KMi Technical Report

KMi–TR–97

Using genre to support active participation in learning communities

Trevor Collins, Paul Mulholland and Stuart Watt

Knowledge Media Institute

The Open University, UK

November 2000

Revised January 2001

To appear in *The Proceedings of the European Conference on Computer Supported Collaborative Learning (Euro-CSCL'2001)*

Using genre to support active participation in learning communities

Trevor Collins, Paul Mulholland and Stuart Watt

Knowledge Media Institute, The Open University, Milton Keynes, MK7 6AA, UK

Email: {T.D.Collins|P.Mulholland|S.N.K.Watt}@open.ac.uk Web: http://kmi.open.ac.uk/projects/elc

ABSTRACT

Many communities exist that learn and share information either partly or wholly online. These (wholly or partially) on-line communities share messages, documents, and other artefacts that contain useful community knowledge. Members of the community learn through this sharing process, and the growing archive they create forms a valuable learning resource for existing and new members of the community. Two main kinds of approach exist to support community members in accessing resources. The first kind associates each communal artefact with a conceptual structure that represents its meaning. This approach requires high levels of maintenance, especially when the community resource grows at a fast rate. The second uses statistical and text analysis techniques to (semi) automatically derive semantics from the resource. There is increasing evidence that artefacts constructed and shared within a community follow genres revealed in the structure of the artefacts and the terminology used. These implicit genres used in the community are invaluable to members in constructing and interpreting artefacts, but existing tools that support members in locating and classifying resources make little or no use of genre. Our preliminary findings demonstrate the potential of genre-sensitive classification and retrieval tools.

Keywords

Communicative genre, learning communities.

INTRODUCTION

It is well accepted that learning does not end with the completion of formal education. Much of the learning that occurs throughout life is supported not by textbooks or formal training programmes, but through community membership. Learning about a topic becomes synonymous with learning to be a member of a community of people who are experts on that topic. These communities range from workplace groupings that share ideas and stories, to hobbyists who form communities of shared interest. Increasingly, these communities are learning and sharing ideas at least partially online. These (wholly or partially) online communities create and share messages and documents that could form a useful learning resource for new and existing members. However, finding appropriate content from a large and growing resource, for example by employing text search engines, can often be inefficient or even fruitless.

Currently, a number of knowledge technologies are available aimed at alleviating this problem. Many involve annotating the documents and other artefacts of interest according to some type of formal representation or ontology of the knowledge and interests of the community. A notable approach is the CommonKADS methodology (Shreiber et al., 2000). These approaches run into difficulty if the turnover of documents necessitates a level of maintenance beyond the available resources of the community. Maintaining a knowledge base involves effort and therefore cost. For the effort to be worthwhile, the cost has to be paid back, ideally with interest, as workers benefit from access to useful resources. In corporate enterprises, even when the knowledge contained in documents is valuable, if the topic of interest is fast moving then individual documents have a limited life span in which to pay back the investment.

A number of research initiatives are investigating how fast moving information archives can be managed with low maintenance, using statistical techniques to analyze the resource (e.g. Berry et al., 1999). These tend to draw on one characteristic of the community, the community vocabulary. In practice this concerns the frequency with which particular words are used in different documents. These approaches tend to ignore other features, such as the formatting and location of words or phrases within documents. Other tools, like the Office Assistant (Horvitz, 1999), are aimed at helping the individual rather than the community as a whole. These approaches fail to make full use of the nature of communications within a particular community, as the messages and documents an individual receives will often originate from a number of communities of which the person is a member. These approaches are therefore more

often used to classify documents into their community of origin, rather than locate particular information from within a community corpus.

We aim to tackle the problem of fast moving information by leveraging the characteristics of the community that use and generate the resource. Our motivation for adopting this approach is that communities have particular characteristics, such as communicative genres, which are revealed in their way of doing things. We hypothesize that if the personality of the community could be understood it could be utilized in the development of low maintenance tools that could provide high accuracy for the search and retrieval of important community artefacts. More specifically, we hypothesize that the perspective of the community (i.e. what the community sees as important) will be reflected in the implicit structures found in the artefacts they create and share. Our approach utilizes and adapts a number of existing tools and methodologies including heuristic analysis, knowledge modelling and statistical analysis.

This paper gives an overview of the related work influencing our approach and our initial methodology drawing on examples from our study of an electronic newsletter. The next two sections provide an overview of communities and how they learn, and existing tools and techniques to support them in finding resources. This is followed by a description of our own approach, illustrated by our preliminary tests. The paper concludes with a summary of our findings.

COMMUNITIES AND HOW THEY LEARN

By way of theoretical foundation, in this section we characterize the nature of communities and how they learn. The kind of communities we are particularly interested in are what Lave and Wegner (1991) call communities of practice. These are groups of individuals who work, learn, or socialize together sharing insights and developing a shared knowledge as a consequence of participation. Communities evolve, develop and merge around shared interests and expertise.

Learning in a community is not so much about learning a topic as learning to be a member. From a learning viewpoint, communities have some key characteristics (Lave and Wegner, 1991). First, members have different roles such as master and apprentice. Second, there is a developmental progression through these roles. Third, access is gained to these communities through peripheral participation, that is active involvement is required. Fourth, communities develop their own language, generally through story telling. Fifth, communities adapt to change, often through newcomers providing new insights. Sixth, different communities need to develop and take perspectives in order to learn from each other (Boland and Tenkasi, 1995). *Learning in a community is therefore about communicating, sharing and discovering through participation.*

Recent research in the area of work-based learning has highlighted the importance of group tacit knowledge within communities (Cook and Brown, 1999; Raelin, 1997). This is tacit knowledge not held in the heads of individual members, but reflected in the artefacts created and shared by the community. As described by Cook and Brown (1999) when members of a community share stories, the stories not only transfer information between group members, but affirm the identity of the community. There are implicit rules about why a story is told, what makes a good story, and how it is told. Cook and Brown (1999) use the term genre to describe the personalizations that occur to group artefacts. Most people are familiar with the memo genre, and have expectations of what kinds of information memos contain (e.g. short directive, name of sender, subject heading), where they are located on the memo and the kinds of language used. A memo is generic and not tied to any particular community but serves to illustrate the kind of knowledge that is tacit and on the level of the group. The genre may include the use of certain words or phrases, the location of words, even the use of formatting constructs (e.g. bold or italics) to signify meanings relevant within that community. Group tacit knowledge does not only apply to stories. Schön (1988) for example describes how designers share models that serve as holding environments for ideas that need not or cannot be articulated.

Communities, therefore personalize the way they communicate and learn from each other. There are two important consequences of this. First, it makes it harder for the outsider who has not previously participated in the community to correctly interpret what is happening. Second, it allows members of the community to share knowledge and learn more effectively. The community genre allows them to share almost in a form of shorthand, where not everything has to be said or made explicit. What is mystical for the observer is obvious for the member. Our aim is to develop tools that can effectively help community members to share and retrieve communal artefacts, by being "tuned" to the genre of the community they are intended to serve.

APPROACHES TO SUPPORTING LEARNING COMMUNITIES

The overriding goal of the methodology and tools we wish to develop is to derive or attach semantics to communal artefacts so that they can be classified, identified and accessed as needed. A number of research endeavors are aiming to achieve this goal. Two of the dimensions along which these endeavors can be described are the nature of the content, and whether the tool is primarily intended for use by an individual or a community. These dimensions are considered in turn.

The nature of the content

In terms of the nature of the content, there are four main kinds of approach; keywords, knowledge modelling, statistics, and text analysis. First, keywords may be associated with documents (or other artefacts) contained within the corpus. Here, the list of keywords represent the meaning of the document. The document itself is not analyzed by the search tools. Although offering a reasonably lightweight solution, it can be difficult to maintain the use of a specified set of keywords, or to devise a set of keywords that adequately cover the domain and appropriately describe each document within the domain. Second, knowledge modelling techniques associate the documents with a conceptual structure of the domain and the resources within it to be described. As above, a conceptual representation is analyzed rather than the document itself. The conceptual structure provides a representation of how different concepts in the domain are related and allows sophisticated reasoning mechanisms to be deployed. It does however require a great deal of effort to develop and maintain, and is therefore better suited to stable knowledge.

Third, statistical approaches analyze the frequency of particular words across documents, a popular example being Latent Semantic Indexing ("LSI") (Berry et al., 1999). LSI and related approaches can offer interesting insights into the word frequencies occurring in different kinds of documents that can be used to generate clusters. These approaches are generally applied to the total number of words in the document, and ignore word patterns, such as phrases, the location of words within documents, and other cues such as formatting. Fourth, text analysis techniques attempt to parse or interpret the document. On the one hand, the approach of Cleary and Bareiss (1996) performs a shallow analysis of the text, and zones in on particular sections of the document such as the headings or abstract. This approach is reasonably scalable and easy to maintain and can be particularly successful if the parts of the document analyzed are representative of its meaning. On the other hand, techniques from a computational linguistic perspective perform a deeper analysis and parse the document to derive semantics (e.g. Osgood, 1994). Although potentially offering a high degree of accuracy, the language parser can be difficult to develop, and inefficient for larger archives.

Some of these approaches can identify and use particular aspects of group genre. For example, LSI can be used to identify the evolution of the terminology used in documents, and shallow analysis aims to identify and use important locations in documents. *However, none of the existing techniques alone fully exploit the genre of the community in the analysis, classification and retrieval of resources.*

Individual or community?

Individual centered approaches work by building or using a conceptual model of the individual in order to group, label or act on documents or incoming messages on behalf of the user. These approaches are generally used to sift, search and categorize received or requested documents or messages. One interesting example is the LookOut agent for supporting calendaring and scheduling (Horvitz, 1999). LookOut automatically parses incoming email messages to identify events and their date and time. If identified, the agent can automatically open the user's calendar and make an entry. The system learns from user input over time with the aim of helping the user without causing irritation. The analysis of messages and user feedback is supported by the Bayesian Belief Network statistical technique (Horvitz, 1999).

One major challenge for tools aimed at the individual is that any person is a member of multiple communities, the individual will take different roles in each community and each community has its own way of doing things. The success of LookOut may be partly due to a large degree of similarity in how meetings events are proposed, regardless of the community from which they originate. Possibly, the information that has to be analyzed in the message (e.g. date, time, location) can only be said in a limited number of ways. This may not be true for different kinds of information received from different communities. For example, within email discussion groups members develop their own terminology and ways of structuring messages. What is of interest to the recipient of the message may have a very different structure and terminology depending on the community from which the message originated. This will no

doubt greatly increase the complexity of an agent employed to act on messages on the user's behalf. Messages (and other artefacts) created within a community have a certain context, consistency and coherence, but any person is a member of multiple communities, each having their own way of doing things. *It may be more effective to develop technology aimed at one community comprising a number of people, than one person who is a member of multiple communities*.

Other approaches are aimed at the level of the community. Interesting examples include the Virtual Participant (Masterton, 1997) and the Vicarious Learner (McKendree and Mayes, 1997). These technologies are deployed to analyze email discussion sites used in an educational setting. They analyze the text content of messages to identify topics being discussed. The identification of a topic leads to the automatic sending of a new message, giving further information about the topic. The content of the additional message may have been authored for this purpose by the lecturer or derived from a discussion occurring on a previous presentation of the course. These approaches analyze words or phrases from the domain to which the discussion applies, but tend to ignore issues such as formatting, or the location of words within the document. One advantage of approaches that work on the community level is that the artefacts are analyzed in the context of the community within which they were created and shared. *Community centered approaches could be even more effective if they were able to more fully recognize and utilize the genres of the group artefacts*.

OUR APPROACH

Our approach to investigating genre to support the identification, analysis and classification of communal artefacts can be divided into five main tasks:

- Elicit the perspective and use scenario of the shared artefacts.
- Represent core, stable, hard to derive knowledge in an ontology.
- Investigate potential genres.
- Genre analysis in terms of the format and structure, and community vocabulary.
- Apply and test within the target community.

Eliciting the perspective must be performed first. The other stages are performed in parallel. These tasks will be illustrated in terms of our preliminary work in accessing and utilizing genre from a departmental electronic newsletter called KMi Planet (Domingue and Scott, 1998) (see figure 1).

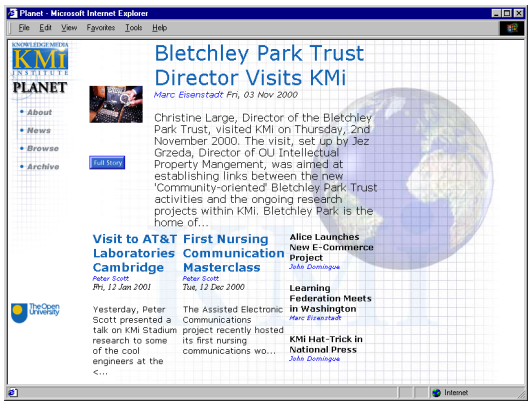

Figure 1. KMi Planet, the electronic newsletter.

The newsletter takes the form of a web page. Whenever an important event occurs in the department any member can submit a story as an email message. The story contained in the email is automatically formatted and turned into a "front page" story. Older stories are stored in an archive. The newsletter has been in operation for approximately four years, and has an archive of around 130 stories. This corpus is relatively small, but was considered sufficient for the preliminary analysis. The aim behind the preliminary analysis was to gain evidence as to whether community genres can be effectively identified and utilized to support the sharing and retrieval of documents or messages within a community. Specifically in this case to determine whether the authors of the stories had developed a genre that could be utilized to allow the correct stories to be retrieved with a high degree of accuracy.

1. Elicit the perspective and use scenario of the shared artefacts

Before performing any detailed analysis of the existing corpus and the existence of genre, it is necessary to determine the perspective of the community, in terms of the scope and nature of their interests, how resources (e.g. documents or messages) are used within the community and how this could be supported. In the case of the KMi Planet archive, previous research aimed at building a knowledge model of the archive (Domingue and Motta, 2000) identified three main types of story: awards, visits, and demonstrations or applications of technology.

We decided to take these as a starting point. This had the added advantage, that the accuracy of the methods we developed could be compared against the explicit categorization of stories within the knowledge model. Locating important information and kinds of stories within the archive was selected as the activity we wished to support. Additionally, we wished to support new members of the institute in gaining access to stories of particular types, in order to support them in constructing their own stories. For example, a new member preparing a story about a particular topic may wish to look at previous similar stories to get a feel for how they should be presented. Given the nature of the archive, we were therefore aiming to use genre to locate stories that help provide answers to questions such as "what projects have been awarded in KMi recently?" and "what do visit stories look like?"

2. Represent core, stable, hard to derive knowledge in an ontology

Although we aim to use genre to automatically classify news stories and identify information within them, it is not necessarily appropriate to derive all information automatically from the stories. Statistical, heuristic, or other techniques that automatically derive information will always have some degree of error. For knowledge that is core or highly valuable to the community any degree of inaccuracy may be unacceptable. This knowledge would be more appropriately represented explicitly in an ontology. Similarly, some knowledge may be important but hard to derive, and therefore more suited to explicit representation. Within our newsletter example, names and roles of staff within the institute were identified as something that it would be more sensible to represent explicitly rather than attempt to automatically derive. Other knowledge in the core ontology included the kinds of funding agencies that support research in the institute, as these are difficult to derive and could ease the identification of stories concerned with new projects.

3. Investigate potential genres

Communicative genres are socially constructed behaviors developed by a community for the fundamental purpose of efficiently conducting communal activities (Yates and Orlikowski, 1992). This definition specifically concerns physical artefacts of the community such as documents, memos and letters. Cook and Brown's (1999) definition of genre, introduced earlier, encompasses a wider range of physical and social artefacts, such as different kinds of meetings conducted in an organisation. Communicative genres are characterized by the form, medium and language used (Orlikowski and Yates, 1994). In the case of (partially or wholly) online communities the medium used is primarily electronic text. Form refers to the easily observable features of the communication, including structural features e.g. text formatting devices such as headings and bullet points, and devices for structuring interactions at meetings such as agenda and chairpersons. Examples of the language used in a genre include the level of formality and any specific vocabulary or terminology used.

In order to investigate the potential genres used in the KMi Planet stories, the stories were examined with respect to both their form and language. By iteratively analysing the stories from a number of different perspectives a range of potential genres were explored and a series of story characteristics were identified. Specifically, it was noted that all of the stories were news stories about the department, they were quite short and direct, and various members of the lab wrote them. In order to explore the extent to which the authors' familiarity with traditional newsprint stories affected the construction of electronic news stories, a comparison was made with newspaper journalism.

In journalism trainees are explicitly taught to use an inverted pyramid heuristic to structure their stories, where the most important information is presented in the 'lead' and is then followed by gradually less important information as the story progresses (Keeble, 1998) (see figure 2). The inverted pyramid principle can be applied to structure an entire story based on a single source, or it can be nested, with pyramids inside pyramids, to structure more complex stories from multiple sources. In all cases it is the lead that contains the most important elements of the story. In newsprint stories the advantage of using the inverted pyramid structure is that the stories can be truncated without loosing the meaning of the story. Not only does this allow the copy editor to trim the length of a story it also facilitates the selective scanning of news stories by the reader.

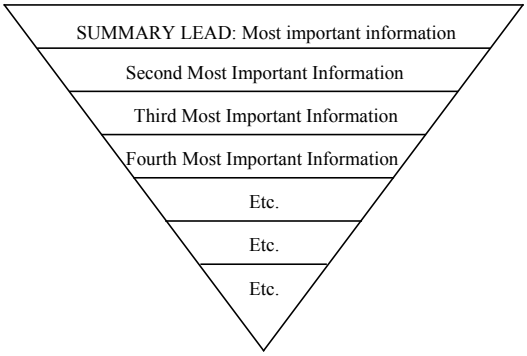

Figure 2. The Inverted Pyramid, adapted from Keeble (1998).

4. Genre analysis

Once a potential genre has been identified, it is possible to analyse the messages from the perspective of the community that is using these genres. There are two aspects to this analysis, messages are analysed in terms of their format and structure, and in their use of vocabulary. These are done in parallel, as they are to some extent interlinked.

Analyse the genre in terms of format and structure

In the 'inverted pyramid' genre, the first part of the story, the 'lead', gives a summary of the whole story, also setting out the most important parts of the story, the 'who', 'what', 'where', 'when', 'why', and 'how'. The headline, on the other hand, has a different role – grabbing people's attention and flagging the most important part of the story. Because the headline isn't actually telling the story, the lead is more important than the headline in conveying the essence of the story. We began by analyzing the stories to look at the relationship between the inverted pyramid structure and the story content in more detail.

In our sample of 127 stories, 111 (87.4%) began with a sentence that summarized the story. A typical summary sentence, for example is:

> "*On Friday 24th October 1997 a team from British Aerospace visited the Open University to talk about the British Aerospace Virtual University initiative*."

This says who visited who, and when and why this happened. The headline for this story was:

> "*The Newest Virtual University Meets the Oldest!*"

which grabs the attention, but doesn't even begin to tell the story.

A further nine stories (7.1%) followed a variety of other standard news genres (e.g., anecdotes, direct address to the reader). A good example is a story which begins:

> "*As the theme of the 8th world conference on Artificial Intelligence in Education, held in Kobe Japan, was 'Knowledge and Media in Learning Systems' I shouldn't have been as surprised as I was that the current interests of the AI & ED community now closely match that of KMi's*."

This follows a more anecdotal pattern common to feature-style stories.

The remaining seven (5.5%) did not clearly fall into any straightforward news story genre, and could not be classified. Many of these were outliers in content and purpose, as well as their format, and were often light-hearted. The story headlined:

> "*Christmas comes early to KMi.*"

was a good example, with the lead:

> "*The delivery of 25 iMacs took on a festive mood when the colourful collection arrived*"

Overall, we found a pattern that is very different from the naïve 'distance from the beginning' heuristic used by some search algorithms. For example, the headline is right at the beginning, but it was rarely useful as a guide to the content of the story, as its role is to grab people's attention. The structure of the genre does influence where in the structure the story content is most effectively represented, so community-independent heuristics like these are bound to be limited.

There was evidence for a gradual convergence onto the story genre. There was a significant increase (1 by 2 chi-square test, $\chi^2 = 5.28$, $p<0.05$) in the use of bullet-pointed links at the end of a story, with 36.5% of recent stories, and only 20% of older ones using this form. Older stories often inserted URLs to related information earlier in the story, although according to the inverted pyramid genre, one would expect to find this background information at the end. This supports the findings of Orlikowski and Yates (1994) that genres change over time as peoples' experiences with the media and community develop. There is some evidence that people were getting more skilled with the news genres in that the number of unclassifiable messages was dropping, and the number of other, perhaps more sophisticated, news genre forms being used was growing, but

the sample size was too small to analyse this in detail.

Analyse the genre in terms of the community vocabulary

We then started to analyze the different structural components of a story, looking for differences in the vocabulary that they used. We began by looking at the first sentence, and looked for evidence of the genre in the terminology. We carried out 1 by 2 chi-square tests to look for those words that were significantly more likely to appear in the first sentence compared to the rest of the document. Some of these high frequency first sentence words (e.g., 'award', 'visit') clearly identify the classes of story identified by Domingue and Motta (2000) in their ontology. Others (e.g., 'part', 'of') are connectives, and can be used to parse the summary sentences and recognize the objects and concepts the story is about.

Let's look at one category of story as an example. Visits are one of the most common stories in the database. There are 30 visits in the database, making up 24% of all stories. Of these, 25 (83%) used the term visit in the first sentence, and 19 (63%) used the term 'visit' in the headline. Overall, 90% of visit stories can be categorized correctly through this single term.

However, we can also use the connectives to parse the most common summary forms. For example, the pattern "[**on** *date*] *person* **visited** *something* [**on** *date* | **as part of** *something* | **to** *something*]" matches 19 (63%) of the stories in a way that allows people, dates, and reasons behind visits to be identified.

For example, if we take the story lead sentence (bold added to emphasize structure):

> "*David Brown, the Chairman of the University for Industry Design and Implementation Advisory Group and Chairman of Motorola*, **visited** *the OU* **as part of** *a fact finding exercise, prior to drafting his initial 100 Days Report to HM Government*"

we can quickly identify what happened, who visited who, and why. In this case, the 'who' part can itself be analyzed to identify the name of the person, and any roles they may have.

To confirm the special role of the first sentence, we carried out a similar test comparing the second sentence to the rest of the story. Many fewer words (5 compared to 13) were significantly more frequent in the second sentence, and these words, (e.g., provide, part, and research) tended to give supplementary information to that in the first sentences. This seems to confirm our hypothesis that stories strongly followed an inverted pyramid story genre.

We also investigated whether the headline played a significant role in the genre, but only four words were significantly more common in the headline: 'KMi', 'at', 'for', and 'of'. This confirms that terms in the headline are less central to the genre than those in the first sentence, and that the headline is not a useful guide to the story's content.

5. Apply and test

In order to test the approach we developed a browsing tool that exploited the community genre by focusing on word patterns occurring in the first sentence (see Figure 3). The tool parses the html stories into a structured database, which can then be queried using the identified story structure. By querying the structured database using terms identified in the analysis of the vocabulary, the stories are categorized according to the topics of interest to the community. For example, stories about awards contained words such as: grant, contract, proposal, and fund.

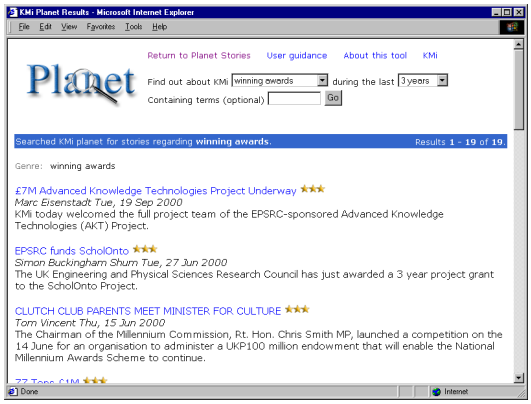

Figure 3. The KMi Planet story browser.

The classification according to first sentence was found to have a surprisingly high accuracy even among new stories not considered in the genre analysis stage.

CONCLUSIONS

Although not trained in journalism the authors of the stories had converged on a style that mimics journalistic practice. This genre could then be used to support the automatic analysis of the stories and retrieval of information. As a consequence of this research we believe that

tools to support a community in sharing and retrieving information must be sensitive to community genre in order to properly reflect what that community sees as important.

ACKNOWLEDGEMENTS

This work was conducted as part of the Marconi Communications funded project, Empowering Learning Communities. Thanks to John Domingue for providing access to the Planet story archive.

REFERENCES

Argyris, C. and Schön, D. A. (1996). *Organisational Learning II: Theory, Method and Practice*. (Addison Wesley, Reading, MA).

Berry, M.W., Drmac, Z. and Jessup, E. R. (1999). Matrices, vector spaces and information retrieval. *SIAM Review*, **41** (2), 335-362.

Boland, R. J. and Tenkasi, R. V. (1995). Perspective making and perspective taking in communities of knowing. *Organization Science*, **6** (4), 350-372.

Brown, J. S. and Duguid, P. (1998). Organizing knowledge. *California Management Review*, **40** (3), 90-111.

Brown, J. S. and Duguid, P. (2000). *The Social Life of Information*. Boston, MA: Harvard Business School Press.

Cleary, C. and Bareiss, R. (1996). Practical methods for automatically generating typed links. *The Seventh ACM Conference on Hypertext (Hypertext '96)*, Bethesda, MD, pp. 31-41.

Cook, S. D. N. and Brown, J. S. (1999). Bridging epistemologies: The generative dance between organizational knowledge and organizational knowing. *Organization Science*, **10** (4), 381-400.

Domingue, J.& Motta, E. (2000) Planet-Onto: From News Publishing to Integrated Knowledge Management Support. *IEEE Expert Systems Special Issue on Knowledge Management and Distribution over the Internet*. **15** (3), 26-32.

Domingue, J. and Scott, P. (1998). KMi Planet: Putting the Knowledge Back into Media. In M. Eisenstadt and T. Vincent (Eds.). *The Knowledge Web: Learning and collaborating on the net*. London: Kogan Press. pp. 173-184.

Horvitz, E. (1999). Principles of mixed initiative user interfaces. *Human Factors in Computing Systems (CHI '99)*, Pittsburgh, PA, May 1999, pp 159-166.

Keeble, R. (1998). *The Newspapers Handbook*. Second Edition. Routledge: London.

Lave, J. and Wenger, E. (1993). *Situated Learning: Legitimate Peripheral Participation*. New York: Cambridge University Press.

Masterton, S. (1997). The Virtual Participant: Lessons to be learned from a case-based tutor's assistant. In *Proceedings of Computer Supported Collaborative Learning*, Toronto, Canada.

McKendree, J. and Mayes, J. T. (1997). "The Vicarious Learner" Investigating the benefits of observing peer dialogues. In *Proceedings of Computer Assisted Learning*, Exeter, UK.

Orlikowski, W. J. and Yates, J. (1994). Genre repertoire: Norms and forms for work interaction. *Administrative Science Quarterly*, **39**, 541-574.

Osgood, R. (1994). The conceptual indexing of conversational hypertext. Unpublished PhD thesis. Northwestern University.

Raelin, J. A. (1997). A model of work-based learning. *Organization Science*, **8** (6), 563-578.

Schön, D. A. (1988). Designing: rules, types and worlds. *Design Studies*, **9** (3), 181-190.

Schreiber, G., Akkermans, H., Anjewierden, A., de Hoog, R., Shadbolt, N., Van de Velde, W. and Wielinga, B. (2000). *Knowledge Engineering and Management*. Cambridge, MA: MIT Press.

Yates, J. and Orlikowski, W. J. (1992). Genres of organisational communication: a structural approach to studying communication and media. *Academy of Management Review*, **17** (2), 299-326.