# Recommendation, Collaboration and Social Search

David M. Nichols
University of Waikato, New Zealand

Michael B. Twidale
University of Illinois at Urbana-Champaign, USA

## Introduction

This chapter considers the social component of interactive information retrieval: what is the role of other people in searching and browsing? For simplicity we begin by considering situations without computers. After all, you can interactively retrieve information without a computer; you just have to interact with someone or something else. Such an analysis can then help us think about the new forms of collaborative interactions that extend our conceptions of information search, made possible by the growth of networked ubiquitous computing technology.

Information searching and browsing have often been conceptualized as a solitary activity, however they *always* have a social component. We may talk about 'the' searcher or 'the' user of a database or information resource. Our focus may be on individual uses and our research may look at individual users. Our experiments may be designed to observe the behaviors of individual subjects. Our models and theories derived from our empirical analyses may focus substantially or exclusively on an individual's evolving goals, thoughts, beliefs, emotions and actions. Nevertheless there are always social aspects of information seeking and use present, both implicitly and explicitly.

We start by summarizing some of the history of information access with an emphasis on social and collaborative interactions. Then we look at the nature of recommendations, social search and interfaces to support collaboration between information seekers. Following this we consider how the design of interactive information systems is influenced by their social elements.

## Background

The history of organized information is, necessarily, one of social relationships. The organizers structured documents to ease retrieval, possibly for themselves in the first instance, but then for users who wished to access the collection. Personal libraries, organizational libraries, then public libraries provided access to information resources hundreds of years before computers arrived. These libraries contained, or indeed embodied, two main social relationships between the librarians and the users. The simplest relationship was a dialogue between the information seeker and the librarian, which has evolved into the reference interview. The other relationship was indirect, via the structure of the collection; where the librarians arranged the items to facilitate retrieval. These 'arrangements' have evolved into classification systems such as the Dewey Decimal Classification. As the literary record grew through books and journals the connections between published works, authors, readers and information seekers have become ever more complex.

In a pre-computer world of entirely physical libraries many kinds of recommendation existed:

- citations
- book, movie and product reviews
- librarians' suggestions
- peers' suggestions
- anthologies
- bibliographies
- best seller charts
- curated collections: special libraries, bookshelves of other people, a course syllabus, other reading lists

- the professor's or the lab filing cabinet of papers, preprints, tech reports
- senior students' literature reviews, boxes of index cards

and no doubt many more. It can be informative to attempt to add to this list. Physical analogues have inspired computational features to facilitate collaboration. This can be a useful heuristic to inspire future innovations.

The advent of digital resources and functionalities to access those resources presented new challenges. In addressing these challenges, aspects of the sociality of information seeking seem to have been overlooked: "Like much of the information-seeking literature, this work [on models of search] overlooks the role of other individuals in search, instead focusing on the search act from a single user's perspective" (Evans and Chi, 2010).

One possible explanation for this individual focus was that, at least implicit in the design of many pioneer applications, there was an idealization of individual use and a view of other people as a 'problem' to be minimized rather than an opportunity to be embraced. Thus, an online database is a shared resource, but other users are treated as a necessary evil in order to achieve economies of scale. These other users are indeed in part a complicating factor - what if there is a contention in more than one person trying to update an item at the same time? What if other people's use of the database degrades performance? In classic engineering style a problem was identified and effort applied at optimization - how to minimize the impact for an individual user of all the other users contending for access to extremely limited computational resources?

Very impressive progress was made in this engineering endeavor. But some of the advantages of sociality were lost. In the physical world, many of these 'social side-effects' are integral to the physical world - you get them 'for free'. In the digital world if they are to exist, they must be deliberately designed into the functionality and the interface. For example, a much-borrowed library book looks battered when on the shelf next to a rarely borrowed one. On picking it up it falls open at the most consulted pages. To the frustration of librarians, it may have pages dog-eared and contain annotations. By comparison, a book in a digital library remains pristine. Bits do not get battered through use, but there is no immediate way to gauge a digital book's popularity by its state of decay.

In recent years, as the challenges of database design have been addressed and as the costs of processing power, storage, interconnection and bandwidth have fallen, we have had the opportunity to reintroduce explicit support for sociality. In many ways this has been a recovery of features already available in a physical setting. However a digital version can allow for greater reach across space, time and resource sets. For example, you may create a set of items (research papers, web pages or photographs) that you organize, classify, rate, tag or make notes on for your own benefit. But various applications make it possible, even easy, to share that with a few other people or with everyone. The organizational work you do mostly for your own benefit has the potential of enabling others to access the information that they want more easily.

Research aimed at explicitly adding different aspects of sociality to information retrieval systems has used a variety of labels, including collaborative filtering, recommender systems, collaborative browsing and collaborative information retrieval (CIR). Dourish, using the broad term 'social navigation', stresses that social concerns should underlie our design decisions:

> Social navigation is one of the most direct expressions of the fundamental principle that the action of a user at a computer is driven not simply by a set of internal goals and cognitive processes, but by the social setting in which people find themselves, by the action of others, individually and collectively, and by the social nature of the work being conducted and the goals sought. … social navigation techniques allow us to be able to capitalise upon the social nature of the everyday world and so to enrich the interface with collaborative support for individual tasks. (Dourish, 1999)

Much of the work in this area has come from computer science; particularly from the fields of human-computer interaction (HCI) and computer supported cooperative work (CSCW). However, there is a Library and Information Science (LIS) perspective which should not be forgotten: "It is interesting, therefore, that library scientists have recognized for some time that other individuals may be valuable information resources during the search process" (Evans and Chi, 2010).

Reich and Weiser (1993) note that "community public libraries provide more than information" and that their social roles should be considered when building digital libraries. The classic reference interview (Bopp and Smith, 2001) is a prime example of a collaborative interaction rooted in the service tradition of libraries. The nature of reference interviews has changed over the past century, from co-located synchronous meetings to embrace a variety of technologies (Barnello, 1996). Technologies used can vary from online text chat to attempts to capture and visualize search processes to enhance virtual reference interactions (Twidale, Nichols and Paice, 1997).

Despite this collaborative heritage, an individual focus can be found in some (but not all) analyses and models of the information-seeking process. Talja and Hansen (2005) ascribe the rarity of studies of collaborative information behavior to an LIS analytic focus on "sources and channels" rather than contextualized, embedded (and social) processes. The individual focus is most apparent in research in computerized IR, which until recently mostly concentrated on one person interacting with a computer.

Karamuftuoglu (1998) seeks to broaden the conceptualization of IR so that "the fundamental theoretical issues of information retrieval are the production and use (consumption) of knowledge". This extension of the remit of IR to also consider information use and production means that it necessarily must consider aspects of collaboration. Hertzum (2008) builds on the work of Karamuftuoglu claiming that: "information seeking is just as much about making coherent sense of information as it is about finding extant information. Hertzum defines collaborative information seeking as "the information-seeking activities performed by actors to inform their collaborative work combined with the collaborative-grounding activities involved in making this information part of the actors' shared understanding of their work".

The refinements of models and definitions have become necessary as studies of information work have shown that in practice people rely on social and collaborative interactions to achieve their goals (e.g., Twidale, Nichols and Paice, 1997; Hansen and Järvelin, 2005; Morris, 2008; Prekop, 2002; Fidel, et al. 2004). In summary, "we can conclude that collaboration is common in information retrieval, this collaboration can occur at any stage of the IR process, and that collaboration occurs in spite of a lack of support for these activities in most IR systems" (Foley and Smeaton, 2010).

Our aim in this chapter is to show different kinds of social aspects in interactive information retrieval. The themes we report use different terms and draw on different methods. We bring them together because we see commonalities in their various uses of the social to help people access information. It may help to think of social approaches belonging on a continuum of the explicit nature of the collaborative interaction. At the explicit end are both the traditional face-to-face library reference interview and collaborative information retrieval using large displays, video conferencing and interactive tabletops. At the other end are recommendations based on the popularity of certain choices derived from aggregations of many actions by other people. These may involve people explicitly rating choices, or information inferred by what people do, such as purchasing decisions or time spent looking at an item. At other points on the continuum are recommendations made by particular people, including citations to be followed; elsewhere recommendations for a particular person, but calculated from explicit and implicit data collected from other users. As these examples show, collaborative explicitness is independent of the kind of technology used, and includes different kinds of collaboration. Traces of the articulated preferences and actions of others can be used in many ways by different applications and also directly by people themselves. When such activity traces are directly used by people to help find information, they are often called social search. Even such a simple piece of information as the relative popularity of ten documents can be useful in helping prioritise what to read.

## Recommendation

Recommendations are important because they are a general mechanism for coping with information. They enable information searchers to cope with greater quantities of data and locate desired material more efficiently. In fact, recommendations are endemic in our lives; almost any time we have a conversation someone is likely to be expressing some form of preference for products, services or actions. A substantial portion of the media we consume involves general recommendations; for books, movies, music, restaurants, travel destinations, recipes, consumer electronics, etc. Try counting how many recommendations you make (or hear) in just one day.

In circumstances of relative information scarcity, recommendation may be less important. It seems to have a greater impact in cases of information abundance. In some contexts there is a particular piece of information that we need and that we believe exists, and the challenge is how to obtain it. It may be buried amidst millions of documents. Much research in information organization and information retrieval has focused on this problem, initially with respect to finding information within paper-based documents, and increasingly focusing on finding information in digital documents. However, although this kind of need is widespread and important, it is not the only information need. Sometimes our need can yield many highly relevant pieces of information, but there are too many to attend to, and even of those we look at, it can be difficult to decide which to act upon. It is in these cases of information abundance that recommendations can play a critical role, both in everyday life and in specialist applications.

The recommendations that we encounter may be broadcast to wide audiences (in newspapers or on television) or may come from friends who already know about our interests. Alternatively, they may come from software applications that have been recording our interactions and use algorithms to generate suggestions. One approach to start to understand recommendations is to separate out the *subject* of the recommendation and the *nature* of the recommender:

What is the subject of a recommendation?

- items, parts of items (quotations, findings, method)
- resources (collections, digital libraries, databases)
- people
- organizations
- techniques (searching strategies)

Who/What is doing the recommending?

- individual (known or anonymous)
- group (known or anonymous)
- salesperson (with an incentive to sell more expensive items)
- organization ( industry group or consumer watchdog)
- computer system

When we consider the person, group or system that provides a recommendation we are naturally led to ask whether they have specific relevant information. A friend who knows that you saw a particular movie last week would be unlikely to recommend it again. On the other hand a national newspaper article recommending the film is distributed to everyone who reads the newspaper, there is no way to tailor or adapt the recommendation to an individual reader. The key part of recommendation in interactive information retrieval is: on what basis is the recommendation made? What information does the recommender system possess?

For broadcast media the model of the user (reader or viewer) may be simply the people that tend to read this newspaper or watch this television program. The demographics may be estimated from surveys and a typical target reader (viewer) may be envisaged. In contrast a personal

recommendation from a friend comes with large amounts of background knowledge of your preferences across a wide range of media over several years. In practice, recommender systems fall somewhere between these points. However, they may have access to information about thousands of other similar people which they can use to supplement individualized information about any particular user.

Human-based recommendations may be based on a mixture of authority and expertise (the degree to which the person recommending is an expert in the domain, such as a renowned movie critic or an automobile expert), and the degree to which they know the interests and preferences of the intended audience (maybe just one person or a particular demographic). The selection of whom to seek recommendations from is a significant issue. Often it can be built up over time based on previous successful recommendations, so that the recipient grows to trust the judgment of a particular recommender. Deciding that one movie critic seems to be better at picking movies that you will like, compared to another critic, allows for the development of a constituency for a particular set of recommendations from a recommender. Certain critics can build a reputation for their opinionated views, and people who agree with those views would tend to stick with and trust their recommendations, whilst others may seek a critic whose worldview seems more in keeping with their own. Selecting a particular recommender can of course be a matter of recommendation itself. These issues of trust, selection, tailoring and personal preference also apply to computational recommendation systems – they are just more familiar and easier to think about in non-computational settings.

## Amazon.com: a case study of social information retrieval

Many of the issues involved in social recommendation are illustrated by the infrastructure created by Amazon.com around the wide variety of products they sell. The simple display of an item and its metadata is common to both library catalogs and ecommerce systems. The social enhancements added by Amazon to this basic metadata provide a very different experience for the user, some examples are shown in Figure 1.

*Editorial Reviews.* Reviews (such as for books, movies, cars or other consumer goods and services) by experts and/or journalists are common across many media. Amazon adds these free text commentaries to the item display, although there is no indication that they are all the available reviews, or even a fair sample.

*Sales statistics.* As with reviews, sales figures have long been prominent in areas such as music, films and books. Amazon only provides its own sales rankings and the user has no clear idea that they are representative of all sales.

*Patterns of sales.* Although many shops display local lists of bestsellers Amazon has mined its previous purchasing data to provide information about other users' actions. Figure 1 shows some examples. Populating the "Customers who bought this item also bought" list requires that Amazon can associate purchases at different times and link them to the same customer; it does this via its account system.

*Customer reviews.* Consumers have always had opinions on the worth of the goods and services they use, though they have not traditionally been integrated into the purchasing environment. A likely reason is that it seems somewhat counter-intuitive that a retailer would publicize a negative recommendation about a product they are selling. Reviews are effectively anonymous and there are a variety of possible biases in allowing anyone to review anything. For example, if early adopters have different preferences then their reviews may have a disproportionate effect on consumer preferences (Li and Hitt, 2008). A separate source of bias occurs when reviewers have a vested interest in products being purchased. Negative reviews of some items can push sales to competing items; one reviewer "is to pay damages and costs to two rivals who launched a libel case after a row erupted over fake reviews posted on the Amazon website" (Topping, 2010). Reviews at Amazon

have a free text component and a star rating from one to five. The distribution of star ratings and a graphical representation of the average are shown to users.

*Review feedback*. Users can provide Boolean feedback on whether a review was "helpful", a review of the review. This feature provides a mechanism to limit the influence of outlier reviews; it is dependent on community use but this feedback is made low-cost through a "one-click" interface.

*Anonymity.* As noted above, the anonymity of reviews can provide a shield for biased reviews, in general we ascribe more trust to identifiable individuals. A "real name" feature is designed to encourage greater trust in these identifiable reviews.

*Added metadata.* Users can add tags to the metadata of an item to enhance its description and aid findability. As with the basic reviews, a low-cost feedback mechanism is provided to identify accurate tags.

*Lists.* Amazon provides a "Listmania" service to allow users to group items into lists. List topics include general recommendations ("Books every Geek Should Read"), specific recommendations ("If You Love Sookie Stackhouse...Part 1"), university courses ("PS 545 Fall 2010") and personal favorites ("Jeff's SciFi List").

## Customers Who Bought This Item Also Bought

---

## Frequently Bought Together

Customers buy this book with Everything Is Miscellaneous

---

## What Do Customers Ultimately Buy After Viewing This Item?

## Customer Reviews

**8 Reviews**

| | | |
|---|---|---|
| 5 star: | ▬▬ | (6) |
| 4 star: | ▪ | (1) |
| 3 star: | ▪ | (1) |
| 2 star: | | (0) |
| 1 star: | | (0) |

**Average Customer Review**
★★★★☆ (8 customer reviews)

---

Help other customers find the most helpful reviews

Was this review helpful to you?  [ Yes ]  [ No ]

---

## Most Helpful Customer Reviews

17 of 17 people found the following review helpful:

★★★★★  **excellent look at information**, November 6, 2007

By **E. Schofield** ⌄ - See all my reviews
REAL NAME

---

## Tags Customers Associate with This Product (What's this?)
Click on a tag to find related items, discussions, and people.

☐ information (1 ⌄)          ☐ scholarship (1 ⌄)

---

☐ information (1 ⌄)

Do you agree this product is related to **information**?

[ Yes ]  ▬▬▬▬  1
[ No ]   ▭▭▭▭  0

**Overall agreement: 1**

You
(Pre

ell a
If y

---

## Listmania!

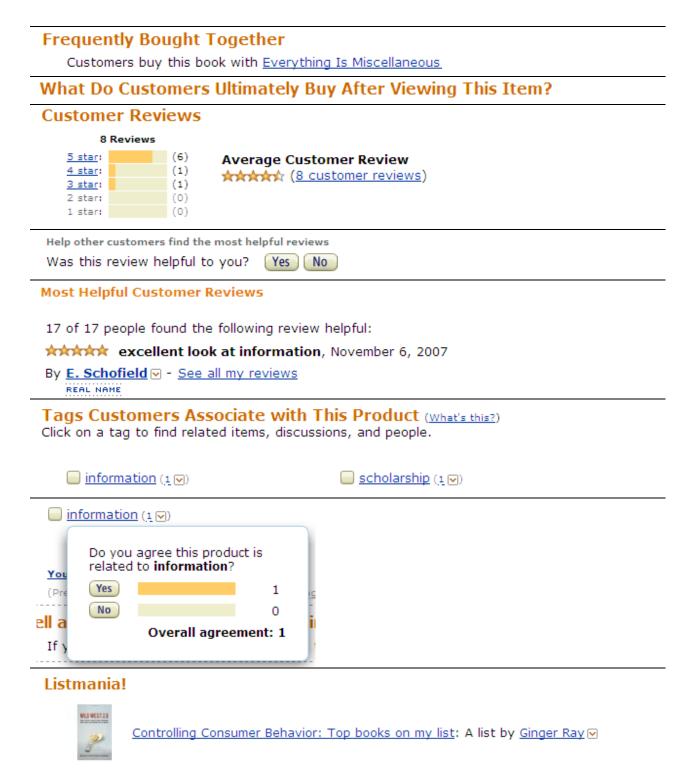Controlling Consumer Behavior: Top books on my list: A list by Ginger Ray ⌄

Figure 1. Excerpts from the socially-enhanced item display at Amazon.com, for the book *Glut: Mastering Information Through the Ages* by Alex Wright

The combination of these elements produces a socially-enhanced experience for Amazon's users. In addition, all of these information sources can be used as input to Amazon's algorithms to provide customized personal recommendations for their users.

## Collaborative Filtering

The utility of recommendation as an information management technique is limited by problems in processing free-text recommendations (such as book reviews). Aggregation, comparison and processing of recommendations are greatly simplified when reviews are reduced to numbers (such as three stars out of five). Although the richness of the textual version is lost, these simpler formats enable computational approaches to recommendation aggregation and led to the emergence of *collaborative filtering* systems in the 1990s.

Shardanand and Maes (1995) describe their "social information filtering" as a complement to traditional content-based approaches. As a system to automate the word-of-mouth recommendations of friends, their technique (and other similar systems) can side-step some of the problems of content-based techniques as it does not need to be able to understand the contents of a document. In fact a purely preference-oriented system can recommend anything – from airlines, to tropical fish, to poetry: as long as the user can understand the subject of the recommendation. The Ringo system of Shardanand and Maes (1995) recommended music and relied on explicit ratings: a user expressed their preferences for various music artists and the system created a user profile. Using these user profiles Ringo then algorithmically combines these preferences to generate recommendations.

One approach to combining preferences is: if you have rated a certain set of items particularly highly, the system looks for other people who have also rated those items highly and recommends to you those items that they have also rated highly that you have not rated. These recommender algorithms are based on the idea of: "people who liked those things you said you liked also liked these extra things – maybe you will too". The exact details of how the algorithms work vary, but the approach is similar. It is robust, because the algorithm does not need to know anything about the nature of what is being recommended (movies, books, music, etc.). It just looks for patterns amongst people's ratings.

Su and Khoshgoftaar (2009) provide a detailed survey of collaborative filtering systems and consider three main approaches:

1. *Memory-based collaborative filtering*. Based solely on user preferences these techniques need not consider the content of documents They typically compute similarity between items or users (using methods such as Pearson correlation and vector cosine similarity) and then make a prediction for a particular user (using weighted averages of other users' ratings). They can have problems generating recommendations involving new users or items lacking any ratings (the 'cold start' problem).
2. *Model-based collaborative filtering*. An alternative approach to using the user-item ratings is to allow a model to learn from the data and then use just that model to make the predictions. A variety of learning algorithms have been used to generate the model, including Bayesian and clustering approaches.
3. *Hybrids*. Combining content-based and preference-based approaches to compensate for the weaknesses of each method. For example, where pure preference-based systems can struggle with items that have not been rated (or users who have not made ratings), this is precisely where a content-based technique (such as full-text querying) performs well.

The field of collaborative filtering received a significant boost in interest with the launch of the $1m Netflix prize in October 2006. Netflix made available over 100 million anonymized customer-generated movie ratings (made by nearly 500,000 users rating nearly 18,000 movies). The challenge was to achieve a 10% improvement in recommendation accuracy compared to the algorithm used by Netflix. The prize was awarded in September 2009. It is interesting to note that teams made progress by collaborating with other teams, producing a better algorithm by carefully

combining features developed separately (Van Buskirk, 2009). It seems that even recommendation algorithms benefit from collaboration and social search of a design space.

## Implicit and explicit rating

Since Shardanand and Maes' (1995) seminal paper there have been many others that look to use alternative mechanisms for processing user profiles to produce a recommendation without analyzing the document content. In addition there has been research in combining preference-based and content-based approaches in hybrid systems. This area of research is now generally known as *recommender systems*.

The Ringo system used explicit ratings and was thus reliant on users taking the time to enter them into the system. The more ratings a user supplies, the more data the algorithm has to find significant correlations with the ratings of others, and so the greater the chances are that the suggestions will be found to be useful. The cost this imposes on users is thus a barrier to adoption and use. Also, acquiring the initial rating values is problematic when the value the system provides is related to the number of its participants (Konstan, et al. 1997). Most recommender systems perform better with increasing numbers of people supplying ratings and this can lead to the *cold start* problem: having low numbers of users can produce poor quality recommendations and poor quality recommendations can fail to incentivize more users to join the system. Social computing systems with network effects often have similar problems.

System designers looked to other sources of information to supplement (or replace) explicit ratings and began to explore implicit ratings such as:

- web browsing records
- item inspection time
- purchase decisions

These measures can be used as proxies for explicit ratings, with the advantage that they can be collected at (virtually) no cost to a user (Fox, et al. 2005). If a user purchases a item then it is reasonable to conclude that they like the item. Although there are clearly exceptions (e.g. purchasing gifts) in general, over a large population of purchases, this proxy approach seems to work sufficiently well to be used to generate recommendations. Kelly and Teevan (2003) provide a good summary of early work on these implicit sources of information. The web interface at Amazon.com shows the use of both explicit ratings (the one to five star ratings) and implicit data sources ("Customers who bought this item also bought"). Bell and Koren, competing in the Netflix competition found that their more conventional approaches suffered due to the large number of people who had only rated a few movies. However they we able to model user preferences better by also taking account of which movies users bothered to rate. This use of more implicit rating information effectively complemented the explicit rating scores users provided. Bell and Koren also note the potential of other hybrid approaches to complement limited amounts of explicit rating with "abundant implicit feedback, which may be purchase/rental history, browsing patterns, search keywords, etc." (Bell and Koren, 2007).

### Interfaces for Collaboration

Recommendations such as those outlined above can draw on the expertise, opinions and actions of others in order to help address an individual's information need. As such they can be viewed as a kind of collaborative activity, even if the participants do not directly interact, or even know each other. There are also much more explicit kinds of collaborative information seeking. An interaction with an expert, such as in a reference interview, is one well established case. But there can also be interactions between peers with a common information need who work together to split the work to get to a solution. Examples include planning a family reunion, or investigating the state of the art in a particular technological market. Collaborative technologies can substantially help these kinds of activities.

When people work together to look for information, what do they share? Obviously the results of searches that they have done, but they also may share other things more related to the process of searching. These can include how they are going to split the work to minimize duplication of effort, which databases (or other resources) to use, useful keywords, particular search techniques, queries, recommendations, tips, plans for future search actions, interim findings, whether they are finding too many things and need to narrow down, or too few and need to broaden their approach.

The research area of CSCW has analyzed and developed technologies that can be deployed to meet the needs of a variety of collaborative settings. CSCW systems analysis and development has emphasized the importance of explicit support for specific aspects of working with others that seem to recur across domains. These include supporting the division of labor, and the availability of appropriate communication channels for both doing the work and discussing the doing of the work. Within CSCW research, the concept of 'awareness' has been found to be central. Many of the examples of things shared noted above may be categorized as awareness information, that help people work together more productively by knowing what the others are doing and how well they are progressing, allowing for actions to be changed in response (Foley and Smeaton, 2010).

Within CSCW, one way to make sense of the large number of collaborative applications available is to consider the extent to which they support collaboration along two dimensions: working together in the same place or at a distance, and working together at the same time or at different times (Twidale and Nichols, 1998). Collaborative information seeking applications likewise can be considered as particularly supporting different points on these axes of time and space. Several applications have been developed to explicitly support certain kinds of information seeking. For example, SearchTogether (Morris and Horvitz, 2007) supports remote collaboration, both synchronous and asynchronous, ViGOR (Halvey, et al. 2010) supports mostly asynchronous remote collaboration while Cerchiamo (Pickens, et al. 2008) supports co-located synchronous collaboration. In contrast, Coagmento (Shah, Marchionini and Kelly, 2009) supports co-located and remote, synchronous and asynchronous collaboration, and CIRLab (Fernández-Luna, et al. 2010) is a toolkit enabling the construction and testing of different kinds of collaborative interactions a groupware framework for CIR research and experimentation.

Another way to understand the variety of different systems is by the interactive hardware they employ. These include the use of large displays to facilitate sharing and discussion, integration of computers with multiple mice, mobile phones (Amershi and Morris, 2008) or PDAs (Paek, et al. 2004; Maekawa, Hara and Nishio, 2006) to enable more people to interact without the need for expensive hardware, and a notable interest in tabletop displays as a way of supporting the kind of close informal collaboration we are used to with paper-based interaction (Shen, et al. 2002; Smeaton, et al. 2007; Morris, Fisher and Wigdor, 2010).

Additionally, people may appropriate available technologies to support their collaborative information seeking. These include telephone, email, instant messaging, shared online documents such as Google Docs, and Twitter. Consequently, collaboration and social search should not be considered to be only those activities undertaken with purpose-built applications, even though these can make certain kinds of collaborative interaction substantially easier or more powerful. For a review of issues around explicit collaboration in information search, see Twidale and Nichols (2010).

## Broader Issues in Information Science and System Design

The social aspects of information search add a layer of complexity to an understanding of the information seeking process. It is not surprising that much work on information retrieval chose to ignore these issues in order to concentrate on the challenge of improving computational access to information for an (assumed) individual searcher. Attempting to optimize this interaction has brought direct benefits to millions of people in thousands of very different contexts.

LIS has not ignored the social aspects of information seeking, for example, Taylor's (1968) classic paper notes that starting from an individual's information need, one possible resolution is to ask a colleague. It is just that Taylor is more concerned with studying and understanding another resolution: going to the library and either helping oneself or to asking an information specialist. Similarly, Wilson (1981) notes that: "the user may seek information from other people, rather than from systems", and further notes that such kinds of interaction have "tended to fall within the sphere of interest of sociologists and organizational theorists rather than within that of information scientists." Many of the theories of information seeking behavior have components of collaborative interaction, although individual information interactions are typically central – see Twidale and Nichols (2010) and Chapter 3 for examples.

Wilson (2008) has proposed Activity Theory as an overarching paradigm that at least has the potential to bring together much research in LIS. Activity Theory explicitly takes into account the interactions between people and artifacts as they get things done. As such, various kinds of collaborative interactions are *central* to the theory rather than extras tacked on to a generalization of an individual-centric analysis. Information seeking is not just social, it is also situated in a social context. People typically want some information so they can use it as part of a larger task. Information access systems are necessarily embedded in a larger user context. Just as information seeking is inherently social, so are the contexts in which it occurs.

The more social approach raises challenges for research. Is it the case that many of our conventional methods for studying information seeking behaviors are more effective in studying the actions of individuals? Do we need new methods to understand different kinds of social search activity? It can be harder to analyze information seeking when it involves using a variety of systems and combining the findings from each; asking people for recommendations face-to-face and online via diverse social media; using dedicated recommendation sites and suggestion features embedded in an interface; searches extended over hours or weeks; and including use of a workplace PC, a laptop at home and a mobile phone on the bus. Approaches such as iterative prototyping may be more appropriate in a world where people are frequently trying out and adopting, adapting or rejecting new features and applications.

Nevertheless, despite the floods of new gadgetry and information services that people confront and cope with, certain classic issues remain. As human beings our attention is limited and precious. Our systems' memory, processing power and search algorithms may continue to improve, but so does the amount of heterogeneous data that we may want or need to access. To most LIS researchers, we are still in the business of saving the time of the user. However we must note that with the abundance of applications and collaborative resources, many of our users are also working at saving their own time by innovating with applications to collaborate with others directly and indirectly. We are no longer the sole custodians of knowledge and sole mediators of access to it – if indeed we ever were. We can't just build technical systems *for* people. We must also co-construct (or refine) *socio*technical systems *with* people.

## Conclusion

As Dourish (1999) emphasized, information activities take place in a social context. New documents are created in the context of previously authored ones. People link documents together by citations, tags and hyperlinks. Searches are performed in collaboration with colleagues for the benefit of teams. Records of user activities are analyzed, shared and re-used.

The information structures that embody these social elements, whether it is a library or the customer database of Amazon.com, force us to consider our information retrieval activities as more than just an algorithm applied to some data. When we approach our retrieval problems through a social lens we often find there is a more effective way to connect people with information. Some design approaches to this social context are algorithmic, using the activity of others to rank or

recommend items. Others are more about supporting direct interaction between people, and yet others are a mixture of the two.

With the wider availability of collaborative technologies and the ability to build on the successes of traditional (individual) IR research, there seems a growing interest in the ways that people continue to invent social solutions to their needs. This applies to the creation, the modification, the search for and the selection of informational resources, ranging from the use of wikis, blogs, and social media to commercial uses of rating and recommending. As computers pervade more of people's lives, the uses of technology for information seeking cross boundaries of work, domestic life, health and leisure, as well as different levels of interaction with other people and their online activities.

As the other chapters in this book show it becomes important to consider information retrieval as an interaction between a person and a system, where better functionality and interfaces allow a more iterative, almost conversational, kind of interaction as opposed to a one shot query response. In just the same way, it becomes important to consider information retrieval as an interaction between several people, their actions and multiple systems.

## References

Amershi, S. and Morris, M. (2008). CoSearch: a system for co-located collaborative web search, *Proceedings of the Conference on Human Factors in Computing Systems (CHI '08)*, ACM. 1647-1656.

Barnello, I. (1996) The changing face of reference: a history for the future. In Stuart, L. and Hutto, D. (eds.), *The Changing Face of Reference*, JAI Press. 3-17.

Bell, R. M. and Koren, Y. (2007) Lessons from the Netflix prize challenge, *SIGKDD Explorations Newsletter*, **9** (2), 75-79.

Bopp, R.E. and Smith, L.C. (eds.). (2001) *Reference and Information Services: An Introduction*, 3rd edn, Libraries Unlimited.

Dourish, P. (1999) Where the footprints lead: tracking down other roles for social navigation. In Munro, A.J., Höök, K. and Benyon, D. (eds.), *Social Navigation of Information Space*, Springer. 15-34.

Evans, B.M. and Chi, E. (2010) An elaborated model of social search, *Information Processing and Management*, **46** (6), 656-678.

Fernández-Luna, J.M., Huete, J.F., Pérez-Vázquez, R. and Rodríguez-Cano, J.C. (2010). CIRLab: A groupware framework for collaborative information retrieval research, *Information Processing and Management*, **46** (6), 749-761.

Fidel, R., Pejtersen, A. M., Cleal, B., and Bruce, H. (2004) A multidimensional approach to the study of human-information interaction: A case study of collaborative information retrieval, *Journal of the American Society for Information Science and Technology*, **55** (11), 939-953.

Foley, C. and Smeaton, A.F. (2010) Division of labour and sharing of knowledge for synchronous collaborative information retrieval, *Information Processing and Management*, **46** (6), 762-772.

Fox, S., Karnawat, K., Mydland, M., Dumais, S. and White, T. (2005) Evaluating implicit measures to improve web search, *ACM Transactions on Information Systems*, **23** (2), 147-168.

Halvey, M., Vallet, D., Hannah, D., Feng, Y. and Jose, J.M. (2010). An asynchronous collaborative search system for online video search. *Information Processing and Management*, **46** (6), 733-748.

Hansen, P. and Järvelin, K. (2005) Collaborative information retrieval in an information-intensive domain, *Information Processing and Management*, **41** (5), 1101-1119.

Hertzum, M. (2008) Collaborative information seeking: The combined activity of information seeking and collaborative grounding, *Information Processing and Management*, **44** (2), 957-962.

Karamuftuoglu, M. (1998) Collaborative information retrieval: Toward a social informatics view of IR interaction, *Journal of the American Society for Information Science*, **49** (12), 1070-1080.

Kelly, D. and Teevan, J. (2003) Implicit feedback for inferring user preference: a bibliography, *SIGIR Forum*, **37** (2), 18-28.

Konstan, J.A., Miller, B.N., Maltz, D., Herlocker, J.L., Gordon, L.R. and Riedl, J. (1997) GroupLens: applying collaborative filtering to usenet news, *Communications of the ACM*, **40** (3), 77-87.

Li, X. and Hitt, L.M. (2008) Self-Selection and information role of online product reviews, *Information Systems Research*, **19** (4), 456-474.

Maekawa, T., Hara, T. and Nishio, S. (2006) A collaborative web browsing system for multiple mobile users, *Proceedings of the Fourth Annual IEEE international Conference on Pervasive Computing and Communications (PerCom 2006).* IEEE Computer Society. 22-35.

Morris, M. R. (2008) A survey of collaborative web search practices, *Proceedings of the Conference on Human Factors in Computing Systems (CHI '08)*, ACM. 1657-1660.

Morris, M. R., Fisher, D. and Wigdor, D. (2010). Search on surfaces: Exploring the potential of interactive tabletops for collaborative search tasks, *Information Processing and Management*, **46** (6), 703–717.

Morris, M.R. and Horvitz, E. (2007) SearchTogether: an interface for collaborative web search, *Proceedings of the 20th Annual ACM Symposium on User interface Software and Technology (UIST 2007)*, ACM. 3–12.

Paek, T., Agrawala, M., Basu, S., Drucker, S., Kristjansson, T., Logan, R., Toyama, K. and Wilson, A. (2004) Toward universal mobile interaction for shared displays, *Proceedings of the Conference on Computer Supported Cooperative Work (CSCW '04)*, ACM. 266-269.

Pickens, J., Golovchinsky, G., Shah, C., Qvarfordt, P. and Back, M. (2008) Algorithmic mediation for collaborative exploratory search, *Proceedings of the 31st Annual International Conference on Research and Development in Information Retrieval (SIGIR '08).* ACM. 315-322.

Prekop, P. (2002) A qualitative study of collaborative information seeking, *Journal of Documentation*, **58** (5), 533-547.

Reich, V. and Weiser, M. (1993). Libraries are more than information: situational aspects of electronic libraries, *Serials Review*, **20** (3), 31-37.

Shah, C., Marchionini, G., and Kelly, D. (2009). Learning design principles for a collaborative information seeking system. *Proceedings of the Conference on Human Factors in Computing Systems (CHI '09),* ACM. 3419-3424.

Shardanand, U. and Maes, P. (1995) Social information filtering: algorithms for automating "word of mouth". *Proceedings of the Conference on Human Factors in Computing Systems (CHI'95),* ACM Press. 210-217.

Shen, C., Lesh, N. B., Vernier, F., Forlines, C. and Frost, J. (2002) Sharing and building digital group histories. *Proceedings of the Conference on Computer Supported Cooperative Work (CSCW '02)*, ACM. 324-333.

Smeaton, A., Lee, H., Foley, C. and McGivney, S. (2007) Collaborative video searching on a tabletop, *Multimedia Systems*, **12** (4), 375-391.

Su, X. and Khoshgoftaar, T.M. (2009) A survey of collaborative filtering techniques. *Advances in Artificial Intelligence,* **2** (2). Article ID 421425, 19 pages. doi:10.1155/2009/421425

Talja, S. and Hansen, P. (2005) Information sharing, In Spink, A. and Cole, C., (eds.) *New Directions in Cognitive Information Retrieval*, Springer. 113-134.

Taylor, R.S. (1968). Question-negotiation and information seeking in libraries, *College & Research Libraries,* **29** (3), 178-194.

Topping, A. (2010) Historian Orlando Figes agrees to pay damages for fake reviews*, The Guardian,* (17 July), 9.

Twidale, M.B., Nichols, D.M. and Paice, C.D. (1997) Browsing is a collaborative process, *Information Processing and Management*, **33** (6), 761-83.

Twidale, M.B. and Nichols, D.M. (1998) Computer supported cooperative work in information search and retrieval, *Annual Review of Information Science and Technology*, **33**, 259-319.

Twidale, M.B. and Nichols, D.M. (2010) Collaborative information retrieval, in (eds.) Bates, M.J. and Maack, M.N., *Encyclopedia of Library and Information Sciences*, 3rd edn, CRC Press. 1080-1087.

Van Buskirk, E. (2009) How the Netflix Prize Was Won*, Wired*, September 22*.* www.wired.com/epicenter/2009/09/how-the-netflix-prize-was-won/

Wilson, T.D. (1981) On user studies and information needs, *Journal of Librarianship*, **37** (1), 3-15.

Wilson, T.D. (2008) Activity theory. *Annual Review of Information Science and Technology*, **42**, 119-161.

---