalysen. Project Web Konsortium. Copenhagen.

Nonaka, I., Takeuchi, H. 1995. The knowledge creating company, Oxford University Press. Oxford.

Shum, S.J.B., Uren, V., Li, GM, Domingue, J., Motta, E. 2002. Visualising Internetworked Argumentation. in P.A. Kirschner, S.J.B. Shum and C.S. Carr (eds.). Visualising Argumentation: Software Tools for Collaborative and Educational Sense-Making. Press. Springer-Verlag. London: 185-204.

W3C. 2002. RDF Vocabulary Description Language 1.0: RDF Schema. W3C Working Draft 30 April 2002. http://www.w3.org/1999/02/22-rdf-syntax-ns#

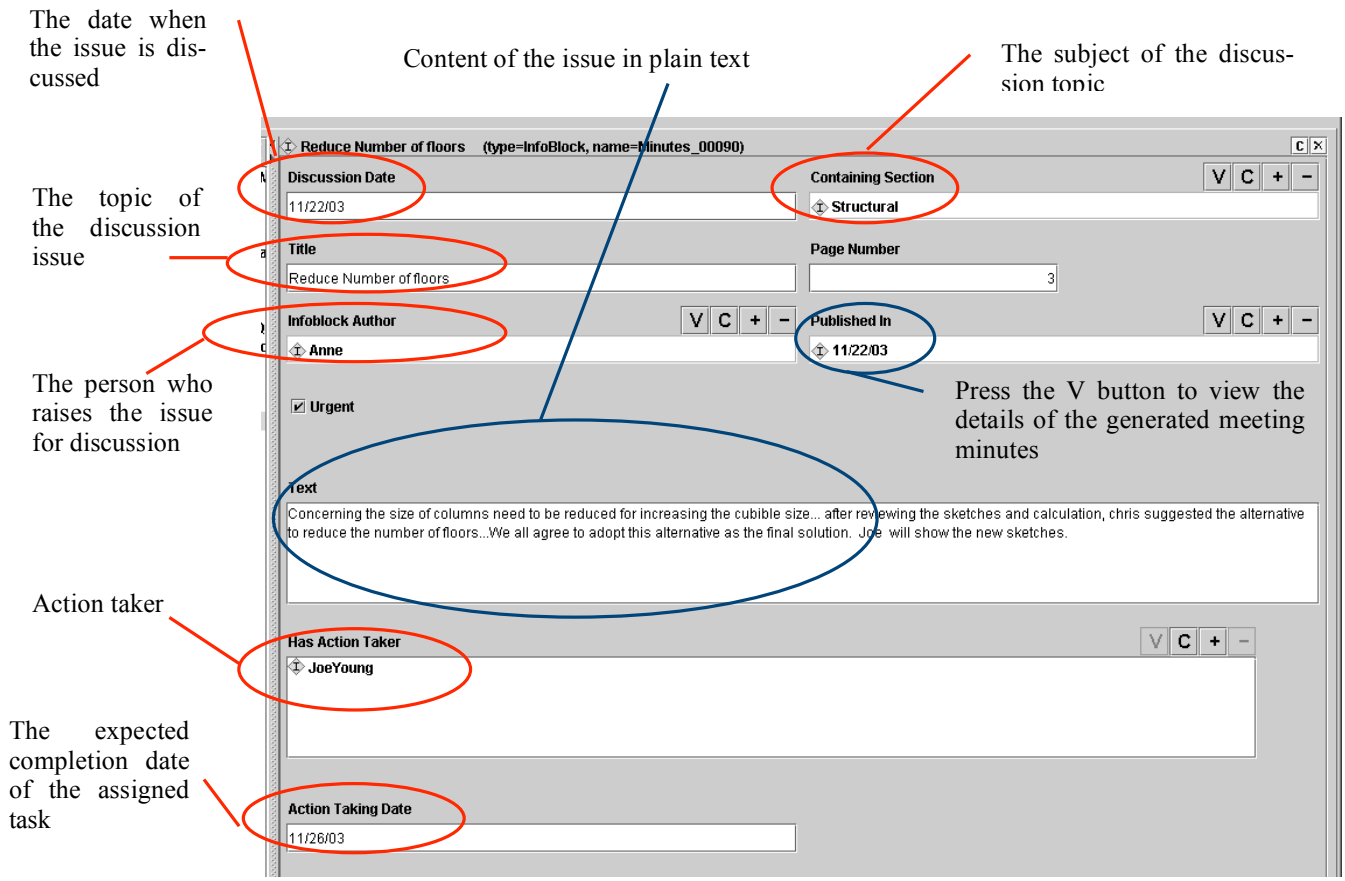


Figure 3. The form filling user interface of the prototype

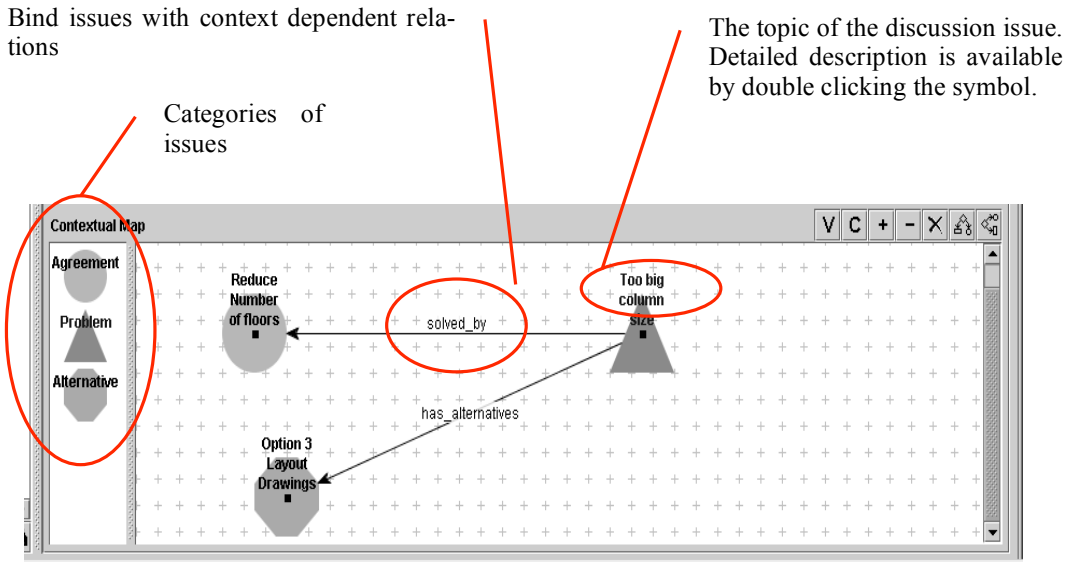


Figure 4. The contextual map feature supported by the prototype