INTERNATIONAL CONFERENCE ON ENGINEERING DESIGN, ICED'09
24 - 27 AUGUST 2009, STANFORD UNIVERSITY, STANFORD, CA, USA

DESIGNWEBS: INTERACTIVE ORGANIZATIONAL MEMORY

Sharad V. Oberoi[1] and Susan Finger[1]
(1) Carnegie Mellon University, USA

ABSTRACT
Knowledge generated during the design process frequently goes uncaptured, and when it is captured, it is usually poorly organized and buried in obscure documents. Effective capture of both semantic knowledge and episodic knowledge can have many benefits for both student and professional design teams. In this paper, we describe DesignWebs, which are dynamic, navigable networks of the documents and conversations created during the design process.

Keywords: Design knowledge, organizational memory

1 INTRODUCTION
Design knowledge is a key asset that companies find difficult to capture, manage, and access [1]. Design knowledge is diffused throughout the organization: in the minds of the design team, in reports and documentation, in sketches, in CAD drawings, and in the artifact itself. Since most design is redesign [2], being able to use prior knowledge effectively is crucial, but knowledge reuse is complicated by the facts that the members of the design team may change, the internal constraints may change, and the external environment almost certainly changes. Therefore, any system for design knowledge capture must take into account not only the diversity of forms in which the knowledge exists, but the diversity of sources and uses.
In this paper, we present the framework for a system that captures the in-process design knowledge that is expressed in written text, both formal and informal. We are in the process of expanding the framework to encompass design team conversations as well as other forms of communication. In creating DesignWebs, our goal is to address the following barriers to effective capture and reuse of design knowledge:

- **Knowledge isolation**: Individuals do not always share their knowledge and experience both because mechanisms do not exist to support sharing and because the local culture often does not reward sharing [3]. “Corporate memory is warehoused in people’s minds in a subjective way” [4]. Such implicit knowledge is the most valuable asset for an organization, but is often communicated informally. According to Pedler et al. [5], collective knowledge becomes “hidden in filing cabinets, in people’s heads, discussed covertly over the coffee machine or, indeed, forgotten.” Orange et al. [3] refer to it as personal, being based on an individual’s perceptions, values and intuition and is a significant part of the knowledge which defines an individual as an expert.
- **Transient team membership:** People frequently move from one team to another, so it is difficult to track who was involved in a decision and who understands the context in which the decision was made and implemented [3]. People may leave taking with them important tacit knowledge, which is lost to the organization [6]. In addition, new members who join a team often are unaware of prior discussions and decisions.
- **Knowledge messiness:** Even when knowledge capture does take place, it is usually limited to formal knowledge stored in a well-defined ontology. Knowledge generated during the design process is often poorly organized and buried in details. Previous research has also shown that although the end products of projects do not capture the contextual nuances, the process can provide such details [7, 8].
- **Lack of knowledge synthesis:** Knowledge is created when people actively reflect on the events represented by the project data. Individuals spend most of their time planning and acting

and much less on observation and reflection and even less on justification of their actions [9]. Reflection is not an individual process, it can be seen as a social process involving a process of reconstructing meaning, which contributes to organizational learning [9,10]. In addition, a lack of design reflection can lead to inefficient processes being perpetuated within organizations [11].

While our framework, called DesignWebs, may aid designers, DesignWebs can also help us as researchers to understand the nature of design knowledge and how designer teams generate and refine it. As team members interact with a DesignWeb, they produce an organization of knowledge that can serve not only as a resource for subsequent design teams, but can also as the foundation of an assessment of how effectively a team has engaged in knowledge access, building and sharing.

In our prior work, we have developed and deployed a system that captures design conversations and documents for project teams [12, 13]. We have captured the process of over 75 student design teams and over 50 research teams. The data from one of these semester-long design projects is used as the basis for the DesignWeb presented in this paper. The data includes chat-like conversations among the team members, successive drafts of required reports and presentations, links to external web sites, as well as uploaded reference documents. Using recent advances made in machine learning and language technologies, we are creating DesignWebs from this data.

2 DESIGNWEBS: A REPRESENTATION OF GROUP KNOWLEDGE

2.1 Visionary scenario

Each semester in an undergraduate project course, teams of students work on the preliminary design of campus-based infrastructure projects. One year, one team of students worked on redesigning the main road through campus, in order to improve pedestrian safety and campus visibility to the community. During their project, they uncovered several design problems with the entrance to the university gymnasium, which mixes pedestrian traffic with delivery trucks, short-term parking, live parking and long-term garage parking; however, redesigning the entrance was beyond the scope of their project. The following year, a new team of students took on the project of redesigning the university gymnasium entrance. Their goal was to improve pedestrian safety, improve the flow of traffic, and improve wayfinding for visitors.

Even though the students were able to access the information from the previous year, they made little use of it because they found it easier to reconstruct the information than to find it within the old reports and presentations. Suppose instead the students had a DesignWeb from the previous year. This web would have included all the documents and presentations generated by the previous year's team, all the documents and websites they referenced and their internal discussions. Most of the DesignWeb would have been about the redesign of main road; however, they could have navigated to the portion of the DesignWeb about the gymnasium entrance. The students would have found a sketch of the redesign from the previous year, some discussions about the problems with the entrance, including a discussion between two students on pedestrian safety. Navigating through the DesignWeb, this discussion would have led the students to the research papers on traffic calming and pedestrian safety that the previous year's students had found. In the DesignWeb, the students would have found information on raised crosswalks and information on predicating pedestrian paths. They would be able to incorporate these papers and discussions into their own DesignWeb. Later in the semester, as the students were starting to gather data for their project, they would have found that not only had last year's group had done traffic counts for the traffic entering and exiting the driveway, they also had a complete set of instructions and templates on taking traffic counts which had been provided by an alumnus working at a local consulting company. Having the previous year's traffic counts would have reduced their data collection requirements and contacting the local alumnus would have given them access to traffic modeling software to help them evaluate their proposed design.

2.2 Approach

A DesignWeb is a dynamic, navigable network of the salient words used during a design process. These words are extracted from the documents and conversations created by a design team. Because the documents often cover many topics, they are divided into topic-sized segments. The salient words – or topics – of each segment become the nodes in the network. The links in the networks show the strength of the co-occurrence of the topics across the segments. Because the documents often have many versions, each segment may have a sequence of versions. The newest segment is always used to create the network, but all versions are available if needed. In addition, because the document – and the artifact – have an inherent structure, the nodes in the network can be expanded and contracted depending on the current focus of the user. Each time a new document becomes available, it is added to the DesignWeb so that over time, the structure of the web will change reflecting the current status of the design. The following section summarizes the research in organizational memory and language technologies that we are using to create the DesignWebs.

3 BACKGROUND

3.1 Organizational memory

Researchers in organizational science and information technologies have done a significant amount of work in capturing collective memories that are created by teams of designers working on an artifact. The notion that memories exist at the group level was first suggested by Halbwachs [14]. What distinguishes collective memory from a historical account is that the latter is objective, while the former is influenced by the context of the situation and its social aspects. Roth et al. [15] have shown that externalizing an organization's experiential knowledge can help future teams be successful with lower transactional costs. Effective capture of organizational memory (also known as corporate memory) can have many advantages for institutions: organizational learning, better knowledge reuse, lower transactional costs and lower resistance to decisions [16, 17].

Organizational knowledge can be divided into semantic knowledge and episodic knowledge, where semantic knowledge in a domain is captured through building consistent and clear ontologies for the concepts in the domain, while the episodic knowledge is observed in contextually situated problems and their successful resolutions [18]. This episodic knowledge is usually stored as project memory that "captures, retains, and indexes project information so that people external to the project can use it later" [19].

Organizational learning in particular has been found to be more effective when organizations store information in shared repositories that allow individuals to interpret them [20]. However, group memory-based systems have been found to suffer from problems related to individuals' differing personal preferences, making navigation of the documents and conversations a tedious task [7].

Although a number of different approaches to creating organizational memory have been suggested in information systems and organizational science [21], most of them are restricted to particular domains or have a high overhead, making them difficult to scale and maintain over time. Researchers have looked at developing organizational memory for specific domains using ontology-based taxonomies, where "entities (such as actors, processes and products) are interlinked to represent the essence of the knowledge in the domain" [4]. While having taxonomy of a subject domain makes it easier to classify entities, its rigid structure makes it difficult to accommodate different points of view. For example, the minutes from a meeting can fall under a number of categories and hence may either be filed under different categories or be segmented to correspond to a single theme. This approach decontextualizes the discussion and an individual who may not have attended the meeting finds it difficult to understand the rationale (and history) behind the decisions made at the meeting [22].

Kwan et al. [23] presented a review of knowledge storing systems, including knowledge repositories, process memory systems and organizational memory information systems. They attributed the lack of adoption of knowledge repositories to four main factors: 1) the extra effort expected of individuals to document their activities is perceived as having no immediate benefit; 2) the documentation is usually after-the-task and so any unsuccessful approaches are not documented; 3) the creators of the knowledge and the eventual users may have different backgrounds and mental models of the knowledge structures leading to difficulties in learning; 4) the emphasis in most organizational memory systems has been on the content aspect (including the final drafts of reports, presentations and memos) with little significance given to the process and context.

The objectives of knowledge management projects have been categorized as [24, 25]: 1) to create knowledge repositories, which store both knowledge and information (classified as external knowledge, structured internal knowledge, informal, internal or tacit knowledge); 2) to improve knowledge access or to facilitate its transfer among individuals; 3) to enhance the knowledge environment, to make it conducive to more effective knowledge creation, transfer and use; 4) to manage knowledge as an asset, and to recognize the value of knowledge to an organization. The proposed DesignWebs will fall under categories 1, 2 and 3, as they are expected to create knowledge repositories by assimilating information scattered in design discussions and documents and making it more accessible to the design team.

Konda et al. [26, 27] viewed shared memory in design teams as the embodiment of both context and shared meaning. Their *n*-dim project categorized shared memory as vertical and horizontal shared memory and used a task-level view for configuring and managing the design process and displays the user's existing design context. Davis et al. [28] used the *n*-dim information modeling tools to create an on-line system that supported integration problems and problem-solving by large and globally distributed design teams working on product development.

The Issue Based Information Systems (IBIS) approach [29, 30] has been widely accepted as a model for capturing design decisions. It is based on the principle that the design process for complex problems is fundamentally a conversation among the stakeholders (e.g., designers, users, implementers, etc.) in which they bring their respective expertise and viewpoints to the resolution of design issues. In IBIS any problem, concern, or question can be an issue and may require argumentation for the design to proceed. However, IBIS-based systems have not been widely adopted. Based on a preliminary analysis of the captured design conversations, we believe that one reason for their lack of adoption is that utterances in design conversations are difficult to classify using the IBIS scheme; that is, design arguments and decisions are diffused throughout design conversations and frequently do not occur at a particular point in time. This hypothesis will be explored in future work.

3.2 Machine Learning and Language Technologies

Advances in machine learning and language technologies provide new opportunities for assimilating and presenting the information contained in documents. These methods include information fusion from multiple text sources [31], incremental hierarchical clustering of text documents [32], clustering based text segmentation [33, 34], co-word analysis [35,36,37], latent semantic analysis [38], document fragment retrieval [39], and interactive navigation mechanisms [40].

Topic Segmentation

Work on automatic topic segmentation can be broadly classified into two types: 1) lexical cohesion models, and 2) content-oriented models. In lexical cohesion models, the segmentation of text is guided primarily by the distribution of terms. So the lexical co-occurrence of thematically-related or synonymous terms indicates continuity in topic and the introduction of new vocabulary refers to a new topic, implying a boundary between the two. Algorithms based on this approach include TextTiling [41] and Latent Semantic Analysis (LSA) [42, 43, 44]. In content-oriented models, the re-occurrence of topic patterns over multiple thematically similar discourses is evaluated. Algorithms based on this approach use hidden Markov models in which states correspond to topics and state transition probabilities correspond to topic shifts [45].

Summarization

Recent work on text summarization of scholarly articles includes using lexical cues to analyze the functional structure of technical papers and using the structure for document retrieval [46], multi-paper summarization using reference information [47]. Cross-document Structure Theory is used for multi-document summarization [48]. This method takes into account the rhetorical structure of clusters of related textual documents and creates taxonomy of cross-document relationships. It also considers user preferences for summary length, information provenance, cross-source agreement, and chronological ordering of facts for creating the summaries.

Cluster Analysis

Cluster analysis is used to group text segments with the goal of maximizing intra-group similarity and minimizing inter-group similarity [49]. Researchers have used dimensionality reduction as a clustering technique to derive useful representations of high dimensional data using a range of techniques: Eigenvalue/Eigenvector decomposition, Factor Analysis, Multidimensional Scaling, Pathfinder Network Scaling, and Self-Organizing Maps. Vector Space Model (VSM) was initially used as a framework for storing, analyzing, and structuring documents originally for information retrieval [48]. VSM, however, suffers from the vocabulary mismatch problem and cannot distinguish between words used in different contexts across documents. To solve this limitation, alternative techniques such as LSA, Lexical Chaining and automatic discovery of vocabulary and thesauri [50] are used.

Co-word Analysis

Co-word analysis uses content analysis to map the strength of association between keywords in textual data [51]. The basic principle of co-word analysis is that it reduces a "space of descriptors (or keywords) to a set of network graphs." These graphs do not display data like other statistical graphs, but construct multiple networks that highlight associations between keywords, and where associations between networks are possible [52].

TouchGraph Navigation Mechanism

TouchGraph is an open source Java environment for the creation and navigation of interactive network graphs [40, 53, 54]. TouchGraph offers several desirable features for visualizing networks, such as high level of interactivity, fast rendering, pan and zoom capability, and locality control [54, 55].

4 CREATING DESIGNWEBS

Various representations have been used in previously published work for characterizing Problem based Learning. Frederiksen used a representation that was modeled on semantic webs to encode how ideas contributed by individual group members in a collaborative discussion build on one another [56]. Frederiksen's work used transactivity, which is the extent to which participants address their contributions directly to the prior contributions of others, as an important aspect of successful collaborative learning dialogue. Frederiksen's analysis was conducted painstakingly by hand, which is only practical on a small scale. In contrast, we propose to automate the analysis of the interaction data we collect.

While it may be beyond the state-of-the-art in automatic language processing to construct a semantic map at the level of detail found in Frederiksen's work, we can approximate the analysis he has conducted in two main ways. First, basic building blocks of dialogue structure, including exchanges and topic segments, are important for identifying the basic units within the representation. Next, we must identify connections between those segments, which we propose to do in our work.

4.1 Data Capture

To illustrate DesignWebs, we have analyzed the student interactions in an undergraduate project class. For asynchronous meeting capture between students in this course, a web-based, asynchronous collaboration tool known as the Kiva (http:/thekiva.org) [13] has been used. The core interaction of the Kiva combines aspects of both email and bulletin boards to keep threaded discussions intact. Students can post documents, diagrams, conversations, meeting notes, notes to self, task assignments, and so on. The discussion pages are designed to feel like a chat session in which students respond easily to one another. Typical Kivas have many thousands of posts organized into hundreds of threads. The example class had 41 students who created over 500 topic threads, each with an average of approximately 10 posts per topic with more than 1000 files posted by the students.

4.2 DesignWeb creation

Seeding the DesignWeb with Research Papers

At the beginning of the semester, the faculty member seeds the DesignWeb with research papers that she finds relevant to the project. These papers provide the basic vocabulary for the initial DesignWeb. In addition, these research papers provide the students with necessary background information for the project.

Topic Segmentation

Most of the posted documents are progress reports that cover many aspects of the project; however, we need to segment the documents so that each segment is about a single topic. This is necessary both for mapping the structure of the artifact and for targeted information retrieval. We divide the documents into segments, each of which deals with a separate topic. Two complementary strategies can be applied: 1) use existing topic segmentation algorithms to divide documents into meaningful segments based on lexical cohesion, with linguistic indicators such as term co-occurrence to signal topic continuity or syntactic features; and 2) use the document structure (i.e., table of contents and internal headings) and any associated metadata. Currently, we are using the TextTiling algorithm based on the first approach [41].

Summarization

Each node in the DesignWeb should give the user a short preview of the document fragments that it points to. This can be done by using extraction-based document summarization techniques: picking the important sentences from the text based on position, content and length information and then combining them. Maximum Marginal Relevancy can be used to reduce redundancy. Our initial experience with extraction-based approaches suggests that they do not work well for document fragments that are short (2-3 sentences). We are currently working to improve the summarization algorithms.

Multi-document summarization

To enable users to focus on a specific subgroup of documents, DesignWebs provide the option of filtering results by sub-topic, time or author and even by version. Research has been done to identify cluster themes from documents using unsupervised clustering techniques. One possible approach is to generate the label of each sub-topic (cluster) automatically by picking the top-ranking k-terms with largest scores within the cluster. Moreover, each cluster can be described by a short snippet that covers the main idea of multiple documents in the cluster, using multi-document summarization techniques.

Clustering

Cluster analysis can be used to group similar objects together, and to determine category boundaries and labels. We use a combination of top-down clustering (Bisecting K-Means) and bottom-up clustering (Agglomerative Clustering). This hybrid clustering approach leverages the best of both approaches.

Information Visualization

DesignWebs use co-word analysis to identify the association between key-words that occur in documents for the proof-of-concept. We apply dimensionality reduction techniques to represent the n-dimensional data using a small number of salient dimensions and thus to display multivariate data on a two-dimensional surface. Several of these algorithms produce ordination that involves a 2-D or 3-D spatial layout in which similar objects are close to one another. We are evaluating these approaches and comparing the results for showing linkages between different concepts, as they occur in discussions or design documents.

4.3 Example DesignWeb

In the following sequence of figures, the interface is a mockup, but the underlying data is derived using the techniques described above using the data from the class. Figure 1 shows the major clusters of topics when the DesignWeb is opened. The left pane contains a summary of the highest level node and the bottom pane displays a scrollable panel of the document segments that have been used to create the DesignWeb.

The user can rearrange the spatial display of the nodes in the DesignWeb, can click on a node to see its summary, can explore a node to find the document segments it represents and what other topics it contains, can see how topics are connected to other topics, can redisplay the web by author or time, and can search the DesignWeb for words or phrases.

In Figure 2, the student has clicked on the Traffic Count node, which is in the lower right in Figure 1. This action opens the node and shows the topics within it. Figure 2 shows the DesignWeb after the Traffic Count node is opened. The student scrolls through the documents at the bottom of the screen and selects the cluster of document segments that were used to create the summary for the Traffic Count node. This cluster contains date-stamped versions of the documents, but since the student is only interested in the conclusions, she looks just at the final version as shown in Figure 3.

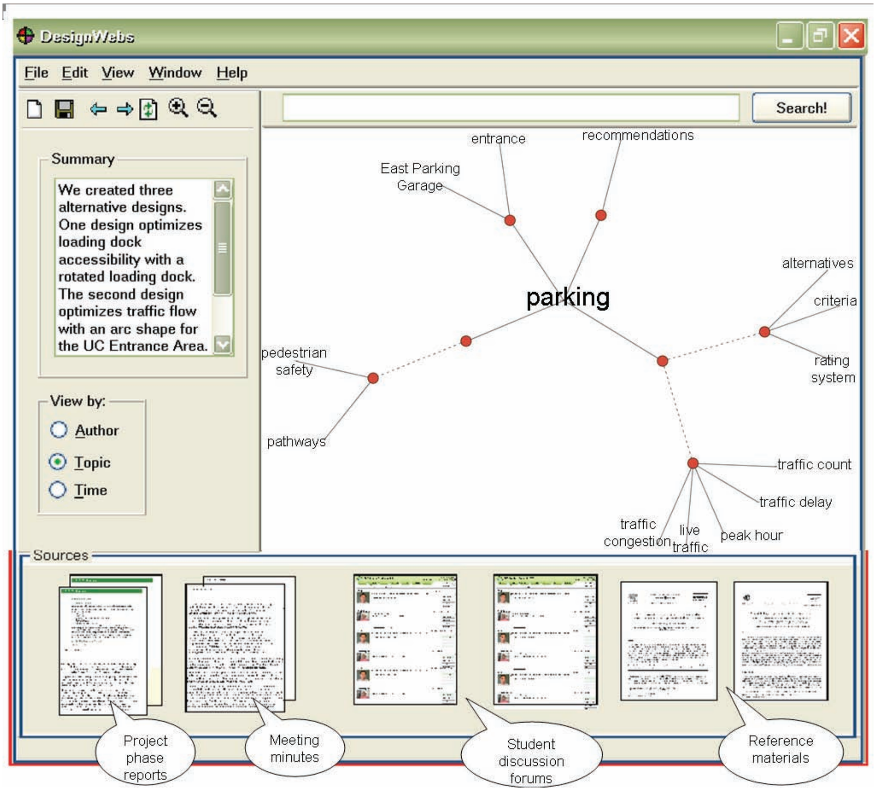

Figure 1. Opening screen of the DesignWeb

5 CONCLUSIONS

DesignWebs provide a robust, dynamic, and automatic method to organize, navigate and synthesize the documents and conversations that occur during a design process. These are expected to support organizational memory by providing a bird's eye-view that is otherwise not possible due to information scattered in design discussions and documents. The proposed research will have the following contributions:

a) Tools that summarize the evolving content of the documents created by a group and enable the visualization and navigation of the ideas and their connections;
b) An interface and interaction metaphor for integrating and summarizing the team communications to track the emergence of the shared solution.

This paper has presented an overview of the DesignWeb framework. Much work remains to test and refine the algorithms used to segment, summarize, cluster and display the DesignWebs, but the proof of concept presented here shows the potential of this approach.

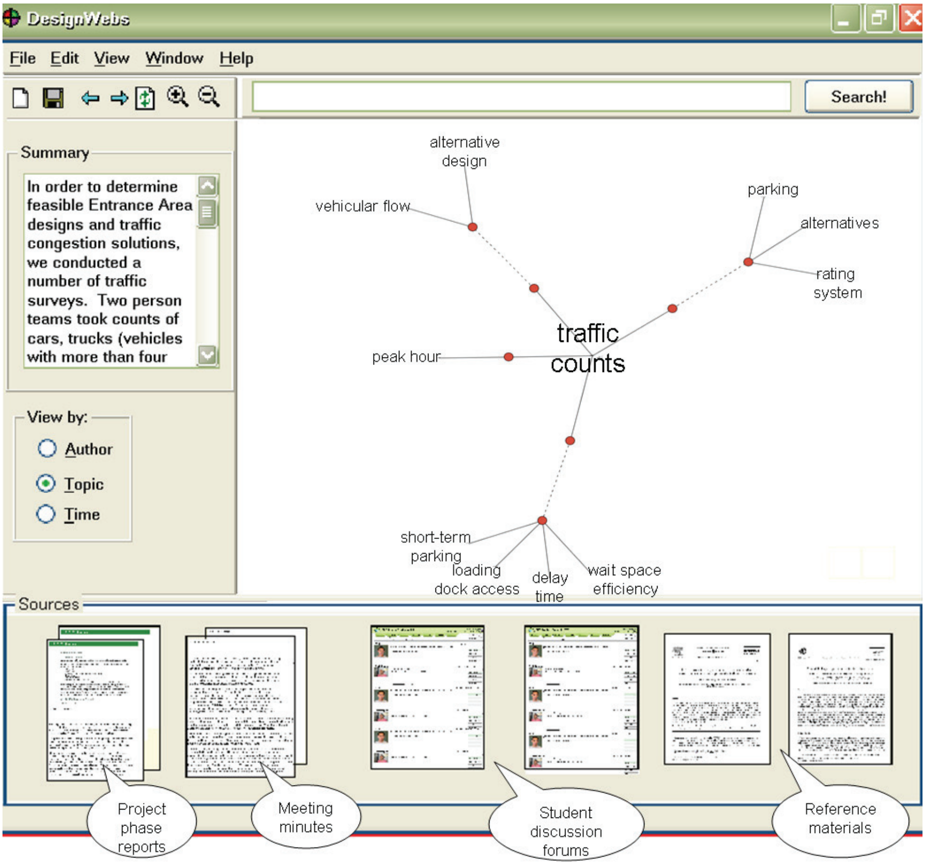

Figure 2. Traffic Counts DesignWeb

REFERENCES

[1] Hult, G.T., An integration of thoughts on knowledge management, *Decision Sciences*, 34(2), pp. 189-195, 2003.

[2] Leifer, L.J., "Instrumenting the design process," *Proc. ICED'91*, Zurich, pp 301-310, 1991.

[3] Orange, G., Cushman, M., and Burke, A., COLA: a cross organizational learning approach within UK industry, paper presented to *4th International Conference on Networking entities (Neties '99), Donau-Universität, Austria,* 18-19 March, 1999.

[4] El-Diraby, T. E., and Zhang, J., A semantic framework to support corporate memory management in building construction, *Automation in Construction*, 15, pp. 504–521, 2006.

[5] Pedler M., Burgoyne J., and Boydell T., *The Learning Company: A Strategy for Sustainable Development*, 1996. (McGraw-Hill)

[6] Argyris, C., and Schön, D. A., *Theory in Practice: Increasing Professional Effectiveness*, 1978. (Jossey-Bass, San Francisco, CA)

[7] Berlin, L. M., Jeffries, R., O'Day, V. L., Paepcke, A., and Wharton, C., Where did you put it? Issues in the design and use of a group memory, *Proceedings of the INTERACT '93 and CHI '93 conference on Human factors in computing systems*, Amsterdam, The Netherlands, pp. 23-30, 1993.

[8] Sim, S. K., and Duffy, A. H. B., Evolving a model of learning in design, *Research in Engineering Design*,15, pp. 40–61, 2004.

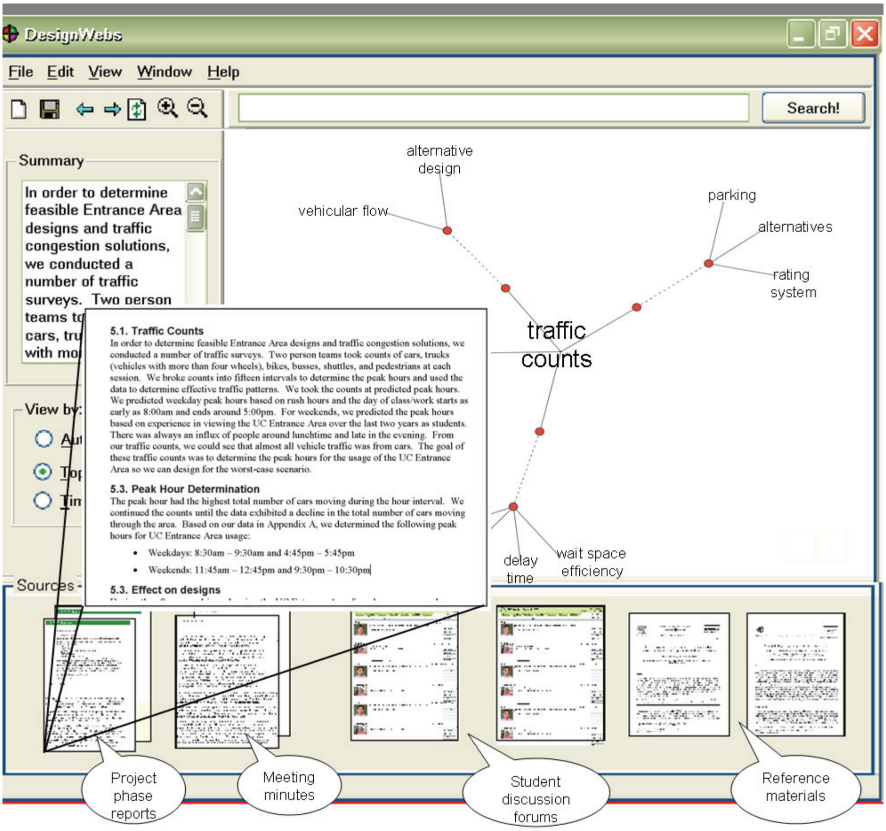

Figure 3. Viewing a document fragment in a DesignWeb

[9] Day, C., Reflection: a necessary but not sufficient condition for professional development, *British Educational Research Journal*, 19(1), pp. 83-93, 1993.

[10] Elton, L. and Pope, M., Research supervision - the value of collegiality, in O. Zuber-Skerritt (ed) *Starting Research - Supervision and Training*, 1989. (The Tertiary Education Institute, Brisbane)

[11] Subrahmanian, E., Classification for communication and communication for classification in engineering design, In: *Workshop on Classification Schemes in Cooperative Work at CSCW2000*, eds. Schmidt, K., Simone, C., Star S. L., Philadelphia, 2000.

[12] Finger, S., Gelman, D., Fay, A., and Szczerban, M., Assessing collaborative learning in engineering design, *International Journal of Engineering Education*, 22(3), pp. 636-644, 2006a.

[13] Finger, S., Gelman, D., Fay, A., Szczerban, M., Smailagic, A., and Siewiorek, D.P., Supporting collaborative learning in engineering design, *International Journal of Expert Systems and Applications*, 31(4), November, pp. 734-741, 2006b.

[14] Halbwachs, M., *Collective Memory*, 1950 (Harper and Row, New York).

[15] Roth, G., and Kleiner, A., Developing organizational memory through learning histories, *Organizational Dynamics*, 27, pp. 43–59, 1998.

[16] Hedberg, B., How organizations learn and unlearn, in: P. Nystrom, W. Starbuck (Eds.), *Handbook of Organizational Design*, pp. 1-27, 1981 (Oxford Univ. Press, New York).

[17] Walsh, J.P., and Ungson, G.R., Organizational memory, *Academy of Management Review*, 16 (1), pp. 57–91, 1991.

[18] Stein, E.W., and Zwass, V., Actualizing organizational memory with information systems, *Information Systems Research*, 6 (2), pp. 85–117, 1995.

[19] Weiser, M., and Morrison, J., Project memory: information management for project teams,

Journal of Management Information Systems, 14 (4), pp. 149–166, 1998.
[20] Hine, M.J., and Goul, M., The design, development, and validation of a knowledge-based organizational learning support system, *Journal of Management Information Systems*, 15 (2), pp. 119–152, 1998.
[21] Paul, S., Haseman, W.D. , and Ramamurthy, K., Collective memory support and cognitive-conflict group decision-making: an experimental investigation, *Decision Support Systems*, 36, pp. 261–281, 2004.
[22] Watts, S., Adoption of mediated knowledge in organizations: source credibility and information usefulness, Doctoral dissertation, Boston University, 1998.
[23] Kwan, M., and Balasubramanian, P., KnowledgeScope: managing knowledge in context, *Decision Support Systems*, 35, pp. 467– 486, 2003.
[24] Davenport, T., DeLong, D. and Beers, M., Successful knowledge management projects, *Sloan Management Review*, 39(2), pp 43-57, 1998.
[25] Metaxiotis K., Ergazakis K., and Psarras J., Exploring the world of Knowledge Management: agreements and disagreements in the academic/practitioner community, *Journal of Knowledge Management*, 9(2), pp 6-18, 2005.
[26] Davis, J. G., Subrahmanian, E., Konda, S. L., Granger, H., Collins, M., and Westerberg, A. W., Creating shared information spaces for collaborative engineering design, *Information Systems Frontier*, *3*(3), pp. 377–392, 2001.
[27] Konda, S., Monarch, I., Sargent, P., and Subrahmanian, E., Shared memory in design: a unifying theme for research and practice, *Research in Engineering Design*, 4, pp. 23-42, 1992.
[28] Westerberg, A. W., and the n-dim group, Distributed and collaborative computeraided environments in process engineering design. In: *Proc. 1st Int. Conf. on Intelligent Systems in Process Engng ISPE'95,* AICHE Symposium Series No. 312, 1996.
[29] Kunz, W., and Rittel, H., Issues as elements of information systems, *Center for Planning and Development Research, Working Paper 131*, Univ. Calif., Berkeley, CA, 1970.
[30] Rittel, H.W.J., and Webber, M.M., Dilemmas in a general theory of planning, *Policy Sciences*, 4, pp 155-169, 1973.
[31] Radev, D. R., Blair-goldensohn, S., and Zhang, Z., Experiments in single and multidocument summarization using MEAD, In *First Document Understanding Conference*, 2001, Downloaded from http://perun.si.umich.edu/~radev/papers/duc01.ps on January 18, 2008.
[32] Sahoo, N., Callan, J., Krishnan, R., Duncan, G., and Padman, R., Incremental hierarchical clustering of text documents. In: *ACM CIKM*, pp. 357–366, 2006. (ACM, New York)
[33] Lamprier, S., Amghar, T., Levrat, B., and Saubion, F., Seggen: a genetic algorithm for linear text segmentation, In: *Proceedings of IJCAI*, pp. 1647–1652, 2007.
[34] Arguello, J., and Rosé, C. P., Museli: a multi-source evidence integration approach to topic segmentation of spontaneous dialogue, *Proceedings of the Human Language Technology Conference of the North American Chapter of the ACL*, New York, pp. 9-12, 2006.
[35] Callon, M., Courtial, J.P., and Laville, F., Co-word analysis as a tool for describing the network of interactions between basic and technological research: The case of polymer chemistry, *Scientometrics*, 22(1), pp. 155-205, 1991.
[36] Chen, J. Q., Lee, T. E., Zhang, R. and Zhange, Y. J., Systems requirements for organizational learning, *Communications of the ACM*, 46(12), pp. 73-78, 2003.
[37] Chung, Y. M., and Lee, J. Y., A corpus-based approach to comparative evaluation of statistical term association measures, *Journal of the American Society for Information Science and Technology*, 52(4), pp. 283-296, 2001.
[38] Foltz, P.W., Kintsch, W., and Landauer, T., The measurement of textual cohesion with latent semantic analysis, *Discourse Processes*, 25, pp. 285-307, 1998.
[39] Liu, S., McMahon, C. A., Darlington, M.J., Culley, S.J., and Wild, P.J., An approach for document fragment retrieval and its formatting issue in engineering information management, *Lecture Notes in Computer Science 3981,* pp. 279–287, 2006.
[40] Shapiro, A., Method of animating transitions and stabilizing node motion during dynamic graph navigation, US Patent No. 20050180330, TouchGraph LLC, 21 Hearthstone Terr, Livingstone, NJ, US, <http://www.freepatentsonline.come/20050180330.html>; http://sourceforge.net/projects/touchgraph, 2005.
[41] Hearst, M., TextTiling: segmenting text into multi-paragraph subtopic passages, *Computational*

Linguistics, 23(1), pp. 33-64, 1997.
[42] Landauer, T.K., and Dumais, S. T., A solution to Plato's problem: The Latent Semantic Analysis theory of acquisition, induction, and representation of knowledge, *Psychological Review*, 104, pp. 211-40, 1997.
[43] Olney, A., and Cai, Z., An orthonormal basis for topic segmentation of tutorial dialogue, *Proceedings of HLT-EMNLP* 2005, 2005.
[44] Barzilay, R., and Lee, L., Catching the drift: probabilistic content models, with applications to generation and summarization, *Proceedings of HLT-NAACL 2004,* 2004.
[45] Kando, N., Text-level structure of research papers: Implications for text-based information processing systems. In *Proceedings of BCS-IRSG Colloquium*, pp. 68-81, 1997.
[46] Nanba, H., and Okumura, M., Towards multi-paper summarization using reference information. In *Proceedings of IJCAI-99*, pp. 926-931, 1999.
[47] Radev, D. R., Hovy, E., and McKeown, K., Introduction to the special issue on summarization. *Computational Linguistics*, 28(4), pp. 399–408, 2002.
[48] Börner, K., Chen, C., and Boyack, K. W., Visualizing knowledge domains, *Annu. Rev. Inform. Sci.*, **37**, pp. 179–255, 2003.
[49] Salton, G., Yang, C., and Wong, A., A vector space model for automatic indexing, *Communications of the ACM,* 18(11), pp. 613-620, 1975.
[50] Mostafa, J., Quiroga, L. M., and Palakal, M., Filtering medical documents using automated and human classification methods, *Journal of the American Society for Information Science,* 49(14), pp. 1304-1318, 1998.
[51] Krsul, I. V., *Software Vulnerability Analysis*, Ph.D. Thesis, Purdue University, 1998.
[52] Coulter, N. Monarch, I., and Konda, S., Software engineering as seen through its research literature: a study in co-word analysis, *Journal of the American Society for Information Science*, 49(13), pp. 1206-1223, November 1998, 1998.
[53] Alani, H., Tgviztab: an ontology visualisation extension for protégé, *Knowledge Capture 03—Workshop on Visualizing Information in Knowledge Engineering*, ACM, Sanibel Island, FL, pp. 2–7, 2003.
[54] Herman, I. Melancon, G., and Marshall, M.S., Graph visualization and navigation in information visualization: a survey, *IEEE Transactions on Visualisation and Computer Graphics*, 6(1), pp. 24-43, 2000.
[55] Wills, G.J., NicheWorks - interactive visualisation of very large graphs, *Proc. Graph Drawing '97,* Rome, Italy, Springer-Verlag, pp. 403-414, 1997.
[56] Fredriksen, C., Learning to reason through discourse in a problem-based learning group, *Discourse Processes*, 27(2), pp 135-160, 1999.

Contact: Susan Finger
Carnegie Mellon University
Department of Civil & Environmental Engineering
5000 Forbes Avenue
Pittsburgh, PA 15213
USA
Tel: +1 412 268 8828
Fax: +1 412 268 7813
Email: Susan.Finger@cmu.edu
URL: http://www.cs.cmu.edu/~sfinger

Dr. Finger is on the faculties of Civil and Environmental Engineering and Architecture at Carnegie Mellon University. She is the co-editor in chief of the journal Research in Engineering Design. Her research interests include representation languages for design, integration of design and manufacturing concerns, and computer-supported collaborative learning and group cognition.

Sharad V. Oberoi is a PhD Candidate in Civil and Environmental Engineering at Carnegie Mellon University. His research concerns the role of language in design, especially in tracking the evolution of ideas within design teams. He is also interested in creating analytical methods based on the theoretical findings of group cognition to study learning in engineering project teams.