

**DETC2001/DTM-21713**

**IDENTIFYING SHARED UNDERSTANDING IN DESIGN USING DOCUMENT  
ANALYSIS**

**Andrew Hill, Shuang Song, Andy  
Dong and Alice Agogino**

University of California, Berkeley  
Department of Mechanical Engineering  
5136 Etcheverry Hall  
Berkeley, CA 94720-1740

**ABSTRACT**

Design as a social activity characterized by information exchange, compromise and negotiation frames much of our understanding of the design process. At the heart of this social activity is the development of a shared understanding of the design problem. The design stakeholders jointly form a shared understanding through a process of defining the problem, exploring the space of solutions and searching for information. A shared understanding is a critical element in successful, collaborative design. This paper describes a formal approach for identifying shared understanding in design by analyzing the documentation. By mining the documentation for signatures of a shared understanding, specifically a common frame of reference and a similar voice, a shared language and vocabulary of the design could emerge. Information management tools built around the shared understanding would be more effective in directing and alerting the design team of relevant information.

**INTRODUCTION**

Design is a word that has been defined in many different ways. Among its many meanings (e.g., design is a formal process, design as a social activity) there runs a common thread, linking the cognitive processes of a designer (or designers) to the tangible entities that are produced when a designed object is experienced. In a collaborative, team-based engineering design environment, design knowledge is generated, codified, and reflected on in an ongoing process

between the various stakeholders. Learning occurs when the designers synthesize concepts, expressed through drawings and models based on their knowledge, and then reevaluate that knowledge based on what they have just learned. Eventually, a shared understanding of the design emerges when the design team establishes a common frame of reference. Developing the shared understanding relies on the exchange of information and mutual agreements as to the relevance and meaning of that information (Citera et al., 1995).

Within a multi-disciplinary collaborative product development environment, in which teams of engineers from multiple organizations work together supported by information services and a computing infrastructure, design occurs as a social process of reaching a “shared understanding” (Toye et al., 1995) of the design problem, the requirements and the design process itself. Social interaction in the design process is a significant determinant of the success of collaborative design (Bucciarelli, 1994).

Because design engineers will bring with them their own language, jargon and perspectives to the design team, or “design identities” (Kilker, 1999), incompatible viewpoints among design team members may result in ineffective collaborative, sub-optimal decision-making and impaired projects. Identifying when design teams have reached a shared understanding, or not, is not only an important management aid, but also an advancement in understanding how design teams acquire and maintain their collective identity.

Design teams succeed when they pool their resources to negotiate different design perspectives and specialties. When these viewpoints break down due to disagreements, collaboration and design suffer.

As a first step towards identifying a shared understanding among design teams *in situ*, we present a methodology for identifying shared understanding in design documentation at the detailed stage of design. Sociological theories of communication and studies in design communication offer the theoretical underpinnings for our methodology, while the implementation draws upon the computational linguistic techniques of natural language processing and latent semantic analysis. The hypothesis of the research is that mining design communication documents for topical similarity and voice similarity, identifiers of a common frame for reference of the design team, can reveal a shared understanding of the design. The underlying aim of the research then is to discover the terminological patterns in design text as a basis for characterizing a shared understanding in design.

## IDENTIFIERS OF SHARED VISION

A shared understanding of design is necessary for the successful completion of a design project. The shared design understanding often manifests through the use of similar jargon in documentation because vocabulary differences can create communication problems. Designs that are riddled with conflicting visions result in conflict and a stalled design process.

According to some prevailing theories of inter-personal communication (Wertsch, 1991), a shared understanding between communicators is comprised of two components: a topical or contextual component and a voice component. The topical or contextual component is comprised essentially of the topic, for example; component X in assembly Y or the final colors of component X share contextual similarity through component X. Components of voice, in collaborative design, can be as simple as the jargon or specific language used in the design team. On a deeper level voice, particularly the voice of designers operating in a team environment, becomes defined more as the ability of a designer to borrow the shared vision of a design team. In group communication, true collective understanding occurs when the “team is on the same page” as it is popularly called. Getting “on the same page” requires achieving group acceptance of a common set of vocabulary. This is particularly important when designers on the team come from different disciplines or backgrounds. Similarities in voice between designers in a group are critical to progressing through the design process.

Perhaps more important is the concept of a collective voice. Bahktin (Wertsch, 1991) wrote extensively that effective group communication occurs when the group shares a voice, and in fact the intra-member dialogues are merely extensions of this shared voice. When a speaker wishes to address a group or a teammate, the speaker borrows a voice from the collective

group. This collective group voice is dynamic and changes as the team progresses, but effective group communication occurs when all members are able to borrow from and relate to this combined group voice.

Thus, one could argue that a common topic and voice forms the basis of shared understanding. Although topical and voice similarity may not be the only indicators of a shared understanding, they are likely sufficient factors. Research in information use in design (Lloyd, 2000; Baird, 2000) have surfaced that words and phrases used by designers in the design process often capture personal experience and contribute to a wider narrative at the team, project or corporation level. Research in small group interactions and the broader socio-cultural environment in which the interaction occurs offers insight into language and the purpose of communication. The actual conditions of the use of the language relates to forming a shared understanding between the communicators. One theory proposed suggests that a voice, or “speaking personality,” exists in any social setting (Wertsch, 1991). The process of understanding the communication between speakers demands that each participant understand not only the vocabulary of the communication but also the corresponding context in which the communication took place.

These theories offer insight into the function of texts as a means of communication between members of team, and in particular, a design team. The purpose of the design communication is to establish a set of coherent ideas, for the stakeholder to “get a thought across.” As opposed to a merely passive purpose, namely the passive transmission and reception of information, the text serves an active role of generating new meaning. Embedded in design documents is information regarding the design vision of an individual designer, a vision the designer is attempting to communicate and advance. Extracting this vision is the primary objective of this research.

## METHODOLOGY

### General Approach

In order to identify design vision in documents using a text analysis approach, it is first necessary to represent the design documents in a matrix format. The rows of the matrix represent all of the unique words found within the corpus, the columns represent each individual document and the cells represent the frequency of each word as it appears in each document. Singular value decomposition is applied to this matrix and the dimensions of the matrix are reduced to an optimal value. Singular vectors represent each document in the reduced space. In this reduced space, the singular vectors for each document are correlated to each other by measuring the cosine between the vectors. Documents that are closely correlated in this manner can be said to share a similar design vision. The following text explains in detail the methods of analysis.

**Document Representation Method**

To apply our methodology, the corpus of text data needs to be represented as a matrix in which the rows of the matrix correspond to the unique words found in the text and the columns represent each text passage and, in this case, each author (Salton and McGill, 1983). The cells of the matrix represent the frequency that each word occurs in each document. A representation of this matrix is shown in Figure 1.

		<i>Doc<sub>1</sub></i>	<i>Doc<sub>2</sub></i>	□	<i>Doc<sub>n</sub></i>	
X	<i>Word<sub>1</sub></i>	0	0	1	1	
	<i>Word<sub>2</sub></i>	1	0	1	0	
	<i>Word<sub>3</sub></i>	1	1	1	1	
	□	1	2	0	1	
	<i>Word<sub>m</sub></i>	2	0	2	0	

**Figure 1 Representative Word Document Matrix**

The log-entropy weights are then computed for this matrix. Log-entropy weights for the matrix are calculated using the following formula (Deerwester et al., 1990).

$$\frac{\log(freq_{ij} + 1)}{\sum_{i,j} ((\frac{freq_{ij}}{\sum_{i,j} freq_{ij}}) * \log(\frac{freq_{ij}}{\sum_{i,j} freq_{ij}}))} \tag{1}$$

**Identifying shared vision using LSA**

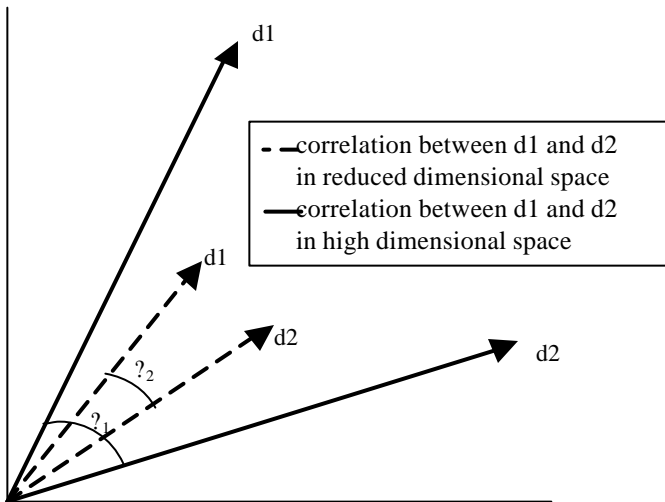
Latent Semantic Analysis (LSA) is a full-text indexing and retrieval method that takes advantage of implicit higher-order structure in the association of terms within documents to detect semantic similarities (Deerwester et al., 1990). In this case, LSA was used to find both contextual similarities between design documents and similarities in the voices of design document authors. LSA was chosen as an analysis tool because of its demonstrated successes in identifying contextual meanings of documents as well as identifying voice of a given documents author (Landauer, et al., 1998). LSA is an analytical tool that extracts the meaning of documents and of words by applying a deep statistical analysis to a large corpus of text and mapping word frequency to a series of documents. The mathematical foundation for LSA lies in Singular Value Decomposition (SVD), a method for reducing the dimensions of a matrix.

The baseline theory suggests that by looking at the entire range of words chosen in a wide variety of texts, patterns will emerge in terms of word choice as well as word and document meaning. LSA is unique in its method to analyze text; in its analysis, there is no consideration of word order or syntax. Furthermore there are no human constructed dictionaries, expert knowledge or domain-specific knowledge bases required. There have been a number of applications for LSA, both practical and

theoretical. LSA has been used with some success to grade college essays automatically (Landauer et al., 1998). The results of this experience show that with proper administration of the LSA the assigned grades from LSA do not markedly differ from those of human graders. Other applications of LSA allow information retrieval to occur more accurately (Landauer et al., 1998). Using LSA to map out the meanings of a series of words, document queries can be run to retrieve contextually similar documents, even when documents contain no common words with the query. The importance and practicality of this is obvious, and it is an improvement over word coincidence based correlation methods.

LSA is able to find similarities in documents that are not always apparent with simple word coincidence based methods. By representing what is by nature a very high dimensional space in an aptly chosen lower dimension, correlations between documents are better able to be determined. For instance take two documents (*d<sub>1</sub>* and *d<sub>2</sub>*) that are very similar in topic, although they use different vocabularies. Comparing these two documents directly using a standard technique such as keyword matching will yield a low correlation because these documents share few words. LSA allows you to compare these documents not only directly against each other, but also with a large corpus of other documents. Similarities in words and documents become more apparent when documents are compared in this larger space. That is to say that the words used in *d<sub>1</sub>* may not be the same as the words used in *d<sub>2</sub>* although how they are used in context would be similar. LSA is able to make these types of comparisons; the ability for LSA to remove the obscuring “noise” makes LSA useful as an analytical tool for discovering the primary conceptual content of documents.

Graphically, the ability to reveal “latent” semantic similarity between documents is illustrated in Figure 2. Suppose that documents *d<sub>1</sub>* and *d<sub>2</sub>* both regard a common subject, but do not share many words. Many information retrieval systems would fail to correlate *d<sub>1</sub>* and *d<sub>2</sub>* together because they compute the relevance between two documents as the cosine between the vector representations of the document (Cooper, 1973a; Cooper 1973b). This means that correlations are merely a measure of the overlap of words between the two documents. Figure 2 shows that in this high dimensional space, the correlation between *d<sub>1</sub>* and *d<sub>2</sub>* is simply the cosine of angle  $\theta_1$ . A large angle between them would indicate low similarity. Applying LSA on a large corpus of text including *d<sub>1</sub>* and *d<sub>2</sub>* would show that when a lower dimensional space is developed the angle between the two documents,  $\theta_2$ , has been reduced and as such the correlation between the documents revealed. This occurs as LSA is able to find similarities in words based upon higher degree similarities, that is if a third document *d<sub>3</sub>* uses several words from both *d<sub>1</sub>* and from *d<sub>2</sub>* LSA will infer the correlation between *d<sub>1</sub>* and *d<sub>2</sub>* through their mutual correlation to *d<sub>3</sub>*.



**Figure 2 Using LSA to Reveal Document Similarity**

SVD operates over the document by term matrix described in Figure 1. SVD takes the original matrix and represents it as the product of three matrices. The first matrix describes the row entries as vectors derived from the orthogonal singular values. The second matrix is a diagonal matrix that represents singular values to ensure that when the other two matrices are multiplied together the result is of the same magnitude as the original matrix. The third matrix is similar to the first, although it represents the original columns of the matrix rather than the rows. It is possible to reconstruct the original matrix without using the entirety of the three derived matrices. All that is required is enough factors to match the smallest dimension of the original matrix. Furthermore, a least squares best-fit approximation of the original matrix is capable of being obtained by reducing the dimensions of the three derived matrices.

By choosing only the first  $n$  dimensions of the three derived matrices, an  $n$ -dimensional projection of the original matrix is obtained. If the first matrix is reduced to three dimensions, only the first three singular vector values will be retained. To correlate documents, the singular vectors for each author are correlated together by measuring their cosine (Deerwester et al., 1990).

It is possible to use LSA to find similarities in documents based on a variety of variables. By thoughtfully choosing how to group documents into the original word-document matrix both contextual similarities and voice similarities, the components of a shared design vision, can be discovered.

**Decoupling voice similarities and topical similarities in documents**

LSA's ability to correlate documents based upon contextual similarities has been well documented. LSA has the ability to correlate documents together by comparing the meaning of the document, not merely word by coincidence. (Landauer et al., 1998) The improved performance allows LSA to accurately

correlate documents together based on content, not merely by word co-occurrence. The advantages of this are obvious; two documents discussing the same design project would not normally be found highly correlated if in one document the project were referred to by project number and the other document referred to the project by name. LSA is able to determine that, in a large corpus of text, the manner in which the project number is used is similar to the manner in which the project name is used and from that the two documents would be identified as being highly correlated. In terms of identifying design vision in a corpus of design documents, it becomes important to find more than simply the contextual correlation between documents. One also needs to identify the voice of the documents author. Experiments show that in performing LSA the information regarding voice shows up and is in fact coupled with the information regarding context. Discovering the voice components of documents within the LSA therefore requires isolating voice components from contextual similarities in those documents.

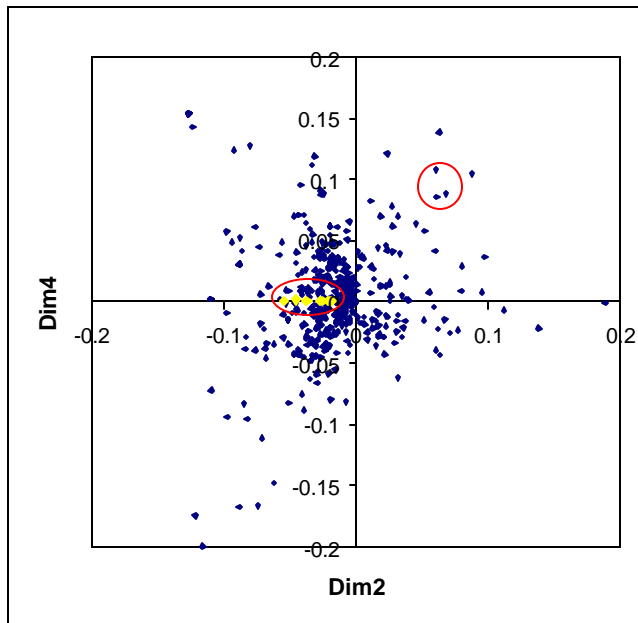
**EXPERIMENT AND RESULTS**

To validate our hypothesis, we applied our algorithm to documents on engineering courseware design from the www.needs.org database. This website hosts an online catalogue of engineering education courseware. Authors submit online courseware to the catalogue for distribution over the Internet. Included with each piece of courseware is a brief description of the courseware. These descriptions serve as the starting point for this experiment. Only those descriptions that were written by the author of the courseware, as opposed to the cataloger or database manager, were included in this study.

These data were chosen for several reasons. First, the descriptions represent a communication of the design vision of the courseware author. The author, for example, may express how a disk drive should be designed. In addition, the descriptions are intended for similar audiences. Finally, the authors exist in a relatively similar space, that is, engineering education. The similarity in context of all of these documents isolates voice variability to the individual author. That is to say, within this confined space each individual author brings an individual voice in, but each author is working within the same contextual space so that the collective voice from which the author would borrow voice from could potentially be the same. Discovering voice in this data set can be reduced to identifying an author for a given document based upon voice similarities. The nature of this data set allows us to validate our hypothesis.

The first component of our hypothesis deals with identifying shared understanding in design documentation through topical similarity. We selected a group of engineering courseware that contained both similar and dissimilar engineering content. If our hypothesis is correct, namely that LSA can discover both contextual and voice similarities, then the LSA analysis should reveal groupings of documents similar

in topic. The results of the LSA analysis are expressed graphically in Figure 3.



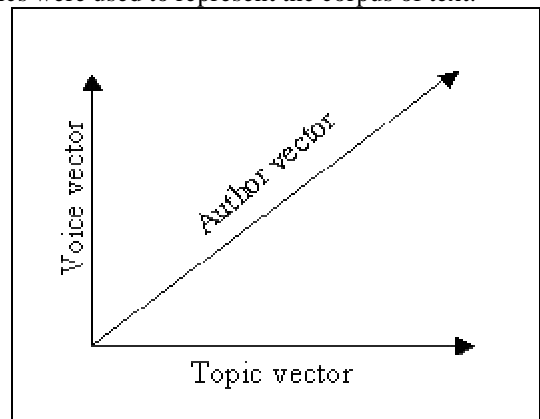
**Figure 3 Similarities in Design Vision Visualized in 2-Dimensional LSA Space**

Figure 3 displays our set of design documents, each represented by a point, as correlated to other design documents in a 2-dimensional LSA space. Note that the two axes of Figure 3 are labeled Dim 2 and Dim 4. As this is a two-dimensional representation of a very high dimensional space, these two axes were chosen simply as a visual representation of the clusters of documents. These two dimensions display a representation of the overall cluster. Dim 2 and Dim 4 represent the second and fourth most significant singular values found in the singular vectors that represent each document in the LSA space. The closer two documents are together in this chart, the closer they share common understanding. It is important to note that while Figure 3 offers an efficient visualization mechanism for what the LSA results are, figure 3 is not an analytic tool. The correlations between documents truly become meaningful when many more dimensions are utilized. To be rigorous, more dimensions are necessary, upwards of 300 dimensions, to correlate documents accurately (Landauer et al., 1998). What can be seen in Figure 3 is a cluster highlighted in the circular region near the origin that consists of 8 documents all dealing with Multimedia Case Studies in Engineering Design. These 8 documents consist of documents dealing with the design of disk drives, printers, automobiles, human powered vehicles and toys. The commonality in the documents is in their focus on the design process and their multimedia nature. It can be seen that in fact these documents share at least a common topical component. In this representation of the LSA space the fact that the 8

documents lie in a horizontal line is a good indicator of higher dimensional correlations. Further more these documents are located near the majority of points in the space. This makes sense because a majority of documents in the www.needs.org database deal with engineering design and analysis topics. The cluster that these documents form is located near the origin in Figure 3. The www.needs.org database catalogs more than just engineering courseware documents. Also included in the database are documents dealing with other types of sciences. For example a cluster of chemistry courseware is located in the top right quadrant of Figure 3. The three documents circled deal with chemistry education resources for teachers. While these documents are correlated tightly with each other they are removed from the core of the database contents. The results from this analysis demonstrate how LSA is able to accurately discern topical similarities in documents as well as sort through documents based upon topical similarities.

However, the above results merely show that documents regarding the same subject matter become highly correlated in the LSA space. These results do not demonstrate what elements of the documents go into forming these correlations. As such these results cannot offer strong evidence that the groupings also indicate voice similarity. These results may only demonstrate that documents that share topical similarities. In this case Engineering Design and Chemistry, appear to be correlated highly in the LSA space. To more closely examine the hypothesis of being able to correlate a voice to an author, one must demonstrate that the topical similarities and voice similarities in documents are independent of each other. If they were not independent, then we would not know if two documents were similar because they were similar in *subject* or if they were similar in *voice*. The latter is the desired result.

In order to isolate similarities in voice from similarities in topic, the original data set was placed into two word document matrices. One matrix contained in its columns each document. The second matrix contained in its columns the documents grouped by topic. The topic grouping was done manually based on expert knowledge of the content. The topics were reflective of the field of study the document was based upon. Eighteen topics were used to represent the corpus of text.



#### Figure 4 Vector Representation of Topic and Voice Components

For the first matrix, we ran LSA to reveal the singular vectors representing each document. This is the standard analysis. We then ran LSA on the second matrix, resulting in the singular vectors representing each topic. We then took the orthogonal projection of the first matrix onto the most significant singular vector from the second matrix to reveal the singular vectors describing each document *independent* of topic. These vectors represent the voice component of the original document vector. Graphically this is shown in Figure 4. Using the resultant singular vectors, we compared their ability to assign an author to a document by calculating the “Average Rank.”

The “Average Rank” indicates the number of authors that the system predicted as potential authors of a document. The lower the rank, the better the prediction, and hence attribution, of voice to author. Out of a possible 110 authors the system was able to rank the correct author in the top 10% as the most likely author of the test document. Figure 4 shows the results of this method. By reducing the LSA space to; 10, 20, 40 and 80 dimensions better results were achieved. The results from these orthogonalized results are compared to results where the author was identified with out isolating the topical components of the document. This means that the singular vectors for each document were compared without first orthogonalizing them with the singular vectors from the topically grouped data set.

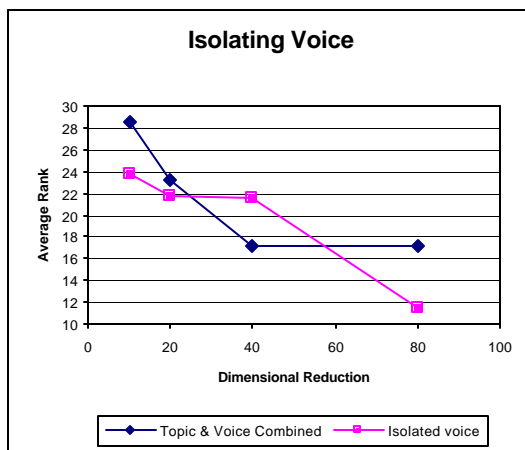


Figure 5 Identifying Voice Independent of Topic

The effects of LSA showed that the voice of an author could be better identified with the isolation from topic, than without. This demonstrates two concepts; one, voice similarity exists within a set of design documents, and two, authors can use a similar voice even when writing different documents. This is significant because a single design team may have a similar voice throughout its projects, even though each project may fall into a different topical category. The improvement in average rank as LSA retains more dimensions demonstrates that higher

dimensional LSA representations allow for better ultimate results. With LSA it is necessary to optimize the number of dimensions to reduce the space too. When this technique is applied to a larger corpus of data, it will become necessary to optimize the number of retained dimensions.

Ultimately many factors may be involved in determining the similarities in these design documents, further research will move towards a more complete understanding of the issues. The results of this work show that similarities between documents based upon topical similarities are able to be ascertained, this research also shows that characteristics exist within a document that allow the author of the document to be ascertained. These characteristics are likely the voice components of the documents author, although further exploration of this is necessary to fully claim that voice similarities are the elements that exist independent of topic.

#### CONCLUSIONS

This paper described a formal methodology for identifying a shared understanding in design by analyzing design documentation. The premise of the paper is that topical similarity and voice similarity are identifiers of a shared frame of reference of the design. Using the computational linguistic tool of latent semantic analysis, engineering courseware design documents written by various authors were analyzed to reveal highly correlated groups of topics and independent voices.

The results demonstrated not only the capability to reveal a shared understanding by analyzing the documentation, but also that by isolating topic from the documents, it is possible to ascertain a common voice. This is an interesting result because it suggests that effective design teams might participate in multiple projects yet communicate with a single voice. Comparing these results with protocol studies of design teams (Ullman, 1988) would be an interesting area for further study.

The extension of this research is to analyze the communication between designers in a design team to verify whether our methodology can also identify a shared understanding in design teams. While these results indicate the ability to find shared understanding at the completion of a design, it would be interesting to understand the evolution of a shared understanding during the full life-cycle of the design.

The implications of this research and future work suggest not only a means for the management of design teams, for example, detecting and diagnosing non-functioning design teams, but also a means for understanding the evolution of information needs in design teams. For example, team meetings might be unproductive if information is brought to bear that is of relevance to only a few individuals. Instead, documents of potential interest to all individuals could be directed to the design team. Some members may use too much jargon in their communication, separating their voice undesirably from the group's voice. The language can be modified or the team educated to get that particular perspective reflected. While one cannot ignore the organizational and social barriers of getting to

a shared understanding among team members, identifying barriers to overcome offers a crucial first step.

For a large organization in which different types of design are occurring in many different areas, being able to determine the design vision of a design team makes possible several potential methods to improve design productivity. With information available regarding the design vision of team members, it becomes possible to group designers together by common vision. This becomes useful in terms of identifying the information needs of a group and then directing useful information to the team. Likewise, providing a convenient means for identifying and classifying information in the personal information of designers offers an attractive means for design information sharing (Court et al., 1997) between team members who share a common interest in topic or voice. On the flip side, a design team in which each member shares too similar a design vision may regress to "groupthink," where the design team is too homogeneous in their thinking to allow the creative process to work (Kilker, 1999).

Many challenges exist in determining what information an engineer wants to see as well as modeling designer information needs (Lowe et al., 1999). The technique shown in this paper for ascertaining topic and voice in design information offers a basis for progressing towards developing solutions for capturing the information needs of design teams and then finding data of interest to them.

Ultimately work in this area may show that as a design team progresses through time, a commonality of design vision can be reached. The premise of the methodology is that the design specifications and solutions as communicated through design documents relate to a shared understanding of the design. This paper has shown the potential for, and the direction of preliminary work in this field. Future investigation will shed more light onto the phenomenon of shared vision and will ultimately lead to a better understanding of the collaborative design process.

## ACKNOWLEDGMENT

The authors would like to acknowledge that the work described in the paper is a result of a research thrust in the Berkeley Expert Systems Laboratory at the University of California at Berkeley. Further more we would like to acknowledge Dr. John Canny whose teachings in Human Centered Computing provided framework for this study.

## REFERENCES

Baird, F., 2000, "An ethnographic study of engineering design teams at Rolls-Royce Aerospace," *Design Studies*, **21**, 333-355.

Bucciarelli, Larry L., 1994, *Designing Engineers*, MIT Press, Cambridge MA.

Citera, Maryalice, McNeese, Michael D., Brown, Clifford E., Selvaraj, Jonathan A., Zaff, Brian S., and Whitaker, Randall D., 1995, "Fitting Information Systems to Collaborating Design

Teams," *Journal of the American Society for Information Science*, **46**(7), 551-559.

Cooper, William S., "On Selecting a Measure of Retrieval Effectiveness," *Journal of the American Society for Information Science*, Arthur W. Elias, (ed.), **24**(2), 87-100, 1973a.

Cooper, William S., "On Selecting a Measure of Retrieval Effectives Part II. Implementation of the Philosophy," *Journal of the American Society for Information Science*, Arthur W. Elias, (ed.), **24**(6), 413-423, 1973b.

Court, A.W., Culley, S.J., and McMahon, C. A., 1997, "The Influence of IT in New Product Development: Observations of an Empirical Study of the Access of Engineering Design Information," *International Journal of Information Management*, **17**(5), 359-375.

Lowe, Alistair, McMahon, Chris, and Shah, Tulan, Culley, S., "A Method for The Study of Information Use Profiles for Design Engineers," *Proceedings of the 1999 ASME Design Engineering Technical Conferences*, September 12-15, 1999, Las Vegas, Nevada.

Deerwester, S., Dumail, S.T., Furnas, G.W., Landauer, T.K., and Harshman, R., 1990, "Indexing By Latent Semantic Analysis," *Journal of the American Society For Information Science*, **41**(6), 391-407.

Kilker, Julian, 1999, "Conflict on Collaborative Design Teams," *IEEE Technology and Society Magazine*, Fall, 12-21.

Landauer, T.K., Foltz, P.W. and Laham, D., 1998, "Introduction to Latent Semantic Analysis," *Discourse Process*, **25**, 259-284.

Lloyd, Peter, 2000, "Storytelling and the development of discourse in the engineering design process," *Design Studies*, **21**, 357-373.

Olsen, Gregory R., Cutkosky, Mark, Tenenbaum, Jay M., and Gruber, Thomas R., 1995, "Collaborative Engineering Based on Knowledge Sharing Agreements," *Concurrent Engineering: Research and Applications*, **3**(2), June, 145-159.

Salton, Gerald and McGill, Michael J., 1983, *Introduction to Modern Information Retrieval*, New York: McGraw-Hill Book Company.

Toye, George, Cutkosky, Mark R., and Leifer, Larry J., and Glicksman, J., 1993, "SHARE: A Methodology and Environment for Collaborative Product Development," *Proceedings of IEEE Infrastructure for Collaborative Enterprises*.

Ullman, David G., Dietterich, Thomas G., and Stauffer, Larry A., 1988, "A Model of the Mechanical Design Process Based on Empirical Data," *Artificial Intelligence in Engineering Design and Manufacturing*, Clive L. Dym, (ed.), **2**(1), 33-52.

Wertsch, James, 1991, *Voices of the Mind, Sociocultural Approach to Mediated Action*, Harvard University Press, Cambridge, MA 47-92.