Discovering Information in Context

Paul Solomon
University of North Carolina at Chapel Hill

Introduction and Overview

This chapter has three purposes: to illuminate the ways in which people discover, shape, or create information as part of their lives and work; to consider how the resources and rules of people’s situations facilitate or limit discovery of information; and to introduce the idea of a socio-technical systems design science that is founded in part on understanding the discovery of information in context. In addressing these purposes the chapter focuses on both theoretical and research works in information studies and related fields that shed light on information as something that is embedded in the fabric of people’s lives and work. Thus, the discovery of information view presented here characterizes information as being constructed through involvement in life’s activities, problems, tasks, and social and technological structures, as opposed to being independent and context free. Given this process view, discovering information entails engagement, reflection, learning, and action—all the behaviors that research subjects often speak of as making sense—above and beyond the traditional focus of the information studies field: seeking without consideration of connections across time.

The chapter, thus, offers a reconceptualization and refocusing of the work of information studies from wondering why people use or do not
use information institutions, systems, or sources toward considering what information is to people, how stuff ends up becoming information, and how information so discovered influences further action. The idea is that through such an understanding in context, a foundation will be set for designing collections (content and contexts), organizational schemes (representations and classifications), retrieval mechanisms, and displays that fit the problems and tasks of life and work. The hope is that this idea of fit might lead to the creation of systems that, through their flexibility, accommodate people at various stages in their discovery of information.

This view examines, in a positive way, such questions as why people do not use information systems when, in the eyes of information professionals, these systems are what people should be using. The aim is to develop an understanding of the difference between what information professionals label as information—the representations of objects or objects themselves that are contained in, for example, databases and libraries—and what becomes information as people move through life and interact socially as they discover what they need to know to function. Another way of labeling this distinction is with the anthropologist's *emic*—the perspective of the outsider—and *etic*—the perspective of the insider (Sandstrom & Sandstrom, 1995). This distinction is made to highlight similarity and difference in order to inform design.

Given the constraints of space and limits on number of references, there was no attempt to be exhaustive in selecting the literature employed here to elucidate the view of discovering information in context. There is a tendency to cite the recent and draw on work from related disciplines. Most attention is given to material that explores the nature of discovering information in context through social (or socio-technical) studies of information.

This is a significantly different approach from that taken by Proper and Bruza (1999), for instance, whose consideration of *information discovery* is logic-based and represents a burgeoning literature emphasizing resource discovery on the Web. One purpose of resource discovery and other approaches that fall under the rubric of information discovery is to support individuals' shaping or making of information as they interact with information systems and technology. The insights gained from research that explores what information means to people in the context
of their lives and work may inform information discovery and assist in information system design. Therefore, a portion of this chapter explores approaches that show promise in helping people discover information in context. Taken together, the social studies of information and the systems that support exploration and learning move toward the integration of research involving people and systems called for by Saracevic (1999).

There are no previous reviews (ARIST or otherwise) that explicitly focus on discovering information in context in the manner of this chapter. Nevertheless, there are several reviews that provide a foundation for the view presented here. These include the work by Faibisoff and Ely (1976), which surveyed studies of information needs and, in the spirit of the time, built upon the research to offer generalizations about information needs and guidelines for the design of information systems. This taking of stock is, perhaps, most significant in that it leads to a realization that such generalizations and principles seldom stand alone, but require the addition of context in order to produce products that support the discovery of information within situations and for particular tasks.

Dervin and Nilan's (1986) review addresses this issue, in part, by calling for a paradigm shift from a focus on system to a focus on user. By understanding situations, gaps, and use, it may be that we can understand what information means to people in context and provide information structures that support the interactions of situation, gap, and use. This work also focuses attention on the process involved in this interaction, which is labeled sense making. Katzer and Fletcher (1992) focus on managers and what constitutes information in managerial situations, where the information that (management) information systems provide and the information that managers use are often incongruent. McKinnon and Bruns (1992) and Auster and Choo (1993) provide reviews related to the Katzer and Fletcher work that focus on internal operations and environmental scanning, respectively. Palmquist (1992) highlights the influence or shaping effect of technology on individuals. Sugar (1995) brings people and systems together in his review. Pettigrew, Fidel, and Bruce (in press) review conceptual frameworks employed in, or resulting from, studies of information behavior.

Additional sources of theory and research relating to this review include the proceedings of the Information Seeking in Context conferences (Vakkari, Savolainen, & Dervin, 1997; Wilson & Allen, 1999), and
the special issue of *Information Processing & Management* (1999, 35[6]) on information seeking in context.

Discovering information in context is an in-the-world idea; thus, one way you can make sense of the process of information discovery is to reflect on what becomes information for you as you read on. You might also consider how you shape what turns out to be information to you in your own way as a reflection of your own interests and concerns. This adventure begins with a consideration of some of the theoretical and conceptual views underlying the discovery of information in context. This is followed by explorations of such discovery in terms of process, structure, contexts, methods, and design.

**Foundations**

The previous section referred to some works that contributed to a conceptualization of individuals' discovery of information as they move through life in terms of structure, process, systems in interaction, and other people. This section adds the contributing theoretical view of structuration as well as the recent contributions of several continuing research streams.

**Structuration**

As discovery of information is viewed as a learning process—although it could also be a particular act within the process—it is useful to recognize that a system of activity functions through the interaction of structure and process. Therefore, ideas about the constitution of socio-technical systems seem basic to identifying and understanding the constituents of information discovery (e.g., making sense, reflection, thinking, learning, engagement). Structures of various sorts exist to focus attention and in doing so sometimes facilitate, inhibit, or prohibit actions. Thus, information systems focus attention on a particular sphere of activity, allow some tasks to be performed, prohibit others, and through these capacities exercise power (Introna, 1997). The act of structuring, then, limits our attention and view of situations.

Anthony Giddens' (1984) structuration theory and Luhmann's (1995) theory of autopoietic systems both provide similar general templates for considering the interaction of structure (resources and rules in Giddens'
terms) and action. Structuration theory, in particular, recognizes the duality of structure and action, where structure encourages certain kinds of actions and prohibits others. The resulting actions reinforce, adapt, or change the facilitating structure.

Orlikowski (1992) uses Giddens’ structuration theory to focus explicit attention on technology as a structure: resources and rules. Her field study, a software consulting firm, illustrates how the separate development and operation of a technological system led to its inability to fully fit the operational environment. This is an empirical expression of what Luhmann (1995) conveys in his treatise on the self-referential nature of social systems (autopoiesis): The structure of the system and the kinds of actions the system recognizes or permits cut the system off from other systems through, for instance, a lack of congruence of terminology or fit with the activities of other systems. Attempts to overcome these self-limiting effects of self-referential systems by incorporating the structures of other self-referential systems result in increased system complexity, which further limits a system’s ability to meet the needs of other systems. This is also an expression of the challenge that we face in gaining an understanding of how people discover information in context and using that understanding to support systems design.

Barreau’s (in press) study of a commercial off-the-shelf software package to support transactions (e.g., ordering, payments) in the publishing industry illustrates this autopoietic characteristic of systems. In each of the three organizational cases studied, the software did not quite fit and the organizations involved made do by forgoing desired functions or creating their own supporting systems, which resulted in a less-than-seamless integration. Bailey’s (2000) analysis of nursing work in relation to an effort to implement a computerized patient record also illustrates how technological structures that are imposed on a task environment (labor and delivery) may not fit with the primary task at hand (patient care) as hospitals try to find ways of meeting secondary (e.g., regulatory, insurance) requirements for information. The lessons learned—both direct and indirect—from such intensive studies of information systems in action provide rich fodder for both general information system improvements and more specific information management structures in context.
**What to Call the Processes and Actions of People as they Discover Information**

One of the continuing struggles within the field of information studies has been the question of how to conceptualize, for the purpose of informing system design, people and the social and technological structures that promote and inhibit access to that which becomes information. There has been a tendency to focus on the individual’s encounter with information systems—often bibliographic. This individual focus has led us to use such terms as *information seeking* and *information behavior* as we struggle to conceptualize the special nature of the interaction of human and system. These terms are, themselves, contentious for various reasons as evidenced by the December 1999 discussion thread on the JESSE listserv (http://listserv.utk.edu/cgi-bin/wa?A1=ind9912&L=jesse). The view taken here of discovering information in context aims at adding social and temporal dimensions to our conceptual models and thinking. This review, therefore, turns to several continuing or emerging streams of research that contribute to the theoretical bases of the discussion.

**Life in the Round**

Chatman’s (1996, 1999) studies of the information worlds of participants in a job training program, janitors, retired women, and women in prison has led her to develop a *theory of life in the round*, where context shapes inhabitants’ definitions of what information is as well as appropriate ways of seeking and using it. Chatman’s (1999) notion of *roundness* captures the difference between a focus on just information seeking, which expresses an ideal developed for scientists and engineers and applies it to all groups, and discovering information in context, which takes notice of the roundness created by the interaction of people, technology, and social structure. Chatman’s (1996) concepts of *secrecy, deception, risk taking, and situational relevance*, when taken together, provide a basis for explaining why people cut themselves off from relevant information sources because the use of those sources may influence what happens in the future in a negative way—as was the case for an elderly woman who was afraid to discuss her dizziness with others for fear that she would lose her independence and end up in a nursing home.
Savolainen's (1995) concepts of *mastery of life* and *way of life* add to the picture by focusing attention on, respectively, the ways by which people maintain orderliness in their lives and balance work and leisure. Solomon (1999) uses the concept of *rounding* to consider from several different perspectives (i.e., work planning, college students' use of the Web, and travel planning) the recurrent patterns of action that structure how people discover information in order to understand what leads someone to break such a pattern.

Overall, this rounding work provides insights into how the social and the individual interact to shape the discovery of information.

**Making Sense and Information Search as Processes**

A primary contributor to understanding the discovery of information in context is the sense-making theory and associated sense-making methodology of Dervin (1999a). This work is important because it focuses attention on situations, information gaps, and the actions that people take to bridge these gaps. This integrated theoretical and methodological approach, thus, highlights people's actions and, consequently, emphasizes *verbings* (acts) over *nounings* (states) in understanding people's making of sense. The early descriptions (cf. Dervin & Nilan, 1986) of the theory and supporting sense-making research tend to emphasize individual sense making. More recently Dervin (1999a) has confronted the social and temporal aspects of sense making. She has also elaborated a theory of design based on sense-making theory that ties design to the possibility of helping people in designing, shaping, and creating information as a support to their making of sense (Dervin, 1999b).

Kuhlthau's (1993a, 1993b, 1997, 1999) *information search process* work is an examination of the ways in which people discover information through a constructive process of learning involving feelings, thoughts, and actions intertwined in a nonlinear progression. There are two significant results of this work that have contributed to an understanding of how people discover information in context. The first is the use of the research findings in developing a normative process model to guide the design of information services and training. This model simplifies the complex, nonlinear behavior of people during the information search into an idealized set of stages (i.e., task initiation, topic selection, prefocus exploration, focus formulation, information collection, search closure,
writing) that require different sorts of actions while producing different kinds of feelings and thoughts. The second result is the recognition that, whatever stage the information search is in, it is influenced by what came before, and, in turn, influences what comes next.

Weick (1995) developed a social framework for understanding sense making in organizations. This framework, in particular, specifies sense making as being (1) grounded in the organization's construction of its identity, (2) an organization's ties to its past (retrospect), (3) what an organization does to enact its environment (enactment), (4) the nature of social interaction within the organization, (5) the emphasis within organizational life on production versus adaptation (ongoing), (6) the cues (or what constitutes information) employed by the organization (extracted cues), and (7) the organization's norms for interpreting cues (plausibility).

Building on the work of Dervin, Kuhlthau, Weick and others, Solomon (1997b, 1997c, 1997d) explores people's information-related behavior over three years of a work planning process in a government agency. This exploration explicitly considers the influence of perspective in understanding what information is in context by viewing the happenings from personal, social, and temporal perspectives. The study also provides insights into the reality of information discovery as a grounded process in which people talk about making sense of their situation and discover information as they engage in work.

Overall, the theoretical and research work related to how people make sense of their situations and construct (or discover) information has focused on social and dynamic, as well as individual, aspects of this process. It has also been exploring the implications of theory and research findings for practice.

Information Encountering

Erdelez (1995, 1997, 2000) has labeled people's accidental discovery of information as information encountering. This term is employed to make a distinction with information seeking, which is seen as a form of information acquisition that is focused on specific needs, sources, and tasks. Erdelez's research in academic and Web environments has led to the identification of the following functional components of information encountering: noticing, stopping, examining, storing, using, and returning. Ross (1999, reading for pleasure), Savolainen, (1995, everyday life
information seeking), and Williamson (1998, older adults) have also studied this "by chance" encountering of information during the course of activities that do not involve direct information seeking. This work taken together suggests a general pattern of behavior by which people discover information as they make sense of what they encounter. This process is apparently a variable one, because much depends on motivations, interests, and connections with life and work. Also, this phenomenon of discovery as people encounter or come into contact with stuff of various sorts seems robust in situations both directed—in the vein of information seeking—and undirected—in the vein of reading books for pleasure or of attending to the daily news program or paper.

**Reflecting, Thinking, Learning, and Knowing**

In the *more, more, more* and *faster, faster, faster* mode of many technology-driven information societies (Davis & Meyer, 1998), consideration of the need for support of reflection, thinking, and learning in the design of socio-technical systems is often lost. Reflection and thinking are not visible and are, consequently, seen by some as the wasting of time. Schön (1983) provides a compelling set of examples of the importance of *reflection in practice* in the development of professional expertise. Using the vehicle of the teamwork involved in the navigation of a naval ship, Hutchins (1995) provides a comprehensive analysis from a cognitive anthropological perspective of cognition in context. A major emphasis of this work is on individual, team, and organizational learning, which together illustrate how individual, social, organizational, and task factors interact to produce learning and change not only in the performance of the navigation task, but also in the broader organizational culture.

After analyzing studies of the impact of new technologies on organizational learning and knowledge management, Blackler (1995) suggests that knowledge is the result of the active process of knowing, which is something that people do as a natural part of their being. Knowing is *mediated* by resources and rules. Knowing is *situated* by time period and task. It is *provisional* in that it is tested and perhaps reconstructed. It is *pragmatic* in that it is goal oriented. Finally, knowing is *contested* through the interactions of daily life and work. Thus, we are left with
dimensions of a process of information discovery that is influenced by the characteristics of the structures in place.

Cole's (1997, 1998) study of information acquisition by history doctoral students illustrates the *in between-ness* in which people often find themselves when they are engaged in activities that result in information discovery. Cole found that his subjects were often *unconscious-unfocused* about what they were looking for because they were in the middle of trying to put the puzzle of their research together. Thus, we may do a disservice to people when we design systems that emphasize finality instead of providing support for reflection, thinking, and learning as options. Limberg's (1998, 1999) study of school children in Sweden adds the additional insight to the reflection, thinking, learning, and knowing in information discovery of a requirement for a substantive basis of content knowledge to support this process.

The review now turns to key aspects of information discovery—structure, process, and context—that were introduced in the preceding discussion.

**Structure as a Facilitator/Inhibitor in Discovering Information**

It is the purpose of this section to highlight the influence of structures of various kinds upon the process of discovering information. In particular, this section focuses on what research tells us about the role of resources and rules in enabling or preventing information discovery.

**Tasks**

Vakkari (1999) reviews studies of problem-solving tasks and concludes that the body of research does not have much to say about how variations in the problem situation influence search and relevance assessment behavior. He goes on to explore a variety of factors that may influence what can happen during the course of people's discovery of information. These factors include the task and its complexity, the nature of the problems (e.g., structured, unstructured) that the task supports (directly or indirectly), and prior knowledge (or experience) relating to the task. Finally, he highlights research findings that shed light
on tasks and information types (e.g., Byström, 1999; Byström & Järvelin, 1995; Kuhlthau, 1993a), tasks and search strategies (e.g., Ellis & Haugan, 1997), and tasks and relevance criteria (Wang, 1997).

In addition to complexity, Marchionini (1995) adds specificity of the goal or motivation, volume and timing of answer, and time-to-completion as other task-related factors. Particularly in work-related situations, the task is a fundamental force that influences how and why people select sources, discover information in sources, evaluate information so discovered in relation to the task, and gain new insights related to completion of the task. As the task is a force for focusing attention, it also limits what workers attend to.

### Communicative Structures

A wide range of social, technological, and socio-technical structures has arisen to support communication. These extend from the traditional approaches of face-to-face conversation, telephone calls, and meetings to real-time chat sessions, discussion forums, and e-mail. Each of these has a range of advantages and disadvantages. Solomon (1997c) considers how the variety of such structures extant in the organization influences communication related to work planning. From the standpoint of discovering information in context, these mechanisms provide situations for possibly obtaining direct access to information in response to a request. It is also possible that the initial request will lead to an interaction among parties to the communicative event that helps all involved discover or learn something they would not have learned with a direct response. Thus, another function of these mechanisms is to provide a context of interaction for gaining insights from others regarding the situation at hand.

### Meetings

The meeting is one event in particular that gets much bad press in the literature of organizational life. Schwartzman’s (1989) study of meetings as communicative events provides numerous examples of situations that lead to negative feelings about meetings (e.g., domination by a few participants, lack of resolution of issues). It also provides considerable evidence of processes that are employed during meetings such as
storytelling, explaining, arguing, and focusing (consensus building) that make meetings one of the primary forums for information discovery, creation, and learning in organizations.

**Texts**

There is something about the shape and structure of texts that people learn to employ as an aid in focusing attention on critical elements as they use a text. This structuring is perhaps clearest in certain disciplinary text types (e.g., the psychological experiment or the clinical drug trial), but it is also something that is learned in elementary school as children, for instance, are taught how to write a letter, book report, or other text format. Imagine trying to read a research report that is just plain text with no headings to mark segments of a text. Cues provided by the headings and other markers in texts aid readers in discovering information. Researchers (e.g., Toms, Campbell, & Blades, 1999; Vaughan & Dillon, 1998) have begun to exploit text structure and shape in the design and use of new text forms, such as digital documents. Digital formats allow a variety of formatting possibilities, including adaptive texts that can be formatted at the command of their users (Hars, 2000), and possibly annotation and feedback to authors to gain some of the benefits of more interactive oral communication formats.

**Terminology, Classification Schemes, and Category Structures**

While many of the structures (e.g., categorizations, index languages, representations, metadata, displays) created by information professionals are designed to aid in access, discovery, and retrieval, they often fail in one or more of these functions. Thus, research in this area assists in our understanding of the role of such structures in facilitating the discovery of information as well as in supporting such discovery in practice.

Haas and Hert (in press) have been exploring the use of terminology both from the point of view of U.S. government statistical agencies (e.g., the Bureau of Labor Statistics) and the terms employed by users of their statistical products. There is great variety on both fronts, which leads to the challenge of understanding and communicating the definitions underlying the statistical terminology employed by the producers as
well as communicating the interests of the users of these statistics to possibly make the contents available in forms and formats that would facilitate their discovery as information.

Bowker and Star (1999) provide an essay on the consequences of terminological and classification infrastructure decisions by institutions in nursing—nursing intervention classification—and medicine—international classification of diseases—among others. They argue that the development of flexible classifications, which incorporate a record of their evolution, is key for meeting the variety of uses of such classifications. Solomon (2000) considers the tension between stability and change that is evident in Bowker and Star’s analysis of the classifications that they studied.

These studies discuss only a small sample of structural elements that may influence how people discover information. There are many others. Pettigrew (1999), for instance, documents the various elements that limit discovery possibilities in a foot clinic for the elderly. These structures include: the physical environment (building layout, weather), clinic activities (waiting and treatment processes), the nurse’s situation (knowledge of local resources and patient, workload), and the patient’s situation (personal circumstances, ability and desire to interact).

Information Discovery as Process

This section focuses on what research tells us about how the discovery of information unfolds over time. Kuhlthau’s (1993a, 1993b, 1999) information search process model has served as a foundation for most of the subsequent research that attempts to either (1) sequence or classify process stages or (2) relate a particular stage to nature, form, or quality of information that comes into play.

Sequencing and Classifying Information-Related Behaviors in Tasks

Algon’s (1999) study of project teams in the drug development process of a large pharmaceutical firm and Cole’s (1997) study of doctoral students in history have both added to our understanding of the sequencing of the discovery process. Algon mapped the relationship of tasks within the drug development process, the form of interaction (with people, ideas
or information, or things), and information-related behaviors such as absorbing, conceptualizing, manipulating, organizing, seeking, providing, and verbalizing. Overall, she found that the type of task and the information-related behaviors applied were related.

Cole focused on patterns of cognitive activity in his study as a basis for specifying five stages in history doctoral students’ discovery process. These stages include (1) the opening, where some anomaly or conflict with previous knowledge raises a question for further investigation; (2) puzzling over or representational activity, where the nature of the anomaly is considered and theories or speculations are offered; (3) searching for corroborating evidence, where support for possible explanations is sought; (4) closing, where the matter has been satisfactorily resolved; and (5) moving on, where the previous information process sets the stage for further activity. In this last, the learning that took place during stages 1 through 4 enables new insight and understanding.

This sort of work, which maps either the physical progress or cognitive aspects of a task, is important because it provides a context for understanding where, when, why, and how information comes into play in the task. This understanding, in turn, may provide insight into how such activity could be supported.

**Specifying the Relation of Information Attributes to Task or Process Steps**

The few studies that take a process view in considering how what is or becomes information changes with the phase or stage of a task or process seem to cluster in three areas: selecting documents, learning a skill, and completing a work task.

**Document Selection**

A series of studies by Wang and White (1995), Wang (1997), and Wang and Soergel (1998) considers the changes in criteria and term use of agricultural economics graduate students and faculty as they decided whether to read and then cite a particular work. Wang and Soergel (1998), in particular, bring this series of research reports together into a model of document selection that lists key information elements of documents (e.g., title, author, date), criteria for evaluating the acceptability
of a document on the basis of its information elements (e.g., topicality, quality, novelty), and values (e.g., social, emotional, functional) that come into play in the ultimate decision to accept, reject, or hold on to a document for reading.

Tang and Solomon (1998) studied the relevance judgment process of a graduate student by focusing on two stages of (1) document selection using retrieved bibliographic records, and (2) document evaluation using the full texts of items selected as possibly relevant on the basis of their bibliographic representations. During the first stage (document selection), the subject followed a three-step process of (1) a slow, measured interaction with the retrieved items on a general topical basis to see what was in the retrieved set by indicating that a document was relevant, not relevant, or possibly relevant; (2) the sudden discovery or realization as a result of the interaction with the bibliographic records of specific criteria that were important to her; and (3) the quick re-evaluation with some well-formed criteria (e.g., recency of publication, document type, topical relevance). The second stage evaluation was more content-based, showing that the learning that had taken place during the subject’s interactions with the texts and the more detailed exposition of the full texts enabled the subject to move to a more sophisticated level of application of criteria.

Tang (1999) followed up on the previous study by comparing and contrasting a laboratory experiment (using undergraduates enrolled in an introductory psychology class) and a naturalistic study (using advanced psychology graduate students enrolled in a meta-analysis class) that again used a two-staged document selection process to understand how relevance judgment criteria changed from stage one to two. Again, there were substantial differences among the criteria employed at the two stages and the sophistication of their use, though those differences varied among the two subject groups in reflection of the tasks at hand (prepare an outline of a presentation versus write up the results of their meta-analyses) and differences in their depth of knowledge. The findings suggest various possibilities for assisting people during document selection through criteria filters that would be available during the document selection process.
Learning Process

Moving beyond the world of information retrieval, Bergeron and Nilan (1991) focused on sources used by people learning to use a word processing program. In doing so, their interest involved understanding how the learning that took place influenced access to information about the software. Among other things, their findings suggest that the lack of fit of various printed and online help sources (prepared by experts) with the needs of the novice learners may have driven them to seek interactions with people, including fellow learners, who were able to help through direct interaction.

Work Process

Garber and Grunes (1992) offer a remarkable look at art directors' processes of discovering information as they searched for and selected images for advertisements. In specifying the process, the authors developed a model that highlighted information-intensive subtasks (i.e., creating the initial concept, preparing compositions, finding photos) in which the art directors naturally employed different sorts of strategies to get what they needed. For example, initially the art directors formulated an artistic concept—a general sense of what the layout is trying to do—and an image concept—the sorts of images that might fulfill the artistic concept. The task process that accompanied this activity started with a general overview of the type of image along with the criteria or restrictions that the art director had in mind. This set the stage for viewing some images that, on the surface, met the initial requirements. As each image was viewed, additional criteria or restrictions were added, or a previous criterion or restriction was altered. This interaction with candidate images was a critical part of the process of selecting images to assess their relevance to the theme of the advertisement. Based on this sort of evidence, the authors developed a prototype system to support image selection that allows its users to search for similar images—ones that meet the basic criteria—or explore alternative images—ones that are contrasts (e.g., a picture of a woman or child if a man is the starting point).

Research efforts that inform our understanding of information discovery from a process point of view are rare, but offer the potential for better understanding how interaction influences human behavior.
Contexts of Information Discovery

This section explores the situational dimension by very selectively considering research relating to information discovery across a variety of contexts. The research on everyday life by Chatman (1996, 1999) and Savolainen (1995) has already been mentioned. The specific context research offered here includes work life, searching the Web, and collaboration.

Work Life

Wenger's (1998) essay on communities of practice develops a social theory of learning that is based on the concepts of practice, meaning, community, learning, locality, and boundary. The theory brings these concepts together to show how they enable (or do not enable) the development of a community of practice that supports the discovery of information through situated learning and situated action. Nardi and O'Day's (1999) work on information ecologies recognizes the importance of fit between technologies and the people, situations, and tasks that they serve. Both of these works are helpful in that they begin to identify what it is about people working together to accomplish some end that is amenable to design.

Architects

Cohill (1993) studied architects at work by focusing on the information that they created, sought, and employed during the life of an architectural design project. While his motivation for doing this was to develop software to support design, what he found was that there was less need for design support than for support on business functions as the generation, coordination, and use of project management information often overwhelmed the design element of the project. Ultimately, he sketched a design for a small set of information-related tools that focused on the management aspects of a design project.

Nurses/Physicians in Clinical Situations

Several studies have either directly or indirectly illuminated the process of information discovery as medical personnel interact with
patients and other staff to either diagnose illness, participate in patient care, or contribute to student or staff training. Forsythe, Buchanan, Osheroff, and Miller (1992) studied information situations in the clinical settings of teaching hospitals. In analyzing these situations they found that it was often impossible to separate the information from the context, as knowledge of the situation was critical for understanding the information problem. The character of information varied considerably across situations and was defined during the interaction. Timpka and Arborelius (1990) studied the nature of information in telephone interactions between nurses and patients. They found a mismatch between the diagnostic emphasis of the nurses in their requests for information from patients and the needs of patients for empathetic support during an illness. Bailey (2000) explored how nurses discover information in connection with the implementation of a computerized patient record. Her work provides insights into what constitutes information for the nurse during the labor and delivery process in contrast to what is collected in the computerized patient record.

Public Defenders

Hara (2000) studied public defenders in two counties—one large and one small—as she tried to understand the forces that contributed to the development of a community of practice. In particular, she identified the ways in which these two very different groups developed shared meaning, employed informal networks, developed and maintained a supportive culture of trust, and engaged in knowledge building. In the smaller county, there was evidence of scaffolding to support people at early stages in their careers, whereas the larger county housed new and established attorneys in different locations because of space issues. This spatial separation led to a lack of mentoring for the newer attorneys. The attorneys in the larger county also had heavier caseloads. The attorneys in the smaller county seemed more willing to share personal experiences and various supportive acts were in evidence (e.g., showing up for a difficult trial to offer moral support and feedback). There was also evidence of sharing of insights regarding the appropriateness of sources of information and the sorts of information that would be needed to develop an argument for a given case. The use of information technology was also explored. The attorneys’ use of the public defenders’ listserv in the larger
county was similar to the attorneys' reliance upon each other for support in the smaller county.

**Securities Analysts**

Baldwin and Rice (1997) studied the influence of individual characteristics (e.g., age, gender) and institutional resources (e.g., staff, budget size) on use of information sources and outcomes. They found no statistically significant influence attributed to personal characteristics, but concluded that institutional resources did influence information sources and communication channels, as well as outcomes. While these are interesting findings, the link between the structuring effects of institutional resources, sources, and outcomes is not presented. It is also not clear exactly what information is to these analysts, or how information is discovered in the environmental monitoring that is a major component of the analysts' work.

Kuhlthau's (1997) study of one analyst is informative as an indicator of the nature of the analyst's work—routine monitoring resulting in reports providing information to clients with a high certainty of what is being reported, and more complex reports with considerable uncertainty as to future events—and the strategies employed to discover the information needed for both types of products. While written sources are consulted, the analyst needs to be ahead of what merely appears in print so that his research focuses on personal sources and statistics at the company level. Information for this analyst is confined to what is happening in the present, or likely to happen in the near future.

**Searching on the Web**

Choo, Detlor, and Turnbull (1999) develop a model of information seeking on the Web based on the external information seeking practices of knowledge workers. This model relates environmental scanning modes of **undirected** and **directed viewing**, and **informal** and **formal search** with the information seeking behaviors of **starting**, **chaining**, **browsing**, **differentiation**, **monitoring**, and **extracting** (Ellis & Haugin, 1997) to array specific discovery behaviors (e.g., following links on initial pages as a form of undirected viewing and chaining). In addition to being helpful in analyzing how people discover information, the arraying of
specific discovery behaviors by sequence of actions focuses attention on modes and functions that might be worthy of design attention.

Fidel, et al. (1999) studied the Web searching practices of eight high school students in connection with a horticulture class. The students' searches were focused on finding information needed to complete an assignment. The students' searching moves were rapid as they skimmed pages relying on cues provided by pictures and other graphical elements to decide if a page would be likely to have the needed information. Also, students would go back a few pages or go to the start screen when they became lost. They were also quick to ask for help from all present. While the authors' original purpose in the study was to improve Web page design, they suggest other needs, including training not only in Web searching but also in the subject area. Both background knowledge and searching skills come together in the search process, especially where there was little stopping for reflection or thinking. Web search pages might better support such learners by suggesting links to authoritative sites (e.g., encyclopedias) and offering structures to guide the search. Also, the students' dependence on graphics as a tool for judging the applicability of a page suggests the importance of graphics as a design element.

**Crossing Boundaries**

The crossing of boundaries during collaboration provides a natural social interface for people to discover information. Yet, there are many collaborative situations in which such communication is inhibited rather than facilitated. The studies referred to here try to understand the nature of information transfer and discovery in collaboration.

Sonnenwald and Pierce (2000) studied collaboration during a simulated battlefield exercise. Three characteristics of the ways the participants discovered information emerged as influencing the success of the exercise: (1) *interwoven situational awareness*—shared (and often incomplete) understandings of the situation as it is unfolding, (2) dense social networks—recurrent interactions among participants, and (3) *contested collaboration*—the arguments that either stand in the way of progress or help people to understand the nature of the rapidly changing situation. These findings highlight some of the mechanisms within a collaborative situation that influence outcomes. Sonnenwald (1996)
highlights roles at the organizational, task, discipline, and personal levels that support the discovery of information by negotiating, filtering, and transferring knowledge across boundaries.

Overall, contextual comparisons are useful in pointing to those patterns of behavior that appear again and again across contexts, as well as those that vary with the structures of particular domains or nature of the tasks performed. For instance, depth of knowledge is critical across contexts. In contrast, the nature and purposes of tasks tend to be variable across contexts, as are the characteristics of the structures that facilitate or inhibit task performance.

**Methods for Studying Information Discovery**

Much of the cited research has the common methodological element of intensive study of people interacting with other people and/or technology in the context of a task or problem. Many of the studies employ naturalistic approaches that incorporate sustained data collection and analysis over time (e.g., ethnography). Others create realistic environments for research purposes (e.g., experiments). Many collect both quantitative and qualitative data. While these research reports have far-ranging purposes, they all illuminate how people discover information in the context of life and work and, in the process, provide insights that could be employed in the design of systems that support interaction with the information resources necessary for discovery.

This section pursues two questions: (1) from a design point of view, what do our data collection methods need to capture to support design, and (2) what methodological strategies are emerging to support the discovery of information about how people discover information?

**Orienting Data Collection and Analysis to Support Design**

Bates's (1986) design model provides some clues for data collection. While her model was initially proposed in connection with subject access in online catalogs, it has also been used as a foundation for research design. The model highlights variety (of terminology employed in queries
and object descriptions), uncertainty (of what the next query will be or how objects are described), and complexity (of docking and use of an information system) as key matters to be considered in design. The model, thus, provides a starting point for orienting data collection and analysis. If we can map the nature of variety, uncertainty, and complexity in a particular situation, we will have gone a long way toward understanding what information is to people with respect to particular tasks. Kuhlthau (1993b) elaborates on uncertainty and applies her conceptualizations to a security analyst (Kuhlthau, 1997).

When taken literally, use of variety, uncertainty, and complexity in the orientation of data collection has a certain biasing effect. Yoon and Nilan (1999) point out that people understand what they do not know through the lens of what they do know. An implication is that, in mapping variety, uncertainty, and complexity, we need to understand what is predictable, certain, and easy to use. Yoon and Nilan also add topic and comment as orthogonal considerations as people develop meaning through their discovery of information. Topic is seen as a first step in focusing; comment establishes context for topic. Comment as a process within the discovery of information may, in turn, lead to refinement in topic as a result of the learning about what information is that results from commenting.

Another design issue is the management of change that results from such processes as commenting, arguing, learning, and sense making during the course of discovering information. Time and timing are key related issues. Kuhlthau (1993a) highlights the dynamic aspects of discovery through her information search process model. Solomon (1997a) considers cycling and changes in organizational life and the influence of timing of actions on outcomes as limiting what is available in formal information systems, as well as what becomes information during performance of a task.

This suggests that methods to illuminate the discovery of information need to address variety/consistency, uncertainty/certainty, complexity/ease of use, topic/comment, and time and timing, along with specific contextual factors related to the purpose of research or design support.
Discovering How People Discover Information

The second question is addressed by highlighting some of the many approaches that aid in the investigation of learning, understanding, and meaning development as a basis for discovering what constitutes information in context.

Activity Theory

Activity theory is theoretical in that it aims at an account of knowing and doing in some situation. It is methodological in that it provides a framework for viewing a situation. It is analytical in that it focuses attention on certain elements, such as agent, object, and community, and their interactions, including instruments, rules, and roles (Blackler, 1995). As it emphasizes interactions to fit some task such as judging (Engeström, 1996), activity theory focuses attention on the struggles that agents, objects, and community go through to specify or shape what information is. Data collection includes specifying the characteristics of elements and mapping their interactions (or lack thereof). Analysis focuses on what makes sense (i.e., smooth flow of events) or not (i.e., anomalies, conflicts), and how and why repairs and revisions are made.

Contextual Inquiry for Contextual Design

Contextual inquiry (Beyer & Holtblatt, 1998) consists of a range of structured data collection (e.g., interviews, observations) and analysis tools (e.g., activity sequence, information flow) that provide some of the specific insights of, for instance, ethnographic methods, in a much abbreviated time frame. The models developed during analysis are, in turn, used to generate design requirements. Coble, Maffitt, Orland, and Kahn (1996) summarize the use of contextual inquiry to discover what information is for physicians and how the primary care services provided by physicians might be supported by a clinical workstation.

Discourse Analysis

Intensive analysis of texts—oral or written—is another approach that is increasingly coming into use as a mechanism for understanding the nature and form of both the interactions (human-human, human-computer,
human-text) and the products of those interactions that are evidence of information discovery. Solomon (1997a) provides a series of analytical conceptualizations that are drawn from the literature on language for special purposes, conversational analysis, and discourse processes. These analytical conceptualizations are used to illustrate what constitutes information in information seeking conversations.

Tuominen and Savolainen (1997) outline an approach for studying information use as a process of social construction that focuses on discursive evidence. Use of this approach shows how information is constructed in the arguing, commenting, and discussing of work and life. Tuominen (1997) illustrates the approach in a study of the interactions of people seeking information in libraries. Findings include the specification of what seem to be problem-centered strategies (e.g., inside-the-head, misunderstanding) that attack the biases of either user-centered or system-centered approaches. Talja, Heinisuo, Pispa, Luukkainen, and Järvelin (1997) employ discourse analysis to address the issues of variety and uncertainty in the design of a Web-based regional information service in Finland. That is, the authors use discourse analysis to identify the variety of starting points of potential users of the system, the kinds of terms used to describe what people are looking for (e.g., family restaurants, exotic restaurants), and the classifications that these potential users employ as a basis for design.

Phenomenographical Analysis

Phenomenographical analysis (Marton & Booth, 1997) aids in the exploration of how people experience, understand, or think about the phenomena that they encounter in the world. This exploration has two dimensions: what people experience, and how they experience it. Data collection is typical of other ethnography-related methods, with analysis to categorize thoughts, feelings, and actions in relation to the focus of the study. Limberg (1998, 1999) recently employed this approach in information studies, comparing variation in learning outcomes with variation in students' information seeking and use. She found that learning was shaped by how these students saw what they were doing: finding facts to support a position, balancing 'facts' in order to choose the 'right' side of the argument, or scrutinizing and analyzing in order to learn about the topic to develop a considered and defensible position.
This chapter has so far emphasized a view of information as being discovered, shaped, and created by people in the context of their lives and work. This might be seen as the front-end work of knowing enough about variety/consistency, uncertainty/certainty, complexity/ease of use, topic/comment, time and timing, and contextual factors to advance the design of interactions that support discovery. This section has suggested the sorts of information that might inform design and considered several recent approaches specifically related to information studies concerns. The next section highlights some of the key matters needing attention in the design of interaction for information discovery.

**Designing to Support Discovering Information in Context**

The rise of the World Wide Web has been a force for the development of tools to support the discovery of information. The enormous scope and scale of the Web creates a tremendous need, from the system end, for interaction approaches that help people narrow down, focus in, and specify what is of likely interest. The term *information discovery* (along with such related terms as *resource discovery*) has been employed on the system end to describe the architectures (e.g., representations, classifications, and displays) needed to help people explore the Web. Lynch (1995) provides an overview of resource discovery issues in a networked world. Desai (1997) explores the indexing and metadata needs of discovery in digital libraries. Earlier Carroll and Rosson (1987) provided a thoughtful analysis of production and assimilation biases of people as active users of information systems, including a variety of strategies for designing interactions that support task performance and discovery or learning.

Designing interactions that provide for discovery requires support for (1) social awareness and navigation, as people become increasingly connected in dispersed networks, and (2) learning, as people shape or redefine what information is during the process of interaction. Brown and Duguid (2000, pp. 173–174) capture the spirit of this challenge in various ways—particularly in Duguid's anecdote of a medical historian's sniffing of letters for the aroma of vinegar in his quest to document outbreaks of cholera.
There is also the fundamental problem of scale. For instance, Arms (2000) considers information discovery in the context of automated digital libraries. His approach emphasizes existing tools (e.g., the search engine Google), which employ the inherent characteristics (e.g., number of incoming links) of the resource (e.g., Web pages) to provide the searcher with an ordered list of information to be discovered. While there are benefits to this approach, it does have limitations: A recent query produced about 834,000 items retrieved in response to a search on "information discovery." Work on clustering (Zamir & Etzioni, 1998) or classifying (Chen & Dumais, 2000) Web search results would contribute to information discovery by providing an overview of possibilities.

Can we translate an understanding of how people discover information as a process into information or resource discovery tools that support this process? The remainder of this section points to some advances in the areas of support of creativity, social navigation, and visualization.

**Creativity Support**

Creativity and support for innovation seem to be major impetus for, and potential outcomes of, discovery as a natural process. Ford (1999) speculates on how traditional information retrieval systems might be transformed to support divergent as well as convergent thinking. His speculation leads to the suggestion of combining higher-order knowledge representations and parallel and fuzzy pattern matching techniques as a design direction. As serendipity is a major force for creativity, it seems that offering the opportunity to see relationships through both high levels of abstraction, using case-based reasoning or argumentation, and relatively broad subject domain categorizations, using text summarization or term clustering, may be fruitful.

Shneiderman (2000) focuses on the capacities that interfaces need to support creativity. He lists four areas for interface support: helping people (1) collect existing work to support basic learning; (2) interact with others as they relate what they know with others (i.e., peers, mentors); (3) explore as they think or create; and (4) share the results of their efforts (donate). He sketches the sorts of tools that might be incorporated in interfaces to promote creativity by supporting (1) searching, filtering, and browsing; (2) collaboration; (3) visualization of data and processes; (4) seeing relationships and thinking about them; (5) what-if analysis of
possible solutions; (6) composing; (7) reviewing previous activity; and (8) dissemination.

Themes for design attention that emerge from Ford's and Shneiderman's suggestions for design to support creativity as well as from research relating to discovery as a process include social navigation and flexible visualization.

**Social Navigation**

Social navigation is key to the phenomenon of interaction. That is, social navigation recognizes that the structure of the experience contributes to what people get out of an interaction. It also recognizes that information about other people and their activities may contribute to the interaction. The literature on computer-supported cooperative work is important to the consideration of social navigation. Overviews of technological developments can be found in the annual proceedings of the ACM conferences on Computer Supported Cooperative Work and Supporting Group Work. Here, attention is given to some recent work that informs design for social navigation.

Munro, Hook, and Benyon (1999) present a collection of fourteen articles on social navigation from a 1998 workshop in Sweden. This is a rich source of conceptualizations of how to support social navigation, of research on how social navigation influences the performance of work, and of examples of technologies that support social navigation.

Davenport (1999) provides a brief summary and analysis of efforts by a group of researchers in Edinburgh, Scotland, to conceptualize a social browser (Partner Lenses—PaLs) to mobilize the expertise within a community as well as provide a vehicle for observing and promoting social interactions. One of the outcomes of the work so far has been the realization of a need for multiple layers, including an information retrieval layer to support the finding of partners, a social layer to capture attributes of interactions (e.g., courtesy, responsiveness), and a trust layer to develop a sense of trust among partners. Trust is increasingly being recognized as an important issue in the social aspects of discovery (Iivonen & Huotari, 2000).

Erickson and Kellogg's (2000) work on social translucence adds three dimensions—visibility, awareness, and accountability—to Davenport's social browser requirements. While there is seemingly some overlap
between the two sets, Erickson and Kellogg seem to get beyond technologies that impose barriers and move toward those that provide windows for sharing, creating, and focusing for the purposes at hand.

**Flexible Visualization**

The wide range of current design work on flexible visualization can be glimpsed through conference proceedings of the ACM's Special Interest Groups on Computer Human Interaction and Information Retrieval. The focus here is on recent efforts to facilitate discovery by putting control in the hands of people. Greene, Marchionini, Plaisant, and Shneiderman (2000) present a framework for the design of displays to support information discovery. Previews and overviews of individual objects (or object collections, e.g., the human body, images, or a medical history) serve as the entry point. These allow the user of the system to develop an overview of the system contents as well as to see how objects are related, displayed, and acted on. This framework has considerable potential in situations in which there is structured data (or unstructured data that is amenable to automatic structuring) to provide flexible interaction enabling people to discover information and perhaps gain knowledge.

Hars (2000) provides a model of an adaptive document that can actually be experienced in the HTML version of the article. The article does a number of things: (1) it introduces a range of knowledge infrastructures available on the Web (e.g., preprint archive, digital library, collaboratory), (2) it ties those infrastructures to modes of scientific interaction (e.g., reading, editing, commenting), and (3) it considers how texts might be structured to allow dynamic adaptation for individual readers at different stages of their interaction with the text. Such an adaptive mechanism would be a useful option as texts become available in full-text form, but where particular pieces of the text (e.g., abstract, methods, references) are needed at different stages of a discovery process.

Nardi and O'Day's (1996) work provides some interesting insights into the possibility of empowering people's discovery of information through the use of intelligent agents. They began their thinking about intelligent agents by studying human agents (e.g., reference librarians) and, in the process, came up with the ideas of (1) focusing software agents on particular purposes, (2) enabling evolution-friendly agents, (3) providing for cooperation between agents, users, and experts, and (4)
using agents to add what they find to information bases. All of these ideas contribute to the design of diverse information ecologies that support a wide range of collaborative interactions. Such information ecologies would meet the requirements for social navigation and flexible visualization.

Davenport and Cronin (1998) extend Nardi and O'Day’s work by considering a variety of just-for-you service views that contribute to the training and practice of information professionals. These include user needs, relevance studies, indexing and classification, professionalism, and the digital environment. While their analysis focuses on developing connections among these varied views in the curriculum, their work also suggests that, regardless of whether clients, customers, or patrons know the details of just how the pieces fit together to create the whole, researchers, faculty, and professionals need to provide the flexibility to help people shape their own discovery processes.

**Conclusion**

This review has considered information as something that is constructed by people in their interactions with other people, technology, and structures as they move through life and work. It also draws attention to the power, in turn, of people, systems, and institutions to shape what information is. The idea is that by understanding something of the consistency/variety, certainty/uncertainty, and ease of use/complexity of information construction, we will be able to develop theory and socio-technical designs that provide users with increased ability to explore, shape, reflect, and learn.

A general design implication is the need to provide flexibility in selecting, ordering, and viewing. Key to this is the design of representations that allows such flexibility, which can be informed by gaining a more holistic understanding of how information is constituted in task situations. This is not to say that gaining insight into information discovery is easy or using that insight to influence design is trivial. Nevertheless, the hope is that such an evidence-based approach will result in designs that fit people’s tasks and problems and help them shape their own discovery processes rather than stand in the way.
Acknowledgments

Many thanks to the anonymous reviewers whose comments have substantially improved this review.

Bibliography


Proceedings of the second international conference on information needs, seeking and use in different contexts (pp. 150–175). London: Taylor Graham.


needs, seeking and use in different contexts (pp. 412–421). London: Taylor Graham.


