



1

Introduction to Social Information Access

Peter Brusilovsky  and Daqing He  

School of Computing and Information, University of Pittsburgh,
Pittsburgh, PA 15213, USA
{peterb,dah44}@pitt.edu

Abstract. This chapter offers an introduction to the emerging field of social information access. Social information access focuses on technologies that organize users past interaction with information in order to provide future users with better access to information. These technologies have become increasingly more popular in all areas of information access, including search, browsing, and recommendation. Starting with a definition of the new field and a brief history of social information access, this chapter introduces a multi-aspect classification of social information access technologies. The two important factors for our classification are the types of information access involved and the source of the social information that has been leveraged to support information access. These two factors are the angles we use in this chapter to create a map of the field, as well as to introduce the book structure and the role of the remaining book chapters in covering social information access topics and technologies.

1 Social Information Access

The social Web (or Web 2.0), through various platforms, such as Wikis, blogs, Twitter, or Facebook, have changed the role of Web users from simply information consumers to their emergence as key information producers for content sharing and community building. Recognizing users as information producers has also attracted attention to novel information access technologies supported by “collective wisdom,” distilled from actions of those who worked with this information earlier. The ideas of so-called “social information access” has been explored by a number of research groups worldwide as part of the effort for developing techniques to help users obtain the right information for satisfying their information needs [9, 15, 33, 54, 56, 59].

In this book, we define *social information access* as a stream of research that explores methods for organizing the past interactions of users in a community in order to provide future users with better access to information. Social information access covers a wide range of different systems and technologies that operate on a different scale, which can range from a small closed corpus site to the whole Web. Although the technologies located on the different sides of this stream may not even recognize each other as being a part of the same whole, the

whole stream is driven by the same goal: to use the power of a user community to improve information access.

Self-organization is a vital feature of all social information access systems. These systems are able to work properly despite little or no involvement of human indexers, organizers, or other kinds of experts. They are truly powered by a community of users. Due to this feature, social information access technologies are frequently considered as an alternative to the traditional (content-oriented) information access technologies. In most cases, social information access can run in parallel with traditional types of information access, and may help users to find resources that would be hard to find in a traditional way. In other cases where traditional information access is hard to organize (for example, in a collection of non-indexed images), social mechanisms (such as tagging) can serve as a handy replacement. However, it has been more and more frequently demonstrated that most of the benefits could be obtained by integrating social and traditional technologies. For example, hybrid recommender systems can integrate collaborative and content-based recommender mechanisms.

As a type of information access that can offer multiple benefits while being relatively easy to organize and maintain, social information access has been attracting more and more attention from researchers and practitioners. The overarching goal of this book is to provide a comprehensive hands-on overview of modern social information access technologies. The book is designed with two audiences in mind. On the one hand, it could help students and young researchers who are interested in learning about this new field; and on the other hand, it can assist more experienced researchers and practitioners in the development of new social information access technologies and applications. To support this goal, each chapter carries a dual nature. To support novices, it provides a review of a specific group of techniques, and to support practitioners, it explains critical algorithms or reviews a set of case studies.

The role of this chapter is to introduce the field of social information access to the readers and to explain the layout of the book itself. We start our introduction with a brief history of social information access and follow with a multi-aspect classification to define the space of social information access. The classification provides the readers with the necessary knowledge for comprehending and distinguishing different kinds of social information access techniques. Based on the same classification, we also introduce the content and the structure of the book to provide the readers with a “big picture” of how different chapters cover various dimensions of the classification.

2 The Emergence of Social Information Access

The ideas that underpin social information access can be traced back to several visionary projects. Vannevar Bush’s seminal paper on *Memex* introduced the idea of “trails” through information space, which Memex users could create and share with others [13]. The *Superbook* project [53] demonstrated the benefits of directly engaging end-users into document indexing. The *Edit Wear and Read Wear* project [32] introduced the concept of a history-rich information space and

demonstrated that social information access could be based on implicit, rather than explicit, actions of past users.

Besides these influential pioneering projects, it was the opportunities and the needs produced by the rapidly expanding World Wide Web (WWW) that led to the emergence of social information access as a research area. On the one hand, the World Wide Web, with its increased volume of users, has enabled the collection of social information at scale. On the other hand, it has opened access to information to many inexperienced users who could benefit from additional support. The “Web push” led to a rapid expansion of social information access research between 1994 and 2000. This period brought many innovations as research teams investigated new approaches to help users in the rapidly expanding information space. In the context of this book, we will examine four main streams of research that established the field of social information access at the turn of the centuries: collaborative filtering, social navigation, social search, and social bookmarking.

Collaborative filtering [24,39,54] attempted to propagate information items between users with similar interests. The emergence of collaborative filtering is typically traced back to the *Information Tapestry* project [24], which coined the term collaborative filtering. Information Tapestry introduced an approach that was later called pull-active collaborative filtering, where users had to actively query the community feedback left by earlier users in order to receive social guidance. Later, several projects expanded the scope of collaborative filtering. For example, *Lotus Notes recommender* [45] proposed push-active collaborative filtering, where users are encouraged to send interesting documents directly to their colleagues. However, the majority of pioneering work in this area has focused on indirect approaches to collaborative filtering based on *automatic* matching users with similar interests and cross-recommending positively-rated items. This stream of work includes such pioneer systems as *GroupLens* [54], *Ringo* [56], and *Video Recommender* [31].

In its early form, *social navigation* [15,18,20] attempted to visualize the aggregate or individual actions of community users. It was motivated by observing users’ navigation in real space where they frequently follow the footprints of others. Proponents of social navigation in information space argued that “digital footprints” could also help future users to navigate through information space [16,62]. Inspired by the “footprint” examples provided in *Read Wear and Edit Wear* system [32] and the concept of social navigation in information space introduced by Dourish and Chalmers [18], early pioneers of social navigation developed and evaluated several well-cited systems, such as *Juggler* [16], *Footprints* [62], and *EFOL* [60].

The work on *social search* also expanded rapidly between 1994 and 2000. At that point, the accumulation of social data embedded in search engine logs and Web structure led to an explosion of creative approaches on how to use this data to help Web searchers [4,8,14,21,37,47,52]. The most influential one was *PageRank* [8], a novel ranking approach based on social data encapsulated in the global structure of Web links. Promoted by the success of the Google search engine, this technology inspired a large number of works on social search [12].

Social bookmarking was another important early avenue of research that happened between 1995 and 2000. It was motivated by the need to support the organization of personal information space as well as to share valuable online resources with others (i.e., the same motivation that encouraged early work on active collaborative filtering [24, 45]). Pioneer systems that focused on social bookmarking, such as *Siteseer* [55], *WebTagger* [38], *WDB* [61], and *PowerBookmarks* [41] explored different ways of organizing and sharing bookmarks. Among these, collaborative tagging, which was originally explored in *WebTagger* [38], emerged as the most efficient way to help new users locate useful information that has already been discovered and classified by others. In fewer than 10 years, social bookmarking and tagging systems, popularized by systems like del.icio.us and flickr.com, grew into a new major Internet technology [25, 28].

There were several attempts to bring together researchers working on different types of social information access during the first decade of research in the field. Several workshops that gathered like-minded researchers, as well as the publications that resulted from these workshops [15, 33, 44, 48] have clearly contributed to the expansion of social information access ideas and the conceptualization of social feedback as a source of knowledge in assisting users. However, these integration attempts were based on a limited volume of work and failed to include work on social search and social bookmarking. Our book represents another attempt to bring together a diverse set of research on social information access. While our main goal is to provide an overview of the current state of each major stream of research on social information access, we base the structure of the book on an integrative multi-dimensional classification of social information access techniques. This classification, which is introduced in Sect. 3, highlights both the similarities between different groups of social information access techniques and the opportunities to support users across multiple kinds of information access.

3 Classifying Social Information Access Technologies

The term “Social information access” contains two parts: “information access” and “social”; as a result, the most natural way to classify social information access technologies is by answering two questions: “What kind of information access is considered?” and “How this information access is made social?” In this section, we expand this idea into a multi-dimensional classification framework. Our goal of designing this framework is to make it compatible with older classification attempts, and at the same time, to make it rich and expressive enough to classify a large variety of modern social information access techniques. In the next section, the suggested framework is immediately applied to introduce and classify the social information techniques that are presented in the remaining chapters of this book.

3.1 Types of Information Access

Following earlier classification attempts introduced in [9, 10], the first dimension of our framework for classifying social information access techniques is the type

of information access. “Access” is a reasonably studied concept that refers to an interactive process, which starts with a user noticing their needs and ends with the user obtaining the necessary information. It is an iterative process with multiple stages and possible back loops. However, there are four different ways to obtain the information, which results in the four core types of information access [9, 10]: *ad-hoc information retrieval*, *information filtering (recommendation)*, *hypertext browsing (navigation)*, and *information visualization*.

In ad-hoc information retrieval (IR), users achieve access to relevant information by issuing a query to an IR system and then analyzing/accessing a ranked list of returned information items (for example, book records). An information filtering (IF) or recommender system also returns a ranked list of information items in response, not to an ad-hoc query, but to a user profile that has usually been accumulated over a longer period of time. Traditional IF systems match a user-provided profile against a flow of incoming documents (for example, news articles) to select the most relevant items for the user. In contrast, modern recommender systems construct and maintain dynamic user profiles by observing user’s interactions to produce new recommendations, even in stable document collections. In hypertext browsing, a user attempts to find relevant documents by browsing links that connect documents in a collection. In information visualization, a set of documents is presented to the user using a certain visualization metaphor in either two or three dimensions; the user observes or (in the case of interactive visualization) interacts with the visualized set to find the most relevant information items.

Since a review [9] suggested to distinguish these four types of information access in 2008, social media have introduced many new ways of accessing information. These new ways could be called *human-driven*, because they focus on automating traditional human ways of information exchange. To differentiate from these human-driven information access, the original four ways could be called *system-driven*. In most cases, these new ways of information access do not introduce new *kinds* of information access, but rather introduce new *sources* of information. For example, a Twitter feed could be searched, browsed, or accessed through a standard information filtering interface. However, there is at least one exception: information access through questions and answers powered by modern Q&A systems. We suggest that this should be considered as the fifth basic type of social information access.

From the point of view of classification, it is important to recognize that the types of information access have strongly influenced the development of certain social information access technologies. For example, browsing-based access encourages research on navigation support systems that can help users to select a link to follow among many links on the current page. The natural approach to using community wisdom in this context is to show “where did the people go” [16, 62] by augmenting links with digital “wear” indicators. The natural approach to collect this knowledge was to track user page visits [11] or link traversals [62]. Consequently, social navigation technologies (history-enriched environments) have been developed for supporting browsing-based access in social context.

Another example is social search technologies that were developed to support traditional IR information access. In this context, users expect to see a ranked list of relevant resources. The natural approach to using community wisdom in this area is by re-ranking results using community wisdom [8, 35, 63] or by inserting community-relevant links into the list or results [59] so that the returned documents reflect not only query relevance, but also the degree of their appreciation by the community. A reliable approach to collecting this wisdom is to track connections between queries and items selected or rated by the community members in the context of these queries [35, 63].

3.2 Making Information Access Social

The “social” aspect of information access stresses its ability to transfer information that comes from one group of the user community (“providers”) to another group (“recipients”). This information transfer is frequently called “collaboration”; although, in most cases, no real collaboration takes place.

By its nature, this information transfer or collaboration could be classified along two important dimensions: *intent* (direct–indirect) and *concurrency* (synchronous–asynchronous). This classification was originally introduced by Dieberger et al. [15] in a social navigation context, which at that time covered both browsing and recommendation. Eight years later, these two dimensions were also introduced by Golovchinsky et al. [26] as a part of a taxonomy for collaborative search. While the latter work used an explicit–implicit dichotomy instead of a direct–indirect dichotomy for classifying intent, it expressed the same meaning (i.e., intent) as the earlier direct–indirect dichotomy in [15]. In the following, we suggest the use of the direct–indirect dimension in its original form to classify intent and reserve the explicit–implicit dichotomy to categorize the types of information traces (see Sect. 3.3).

Direct vs. indirect information transfer determines whether the transfer of information (or a collaboration) is intentional.

In the case of a *direct transfer*, the “providers” directly communicate information to the recipients (or guide the recipients to the appropriate information) with the goal of assisting others. In many cases, to initiate this transfer, a recipient with an information need is expected to also directly solicit information from “providers”. However, indirect approaches could be used to determine the correct “provider” to ask.

In case of *indirect transfer*, “the providers” do not directly provide information to “recipients”. In fact, their work with information only aims to satisfy their own needs. It is the traces of their own work with information that could be processed and used to help the recipients in finding the most relevant information. In other words, information is indirectly collected from the community to help other users. In social information access, indirect transfer is much more common. It also can leverage a larger diversity of “social wisdom” than a direct transfer.

Synchronous vs. asynchronous transfer determines whether providers and recipients coexist in time.

With *synchronous information transfer*, providers and recipients work on their information access tasks at the same time, and information directly or indirectly generated by the providers is immediately used to help the recipients. In synchronous context, the same user frequently works as both a provider and a recipient: they use the information produced by others, and they also generate information to help others.

With *asynchronous information transfer*, the recipients are supported in their information access tasks by “social wisdom” produced by those providers who worked with information earlier in the process. Among the two alternatives, this is the more commonly found case in social information access. Unless a collection of social information has just been started, the fraction of users who work with information at any given time and the volume of social wisdom provided by these users are many times smaller than the volume of all past users and the information that they directly or indirectly contributed.

While this two-dimensional scheme has been useful to the research community for many years, it creates an imbalanced classification, because the vast majority of social information access techniques fall into the *indirect-asynchronous corner*. For example, among the technologies represented in this book (see Table 1), Social Q&A (Chap.3 [50]) is an example of direct-asynchronous access, collaborative search (Chap.4 [64]) offers examples of direct and indirect synchronous access, and social navigation (Chap.5 [20]) provides examples of indirect-synchronous, direct-asynchronous, and indirect-asynchronous groups. The rest of the chapters all focus solely on indirect asynchronous technologies. Consequently, there should be another classification that focuses on the “social” aspect of information access that can specifically help to distinguish various indirect-asynchronous technologies from one another. We offer such a classification in Sect. 3.3.

Table 1. Classification of social information access techniques by intent and concurrency

		Concurrency	
		Synchronous	Asynchronous
Intent	Direct	Collaborative search [64]	Social Q&A [50] Social navigation [20] Recommendation [39]
	Indirect	Collaborative search [64] Social navigation [20]	Social search [12, 29, 49] Social navigation [20] Tag-based navigation [17] Recommendation [6, 7, 27, 34, 39, 40, 51]

3.3 Types and Sources of Social Information

3.3.1 Explicit and Implicit Traces of User Activity

As mentioned in the previous section, it is important to classify social information access techniques by using the concepts of intent and concurrency in the process of collecting social information traces and passing them to new users; but only these two concepts are not sufficient to differentiate a wide variety of modern social information access techniques. To overcome this, Brusilovsky [10] suggested to further classify users' past actions leveraged by social information access. These actions are called *users' feedback* in the field of personalized and social systems. User feedback can be explicit, in which the users explicitly express some opinions about an information item. An example of explicit feedback is a user's rating. Although user ratings are still popular sources of information in some social information access systems (e.g., collaborative recommendation systems), it has long been recognized that user ratings form a comparatively small fraction of user interactions with information. Consequently, recent work has focused on implicit feedback [34], where various users' actions are collected and analyzed to infer their attitudes. The most popular source of implicit feedback is search or browsing logs with a sequence of clicks and dwell time (also known as a clickstream). Although there is a risk that clickstreams and other implicit sources of evidence might be less reliable, they are more readily available in various contexts.

Explicit ratings and clickstreams are the two extreme ends of the implicit-explicit continuum of social information. Nowadays, the gap between these two extremes has been filled by a whole range of user actions collected in social information access systems. For example, at the explicit part of the spectrum, users can take actions such as annotation, commenting, and tagging. Yet, unlike ratings, these actions usually do not quantify the degree of the match between the user's need and the annotated item. On the implicit side, it is possible to have actions, such as purchasing a product online, listening to a digital music track, or eating at a local restaurant. All these types of implicit feedbacks can provide more reliable evidence about a user's interests than a clickstream, because each action is associated with a larger commitment of time and/or money.

Due to the rapid increase of the variety of information traces collected by modern social systems, it has become harder and harder to offer an extensive classification. Therefore, this book attempts to separately discuss and classify explicit and implicit information sources in three broad contexts: browsing, search, and recommendation, which correspond to Chap. 5 [20], Chap. 7 [12], and Chap. 14 [34], respectively. We urge the reader to examine these chapters for detailed discussion of each issue.

In the remainder of this section, we will briefly examine the connections between the main *types* of social information and the main groups of *sources* (i.e., type of systems), and we will use Table 2 to highlight our analysis. The list of types and the classification of sources do not pretend to be exhaustive because it is evident that social information access systems will continue exploring new sources and new kinds of social information. Our goal here is to make this list helpful in distinguishing and classifying the majority of existing techniques.

Table 2. Sources of social information in web and social systems

Social information	Search engines	Hypertext / Web	Communication	Annotation systems	Curation systems	Social networking	Social bookmarking	Recommender systems	Consumption systems	Location based
Queries & SERP clicks	✓						✓		✓	
Browsing trails		✓	✓	✓	✓	✓	✓	✓	✓	✓
References (links)		✓	✓		✓	✓	✓			
Annotations, comments				✓	✓	✓	✓		✓	✓
Tags, categorizations			✓		✓		✓	✓	✓	✓
Social links			✓			✓	✓	✓	✓	✓
Ratings				✓				✓	✓	✓
Consumption actions									✓	
Real world trails										✓

3.3.2 Search Engines

Search engines receive users' search queries and generate search engine result pages (SERPs). By tracking individual users through their search sessions, a search engine can archive successful sequences of *queries and SERP clicks*; namely, those search results that users decided to explore further. Both queries and SERP clicks are useful social information, and mining their accumulation can be used to generate social wisdom. Currently, this is one of the most powerful approaches for improving search.

3.3.3 The Linked Web

The open Web and many specialized Web-based systems (i.e. Wikipedia) allow users to create information pages and link them to one another. They offer activity traces of two kinds of users - page authors and Web surfers. Page authors extensively use *references* (i.e., Web links) to other pages when creating their pages. These links were one of the earliest sources of social information that was used to improve information access through the better ranking of search results [8]. Web surfers leave *browsing trails* as they navigate, where each click offers a small evidence to indicate that the selected link is the most attractive for the given user on the traversed page. Within a single Web site or Web system, clicked links are easily accumulated in Web logs. Across sites and systems, traces of user Web browsing behavior can be aggregated by using browsing agents [42], intermediaries [3], browser plugins [58], and other approaches (see [23] for a review). An advanced user tracking approach could augment browsing trails

with *in-page behavior* (such as scrolling or mousing, among others) - a valuable source of social wisdom that could be used for both distinguishing the most useful pages and guiding users to the most relevant parts of a page [30].

3.3.4 Communication Systems

Various communication systems, such as bulletin boards, discussion forums, e-mail, chat, blogs, and microblogs accumulate large volumes of social wisdom. Open discussion sites, such as bulletin boards, forums, and blogs can be easily crawled and mined. Private e-mail and chat traces are generally harder to use, but given that many Web mail systems are maintained by companies that also operate search engines, the e-mails also emerge as a valuable source for search improvement. Communication systems are also used as sources for useful *references* and implicit *social links*. A reference to a Web link in any kind of message or discussion is a good evidence of the importance of the link. These links could be simply extracted or associated with a discussion topic and surrounding text. Replying to, commenting, or forwarding actions all offer the evidence of a social link between users. Modern blogs and microblogs could also serve as sources of *explicit social links*: users can establish social links in the form of “watching” other users’ updates. In these systems, posts could be also extended with social tags.

3.3.5 Annotation Systems

The Web was originally envisioned as having the ability to provide comments and annotations for every Web page, but this infrastructure has never been fully implemented. Instead, the task of Web page annotation has been taken over by various Web annotation systems including the original Annotea project from WWW Consortium [36]. These systems allowed every Web user to add comments for a Web page or its fragment, or simply to mark up the most valuable fragments. Motivated by the research on Web page annotations, page annotation functionality was implemented in a number of Web systems, including Web-based books, textbooks, and digital libraries [19, 43, 46]. In modern social systems, commenting and annotation functionality have been applied to a broad range of items beyond Web pages. With these systems, users can add comments, annotations, and reviews to hotels, movies, books, and many other items. Web annotation systems and “item-focused” social systems could offer three types of social information. Firstly, an annotation can be treated as a sign of the item’s importance, which could be used to attract attention to it on SERPs [2] or on a Web site. Moreover, a within-page annotation system can collect user in-page behaviors, which are used in guiding future users to the most valuable part of the page [19]. Secondly, the content of page annotations or comments describes a page or an item from the prospect of the annotation author. These comments could be used for search and recommendation [51]. Thirdly, many modern “item-focused” systems collect not only item annotations, but also *item ratings*.

3.3.6 Social Networking Systems

Facebook and LinkedIn started as platforms to connect people, but they have gradually included elements from microblogs and social bookmarking systems. Modern social linking systems serve as the primary source of *social links* while also contributing *item links* and *comments*.

3.3.7 Curation Systems

Since the early days of the Web, there have been multiple attempts to engage Web users into adding additional levels of organization to the Web. The most remarkable among these projects are those that intend to build a hierarchically organized directory of Web pages, pioneered by Yahoo.com and expanded by the Open Directory Project (dmoz.org). Another important group of Web organization systems are various guided path systems [22]. Guided path systems allowed their users to build and publish Web paths, which are annotated sequenced of Web pages. Pages connected by the path are usually conceptually similar to each other and deliver a common narration. In addition to path systems, there are other simpler social systems that allow users to contribute social wisdom by grouping together similar pages without the need to provide comments or sequences [1, 5, 57]. All these Web organization systems offer a good source of *references*: the very fact of page or item sharing is a usual sign of its value. In addition, curation systems could provide other social information: page *comments*, page *categorization*, and user-judged *similarity* between pages, where the similarity criteria can be pages that were contributed under the same category, group, or path.

3.3.8 Social Bookmarking Systems

Social bookmarking systems could be considered to be one of the most successful curation systems. Integrating ideas from several earlier streams of research, including Web annotations, bookmark lists, classification systems, and recommender systems, social bookmarking systems have introduced a new way to organize and navigate socially contributed Web information [25, 28]. Social bookmarking systems allow their users to openly share information items (Web pages, photographs, research papers) while providing text *comments* and annotating these items with a set of free *tags*. These tags offer a nice balance for resource organization between unstructured comments and formal hierarchical classification systems, like the Open Directory Project. Modern social bookmarking systems, such as CiteULike, Flickr, or Pinterest also support one or more types of social connections; usually an ability to watch other users and form groups or communities. As a result, these systems have become a valuable source of various *social links*.

3.3.9 Consumption Systems and Recommender Systems

Consumption systems and recommender systems are two related groups of online systems. Consumption systems refer to all systems where users can access and

“consume” content. Some examples include online shopping systems where users can purchase goods, online journals sites/digital libraries where users can download content to read, and online music and video services that allow users to stream selected content. These systems differ from other Web systems, since obtaining an item in a consumption system requires a higher level of commitment from the user than a simple click (i.e., purchasing or downloading). These *consumption* actions left by past users allow the systems to accumulate more reliable social evidence of the value of the item. To help future users with higher-commitment decisions, consumption systems usually encourage post-consumption items *ratings* and *comments*. The last two aspects make consumer systems similar to classic recommender systems. The difference between these two types of systems is small: ratings and comments in consumption systems are directly used by end users, whereas such information in a recommender system is used by the recommender engine to proactively suggest relevant items for users to explore. In addition, classic recommender systems rarely offer the immediate ability to consume (purchase, play) recommended items. Nowadays, this difference has nearly disappeared, with most recommender systems being integrated into consumption systems and most consumption systems offering some form of recommendation. In addition, modern recommender and consumption systems frequently support certain forms of *social links*, which allow users to watch each other or form *groups*.

3.3.10 Location-Based Systems

The newest group of social systems are location-based systems, where users leave various feedback about objects located in a real space, such as restaurants, stores, cafes, or other physical objects. The feedback may range from simple check-ins to extensive reviews. Location-based offer several traditional types of social wisdom explored by other social systems, such as establishing *social links*, using *tags*, and others. However, they also add a unique new source of social information – real-world user traces. This information could be used to generate a whole new type of social recommendations [7].

4 The Book Structure

When assembling this book, our goal was to provide a broad overview of modern research on social information access. To ensure good coverage, we followed the classification of social information access techniques introduced in the previous section. In other words, the book represents an attempt to provide examples for every aspect of the introduced classification, and often covers the most important combinations of the aspects as well.

One particularly important goal for us is to provide sufficient coverage of social approaches for the three main types of information access: search, browsing, and recommendation. As shown in Table 3, four of the book’s chapters focus on social search, two focus on social navigation and browsing, and seven focus on recommendation. While we are not able to offer a chapter dedicated to social

visualization, we ensured that social browsing chapters address some visualization techniques. We also provided a chapter that focuses on social Q&A.

Table 3. Book chapters organized by the type of information access

Access type	Groups of technologies
Search	Chapters 4, 7–9 [12, 29, 49, 64]
Browsing	Chapters 5, 6 [17, 20]
Recommendation	Chapters 10–16 [6, 7, 27, 34, 39, 40, 51]
Visualization	Chapters 5, 6 [17, 20]
Q&A	Chapter 3 [50]

While the majority of modern social information access techniques could be classified as indirect-asynchronous, we also want to ensure that the book provides examples of direct and synchronous social information access (Table 1). For a browsing type of access, Chap. 5 [20] specifically discusses examples of direct-asynchronous, indirect-synchronous, and indirect-asynchronous social navigation. For search-based access, Chap. 4 [64] focuses on synchronous approaches in social search (more commonly known as collaborative search) and covers both direct and indirect collaboration. Chapters 7 [12], 8 [29], and 9 [49] focus on indirect-asynchronous techniques. There is no chapter focused on direct-asynchronous social search, but this area is covered by the chapter on Social Q&A [50] and some brief discussions in Chap. 4 [64]. For modern recommender technologies, which are asynchronous by their nature, Chap. 10 [39] focuses on classic rating-based recommendation, whose coverage is at the direct-asynchronous corner. Traditional classifications consider ratings as directly provided social feedback, however, it is less obvious nowadays since ratings in modern recommendations are frequently provided to get better recommendations rather than to recommend items for other users. Chapter 14 [34] offers a good discussion on this issue. The remaining recommendation chapters all focus on indirect-asynchronous approaches.

Our last goal in respect to the coverage is to ensure that the chapters cover the major types of social traces, both explicit and implicit. Including this aspect helps to uncover deep similarities among approaches from different groups when the comparison of different social information access is based on the same type of social traces. Table 4 explains how the book chapters cover most of the popular types of social traces.

For the search and browsing types of information access, Chaps. 5 [20] and 7 [12] provide coverage of most types of explicit and implicit sources. Both chapters offer a useful discussion and classification on covered sources (see Table 4 for more details). Similarly, on the recommendation side, Chap. 15 covers a range of information sources for *people recommendation*. Among the explicit types of social information, this book pays special attention to tags and links, due to

their historical and practical importance. Three separate chapters focus on using tags for search (Chap. 9 [49]), navigation (Chap. 6 [17]), item recommendation (Chap. 12 [6]), and people recommendation (Chap. 15 [27]). The book also offers a dedicated chapter on using information and social links for search (Chap. 8 [29]), as well as using social links for item recommendation (Chap. 11 [40]) and people recommendation (Chap. 15 [27]). The recommendation side of the book also provides dedicated chapters for two other types of explicit traces: ratings (Chap. 10 [39]) and text-based feedback (Chap. 13 [51]). The former represents the classic stream of research on collaborative filtering, while the latter focuses on a source that is rapidly increasing in both volume and practical value.

The use of all implicit information sources for recommendation is covered in Chap. 14 [34]. At the same time, there is a dedicated chapter for recommendations based on “real world trails” (Chap. 16 [7]). This chapter plays a special role in the book, as it also serves as the closing chapter. As mentioned in Sect. 2, user navigation in the real world served as a motivation for the pioneers of social navigation, who wanted to visualize traces and the presence of other users in an information space, just as they are visible in real space. It could be considered a sign of the field’s maturity that social information access techniques developed to help people navigate in information spaces have now been brought back to real spaces and are able to help guide users to the most relevant places.

Table 4. Main sources of social information (from implicit to explicit) and their coverage in the book

Source of social information	Search	Browsing	Recommendation
Search engine logs	Chapter 7 [12]	Chapter 5 [20]	
Browsing trails	Chapter 7 [12]	Chapter 5 [20]	Chapter 14 [34]
Real world trails			Chapter 16 [7]
Information links	Chapters 7, 8 [12, 29]	Chapter 5 [20]	
Annotations and comments	Chapter 7 [12]	Chapter 5 [20]	Chapter 13 [51]
Tags	Chapters 7, 9 [12, 49]	Chapter 6 [17]	Chapters 12, 15 [6, 27]
Social Links	Chapters 7, 8 [12, 29]		Chapters 11, 15 [27, 40]
Ratings			Chapter 10 [39]

As previously stated, the book chapters align nicely with our classification of social information access techniques, which provides a sound guide for reading and finding information in the book. In addition, to cover each topic, we invited top experts in the field with extensive knowledge on specific types of social information access techniques. In the process of preparing this book, each chapter went through several cycles of review and feedback among the editors, the authors of other chapters, and a team of PhD students who served as “pilot readers”. We hope that this book will serve as a good reference to the literature of social information access, as well as a handbook that can help readers in developing their own social information access approaches.

References

1. Abel, F., Frank, M., Henze, N., Krause, D., Plappert, D., Siehdnel, P.: GroupMe! - where semantic web meets web 2.0. In: Aberer, K., et al. (eds.) ASWC/ISWC-2007. LNCS, vol. 4825, pp. 871–878. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-76298-0_63
2. Ahn, J., Farzan, R., Brusilovsky, P.: Social search in the context of social navigation. *J. Korean Soc. Inf. Manag.* **23**(2), 147–165 (2006)
3. Barrett, R., Maglio, P.P.: Intermediaries: an approach to manipulating information streams. *IBM Syst. J.* **38**(4), 629–641 (1999)
4. Beeferman, D., Berger, A.: Agglomerative clustering of a search engine query log. In: Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 407–416 (2000)
5. Bernstein, M.: Web research: the Eastgate Web Squirrel. *SIGWEB Newsl.* **5**(1), 6 (1996)
6. Bogers, T.: Tag-based recommendation. In: Brusilovsky, P., He, D. (eds.) *Social Information Access*. LNCS, vol. 10100, pp. 441–479. Springer, Cham (2017)
7. Bothorel, C., Lathia, N., Picot-Clemente, R., Noulas, A.: Location recommendation with social media data. In: Brusilovsky, P., He, D. (eds.) *Social Information Access*. LNCS, vol. 10100, pp. 624–653. Springer, Cham (2017)
8. Brin, S., Page, L.: The anatomy of a large-scale hypertextual (web) search engine. In: Ashman, H., Thistewaite, P. (eds.) *Seventh International World Wide Web Conference*, vol. 30, pp. 107–117. Elsevier Science B.V. (1998)
9. Brusilovsky, P.: Social information access: the other side of the social web. In: Geffert, V., Karhumäki, J., Bertoni, A., Preneel, B., Návrát, P., Bieliková, M. (eds.) *SOFSEM 2008*. LNCS, vol. 4910, pp. 5–22. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-77566-9_2
10. Brusilovsky, P.: The other side of the social web: a taxonomy for social information access. In: *The 18th Brazilian Symposium on Multimedia and the Web*, pp. 1–4. ACM Press (2012). <http://dl.acm.org/citation.cfm?id=2382638>
11. Brusilovsky, P., Chavan, G., Farzan, R.: Social adaptive navigation support for open corpus electronic textbooks. In: De Bra, P.M.E., Nejd, W. (eds.) *AH 2004*. LNCS, vol. 3137, pp. 24–33. Springer, Heidelberg (2004). https://doi.org/10.1007/978-3-540-27780-4_6
12. Brusilovsky, P., Smyth, B., Shapira, B.: Social search. In: Brusilovsky, P., He, D. (eds.) *Social Information Access*. LNCS, vol. 10100, pp. 213–276. Springer, Cham (2017)
13. Bush, V., et al.: As we may think. *The Atl. Mon.* **176**(1), 101–108 (1945)
14. Chakrabarti, S., Dom, B., Raghavan, P., Rajagopalan, S., Gibson, D., Kleinberg, J.: Automatic resource compilation by analyzing hyperlink structure and associated text. In: *Proceedings of the Seventh International Conference on World Wide Web 7*, vol. 30, pp. 65–74. Elsevier Science Publishers B.V. (1998)
15. Dieberger, A., Dourish, P., Höök, K., Resnick, P., Wexelblat, A.: Social navigation: techniques for building more usable systems. *Interactions* **7**(6), 36–45 (2000)
16. Dieberger, A.: Supporting social navigation on the World Wide Web. *Int. J. Hum Comput Stud.* **46**(6), 805–825 (1997)
17. Dimitrov, D., Helic, D., Strohmaier, M.: Tag-based navigation and visualization. In: Brusilovsky, P., He, D. (eds.) *Social Information Access*. LNCS, vol. 10100, pp. 181–212. Springer, Cham (2017)

18. Dourish, P., Chalmers, M.: Running out of space: models of information navigation. In: SIGCHI Conference on Human Factors in Computing Systems, pp. 23–26 (1994)
19. Farzan, R., Brusilovsky, P.: Annotated: a social navigation and annotation service for web-based educational resources. *New Rev. Hypermedia Multimed.* **14**(1), 3–32 (2008)
20. Farzan, R., Brusilovsky, P.: Social navigation. In: Brusilovsky, P., He, D. (eds.) *Social Information Access*. LNCS, vol. 10100, pp. 142–180. Springer, Cham (2017)
21. Fitzpatrick, L., Dent, M.: Automatic feedback using past queries: social searching? In: *Proceedings of the 20th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 306–313 (1997)
22. Furuta, R., Shipman III, F.M., Marshall, C.C., Brenner, D., Hsieh, H.: Hypertext paths and the World-Wide Web: experience with Walden’s paths. In: Bernstein, M., Carr, L., sterbye, K. (eds.) *Eight ACM International Hypertext Conference (Hypertext 1997)*, pp. 167–176. ACM (1997)
23. Gauch, S., Speretta, M., Chandramouli, A., Micarelli, A.: User profiles for personalized information access. In: Brusilovsky, P., Kobsa, A., Nejdl, W. (eds.) *The Adaptive Web*. LNCS, vol. 4321, pp. 54–89. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-72079-9_2
24. Goldberg, D., Nichols, D., Oki, B.M., Terry, D.: Using collaborative filtering to weave an information tapestry. *Commun. ACM* **35**(2), 61–70 (1992)
25. Golder, S.A., Huberman, B.A.: Usage patterns of collaborative tagging systems. *J. Inf. Sci.* **32**(2), 198–208 (2006)
26. Golovchinsky, G., Pickens, J., Back, M.: A taxonomy of collaboration in online information seeking. In: *1st International Workshop on Collaborative Information Seeking* (2008). <https://arxiv.org/abs/0908.0704>
27. Guy, I.: People recommendation on social media. In: Brusilovsky, P., He, D. (eds.) *Social Information Access*. LNCS, vol. 10100, pp. 570–623. Springer, Cham (2017)
28. Hammond, T., Hannay, T., Lund, B., Scott, J.: Social bookmarking tools (I): a general review. *D-Lib Mag.* **11**(4) (2005). https://urldefense.proofpoint.com/v2/url?u=http-3A_www.dlib.org_dlib_april05_hammond_04hammond.html&d=DwIBaQ&c=vh6FgFnduejNhPPD0fl_yRaSfzY8CWbWnIf4XJhSqx8&r=UyK1_569d50MjvIUSODJYRW2epEY0RveVNq0YCmePcDz4DQHW-CkWcttrwneZ0md&m=8USpmwWQapLg5kc22S3fbLNcKx4X7l5lz5l3P9_xg1A&s=TNPKGVSVP4iidBHHcJ7b-Qh6slsuJGc9nB83u6Uh1_c&e=
29. Han, S., He, D.: Network-based social search. In: Brusilovsky, P., He, D. (eds.) *Social Information Access*. LNCS, vol. 10100, pp. 277–309. Springer, Cham (2017)
30. Han, S., He, D., Yue, Z., Brusilovsky, P.: Supporting cross-device web search with social navigation-based mobile touch interactions. In: Ricci, F., Bontcheva, K., Conlan, O., Lawless, S. (eds.) *UMAP 2015*. LNCS, vol. 9146, pp. 143–155. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-20267-9_12
31. Hill, W., Stead, L., Rosenstein, M., Furnas, G.: Recommending and evaluating choices in a virtual community of use. In: *SIGCHI Conference on Human Factors in Computing Systems, CHI 1995*, pp. 194–201. ACM (1995). <http://portal.acm.org/citation.cfm?id=223904.223929>
32. Hill, W.C., Hollan, J.D., Wroblewski, D., McCandless, T.: Edit wear and read wear. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 1992*, pp. 3–9. ACM Press (1992)
33. Höök, K., Benyon, D., Munro, A.J.: *Designing Information Spaces: The Social Navigation Approach*. Springer, Berlin (2003). <https://doi.org/10.1007/978-1-4471-0035-5>

34. Jannach, D., Lerche, L., Zanker, M.: Recommending based on implicit feedback. In: Brusilovsky, P., He, D. (eds.) *Social Information Access*. LNCS, vol. 10100, pp. 510–569. Springer, Cham (2017)
35. Joachims, T.: Optimizing search engines using clickthrough data. In: *Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2002)*, pp. 133–142. ACM (2002)
36. Kahan, J., Koivunen, M.R., Prud’Hommeaux, E., Swick, R.R.: Annotea: an open RDF infrastructure for shared web annotations. *Comput. Netw.* **39**(5), 589–608 (2002)
37. Kantor, P.B., Boros, E., Melamed, B., Mekov, V., Shapira, B., Neu, D.J.: Capturing human intelligence in the net. *Commun. ACM* **43**(8), 112–116 (2000)
38. Keller, R.M., Wolfe, S.R., Chen, J.R., Rabinowitz, J.L., Mathe, N.: A bookmarking service for organizing and sharing URLs. In: *Sixth International World Wide Web Conference*, pp. 1103–1114 (1997). *Computer Networks and ISDN Systems* **29**(8–13) (1997). <http://www.ra.ethz.ch/CDstore/www6/Technical/Paper189/Paper189.html>
39. Kluver, D., Ekstrand, M., Konstan, J.: Rating-based collaborative filtering: algorithms and evaluation. In: Brusilovsky, P., He, D. (eds.) *Social Information Access*. LNCS, vol. 10100, pp. 344–390. Springer, Cham (2017)
40. Lee, D., Brusilovsky, P.: Recommendations based on social links. In: Brusilovsky, P., He, D. (eds.) *Social Information Access*. LNCS, vol. 10100, pp. 391–440. Springer, Cham (2017)
41. Li, W.S., Vu, Q., Agrawal, D., Hara, Y., Takano, H.: PowerBookmarks: a system for personalizable web information organization, sharing, and management. In: *8th International World Wide Web Conference*, pp. 297–311. Elsevier (1999). <http://www8.org/w8-papers/3b-web-doc/power/power.pdf>
42. Lieberman, H.: Letizia: an agent that assists web browsing. In: *The Fourteenth International Joint Conference on Artificial Intelligence*, pp. 924–929 (1995)
43. Liesaputra, V., Witten, I.H.: Seeking information in realistic books: a user study. In: *Joint Conference on Digital Libraries, JCDL 2008*, pp. 29–38 (2008)
44. Lueg, C., Fisher, D.: *From Usenet to CoWebs: Interacting with Social Information Spaces*. Springer, Berlin (2003). <https://doi.org/10.1007/978-1-4471-0057-7>
45. Maltz, D., Ehrlich, K.: Pointing the way: active collaborative filtering. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 1995*, pp. 202–209. ACM (1995)
46. Marshall, C.C., Bly, S.: Sharing encountered information: digital libraries get a social life. In: *The 4th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL 2004)*, pp. 218–227 (2004)
47. McBryan, O.A.: GENVL and WWW: tools for taming the web. In: *the 1st International World Wide Web Conference*, p. 7990 (1994)
48. Munro, A.J., Höök, K., Benyon, D.: *Social Navigation of Information Space*. Springer, Berlin (1999). <https://doi.org/10.1007/978-1-4471-0837-5>
49. Navarro Bullock, B., Hotho, A., Stumme, G.: Accessing information with tags: search and ranking. In: Brusilovsky, P., He, D. (eds.) *Social Information Access*. LNCS, vol. 10100, pp. 310–343. Springer, Cham (2017)
50. Oh, S.: Social Q&A. In: Brusilovsky, P., He, D. (eds.) *Social Information Access*. LNCS, vol. 10100, pp. 75–107. Springer, Cham (2017)
51. O’Mahoney, M., Smyth, B.: From opinions to recommendations. In: Brusilovsky, P., He, D. (eds.) *Social Information Access*. LNCS, vol. 10100, pp. 480–509. Springer, Cham (2017)

52. Raghavan, V.V., Sever, H.: On the reuse of past optimal queries. In: 18th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 344–350. ACM (1995)
53. Remde, J.R., Gomez, L.M., Landauer, T.K.: SuperBook: an automatic tool for information exploration hypertext? In: The ACM Conference on Hypertext, Hypertext 1987, pp. 175–188 (1987)
54. Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., Riedl, J.: GroupLens: an open architecture for collaborative filtering of netnews. In: ACM 1994 Conference on Computer Supported Cooperative Work, pp. 175–186. ACM Press (1994). <https://doi.org/10.1145/192844.192905>
55. Rucker, J., Polano, M.J.: SiteSeer: personalized navigation for the web. *Commun. ACM* **40**(3), 73–75 (1997)
56. Shardanand, U., Maes, P.: Social information filtering: algorithms for automating “word of mouth”. In: Katz, I., Mack, R., Marks, L. (eds.) The SIGCHI Conference on Human Factors in Computing Systems, CHI 1995, pp. 210–217. ACM (1995)
57. Shipman, F.M., Marshall, C.C.: Spatial hypertext: an alternative to navigational and semantic links. *ACM Comput. Surv.* **31**(4es), 1–5 (1999)
58. Smyth, B., Coyle, M., Briggs, P.: HeyStaks: a real-world deployment of social search. In: Sixth ACM Conference on Recommender Systems, RecSys 2012, Dublin, Ireland, 9–13 September 2012, pp. 289–292 (2012)
59. Smyth, B., Freyne, J., Coyle, M., Briggs, P., Balfe, E.: I-SPY — anonymous, community-based personalization by collaborative meta-search. In: Coenen, F., Preece, A., Macintosh, A. (eds.) *Research and Development in Intelligent Systems XX*, pp. 367–380. Springer, London (2004). https://doi.org/10.1007/978-0-85729-412-8_27
60. Svensson, M., Höök, K., Laaksolahti, J., Waern, A.: Social navigation of food recipes. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 2001, pp. 341–348. ACM (2001)
61. Takano, H., Winograd, T.: Dynamic bookmarks for the www: Managing personal navigation space by analysis of link structure and user behavior. In: Grnbk, K., Mylonas, E., Shipman III, F.M. (eds.) *Ninth ACM International Hypertext Conference (Hypertext 1998)*, pp. 297–298. ACM Press (1998)
62. Wexelblat, A., Maes, P.: Footprints: history-rich tools for information foraging. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 1999, pp. 270–277. ACM (1999)
63. Xue, G., Zeng, H., Chen, Z., Yu, Y., Ma, W., Xi, W., Fan, W.: Optimizing web search using web click-through data. In: Proceedings of the Thirteenth ACM International Conference on Information and Knowledge Management, pp. 118–126. ACM (2004)
64. Yue, Z., He, D.: Collaborative information search. In: Brusilovsky, P., He, D. (eds.) *Social Information Access. LNCS*, vol. 10100, pp. 108–141. Springer, Cham (2017)