

**LANCASTER
UNIVERSITY**

**Computing
Department**



Browsing is a Collaborative Process

Michael B Twidale, David M Nichols and Chris D Paice

Cooperative Systems Engineering Group

Technical Report : CSEG/1/1996



CSEG, Computing Department, Lancaster University, LANCASTER, LA1 4YR, UK.
Phone: +44-524-65201 Ext 3799; Fax: +44-524-593608; E-Mail: julie@comp.lancs.ac.uk

BROWSING IS A COLLABORATIVE PROCESS

MICHAEL B. TWIDALE, DAVID M. NICHOLS and CHRIS D. PAICE
Computing Department, Lancaster University,
Lancaster LA1 4YR UK

Abstract — Interfaces to databases have traditionally been designed as single-user systems that hide other users and their activity. This paper aims to show that collaboration is an important aspect of searching online information stores that requires explicit computerised support. The claim is made that a truly *user-centred* system must acknowledge and support collaborative interactions between users. Collaborative working implies a need to share information: both the search *product* and the search *process*. Searches need not be restricted to inanimate resources but people can also search for other people. The ARIADNE system is introduced as an example of computerised support for collaboration between browsers. A number of systems offering varied approaches to supporting collaboration are surveyed and a structure for analysing the various aspects of collaboration is applied.

... support for communication and collaboration is as important as support for information-seeking activities, and ... indeed, support for the former is needed to support the latter.

(Levy & Marshall, 1994)

1. INTRODUCTION

Interfaces to databases have traditionally been designed as single-user systems. The existence of other users and their activities have been implicitly assumed to be an attribute of the system that should be hidden from end-users (Rodden *et al.*, 1992).

In recent years the emergence of the field of Computer Supported Cooperative Work (CSCW) has highlighted the importance of collaborative approaches in many diverse activities. This paper aims to show that social activities are an important aspect of searching online information stores and that they require explicit computerised support.

This paper discusses the role of collaboration in information seeking and the implications of the trend towards increased remote searching of information systems. Collaboration is used in a wide sense; that of people, or agents, together achieving tasks which they couldn't otherwise. The claim is made that a truly *user-centred* system must acknowledge and support collaborative interactions between all users.

The literature is first reviewed to bring together the disparate references to social aspects of information seeking. CSCW terminology is introduced to provide a framework for analysis. The framework is used to consider the impact of the move to digital libraries on existing forms of

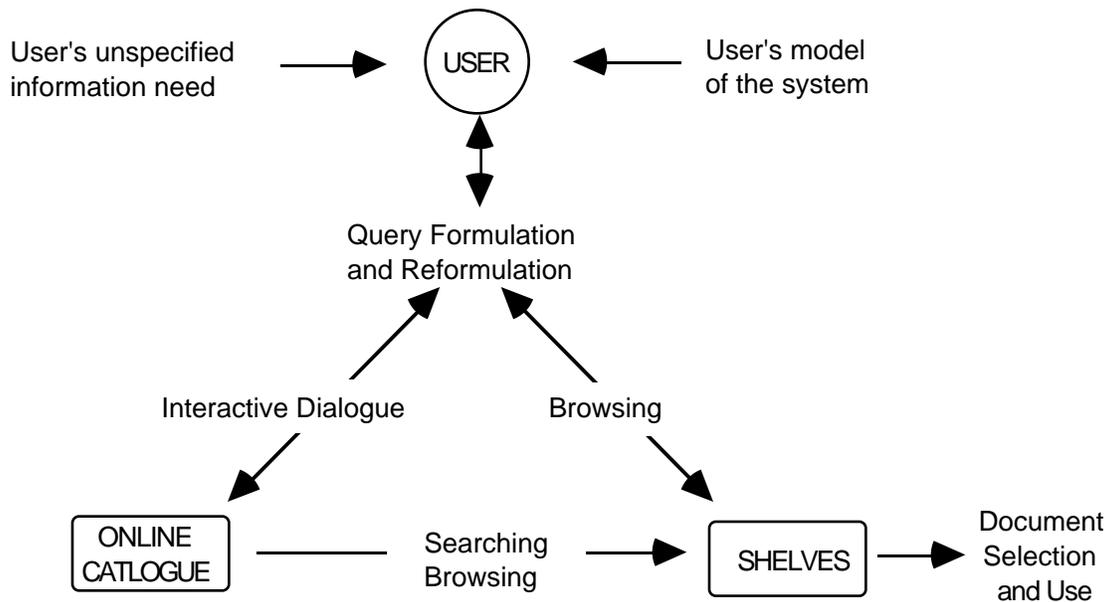


Fig. 1. Information-seeking activity from (Hancock-Beaulieu, 1989)

collaboration. The products of searching and the process of searching are examined in the context of recent systems in information retrieval (IR) and computer science that exploit social information to aid searching. The ARIADNE system is introduced as an example of computerised support for collaboration between browsers.

2. COLLABORATION: A NEGLECTED ISSUE

The trend from a system-centred view of information systems to a user-centred paradigm is now firmly established (e.g. (Dervin & Nilan, 1986; Kuhlthau, 1991; Morris, 1994; Watters & Shepperd, 1994)). This constructivist approach to searching recognises that users seek information from a variety of sources to satisfy their information needs. These needs are often initially vague and evolve during the search process, so that browsing is a more accurate description of users' behaviour than searching (Bates, 1989; Kuhlthau, 1993). Browsing may be defined as information searching that is opportunistic, reactive and unplanned, and for which the 'goal' may be fuzzy and dynamic (Marchionini, 1995) p. 100.

It is commonly assumed that browsing is all about locating physical documents or their electronic records; our view however is that other people should be seen as an integral part of the whole browsing activity. That people can be an important source of information is acknowledged in the importance placed on the user-intermediary reference interview (e.g. (Dervin & Dewdney, 1986; Kuhlthau *et al.*, 1992; Taylor, 1968)). However, such interactions are only a subset of the collaborative activity that occurs in libraries, and form an even smaller subset of the possibilities for collaboration within the virtual libraries of the future.

Table 1. Tactics with social aspects from (Bates, 1979a; Bates, 1979b)

Tactic Name	Description of Tactic
CONSULT <i>Idea tactic</i>	To ask a colleague for suggestions or information in dealing with a search. Comments by practising librarians indicate that this is a valuable and much-used tactic.
WANDER <i>Idea tactic</i>	To move among one's resources, being receptive to alternative sources and new search ideas triggered by the materials that come into view. In our field ... one may hypothesise that to WANDER promotes serendipity and enables useful sources that would not otherwise be discovered.
BRAINSTORM <i>Idea tactic</i>	To generate many ideas and to suspend critical reactions until the ideas are well-formed and can be fully evaluated. ... It is the anticipation of semiautomatic reactions, either from within or from others, that commonly blocks creative ideas.
BIBBLE <i>Information search tactic</i>	One way to cope with the file structure is to find a way to do without it altogether. ... BIBBLE is based on the abbreviation "bibl" for "bibliography." To BIBBLE is to look for a bibliography already prepared, before launching oneself into the effort of preparing one. More generally, to BIBBLE is to check to see if the search work one plans has already been done in a usable form by someone else.

The social aspects of information seeking have been noted in passing by the information science community although (to our knowledge) no systematic investigation of the issues involved has been undertaken. Surveys of browsing (e.g. (Ayris, 1986; Chang & Rice, 1993; Mann, 1986)) show that the predominant scenario is one of *individual* searchers accessing either physical books or electronic records. However, a searcher could also browse for people and the results of their activities.

Unfortunately, very little attention has been paid to the human channel of information exchange in both research and practice of information retrieval. (Zhao & Kantor, 1993)

As an example, Fig. 1 shows a model of information-seeking activity in a library which is intended to provide a conceptual framework for observing user behaviour (Hancock-Beaulieu, 1989). This model contains no representation of people other than a single user; both other users and librarians are omitted. Social interactions could reasonably appear at several places in this model: in informing the user's model of the system, in influencing the information need, in reformulating the query, in physically browsing the shelves and in using the documents. This model is however typical of the single-user stereotype that appears to account for the vast majority of work on browsing (e.g. (Apted, 1971; Bates, 1989; Beheshti, 1992; Cove & Walsh, 1988; Greene, 1977; Hildreth, 1982; Morse, 1970; O'Connor, 1993; Ross, 1983)).

2.1 Collaboration in conventional libraries

Although most browsing is viewed as occurring within a single-user context, social elements are present. (Bates, 1979a; Bates, 1979b) identifies tactics and ideas to aid information searchers – Table 1 shows the four which have explicit social aspects.

To *consult* is described as asking a colleague for help, although it can obviously include seeking help from library staff or from obliging strangers¹. In the future it is possible that this tactic might also include asking an intelligent computer-based system (Paice, 1986).

To *wander* can be viewed in a similar way to *consult*, as the resources available to a searcher are not limited to physical items but can include people and computerised systems.

To *bibble* is to take advantage of searches that have been done in the past and not waste time and resources re-inventing the wheel. A bibliography is a structured version of the results of a past search. The results of most searches are not published as bibliographies but are private, local and temporary and consequently, from the perspective of future users, the information is lost. This means that the great majority of searches that are conducted fail to *bibble* properly; they fail to take advantage of previous results because there is no mechanism to support the sharing of this information. *Brainstorming* can be a personal activity but is more commonly used by groups.

It may be observed that these activities are not at the same level of generality. From our viewpoint, wandering is an all-embracing activity, of which consulting, bibbling and brainstorming form possible components. Wandering is in fact similar to the concept of 'social browsing' (Root, 1988); unplanned informal communication that promotes serendipity (Kraut & Galegher, 1990).

The social element of searching that has received the most attention is the reference interview – where a member of the library staff collaborates with the user to help them find the desired information (Taylor, 1968). The aim of this information intermediary is to obtain a clear picture of the information needs of the user so that appropriate help can be given.

In an online situation involving intermediary searching, a definition of the user problem, goals and intentions emerges in the interaction between users and intermediaries during the pre-search interview. (Kuhlthau *et al.*, 1992)

However, the reference interview is just one element in the information search process, and can occur at different stages of the search process for different searchers. Some library users may never ask for professional assistance with their search, and instead may seek 'unofficial' help or just fumble on alone. It is an open question how serious this is in terms of useful information not found.

¹ For students there are three different types of people to consult: other students, academic staff and library staff.

The importance of other users has been noted by several authors as being a valuable source of information:

- At the beginning of the search process ‘the inquirer decides whether to discuss his problem with a colleague or to go to whatever literature or information center may be available’ (Taylor, 1968).
- (Wilson, 1981) has presented a model of information seeking in which other people are seen as an important resource.
- During the ‘selection’ stage of the Information Search Process ‘typical actions are to confer with others’ (Kuhlthau, 1991). In a study of high school students:

Although the role of the librarian was considered as insignificant, the participants frequently turned to informal mediators, including parents, siblings and friends. Many students considered talking about their topics as an important strategy both during and after topic selection. (Kuhlthau, 1993) p. 130

- Consultation with colleagues is used as a means of finding references, e.g. (Sheperd, 1983). As part of the design process for a digital library system, Fox *et al.* conducted intensive interviews with 12 computer science and IR professionals, and found that

... colleagues are especially helpful in providing pointers into the literature, that is, specific references to works likely to be helpful ... colleagues serve as valuable filters; they point to the few best works in an area without providing an exhaustive list of less valuable materials. (Fox *et al.*, 1993)

In Section 3 below we outline an informal study which reveals the existence of a significant level of social interaction in our library at Lancaster University. It is our belief that collaborative actions are central to the information search process and that they deserve to be taken seriously by the information science community. This is indeed becoming an urgent issue with the current move towards the establishment of digital libraries.

2.2 *Digital libraries: threats and opportunities*

Digital libraries are revolutionary in two distinct ways. Firstly, the documents, catalogues, thesauri and searching tools they contain are represented electronically, and can be accessed remotely over computer networks. Secondly, the documents in them are not held only (or perhaps at all) as text strings, but as digitised pages, and may include not only diagrams and tables but also multimedia objects such as animations, video-sequences and executable software. They also contain internal (hypertext) links cross-referencing ideas within a document, and external links to other documents. Many of these features are already familiar to users of the World Wide Web (WWW) (Berners-Lee *et al.*, 1992).

From our point of view, the critical aspect is the first of these, since any trend towards remote searching will make traditional collaborative interactions rarer by losing physical proximity to other searchers. Even today, access from one's own room to the local online public access catalogue (OPAC) system is reducing the opportunities for social interaction, and with the

development of full-blown digital libraries this tendency will be intensified (Ackerman, 1994b; Levy & Marshall, 1995; Marshall *et al.*, 1994; Reich & Weiser, 1993). Unless steps are taken to preserve the social aspects of information searching then much of the (presently unrecognised and under-valued) collaboration amongst library users and staff may be lost. The seriousness of this is highlighted by Levy & Marshall:

... support for communication and collaboration is as important as support for information-seeking activities, and ... indeed, support for the former is needed to support the latter. (Levy and Marshall, 1994)

Nonetheless, although the digital library threatens existing collaborative activities, it also opens up new opportunities that are presently prevented by the physical constraints of traditional libraries. For example, library staff will be enabled to help searchers in any part of the world rather than only those who can make their way to the enquiries desk (Wielhorski, 1994).

Authors writing about digital libraries sometimes comment on how their (proposed) systems may facilitate interaction between users, though none of them appear to regard this as a key issue:

... such features should encourage online "discussion" and alerting among colleagues ... (Hoffman *et al.*, 1993)

Also, readers (of electronic books) can have ... electronic contact with all other readers of each book, thereby sharing ideas and reactions more rapidly and with more people. (Rawlins, 1993)

Kahle *et al.*, interviewing practising accountants who had been trying out the prototype WAISStation system, found that:

Users are very interested in being able to see what queries other users are conducting, and what information servers and articles are most popular. (Kahle *et al.*, 1993)

Most OPACs and databases do not provide mechanisms to support social activities. This can be attributed to an implicit general belief on the part of system designers: *you can only browse for inanimate objects*. We believe that browsing for people, their electronic representations or representations of their activities, is a neglected and important area. A primitive version of this already occurs in use of the WWW where one information seeking tactic is to browse the personal pages of people likely to have pointers to the desired topic. (Chang and Rice, 1993) mention the social aspects of browsing only in informal situations such as 'hallway chatting or after-meeting discussions.' This 'social browsing' (Root, 1988) is held to be an important part of knowledge creation and collaborative work. These social elements are crucial to the design of digital libraries (Kling & Elliot, 1994; Marchionini, 1995) p. 47 but are sometimes submerged by discussions of full-text online documents and multimedia applications.

A further advantage of the switch to a digital medium is the ease with which information can be recorded, stored and re-used. The activities of searchers can be saved and used to aid subsequent searches. The ease of recording of digital activities also enables information to be passed to others who can make use of it even though they were not the original recipient (Maltz,

1994). This information can also be collated into large anonymous collections that can aid in searching (see Section 4).

3. COLLABORATION IN PHYSICAL LIBRARIES

This section examines how collaboration currently occurs within existing systems such as physical libraries. The concept of serendipitous altruism is introduced as an example of collaboration and a CSCW framework is introduced to clarify the issues involved in moving to digital libraries.

In addition to their collection of items, libraries also contain people: lenders, reference users, enquiry staff, technical support staff, bindery staff etc. One can observe and learn from the browsing techniques of others (both at the bookstacks and at the OPAC terminals), discuss issues with co-learners or with subject experts, and also be aware of the activities of others that may be of interest and relevance to one's own work. For example, upon seeing a colleague in an unexpected part of the library, you might choose to ask what she has found there. Similarly, upon seeing someone in 'your' area, you may decide to introduce yourself as someone also interested in that field (Chalmers, 1995; Twidale *et al.*, 1995a).

We can consider three kinds of collaboration:

- collaboration between a user and a member of staff,
- collaboration between library staff, and
- collaboration between library users.

The first, typified by the reference interview, has been heavily investigated in the literature (e.g. (Dervin and Dewdney, 1986; Kuhlthau *et al.*, 1992; Taylor, 1968)). These studies have concentrated on co-located/synchronous interactions, and further work needs to be done to examine how library staff can asynchronously support remote users (Fowell & Levy, 1995; Twidale *et al.*, 1995b; Wielhorski, 1994).

The second kind, between library staff, has had some consideration, but not as much as it should. We are intending to investigate it in future research.

Collaboration between library users has received little attention in physical libraries but is generating some interest in the digital library community (e.g. (Goldberg *et al.*, 1994; Maltz & Erlich, 1995)).

3.1 *User Collaboration Observed*

We have undertaken some preliminary observations of collaboration between users in Lancaster University Library (Twidale & Nichols, 1995). We chose a bay of 12 OPAC terminals in the main reception area of the library. This is a large area containing photocopying machines, and is adjacent to the issue desks, the help desk, the inter-library loan room and the CD-ROM workstations. There are usually considerable numbers of people passing it as it lies between the

entrance and the current serials/short loan section of the library. An observer stood at one of the terminals unobtrusively noting the other 11 terminals; how long a person used one for and particularly checking for instances of collaboration. Ten observation periods of approximately 15 minutes were made at different times over several days during term time.

We observed work-related collaborations (excluding social chat) in 10% of all usage of the OPAC terminals (that is, 10% of sessions of users walking up to a terminal, interacting with it and walking away also involved some collaboration. A group of people all walking up to a single terminal were counted as a single session for this purpose). Similar activity was observed around the CD-ROM bay although at lower rates, possibly because the searchers were each using different resources with different interfaces. These observations are not intended as an accurate measure of such collaboration but merely an indication that it does in fact occur.

We believe that further informal studies in different locations would provide more convincing evidence for collaboration (by triangulation (Hammersley & Atkinson, 1983)) than a formal study situated in the same library.

Several kinds of collaborative interactions between users were observed. We give some scenarios representative of those kinds:

- *Joint Search*

A group of students (2-4) work around a single terminal, discussing their ideas and planning their next actions. The interaction involves frequent pointing at the terminal screen. They are involved in a group-based problem solving task, either one set by their lecturer as a group task or one where they have chosen to collaborate on searching the literature and sharing the material before working on their individual assignments.

- *Coordinated Search*

A group works on two or more adjacent terminals, discussing what they are doing, comparing results, sometimes seemingly competing to find the information. Much leaning over terminals occurs and they may occasionally all cluster around one terminal as in the former case.

- *Free Query*

Individuals working at separate but adjacent terminals occasionally lean over and ask their neighbour for help. Many such requests start "Excuse me, how do I....?" These are the kinds of questions that could just as well be asked of a member of the library staff, but it is much more convenient (and perhaps less embarrassing) to ask a neighbour rather than to stand in a queue at the enquiries desk. In addition of course, the helper can see the context of the questioner's state of working, something that is lost (with existing technology) if the questioner leaves her terminal.

- *Directed Query*

Individuals working at separate terminals can monitor the activity of others. There is a substantial degree of awareness while working in the library. Much of this is social (e.g. noticing friends walking past), but some appears to be an informal monitoring of the activity of others. Occasionally, this leads to a query of the form "How did you do that?" This kind of cooperation seems to be much rarer than the other kinds noted above and naturally occurs most often among friends and acquaintances. For example, when using the CD-ROM terminals, it is easy to see when someone is using the printer, and a novice may take the opportunity to ask how that is done.

- *Chance Contact*

A related situation occurs when patterns of work intersect through a communal resource such as a printer or a photocopier. One observed example occurred when a student printing search results found an uncollected printout and inquired whose it was – when the owner was identified they proceeded to use the results to discuss aspects of the CD-ROM system.

This degree of collaboration is notable not merely because it appears to be largely ignored in the library and information science literature, but also because the context in which it occurs might be expected to reduce the likelihood for collaboration. Libraries are perceived as quiet places where talking is frowned upon and where people go to study on their own. Furthermore, the computerised information systems are designed to support single users. Despite these social and technological constraints, collaborations still do occur (Erlich & Cash, 1994; Levy and Marshall, 1995).

On the other hand there are features of the situation that encourage collaboration. Library staff do endeavour to make libraries more welcoming (Reich and Weiser, 1993). In the case of the Library at Lancaster University, part of this involves the internal layout by providing a large public space where people queue for services (photocopying, book issuing, help desks etc.). As this is near the entrance there is a constant traffic of people in and out of the library affording opportunities to meet friends and colleagues. The OPAC terminals that we observed were in this relatively busy area.

The sharing of knowledge has been observed in several contexts including the learning of spreadsheet skills throughout an office (Nardi & Miller, 1991) and support staff answering customers problems (Erlich and Cash, 1994). Typically these sorts of social interaction are related to physical proximity (Chang and Rice, 1993) and are the first to be threatened by the move to remote searching.

4. COLLABORATION IN INFORMATION SEARCHING

In the previous sections, we have gathered some evidence for the existence and significance of collaboration in libraries. We now present a fuller description of these collaborative activities, both as they presently exist and as they may be affected by future developments.

4.1 *Information searching*

An information search may be regarded as an attempt by a user to move from an initial information state to an enhanced information state. At the outset, the desired enhanced state may be clear and specific, or it may be vague and imprecise (Bates, 1989). In the latter case the searcher expects the aim - that is, the anticipated final state - to become clearer, and perhaps even to change in 'direction', in the course of the search. A search may be said to be complete when the searcher feels no pressing need or desire to continue searching in that direction (there may be a *general* desire, but it is for the time being outweighed by other priorities).

An information search involves a series of interactions between the searcher and any of a variety of information sources, some human and some inanimate (Marchionini, 1995; Wilson, 1981) p. 38. Each useful interaction results in a change in the searcher's information state which tends in the general direction of the search aim. Even during a highly specific and directed search, the searcher is likely to gather information which is not directly connected with the desired aim. Much of this contingent information will be quickly forgotten, but some may be retained to enhance the general knowledge and the searching skills of the searcher.

On occasion, some item of contingent information will be of great interest. We refer to this as a *serendipitous discovery*, which we may define as an unexpected outcome which the searcher finds highly interesting or gratifying. Obviously serendipity cannot be programmed, but in designing new information systems it is important to ensure that it is not actually inhibited (Rice, 1988).

Many searches are *individual* - that is to say, the searcher's aim is not (so far as is known) shared by anyone else. *Group searching* takes place when two or more people share a common aim, and choose to coordinate their searching efforts. For group searching to be efficient, frequent interchanges need to take place between the members concerned. A slightly different case occurs when members of a group are working in the same area - perhaps on the same project - but their specific searching aims are different; we can call this *differentiated group searching*.

4.2 *Interactions during searching*

Some of the interactions which take place during a search are 'inanimate' interactions in which the searcher uses a computer terminal, a book index, a personal notebook or even a card catalogue. But a variety of human interactions can also occur, and it is our contention that these are important. Some of them involve library staff, but some involve other users or friends; interactions of the latter kind can occur entirely outside the library.

In order to focus the discussion, we first list some of the kinds of human interaction which occur during an information search:

- asking a friend or colleague if she knows anything about the topic of interest
- asking if a colleague has a copy of some particular book or journal
- asking an acquaintance who is found browsing in 'one's own' area of the library whether they have an interest in common
- asking a colleague in a group search whether they have found anything of interest
- offering information to a colleague who is known to have a certain interest
- asking a friend or passer-by for help in using, say, an OPAC terminal
- offering advice to a fellow searcher who is seen to be having difficulties
- leaning over to a neighbour to ask how they did something
- observing another user (whether covertly or by invitation) to see how they do something
- asking a member of the library staff how to use some facility
- consulting a search intermediary who will perform a search on one's behalf
- registering as a user, borrowing or returning a book, paying a fine etc.
- asking at the reception desk for directions to some part of the library
- joining a guided tour of the library
- attending a training session on using library systems:

All of these interactions require the sharing or exchange of information, and it is vital that the digital libraries of the future continue to support most of these activities. Roughly the first half of the list involves interactions with fellow users; a computerised library that is accessed remotely will prevent or inhibit many of these interactions unless steps are taken to re-introduce them into the system.

That so many kinds of personal interaction occur in libraries does not imply that existing libraries are ideally adapted to personal interchange. They in fact limit social activities in several ways:

- people must be in the same part of the same physical library
- people must be present at the same time
- people must either already know each other, or be prepared to approach strangers
- the accidental nature of meetings means that conversations may be curtailed: “Sorry, I’ve got to be in a meeting in five minutes”
- some useful items of information may not be to hand: “I’ve got something interesting back in the office”; this increases the acquisition cost of the item

– a follow-up may not be possible because a person fails to note down the name of the person they just met.

– the library environment does not encourage conversation: “please be quiet in the library”

All of these factors contribute to a lack of mutual knowledge amongst browsers; which in turn contributes to a lack of information exchange (Krauss & Fussell, 1990) and numerous missed opportunities for effective collaboration.

4.3 *Serendipitous Altruism*

Colleagues in a community may be willing to help each other's information searching even if they are not directly involved in the project. In the field of software engineering this has been noted for some time as the 'coffee machine phenomenon': software developers stand round the coffee machine discussing their work and colleagues offer suggestions of how to solve a problem (Root, 1988; Weinberg, 1971). With respect to IR activity we believe the equivalent effect can occur, provided structures are in place to support it. We refer to the idea as *serendipitous altruism*.

Serendipity has been long acknowledged within IR as a phenomenon that skilled searchers know how to allow for and to exploit (Cove and Walsh, 1988; Rice, 1988). However it need not just be personal. If your colleagues know what you are working on and happen by accident, in the process of undertaking their own searches, to come across something that may be of interest, they may altruistically pass the information on to you. The expectation is that, in similar circumstances, you will in future return the favour. It is important to note that the cost (in effort) to the benefactor must be minimal. That is why offering advice in a social context over a coffee is effective and why electronic mail and related systems facilitate the forwarding of potentially useful information to those whom the benefactor believes might be interested.

(Maltz and Erlich, 1995) describe three key features around which they developed their system for promoting what they term *active collaborative filtering* within a hypertext system: flexibility, ease of use and contextual packaging (including extra information along with the link itself).

4.4 *A descriptive framework*

We have already distinguished between inanimate and human interactions in information searching. There are in fact several other dimensions along which interactions may be classified, and these are now described.

Research in CSCW classifies collaboration along spatial and temporal dimensions (Rodden, 1991). Thus, collaboration may be classed as *remote* or *co-located*, as well as being *synchronous* or *asynchronous*. In conventional libraries, we can consider most interactions to be co-located and synchronous (see Section 4.3), but the computerisation process makes the other permutations possible, while offering new opportunities for the first. Figure 2 positions a range of typical information-exchange activities in terms of the temporal and spatial dimensions.

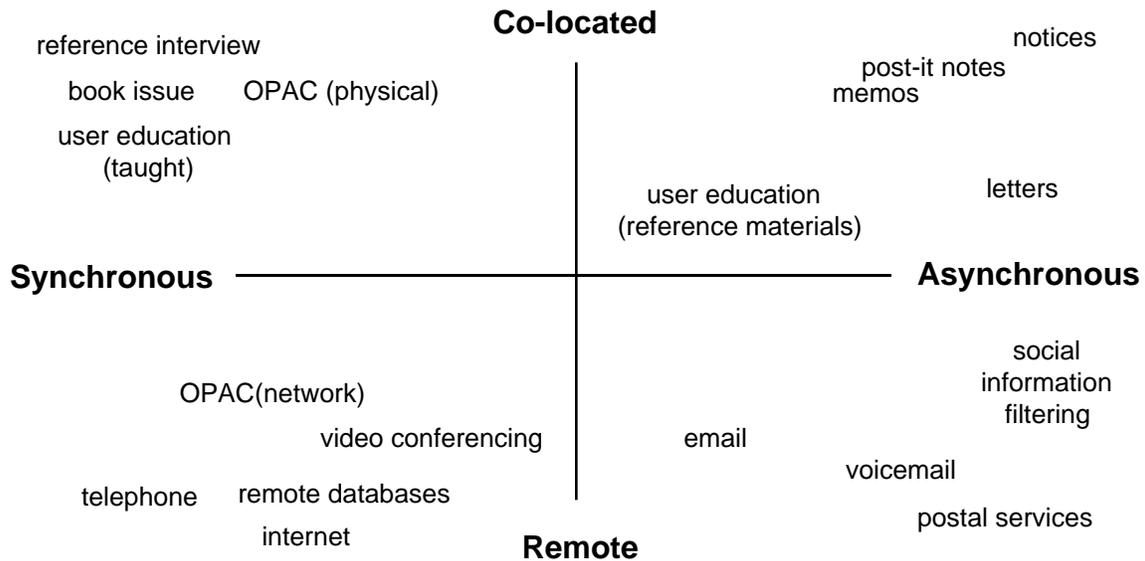


Fig. 2. Typical and potential library activities on spatial and temporal dimensions

Most of the activities in a conventional library take place in the co-located/synchronous quadrant, whereas digital libraries will move many interactions into the remote half. Note that although *access* to remote databases can be regarded as almost synchronous, the whole process of information transfer using them is asynchronous, since the information may have been placed there many years before. Social information filtering systems (Section 5.2) exhibit the same characteristics: information that is used to aid searches was probably gleaned from a remote user a significant time in the past.

Although collaboration is usually regarded as a synchronous activity the ease with which digital information can be stored highlights the importance of asynchronous collaboration. In a sense, asynchronous collaboration is what libraries have always done - selecting and categorising information for their patrons. Librarians have always been implicitly rating and structuring information (Marchionini, 1995) p.189. When a user browses the shelves they are working within the subject classification (e.g. Dewey Decimal) selected by the library. The items they are browsing are those selected by the staff as being worthwhile to acquire: the communication is asymmetrical and asynchronous. The task of finding an item without the structure would be virtually impossible. The satisfaction of the user's goal requires both the user's knowledge of what they are looking for and the librarian's knowledge of structure. This structural knowledge maybe expressed in a face-to-face interview or in the passive structuring of the books..

It is further useful to classify interactions according to whether they are *product-related* or *process-related*; in other words, whether what is exchanged is actual information that the searcher needs, or whether it concerns how to go about finding that information.

We may also classify an interaction according to whether it is *mutual*, which applies when both or all of the people involved hope to obtain some benefit, as when they are involved in some common enterprise, and *instructional*, where help is obtained from some third person, be it a professional librarian or just a friend or passer-by.

The move from physical to digital libraries involves a general move from co-located to remote and from synchronous to asynchronous. The mix among the other factors may remain roughly unchanged, but all must be taken into account in designing future systems, as well as in further formalising models of the browsing process.

5. COLLABORATION IN COMPUTERISED SYSTEMS

The main issue to emerge from our study is that collaborative working implies a need to share information, and that this should include both the *end product* (the hits) and the *process* (the search strategy/tactics).

Computers enable new forms of activity through reducing the costs of communication and facilitating the storage of information that would previously have been lost. These systems have undergone a series of evolutionary steps from single-user systems, to multi-user systems, to networked systems and to CSCW systems. Many diverse collaborative activities can now be supported by computers including some for which this technology is a *necessary* pre-requisite, e.g. (Brewer & Johnson, 1994; Maltz and Erlich, 1995; Shardanand & Maes, 1995).

In this section we discuss various forms of collaborative activity which might be supported in the digital libraries or other extensively computerised information systems of the future. We first consider personal interactions which may occur during the searching process itself (Section 5.1), then we discuss the sharing of the search product (Sections 5.2 and 5.3), and finally the sharing of the search process (Section 5.4). Several of the systems referred to in this section use the term *information filtering* rather than IR, but 'most of the issues which appear at first to be unique to information filtering, are really specializations of IR problems' (Belkin & Croft, 1992).

(Malone *et al.*, 1987), anticipating some of the developments described in this section, describe three different approaches to information filtering: cognitive (or content), economic and social – of which the last is of special interest here:

The social approach to information filtering works by supporting the personal and organizational interrelationships of individuals in a community. (Malone *et al.*, 1987)

5.1 *Performing a search*

'Do you know?' At any time before or during a search, the searcher can ask a known individual for the required information, or for advice or assistance on how to get it. The individual might be a friend or colleague, or might be a member of the library staff. Electronic mail is increasingly used for this kind of request. As well as sending a reply, the respondent may transmit actual documents; at the present time these are likely to be in the form of ASCII texts,

but in future systems it will become normal to send multimedia objects, including executable code.

Does anyone know? If the searcher does not know whom to ask, she may broadcast an appeal to a group of people, many of whom she may not know individually. As in the previous case, electronic mail is suitable for this.

Who might know? If neither of the above tactics can be used, or are unsuccessful, the searcher can try to identify a person who may be able to help.² In this case, the system may hold a profile summarising the interests of a range of human experts. Searching of personal profiles could be provided as a special searching tool, or alternatively the profiles could be held along with representations of inanimate resources in a multimedia searching system.

The Who-Knows system (Streeter & Lochbaum, 1988) allows users to search for groups of people (actually groups of 5-10 people in a large research organisation). Each group is characterised by terms extracted from annual reports produced by the group and the terms from a search statement are compared with the group representations to find the best fit.

Remote help. Problems encountered during a search might be resolved by a special 'help' key to put the user into direct contact with a human advisor. The advisor would need to be able to view not only the current state of the user's screen, but also the sequence of events which led to the current impasse (see Section 5.4). Collaboration with a remote help desk may be by a video link (Sugimoto *et al.*, 1995) or by similar technologies (Donath, 1994).

Coordinated searching. In our observations of the use of OPACs (Section 3.1) we noted cases where two or more users on adjacent terminals were obviously working on related tasks. The individuals concerned would break off to exchange 'news' or to ask questions by word of mouth. Similar interchanges might be allowed between users searching simultaneously on remotely-sited terminals. This would involve not only the passing of messages, but also of screen displays or windows.

(Hoppe & Zhao, 1994) describe a system, C-TORI, which allows synchronous collaborative searching, including cooperative query formulation, cooperative browsing of results and sharing of search histories. This form of collaboration can be extended by interacting in a virtual reality environment where both information and users can be visualised (Benford & Mariani, 1994; Chalmers, 1995).

Brainstorming. This approach might be undertaken by a group when information they needed could not be easily found. Ideas and associations (even quite bizarre ones) are posted to a blackboard or window, and are then rearranged in the hope of obtaining new insights into the problem (Section 2.1). Contribution of the ideas might take place asynchronously, but the critical

² Obviously, the response to an individual query or a broadcast appeal might include a suggestion about who to ask, but what is more interesting here is the possibility of carrying out a computer-based search for a person with specified expertise.

idea generation stage would be most effective as a live session. A number of CSCW tools have been developed to support brainstorming (Applegate *et al.*, 1986; Foster & Stefik, 1986; Hymes & Olson, 1992).

Making contacts. Besides needing help with a currently pressing information need, most professionals and academics are on the look-out for people with similar interests, with whom ideas may be discussed and exchanged. Obviously, broadcast e-mail messages, or personal profile searching, could be used to identify such potential contacts. However, given the continuing nature of this requirement, it makes more sense for the library system itself to keep a look-out for users with shared interests. For example, if a record is kept of books borrowed, or electronic documents inspected, then two users with overlapping 'browsing habits' might be put in contact. Alternatively, if the system maintains 'interest profiles' for many of its users, these could be grouped together using a clustering algorithm. Obviously, identities should only be revealed with the agreement of both the parties concerned.

5.2 *Sharing the search product by recommendation*

In existing libraries and IR systems, the product of a successful search is rarely made available to other users, except perhaps one or two close colleagues. The inescapable result is that many searches are essentially identical to searches which have already been done, perhaps many times over; each time the information is lost, and every new searcher has to 're-invent the wheel' (Bates, 1979b). Users would stand to benefit greatly if the products of past searches could be made available to other interested users.

A system which automatically passes on interesting items to other appropriate users could actually cause annoyance by repeatedly bombarding a user with items that she is already familiar with. Therefore in general, a recommendation should not be made if the prospective recipient is known to have seen the item before.

So far as the beneficiary is concerned the source of shared information may be *identifiable*, *anonymous* or *dispersed*. An identifiable source is an individual whose name is either known to the recipient, or can easily be found out. An anonymous source is an individual whose identity is concealed, for whatever reason (see *Similar searchers* below). The term 'dispersed source' refers to recommendations which are based on aggregated information gathered from a number of unknown individuals (see *Links* and *Use of Ratings* in Section 5.3).

Personal recommendation. A searcher who comes across some interesting items can pass them on to known individuals (Section 4.3) who might be expected to be interested in them. Recommendations of this kind occur today, increasingly through the medium of e-mail, but the fact that the searcher has to remember and take the trouble to do this makes it an unreliable and limited channel of communication.

The use of personal recommendation could be improved by making it an integral part of the searching system. Thus, at the end of an online search, the system could display a list of

associates, and ask if any of them should receive information about the search result. This should avoid the risk of simply forgetting, but still limits the forwarding of information to a short list of possible recipients.

(Maltz and Erlich, 1995) describe a system for filtering Usenet Netnews articles. A reader who finds an interesting document can create a special representative called a 'pointer' which, besides descriptive information, contains a button enabling the full article to be accessed. These representatives can be distributed by e-mail, and also gathered into structured bibliographies called 'information digests'. They refer to this approach as *active collaborative filtering*.

Group recommendation. Passing the information on to all the members of a named interest group increases the coverage, but then it may be of scant interest to many of the people who receive it. There is a clear tradeoff here between coverage and relevance; passing information on to a small group, many of whom are personally known to the recommender, is likely to be more successful than distributing it to thousands of strangers.

Similar searchers. Retrieved information may be passed on to other unknown users whose interest profiles resemble the present searcher. Interest profiles can either be compiled as a deliberate exercise (as in setting up a profile for a old-style current awareness system), by capturing search keys during routine searching activities, or by extracting descriptors from retrieved documents which have been rated highly. This is clearly related to the idea of using user profiles for making contacts (Section 5.1), but in this case the people involved need never know one another. Kantor mentions the possibility of discovering similarities between library users based on their activities in the Adaptive Network Library Interface (ANLI) system (see below) (Kantor, 1993).

The use of profiles containing subject descriptors may seem the obvious method from an IR perspective, but three experimental systems mentioned below use profiles based on user ratings of information items.

5.3 Adding value to documents

Besides systems to facilitate the passing of document references immediately after a search, longer-term benefits may be realised by organising or enhancing the documents themselves. This includes listing documents in bibliographies and adding various kinds of information to the document records.

Releasing an Information Package. Information retrieved during a search may be made available for retrieval by others. The 'retrieved information' could be the product of one search, or it could be a bibliography compiled and updated over a long period. It need not consist only of bibliographic references and text, but could be composed of multimedia objects: hence the more generic phrase *information package*.

An information package can be released by inserting a representative object into an appropriate database. This object must contain a set of searchable descriptors; these may be assigned by the supplier, or derived automatically by analysis of the objects themselves.

The personal bibliographies of a group of cognate users may be merged and shared within the group. This approach underlies the Group Asynchronous Browsing system in which a server merges a set of WWW bookmark files, thus creating an enriched browsing structure for the participating users (Wittenburg *et al.*, 1995).

Links. An information package or bibliography contains a set of documents which are in effect all linked to one another. A different approach is to allow users to insert explicit links from one document to another, which can then be used by others during browsing. Links which are suggested or endorsed by many users, or by users of higher status, can receive higher weights. An example is the ANLI system which provides an additional facility in an online catalogue and allows users to insert links between books (Kantor, 1993; Zhao and Kantor, 1993). The links correspond to a human interaction in which one person expresses interest in book A, and the other replies "In that case you will probably be interested in B as well."

Priming the database by getting a useful set of links into it early on is a problem if only explicit links are used. Possible solutions, which can be used before the system is 'made public', include establishing all-to-all links between documents selected during the same search, or even between books borrowed by a user at the same time (Kantor, 1993).

The ANLI approach of weighted links is similar the NetAgent system (Park & Chon, 1994) where several *trading agents* communicate with users' *personal agents* and resource-specific *information agents*. As queries are entered into the system (via personal agents) the network modifies the associations between the agents to reflect the success or failure of the searches. Trading agents aim to become topic-specific and (Park and Chon, 1994) describe the resulting network as a form of collaborative indexing. Such an agent-based approach allows for the possibility of collaboration between agents that are attempting to find information on behalf of a human user (Lashkari *et al.*, 1995; Maes, 1994).

Annotation. Document representatives in the older established databases are generally prepared to a fairly high standard, and normally include reasonable sets of index terms and other descriptors. By contrast, documents which arrive via electronic networks are often informal in nature and are indexed (if at all) by a crude automatic procedure. One way in which their value and accessibility can be enhanced is by allowing users to attach *annotations* to the documents. These could be free-text comments (Cohen, 1994), additional index terms and judgements of usefulness, or ratings.

In the Tapestry system, developed at Xerox PARC, users can attach annotations to the items they view, including ratings, free text comments and other indicators (Goldberg *et al.*, 1994). Because annotations can be supplied at any time - perhaps even years after receipt of the

document itself - each annotation is stored as a separate document containing a link back to the original document. Users set up standing queries which can refer to annotation fields; thus, they can ask to receive documents which have been endorsed by other named users. They refer to their approach as *collaborative filtering*.

In Tapestry, annotations are attached to a document as a unit. In the ComMentor system (Röscheisen *et al.*, 1995) annotations may be attached at points *within* electronic documents. The annotations are accessed by buttons positioned within the text of the document. Moreover, annotations are associated with groups of users, so that a given reader only sees those annotations relevant to her own group.

In Brewer and Johnson's URN system, 'keywords' are extracted automatically from the title fields of incoming UseNet Netnews articles. Users are able to insert new keywords and delete keywords which they consider unsuitable (Brewer and Johnson, 1994). Annotations on documents can be used for a variety of purposes; including facilitating discussions about the content of the documents (Davis & Huttenlocher, 1995). Annotations can also be used to provide indications of the 'authoritativeness' of a document, as suggested by (Koenig, 1990). Where such indications are placed on a numerical scale they are known as ratings.

Use of Ratings. A rating is an indication of the usefulness, interest or quality of a document as viewed by a user. Ratings may be supplied deliberately by a user (*explicit ratings*), or may be computed by the system on the basis of the 'interest' which users show in a document (*inferred ratings*).

Explicit ratings are usually supplied on a several-point scale (e.g., from 1 to 7); e.g. (Maltz, 1994) uses a scale of *terrible*, *ok*, *good* and *great*. The mechanism for providing the ratings should be as simple as possible, and the user needs to feel that there is some purpose in providing the feedback. The early systems identified the priming problem noted above in *Links*: the benefits to users remain invisible until a large number of ratings have been supplied. Users therefore tend to conclude that the system is ineffective, and stop supplying the ratings. One approach to this problem is to create a population of 'virtual users' who rate one specific topic (e.g. books on cyberspace) highly and ignore everything else (Maes, 1994). In a system in which relevance feedback is routinely used for query enhancement, explicit ratings (albeit usually on a 2-point scale) are available automatically.

Inferred ratings have the advantage that they do not rely on special action by users, but the clues they use are indirect and therefore rather ambivalent. Rating scales cannot be used; instead, the system counts positive actions by the user. Thus, the system may count the number of times a document is opened for reading, relative to the number of times its summary is displayed, or else the average time spent reading a document may be recorded.

One way to use document ratings is to aggregate the ratings for each document into an overall quality score. The rating could be displayed during online inspection of document details, or it

could be used to adjust the likelihood of the item being retrieved during a search. This approach may be useful within a group of people who all share a similar interest, but will be virtually useless for users at large, since many documents are of great interest to a few people and of zero interest to the rest. Another use of ratings is to create filters such as 'show me the articles that Jane Doe liked' (Maltz, 1994); this allows a user to export their expertise to other users.

A much more effective option is to use ratings values to construct a profile for each user, and to use these to identify similar searchers (see above). An example is the GroupLens system (Resnick *et al.*, 1994) which allows collaborative filtering of Usenet Netnews articles by allowing users to assign ratings on a 5-point scale. A profile is constructed which contains all of the document ratings contributed by a user. Users with highly similar profiles are identified, and an item can then be recommended to a user if other similar users have already approved of it. The Ringo³ system uses a similar approach to make personalised recommendations about music albums and artists (Shardanand and Maes, 1995) and uses the term *social information filtering*. (Hill *et al.*, 1995) prefer the term *community of use* in describing a similar explicit rating scheme for recommending videos. (Karlgrén, 1994) has defined a simple 'recommendation algebra' for computing proximities between documents based on user ratings, and has illustrated his method by generating clusters in a collection of Usenet Netnews newsgroup files.

The URN system takes a different approach, by associating the ratings with the individual 'features' (keywords and authors' names) of the documents concerned (Brewer and Johnson, 1994). A user profile consists of a set of features, each with a weight computed from the ratings the user has given to the various documents containing that feature. This means that a new document can be assigned a score reflecting the weights of any profile features it contains. A positive weight means that the features found were mainly favourable, negative that they were unfavourable; a zero weight often means that none of the features in the new document are present in the profile.

Inferred ratings can be regarded as document *wear*; physical objects naturally show signs of wear caused by usage, but many digital objects do not reflect whether (and by whom) they have been used. A variety of *history-enriched digital objects* have been described including *edit wear* and *read wear* (Hill & Hollan, 1992; Hill & Hollan, 1994). That is, a document can record and display information on its history of edits and reading patterns; metadata of the document's history (Böhm & Rakow, 1994). A library database that records and uses the searches of its users can be regarded as a history-enriched digital object⁴. Concepts such as *browse wear* and *borrow wear* could be added to existing history-of-use information (Burrell, 1985) to facilitate user searching rather than just collection management. The open question is: does such an addition 'add more value than it costs to build and maintain?' (Kantor, 1993).

³ Ringo has subsequently changed its name to *HOMR* and then to *firefly*.

⁴ [Hill and Hollan, 1994] also raise some interesting questions regarding the ownership of search history information and the privacy of users, but these are outside the scope of the present review.

5.4 *Sharing the search process*

It is not only the products of searching that need to be saved and shared; in many cases the actual process of searching needs to be preserved. This permits offline debugging of an unsuccessful search, and modification of a search to serve a slightly different purpose. It can also serve a valuable role in the education of searchers.

Saving the search. If a searcher is having difficulties, and instant help is not available, the search may be suspended and a listing of the search session may be saved for later review or 'debugging'. The user can then reflect at leisure on what was done, without the pressure that she may feel during an online search.

A listing of a failed search can also be shown to a librarian, who can offer advice about correcting or modifying it. Without this support, the librarian needs to look over the shoulder of a user as they undertake a search. It is very hard for a novice accurately to recall the steps taken in a search some while earlier. By contrast, if a record of the history of the search is available, the user can approach a librarian and say "I did this (pointing) and I didn't seem to get very much....I'm sure that there must be more stuff there....what am I doing wrong?". The librarian can make quite sophisticated use of a search record. It can reveal gaps in the user's browsing techniques and also provide an indication of their degree of searching sophistication, which can be used in phrasing an explanation at the appropriate level of detail.

Though a printout of a search log is undoubtedly useful, an electronic representation is far superior since, with suitable system support, it can be:

- edited, to delete or correct errors, or to modify its effect; a search can also be edited for educational use, e.g., by eliding portions which are not germane to the point being illustrated;
- annotated, by attaching comments at relevant points in the search object; annotations are valuable for educational and online help applications, and to future users of the search;
- cloned, to provide a separate copy which can be adapted for a new purpose;
- processed, in order to gather statistics and to construct useful visualisations of the search process (see below);
- stored for possible future use;
- communicated to other potential users;
- reactivated, whenever a new search needs to be performed.

The record of a search thus constitutes a structured *search object*, containing not only text but also, perhaps, graphical representations of the search process.

Visualising a search. Whilst saving the mechanical details of a search is obviously important, it may be valuable also to extract and display the broader logical structure, thus providing a

visualisation of the search. (Lin *et al.*, 1991) note that interactive systems have led to study of the processes of searches as well as the products and that:

A major problem for researchers is to represent these data in forms that allow patterns to be easily recognized and analyzed. (Lin *et al.*, 1991)

Whilst this is certainly true, problems of interpretation and analysis exist firstly for *users* and those who seek to help them. Thus, whilst graphical representations may help researchers to understand the problems of users, providing those users with similar representations may well help to eliminate some searching errors in the first place.

The visualisation of a process enables searchers to concentrate on the specific sub-task at hand and removes the cognitive load of maintaining a record of the search history. Cognitive load is held to be an important reason why much user behaviour is inefficient (Rudd & Rudd, 1986). A well designed visual representation could help the searcher to detect and repair errors and reflect on aspects of the search process. Such a facility would also be useful as an educational tool, giving users a clearer picture of the overall 'anatomy' of a successful (or unsuccessful) search (see Section 6 for an example).

Sharing a profile. A profile can be regarded as an abstract from and partial form of a search process. Although the end-result (a set of 'hits') may be the same, the profile usually achieves this without explicit user input. Re-running a complete search sequence which was originally performed in interactive mode is not necessarily an efficient method if the search is to be executed many times. In this case, the search may be consolidated into a stored query or profile.

Computer-based current awareness systems, in which an established user profile is matched against each new release of a database, have been in use for more than 35 years (Tritschler, 1962). Traditionally, profiles were constructed during an interview with an intermediary, or by having the searcher fill in a form. In a modern system, an initial profile can be constructed by capturing and storing the searching activities of a new user. The effectiveness of a profile can be improved, and maintained over time, by using feedback judgements from the user (Salton & McGill, 1983).

Current awareness searching is not in itself a collaborative activity. However, a team who are working on the same project could share a single profile, so that all would be alerted simultaneously when the system found a new relevant item. Moreover, a new researcher starting work in an area closely related to an existing project could be supplied with a copy of the existing profile, which could then be 'personalised' using relevance feedback.

5.5 *Comments*

In the preceding paragraphs we have enumerated various information-sharing activities which might be supported in a computer-based library system. We have not attempted to offer detailed prescriptions, since at present not enough is known about which methods are effective, nor about how the various functions relate to one another. For instance we do not know whether it is more

effective to share the search product through user annotation of documents, or by identifying similar searchers (Section 5.2). It may well be that the answer varies from case to case, depending on the source, type, subject, amount and heterogeneity of the information, and the nature of the user population. It may be also that different methods need to be somehow integrated to give the best overall performance.

The design and introduction of computer supported collaboration systems are subject to two kinds of cost-benefit tradeoff. First, from the institutional viewpoint the question to be asked is whether the extra cost of collaboration technology is warranted by the putative benefits. Since the benefits cannot be predicted with any accuracy, the likelihood is that innovations will be introduced in a limited and piecemeal fashion. This makes it important that these systems be designed with extensibility in mind, so that each advance does not later become a cause of regret and frustration.

From the system designer's viewpoint, a critical issue is whether the user will feel that the effort required to supply recommendations and ratings is compensated by the extra benefit obtained from the system. A particular problem is to ensure that a new, possibly sceptical, user obtains benefits almost immediately.

The ideal is to allow effortless participation based on inferred ratings and search histories (Hill and Hollan, 1994) but it is likely that in general more accurate assistance can be provided by a system which uses explicit user feedback.

No doubt the most effective method for computer supported collaboration will vary according to the nature of the information concerned. Thus, a system based on accumulated ratings requires:

- that the information will remain valid over a significant period of time, and
- that the number of information items is not too large in relation to the number of users

With long-lived databases, the question arises whether annotations and ratings should be 'forgotten' after a certain period of time. At the other extreme, items which remain valid for only a few days (such as Usenet NetNews articles) probably need to be handled by feature-based systems (Foltz & Dumais, 1992).

Evaluation of the various tools and methods is obviously a major issue. For the few experimental systems described in the recent literature, the 'evaluation' has been rather limited. There is a serious need for investigations in which different approaches can be compared within a fixed setting, using objective performance criteria. (Resnick *et al.*, 1994) provide some relevant suggestions.

Many of the points presented in 5.2 and 5.3 fall under the theme of *community* (or *organisational*) *memory* (Ackerman, 1994a; Walsh & Ungson, 1991): the preservation and dissemination of the knowledge (and activities) of members of the community. At present much information is simply lost as no steps are taken to preserve it; thus, many activities simply repeat previous actions and fail to utilise anything learnt about the activity.

6. ARIADNE

Based on our studies of existing, spontaneous collaboration we have developed, and are refining, a system to support the collaborative learning of database browsing skills (Twidale *et al.*, 1995b). The system, ARIADNE, was intended to serve two purposes:

- To allow us to observe, record, analyse and experiment with the collaborative learning process. Studies of the use of ARIADNE such as those outlined below will provide us with more detailed information about the nature of the process we wish to support. This information will be complementary to that obtained from situated observations of collaboration in libraries as exemplified by the informal studies noted above.
- To provide a system which enhances the opportunities and effectiveness of the collaborative learning that already occurs. We also want to include facilities that will allow collaborations to persist as people increasingly search information remotely and so have less opportunity for spontaneous face to face collaboration.

6.1 System Description

The ARIADNE system captures the users' input and the database's output and forms them into a search history consisting of a series of *command–output* pairs. This means data capture is done transparently, so that users can work as if they were interacting directly with their chosen database. It is only in the subsequent playback phase that the new form of working becomes available.

The simplicity of the approach ensures that it can be used for *any* text based interface for *any* library for which remote access via Telnet is possible. This separation between capture and display is necessary because of the lack of separation between the user interface of a database and the database functionality itself. Therefore, the separation we provide allows the graphical ARIADNE interface to work both with data captured in this way and (potentially) through other methods (e.g. by the Z39.50 protocol). Although the current version of the system only produces the visualisation after a search has been completed, we intend to extend it to allow the history to be dynamically created and presented in real time.

The visualisation of the search process (as illustrated in Fig. 3) consists of a sequence of the command-output pairs looking like small playing cards and containing thumbnails of the screendump of the outputs. These can be expanded to full size by clicking on them. The sequence of cards through time progresses horizontally while the vertical position of the card on one of three rows gives information on the semantics of the action it represents: choices from top level menus on the top row, specifying a search in the middle row, looking at particular book details on the bottom row. The aim is to give an impression of 'diving' into a database by composing queries and going down to actual data entries. A session consists of numerous 'dives' into detail, interspersed by 'higher level' activities of composing and combining searches, selecting display options etc. This visualisation makes it easy to pick out certain characteristic patterns of

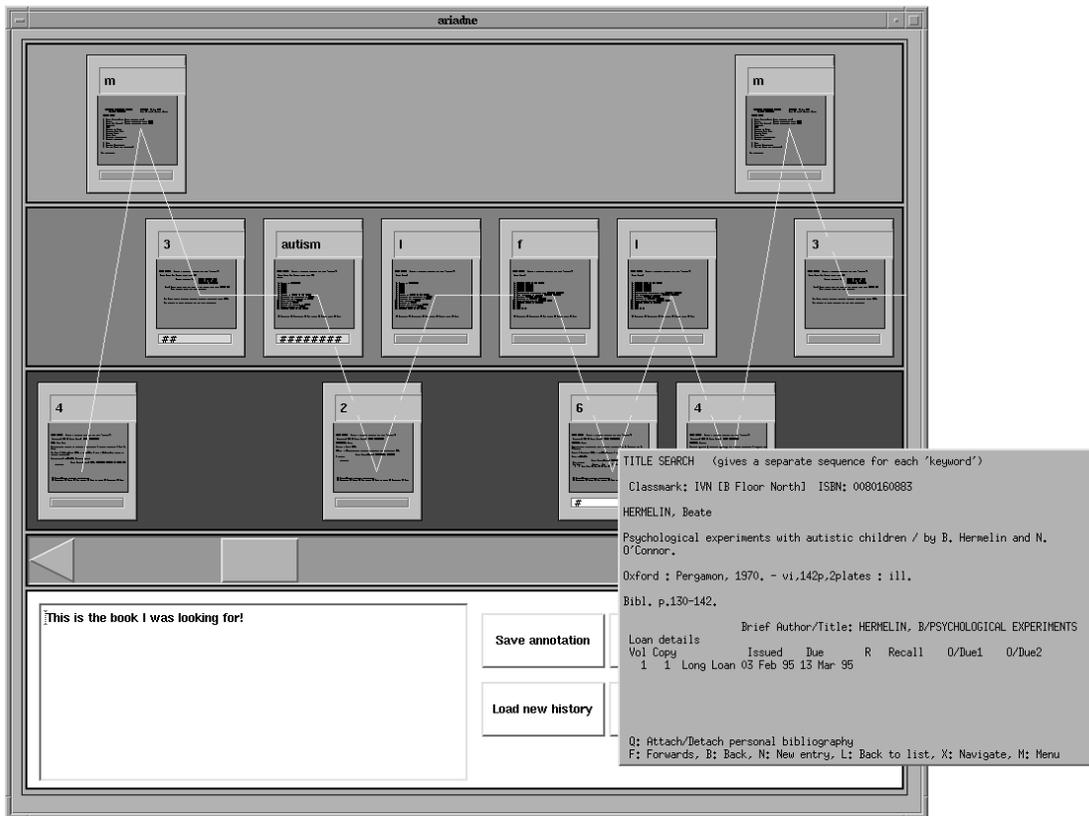


Fig. 3. A search visualisation in ARIADNE

behaviour such as reading through large numbers of hits rather than redefining a query. This externalisation of the search process into a digital object allows it to be annotated, discussed with colleagues around the screen and distributed to remote collaborators for asynchronous commenting. ARIADNE thus provides features for saving and sharing the search process and visualising a search as described in Section 5.4.

A session with a database naturally may yield a large number of command-output pairs. In addition to the use of scrolling, ARIADNE also provides the opportunity to elide parts of the search to allow users and collaborators to focus on particular areas of a search. A sequence of the search history may be 'folded' by selecting the first and last items. All the cards in the sequence are then replaced in the visualisation by a special fold card. Multiple folds are permitted and folds may be nested within further folds to create hierarchies serving a similar simplifying purpose to the use of folders in file systems. A card (including a fold card) may be annotated conferring the benefits noted in Section 5.3.

Testing of the system has involved the use of volunteers who were invited to bring a searching problem that they already had to solve, rather than our imposing a standardised problem upon them. We believe that this more authentic form of testing avoids the dangers of implicit assumptions about envisaged use being embodied both in the system design and in the evaluation

Table 2. Terms and distinctions (with Section numbers)

Searching may be performed by:

- An individual or a group or a differentiated group (4.1)

An information resource may be:

- human or inanimate (4.2)

Search interactions may be:

- remote or co-located (4.4)
- synchronous or asynchronous (4.4)
- product related or process related (4.4)
- mutual or instructional (asymmetrical) (4.4)

The source of the information may be:

- identifiable or anonymous or dispersed (5.2)

Ratings of documents may be:

- explicit or inferred (5.3)
-

task. Students from a wide range of academic backgrounds (including Psychology, Computing, Women's Studies, Chemistry, Religious studies and Environmental Science) have used the system. The testing has informed the iterative development of the system. The results were promising, not least in that complete novices were able to understand the concept of the visualisation and found it useful in informing their understanding of the search process, and in discussing this understanding with colleagues. The visualisation reduces the potential for misinterpretation and makes it easier to check for remaining misinterpretations and to remediate them through pointing and discussion.

To date the system has been chiefly tested for the support of synchronous co-located collaborations. Preliminary investigations have been made of other forms of collaboration. An unexpected advantage of the visualisation was in serving as a recap for individuals and groups returning for subsequent searching sessions of what they had done earlier. An experimental system has been developed for allowing annotated search histories to be sent to other participants for comment and further annotation.

7. CONCLUSION

In the course of this review we have introduced a number of relevant terms and distinctions, these are summarised in Table 2.

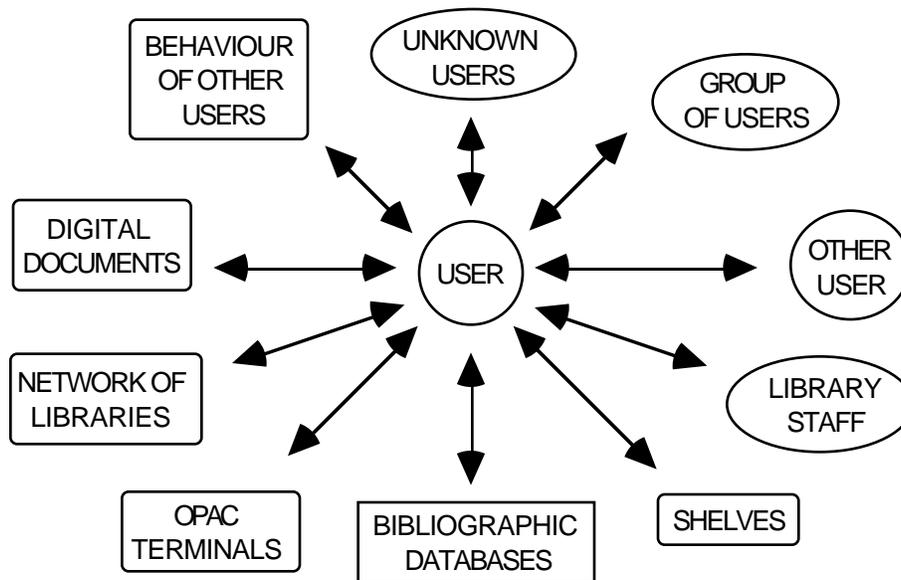


Fig. 4. The social context of information seeking activity

The main conclusions drawn from our investigation are:

- collaboration between users, and between users & system personnel, is a significant element of searching in current information systems.
- the development of electronic libraries threatens existing forms of collaboration but also offers opportunities for new forms of collaboration.
- the sharing of both the search *product* and the search *process* are important for collaborative activities (including education of searchers)
- there exists great potential for improving search effectiveness through the re-use of previous searches; this is one mechanism for adding value to existing databases.
- browsing is not restricted to browsing for inanimate objects; browsing for people is also possible and could be a valuable source of information
- searchers of databases need externalised help to reduce their cognitive load during the search process. This can be provided both by traditional paper-based technology and through computerised systems.

Figure 4 shows the variety of resources used by information seekers; note how it is much less structured than Fig. 1. We are still learning how information seeking activity happens in the digital library and how it co-evolves with the available functionality and as users learn new techniques and develop new ways of working. However the complexity of the task implies that even greater cooperation will be needed than in traditional libraries.

The information supply industry is in a state of ferment, and no doubt this review will be out of date as soon as it is published. Nonetheless, we hope that the discussion here will serve both to emphasise the importance of considering the social use of information resources and to give an initial order to analysing the many varied and creative ways that system developers are addressing the collaborative process of browsing.

Acknowledgements — Financial support for this research came from the Joint Information Systems Committee under the New Technologies Initiative programme. The ARIADNE system was developed by Gareth Smith and Jonathan Trevor.

REFERENCES

- Ackerman, M.S. (1994a). Augmenting the organizational memory: a field study of Answer Garden. *Proceedings of Conference on Computer Supported Cooperative Work (CSCW'94)*, (pp. 243-252). Chapel Hill, NC.
- Ackerman, M.S. (1994b). Providing social interaction in the digital library. *Proceedings of Digital Libraries '94*, (pp. 198-200). College Station, TX.
- Applegate, L.M., Konsynski, B.R. & Nunamaker, J.F. (1986). A group decision support system for idea generation and issue analysis in organizational planning. *Proceedings of the Conference on Computer Supported Cooperative Work*, (pp. 16-34). Austin, TX.
- Apted, S.M. (1971). General purposive browsing. *Library Association Record*, 73(12), 228-230.
- Ayris, P. (1986). *The Stimulation of Creativity: a Review of the Literature Concerning the Concept of Browsing, 1970-85*. CRUS Working Paper, No. 5, Consultancy and Research Unit, Department of Information Studies, University of Sheffield, UK.
- Bates, M.J. (1979a). Idea tactics. *Journal of the American Society for Information Science*, 30(5), 281-289.
- Bates, M.J. (1979b). Information search tactics. *Journal of the American Society for Information Science*, 30(4), 205-214.
- Bates, M.J. (1989). The design of browsing and berrypicking techniques for the online search interface. *Online Review*, 13(5), 407-424.
- Beheshti, J. (1992). Browsing through public access catalogs. *Information Technology and Libraries*, 11(3), 220-228.
- Belkin, N.J. & Croft, W.B. (1992). Information filtering and information retrieval: two sides of the same coin? *Communications of the ACM*, 35(12), 29-38.
- Benford, S. & Mariani, J. (1994). Virtual environments for data sharing and visualisation -- populated information terrains. In Sawyer, P. (Ed.), *Interfaces to Database Systems (IDS94): Proceedings of the 2nd International Workshop on Interfaces to Database Systems*, (pp. 168-182). New York: Springer-Verlag.
- Berners-Lee, T.J., Cailliau, R., Groff, J.-F. & Pollerman, B. (1992). World-Wide Web: the information universe. *Electronic Networking: Research, Applications and Policy*, 2(1), 52-58.
- Böhm, K. & Rakow, T.C. (1994). Metadata for multimedia documents. *SIGMOD Record*, 23(4), 21-26.
- Brewer, R.S. & Johnson, P.M. (1994). *Toward Collaborative Knowledge Management within Large, Dynamically Structured Systems*. Technical Report CSDL-TR-94-02, Collaborative Software Development Laboratory, Department of Information and Computer Sciences, University of Hawaii.

- Burrell, Q. (1985). A note on ageing in a library circulation model. *Journal of Documentation*, 41(2), 100-115.
- Chalmers, M. (1995). Design Perspectives in Visualising Complex Information. In Spaccapietra, S. and Jain, R. (Eds.), *Visual Database Systems 3: Visual Information Management*, (pp. 103-111). New York: Chapman & Hall.
- Chang, S.J. & Rice, R.E. (1993). Browsing – a multidimensional framework. *Annual Review of Information Science and Technology*, 28, 231-276.
- Cohen, E.S. (1994). Review-based information services: lessons learned from the Boston restaurant list. *Proceedings of the Second International World Wide Web Conference '94*, Chicago, IL.
- Cove, J.F. & Walsh, B.C. (1988). Online text retrieval via browsing. *Information Processing & Management*, 24(1), 31-37.
- Davis, J.R. & Huttenlocher, D.P. (1995). Shared annotation for cooperative learning. *Proceedings of Computer Support for Collaborative Learning (CSCL 95)*, (pp. 84-88). Bloomington, Indiana.
- Dervin, B. & Dewdney, P. (1986). Neutral questioning: a new approach to the reference interview. *Reference Quarterly*, 25(4), 506-513.
- Dervin, B. & Nilan, M.S. (1986). Information needs and uses. *Annual Review of Information Science and Technology*, 21, 3-33.
- Donath, J.S. (1994). Casual collaboration. *Proceedings of the International Conference on Multimedia Computing and Systems (ICMCS)*, Boston, MA.
- Erlich, K. & Cash, D. (1994). Turning information into knowledge: information finding as a collaborative activity. *Proceedings of Digital Libraries '94*, (pp. 119-125). College Station, TX.
- Foltz, P.W. & Dumais, S.T. (1992). Personalized information delivery: an analysis of information filtering methods. *Communications of the ACM*, 35(12), 51-60.
- Foster, G. & Stefik, M. (1986). Cognoter: theory and practice of a Colab-orative tool. *Proceedings of the Conference on Computer-Supported Cooperative Work (CSCW '86)*, (pp. 7-15). Austin, TX.
- Fowell, S. & Levy, P. (1995). Developing a new professional practice: a model for networked learner support in higher education. *Journal of Documentation*, 51(3), 271-280.
- Fox, E.A., Hix, D., Nowell, L.T., Brueni, D.J., Wake, W.C., Heath, L.S. & Rao, D. (1993). Users, user interfaces, and objects – Envision, a digital library. *Journal of the American Society for Information Science*, 44(8), 480-491.
- Goldberg, D., Nichols, D., Oki, B.M. & Terry, D. (1994). Using collaborative filtering to weave an information tapestry. *Communications of the ACM*, 35(12), 61-70.
- Greene, R.J. (1977). The effectiveness of browsing. *College & Research Libraries*, 38(4), 313-316.
- Hammersley, M. & Atkinson, P. (1983). *Ethnography Principles in Practice*. London: Routledge.
- Hancock-Beaulieu, M. (1989). Online catalogues: a case for the user. In Hildreth, C.R. (Ed.), *The Online Catalogue: developments and directions*, (pp. 25-46). London: The Library Association.
- Hildreth, C.R. (1982). Online browsing support capabilities. *Information interaction: Proceedings of the the 45th ASIS Annual Meeting*, (pp. 127-132). White Plains, NY.
- Hill, W.C. & Hollan, J.D. (1992). Edit wear and read wear. *Proceedings of the Conference on Human Factors in Computing Systems. CHI'92*, (pp. 3-9). Monterey, CA.
- Hill, W.C. & Hollan, J.D. (1994). History-enriched digital objects: prototypes and policy issues. *The Information Society*, 10(2), 139-145.
- Hill, W.C., Stead, L., Rosenstein, M. & Furnas, G. (1995). Recommending and evaluating choices in a virtual community of use. *Proceedings of the Conference on Human Factors in Computing Systems (CHI'95)*, (pp. 194-201). Denver, CO.

- Hoffman, M.M., Ogorman, L., Story, G.A., Arnold, J.Q. & MacDonald, N.H. (1993). The Rightpages (TM) service – an image-based electronic library. *Journal of the American Society for Information Science*, 44(8), 446-452.
- Hoppe, H.U. & Zhao, J. (1994). C-TORI: an interface for cooperative database retrieval. In Karagiannis, D. (Ed.), *Database and Expert Systems Applications*, (pp. 103-113). Berlin: Springer-Verlag.
- Hymes, C.M. & Olson, G.M. (1992). Unblocking brainstorming through the use of a simple group editor. *Proceedings of Conference on Computer-Supported Cooperative Work (CSCW'92)*, (pp. 99-106). Toronto, Canada.
- Kahle, B., Morris, H., Goldman, J., Erickson, T. & Curran, J. (1993). Interfaces for distributed systems of information servers. *Journal of the American Society for Information Science*, 44(8), 453-467.
- Kantor, P.B. (1993). The adaptive network library interface: a historical overview and interim report. *Library Hi-Tech*, 11(3), 81-92.
- Karlgren, J. (1994). *NewsGroup Clustering Based on User Behavior – a Recommendation Algebra*. Technical Report T94-04, Swedish Institute of Computer Science.
- Kling, R. & Elliot, M. (1994). Digital library design for organizational usability. *SIGOIS Bulletin*, 15(2), 59-70.
- Koenig, M.E.D. (1990). Linking library users: a culture change in librarianship. *American Libraries*, 21(9), 844-849.
- Krauss, R.M. & Fussell, S.R. (1990). Mutual knowledge and communicative effectiveness. In Galegher, J., Krauss, R.M. and Egidio, C. (Eds.), *Intellectual Teamwork: social and technological foundations of cooperative work*, (pp. 111-145). Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Kraut, R.M. & Galegher, J. (1990). Patterns of contact and communication in scientific research collaboration. In Galegher, J., Krauss, R.M. and Egidio, C. (Eds.), *Intellectual Teamwork: social and technological foundations of cooperative work*, (pp. 149-171). Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Kuhlthau, C., Spink, A. & Cool, C. (1992). Exploration into stages in the information search process in online information retrieval: communication between users and intermediaries. *Proceedings of the ASIS Annual Meeting (Vol. 29)*, (pp. 67-71).
- Kuhlthau, C.C. (1991). Inside the search process: information seeking from the user's perspective. *Journal of the American Society for Information Science*, 42(5), 361-371.
- Kuhlthau, C.C. (1993). *Seeking Meaning: a process approach to library and information services*. Norwood, NJ: Ablex.
- Lashkari, Y., Metral, M. & Maes, P. (1995). Collaborative interface agents. *Proceedings of AAAI'94*, Seattle, WA.
- Levy, D.M. & Marshall, C.C. (1994). What color was George Washington's white horse? A look at assumptions underlying digital libraries. *Proceedings of Digital Libraries '94*, (pp. 163-169). College Station, TX.
- Levy, D.M. & Marshall, C.C. (1995). Going digital: a look at assumptions underlying digital libraries. *Communications of the ACM*, 38(4), 77-84.
- Lin, X., Liebscher, P. & Marchionni, G. (1991). Graphical representations of electronic search patterns. *Journal of the American Society for Information Science*, 42(7), 469-478.
- Maes, P. (1994). Agents that reduce work and information overload. *Communications of the ACM*, 37(7), 31-40.
- Malone, T.W., Grant, K.R., Turbak, F.A., Brobst, S.A. & Cohen, M.D. (1987). Intelligent information sharing systems. *Communications of the ACM*, 30(5), 390-402.
- Maltz, D. & Erlich, K. (1995). Pointing the way: active collaborative filtering. *Proceedings of the Conference on Human Factors in Computing Systems. CHI'95*, (pp. 202-209). Denver, CO.
- Maltz, D.A. (1994). *Distributing Information for Collaborative Filtering on Usenet Net News*. MSc Thesis, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology.

- Mann, M. (1986). *Browsing: an annotated bibliography*. Centre for Library and Information Management Report, No. 53, Department of Library and Information Studies, Loughborough University.
- Marchionini, G. (1995). *Information Seeking in Electronic Environments*. Cambridge, UK: Cambridge University Press.
- Marshall, C.C., Shipman, F.M. & McCall, R.J. (1994). Putting digital libraries to work: issues from experience with community memories. *Proceedings of Digital Libraries '94*, (pp. 126-133). College Station, TX.
- Morris, R.C.T. (1994). Toward a user-centered information service. *Journal of the American Society for Information Science*, 45(1), 20-30.
- Morse, P.M. (1970). Search theory and browsing. *Library Quarterly*, 40(4), 391-408.
- Nardi, B.A. & Miller, J.R. (1991). Twinkling lights and nested loops: distributed problem solving and spreadsheet development. *International Journal of Man Machine Studies*, 34(2), 161-184.
- O'Connor, B. (1993). Browsing: a framework for seeking functional information. *Knowledge: creation, diffusion, utilization*, 15(2), 211-32.
- Paice, C.D. (1986). Expert systems for information retrieval. *Aslib Proceedings*, 38(10), 343-353.
- Park, T. & Chon, K. (1994). Collaborative indexing over networked information resources by distributed agents. *Distributed Systems Engineering*, 1(6), 362-374.
- Rawlins, G.J.E. (1993). Publishing over the next decade. *Journal of the American Society for Information Science*, 44(8), 474-479.
- Reich, V. & Weiser, M. (1993). Libraries are more than information: situational aspects of electronic libraries. *Serials Review*, 20(3), 31-37.
- Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P. & Riedl, J. (1994). GroupLens: an open architecture for collaborative filtering of netnews. *Proceedings of Conference on Computer Supported Cooperative Work (CSCW'94)*, (pp. 175-186). Chapel Hill, NC.
- Rice, J. (1988). Serendipity and holism: the beauty of opacs. *Library Journal*, 113(3), 138-141.
- Rodden, T. (1991). A survey of CSCW systems. *Interacting with Computers*, 3(3), 319-354.
- Rodden, T., Mariani, J.A. & Blair, G.S. (1992). Supporting cooperative applications. *Computer Supported Cooperative Work*, 1(1), 41-67.
- Root, R.W. (1988). Design of a multi-media vehicle for social browsing. *Proceedings of the Conference on Computer-supported Cooperative Work (CSCW '88)*, (pp. 25-38). Portland, OR.
- Röscheisen, M., Morgensen, C. & Winograd, T. (1995). Beyond browsing: shared comments, SOAPs, trails, and online-communities. *Computer Networks and ISDN Systems*, 27(6), 739-749.
- Ross, J. (1983). Observations of browsing behaviour in an academic library. *College & Research Libraries*, 44(4), 269-276.
- Rudd, J. & Rudd, M.J. (1986). Coping with information load: strategies and implications for librarians. *College & Research Libraries*, 47(4), 315-322.
- Salton & McGill (1983). *Introduction to Modern Information Retrieval*. New York: McGraw-Hill.
- Shardanand, U. & Maes, P. (1995). Social information filtering: algorithms for automating "word of mouth". *Proceedings of the Conference on Human Factors in Computing Systems (CHI'95)*, (pp. 210-217). Denver, CO.
- Sheperd, J. (1983). *The Information Needs and Information Seeking Behaviour of Polytechnic History Lecturers*. British Library Research and Development Department Report No. 5743.
- Streeter, L.A. & Lochbaum, K.E. (1988). Who Knows: a system based on automatic representation of semantic structure. *Proceedings of RIAO 88 Conference on User-Oriented Content-Based Text and Image Handling*, Vol 1, (pp. 380-388). Cambridge, MA.
- Sugimoto, S., Gotou, S., Zhao, Y., Sakaguchi, T. & Tabata, K. (1995). Enhancing usability of network-based library information system – experimental studies of a user interface for

- OPAC and of a collaboration tool for library services. *Proceedings of Digital Libraries '95*, Austin, TX.
- Taylor, R.S. (1968). Question-negotiation and information seeking in libraries. *College & Research Libraries*, 29(3), 178-194.
- Tritschler, R.J. (1962). A computer-integrated system for centralized information dissemination, storage and retrieval. *Aslib Proceedings*, 14, 473-503.
- Twidale, M.B. & Nichols, D.M. (1995). *Situated Observations of Library Users*. Internal Report, Computing Department, Lancaster University.
- Twidale, M.B., Nichols, D.M., Mariani, J.A., Rodden, T. & Sawyer, P. (1995a). Supporting the active learning of collaborative database browsing techniques. *Association for Learning Technology Journal*, 3(1), 75-79.
- Twidale, M.B., Nichols, D.M., Smith, G. & Trevor, J. (1995b). Supporting collaborative learning during information searching. *Proceedings of Computer Support for Collaborative Learning (CSCL 95)*, (pp. 367-374). Bloomington, Indiana.
- Walsh, J.P. & Ungson, G.R. (1991). Organizational memory. *Academy of Management Review*, 16(1), 57-91.
- Watters, C. & Shepperd, M.A. (1994). Shifting the information paradigm from data-centered to user-centered. *Information Processing & Management*, 30(4), 455-471.
- Weinberg, G.M. (1971). *The psychology of computer programming*. New York: Van Nostrand Reinhold.
- Wielhorski, K. (1994). Teaching remote users how to use electronic information resources. *The Public-Access Computer Systems Review*, 5(4), 5-20.
- Wilson, T.D. (1981). On user studies and information needs. *Journal of Documentation*, 37(1), 3-15.
- Wittenburg, K., Das, D., Hill, W.C. & Stead, L. (1995). Group asynchronous browsing on the World Wide Web. *Proceedings of the Fourth International World Wide Web Conference*, Boston, MA.
- Zhao, S.Y. & Kantor, P.B. (1993). Development of an adaptive network library interface - progress report and system design issues. *Proceedings of the ASIS Annual Meeting (Vol. 30)*, (pp. 211-216).