A System for Organizing Situational Knowledge in the Workplace that is Based on the Shape of Documents

Abstract: The authors propose a system for organizing situational knowledge, or knowledge of appropriate conduct, in workplaces that rely on web-based interaction. The conceptual framework that underlies the system is based on five propositions. First, recurrent and routine practice in organizations is articulated in, and co-evolves with distinct documentary forms, or genres. Second, the presence of sets of documentary genres in a group or other form of organization is indicative of activities that characterize such organization. Third, such indexicality may be observed at different levels of organization (the project, the unit, the firm), and clusters of genres at different levels of aggregation may provide profiles of activities at those different levels. Fourth, a notation (such as XML) which captures the 'shape' of documents may be used to model flexible documentary 'compounds' that capture situational knowledge, or knowledge of appropriate activity in an organization. Fifth, such encodings may be used compare organizations and sort them on the basis of their genre and activity profiles; visualization may accelerate the sorting process. An activity classifying system that integrates these proposals might improve organizational experience in a number of evaluative contexts (like benchmarking, team formation, or merger).

1. A Framework for Classifying Activity in the Workplace

The authors propose that a framework for organizing knowledge of how to act in the workplace ('situational knowledge') may be derived from a synthesis of XML markup language, genre analysis and activity analysis. The latter are linked by the fact that documentary genres reflect recurrent activities in communities of practice, and that recurrent activities are moderated by generic documentary forms ('genres') which orient agents in a community, and consolidate the distinctive practices of that community. The proposed framework has emerged from published work, and discussion with colleagues working in a number of different literary traditions, whose interests converge around studies of documentary genres, documentary 'shape,' and visualization techniques for representing social interaction. We have also drawn on material on information retrieval and structured documents, and, specifically, on the use of XML in e-commerce.

2. Genre, Common Sense and Organizational Order

Recent work suggests that, in the workplace, documentary genres provide an efficient articulation of what to do and why to do it. Key texts by Yates and Orlikowski and their colleagues in the past decade have demonstrated the mutually reinforcing power of domain practice and documentation in stabilising activities at work. Arguing that 'genre repertoires' play an important role in making the workplace manageable, Orlikowski and Yates (1994) identify three elements of genres that can explain their role as stabilisers: a recurrent situation, substance ('social motives...themes...topics'), and form (structural features, communication medium, and symbolism). Genres are enacted through rules which associate appropriate elements of form and substance with situations; to engage with a genre is to 'implicitly or explicitly draw on genre rules', and also to 'reinforce and sustain the legitimacy of those rules'. (Yates and Orlikowski, 1992, 301-302). Genres exist at different levels of abstraction, and are defined differently 'in different cultures and at different times.(op. cit, 303) What is interesting about them is their dual status as (1) an articulation of what has emerged as
appropriate behavior (their role as a 'categorizing' device) and (2) as a prescription for activity in a community of practice (their role as a 'regulatory' device). Genres thus articulate common understandings and shape the 'common sense' of groups that engage with them.

At the micro-organizational level, Bergquist and Ljungberg (1999) have been engaged for a number of years on a longitudinal study of a large Swedish manufacturing firm, and of the role of documentary genres at different eras in the firm's history. They conclude that genre is a 'fruitful concept for how communication enacts organization', and that 'many of the most interesting genres are related to project work in some form or another', and that genres are 'dynamic and living phenomena' as there is much discussion and negotiation of their use. Genres in their case study both 'order' core activities, and 'regulate' them by means of legitimizing dialogue. Davenport (1999) in a review of this and other empirical studies, concludes that genres are, in effect, 'ordering' devices which both 'regularize' and 'regulate' activities.

Subsequent work (Yates and Sumner, 1997; Yates, Orlikowski and Rennecker, 1997) suggests that such insights, derived from studies of the traditional workplace, are valid in the online world. In this particular case, analysis of a discussion list reveals a complex interplay of speech and written genres, some of them signalled by typographic devices (the standard netiquette icons). Erickson (1997) in a discussion of genre theory and on-line discourse, notes the importance of shared community goals and expectations, and of the properties of the recurrent situations in which genres are employed, including the institutional, technological, and social forces that give rise to the regularities of the discourse. Other recent studies (Crowston and Williams 1997; 1999; Procter et al., 1998) have explored the production of genres in a number of digital contexts: analyses of discussion lists, newsgroups, FAQs, and home pages. These also suggest that genres play an important role in regulating activity in webspace.

3. Recognising the Shape of Documentary Genres and their Related Activities

A body of work in both LIS and business literatures has explored how genres articulate and classify collective understanding. A useful summary is given in Vaughan and Dillon (1998). Work by Dillon and his colleagues suggests that genres derive much of their authority from their structure or 'shape' (Dillon and Vaughan, 1997, Dillon and Schaap, 1996). At a deep level of reading, genres both reinforce and activate familiar mental models of text structure (Van Dijk and Kintsch, 1983); experienced readers can thus rapidly appropriate the content of text once a genre is recognized. Whatever the underlying process, readers subjected to direct observation can recognize genres on the basis of stylistic features (Toms and Campbell, 1999). Readers clearly know where they are and what to do in the presence of genres, paper-based or web-based (Vaughan and Dillon, 1998; Dillon and Gushrowski, 1999). Dillon suggests that a 'socio-spatial' semantics (Dillon, 2000) is necessary for understanding how people and documents intersect on the web, a space that affords few physical cues about the shape and size of textual material.

Though these authors demonstrate that readers do exploit perceptions of document shape, much of the fieldwork has been narrow (academic texts), and has focused on individual texts. Participants had already been to some extent 'schooled' in the appropriate genres, and could clearly recognize when 'pieces' were out of place. Two areas remain largely unexplored in the LIS literature: the 'orientation potential' of genres in cases where novices must appropriate cultural norms, and the representation of 'compound information' (Paepcke, 1996), though the latter has recently been put on the 'research agenda' in work by Bishop and others in DL1 (Bishop, 1999; Bishop et al., 2000), a point we return to below. There is, however, a body of work in the HCI and business literatures that addresses the provision of a framework that might draw on shape and genre recognition to explore either traditional documents outside the academic domain, or non-traditional digital documents representing activities like meetings, workflows or business processes (Bargiella-Chiappini and Nickerson, 1999).
To broaden the discussion of 'document shape,' we have looked at a number of studies of visualization and social interaction which demonstrate that recurring actions can be codified (on the basis of analysis of turn-taking or of conversation threads) and re-presented in visual surrogates to provide rapid insight into the patterns of activity that define a group. The infrastructure where such data may be captured is in place in many organizations (email lists and archives, IRCs, newsgroups, and electronic meeting systems); and recent work has demonstrated that interfaces based on visualizations of interaction analysis can offer rapid insight into the group process. Examples may be found in the work of Donath on the visualization of threads in online conversations (Donath et al, 1999; Viegas and Donath, 1999), and visualizations of turn-taking in recent work by Erickson and his colleagues on participation in conversation circles (1999).

Just as a taxonomy of thread diagrams may represent the interactions of a group (independent of content) so to, a taxonomy of genre diagrams may represent the kinds of activity that characterize a group. To make the move from notation for micro-level informal social activity, to a notation for more formally structured activities in the workplace, a comparable set of visualizations or representations may be provided which can function as visual ontologies (Weinstein and Alloway, 1997; Karlgren and Straszheim, 1997) that would act as a ‘fast inscription’ and allow participants in workgroups to quickly grasp a group's activity repertoire.

A mechanism for doing this, we suggest, is markup language, specifically Extensible Markup Language (XML), because existing work on XML for individual document retrieval (O’Donnell et al, 1999) may be extended to retrieval, or sorting and matching, at a higher level of aggregation: 'genre' retrieval. In addition to providing a notation for document 'shape', XML metatags may capture, to some extent, the situational semantics of a given text, and structural, cultural and other characteristics of organizations. By accommodating contextual details (such as cost, constraints, associated experts and so on) XML templates can indicate where a particular text fits into the working habits ('situational knowledge') of a group.

4. Markup: Capturing the Relationship of Compound Documents and Collective Activities

The desire to have a closer understanding of how documents fit into work activity patterns is not new in LIS. Rayward (1996) offers a historical review of attempted solutions. Heine (1996) and Kircz (1998) have both suggested that markup (SGML in this case) might offer a way forward, and the former noted that professional standards for markup in different domains would be required. With the notable exception of work on citation indexing and citation mapping (Wouters, 1999), LIS has focused on individual retrieval of individual documents. Activities in groups are extensible, and heterogeneous, and any representation of workflow will need to take this into account by addressing documentary 'collectives'. Bishop and her colleagues (op. cit.) have identified this as an issue for digital library design, in recent discussions of 'assemblages' (the artefacts, knowledge, practices and community influences that shape library use) and infrastructure, and of the 'document streams' that characterize much of the work in digital libraries, where the distinction between genres, documents and document surrogates may be blurred, and where new document structures are likely to emerge. Bishop raises the issue of the individual document stream and although individual profiles may be useful if benchmarked against a collective profile, we reiterate that a priority for organizational research is the exploration of documents, shape and activity at the collective level.

XML can address this priority. Standards are available which take account of domain differences, and XML has already been used as an object oriented retrieval tool. We suggest that since genre and activity are tightly coupled in the world of business documentation, where many core genres are already agreed upon, a retrieval mechanism based on XML codes for genres, will, in effect, function as a retrieval mechanism for sets of activities, and that
organizations matched on this basis will have common situational knowledge. A genre-driven notation is thus a system for organizing situational knowledge. Our warrant for these claims is to be found in recent business literature, specifically, work on XML and e-commerce. Bryan (1999) points out that XML is flexible enough to describe any logical text structure, whether it is a form memo, letter, report, book, encyclopedia, dictionary or database, and suggests that emerging Business Object Libraries like CBL (from CommerceOne/Veo) or Biztalk will accelerate the adoption of XML as an e-commerce standard. Glushko et al (1999) explain the design rationale for the CBL, based on an extensible public collection of generic business interface definitions that tell potential trading partners what online services a company offers and what documents to use when making those services. The library's strength lies in the fact that it is easier to interconnect companies in terms of the documents they exchange as they already largely agree on these. Services are, in effect, defined by the documents they accept and produce.

The responsibility for 'document definition' will lie with communities of practice: early adopting communities are mathematicians, genealogists and chemists (Green, 1999). Many companies already recognize the need for information-exchange standards, uniting in several initiatives focusing on XML standards for particular industries or business processes (Wall, 2000). Proponents of XML predict it will allow online businesses to build on one another's published content and services to create innovative virtual companies, markets and trading communities, leveraged by 'comparison-shopping' agents, who can exploit the affordance of XML to rapidly configure appropriate resources. (Maes et al., 1999; Glushko et al, 1999). An appropriate feature of these extensible systems will be visualisation.

5. Compound Documentary Genres and Volatile Forms

The CBL handles long-established genres (the invoice, the catalog) whose role is so clearly understood that any ensuing action is intuitive: the 'situational knowledge' conveyed in such forms is implicit. Visualisation in such cases may not be essential. What about the trickier cases of compound or novel genres, the kinds of material studied at the micro-level by Erickson, Crowston and others that were mentioned above? In the domain of ecommerce, Rosenbaum (forthcoming) argues that there are identifiable genres of web based customer service. Another case is Tyrvainen and Paivarinta's (1999 study of a Finnish 'hi-tech' products company which provides a realistic assessment of the resource requirements for an Electronic Document Management System (EDMS) based on genre analysis. This system must accommodate multiple experts (practitioners, domain experts, organization designers, IS specialists), who routinely handle:

- heterogeneity among document genres now and in the future
- variations between document genres within a genre system
- variations between observers
- the fact that no single observer is able to comprehend all the document genres in the organization.

We argue that XML, coupled with visualisation, can help by handling variation; and it can handle 'interdomain' issues by accommodating bridging elements (Maes, 1999). As another example of a 'tricky' problem, we offer the case of an architect's office, and the role of a complex 'planning document' where a mixture of modalities and intersections must be managed in carrying the project to completion (Hildreth & Kimble, 1999). The negotiation of tasks and responsibilities in a complex project is a recurring situation, and the 'planning document' or associated document stream is a recognisable genre. Though the actors in this scenario clearly recognise the 'shape' of the complex document, and take appropriate action, we would suggest that its potential role as a guide to later project work is difficult to exploit without visualisation that might allow 'new' partners to assess where their own practice
(visualised as a 'shape') may fit. A number of visualization applications are available. Greene et al. (2000) describe an 'overview surrogate' that can become a complete application, such as a dynamic interface coupling a zoomable overview with real-time filtering mechanisms using sliders and buttons.

6. Conclusion:

The 'system,' then, referenced in the title is an XML-based representation of genres established in communities of practice. It is supported by visualisation techniques that provide rapid insight into the 'docuverse' that captures the activities of an online workgroup or organization. It will support the representation of the shapes of document collectives and their associated activities and will allow groups to be organised on the basis of situational knowledge that is captured in documentary genres. As we noted above, ecommerce firms are experimenting with a variety of web-based customer service options; can these identifiable genres be adequately represented using XML?

How do we propose to test a system for classifying activity at work that is based on document shape? An analysis of websites is a starting point; these must satisfy certain criteria to qualify as instances that are appropriate for genre analysis (Davenport, 1999). A series of simple 'sorting' exercises will be designed. First, subjects will be asked to infer the activities of a group or organization from an inventory of the 'genres' that are represented on a series of webpages. Second, they will be asked to rank groups or organizations in terms of similarity on the basis of the composite of genres that articulates practice in each of a series of web pages. Third, they will be asked to role play in a scenario where they wish to find a group who will be compatible to work with. The scenario might have two versions: the search for commensurate interests (in this, they may match the 'activity profile' reflected in genre analysis of their own group against that of other websites), and the search of complementary interests, in which they will search with a synthetic profile that reflects a need to exploit an extended repertoire.

With these data, an XML DTD will be developed which will define a markup language for these genres. Documents will then be created based on this genre markup language and the sorting exercises will be repeated. We hypothesize that we will see an improvement in performance in these exercises when using an XML-based representation of organizational genres based on communities of practice.

References:


collaborative retrieval in the virtual library. *Interacting with Computers*, 10, 157-175.


http://www.slis.indiana.edu/rosenba/www/Papers/ASIS971.html

Rosenbaum H. (Forthcoming) After the sale: understanding web-based customer service. *Information Technology and People*.


