

## Beyond Search



# POWERS OF 10: MODELING COMPLEX INFORMATION- SEEKING SYSTEMS AT MULTIPLE SCALES

Peter Pirolli, *Palo Alto Research Center*

**New models of information-seeking support systems offer two advantages: They move us from prescientific conceptual frameworks about information seeking to more rigorous scientific theories and predictive models while, at the same time, expanding the kinds of things we study and develop.**

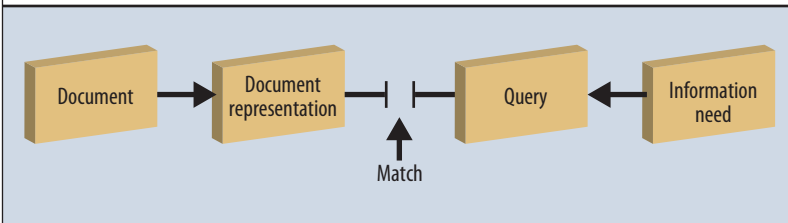
**T**hese are exciting times for scientists, designers, and engineers in the field of information-seeking support systems (ISSSs). New technology continues to fuel a staggering growth of available information, which in turn affects our lives by providing resources for adapting to an increasingly complex world, as well as new ways of being entertained.

We are witnessing an efflorescence of new ways to interact with—and produce—rich content. National efforts, such as the US Cyberinfrastructure initiative ([www.nsf.gov/news/special\\_reports/cyber/index.jsp](http://www.nsf.gov/news/special_reports/cyber/index.jsp)), aim to produce even more fertile platforms for information. This evolving domain offers science a great opportunity because there

is so much new territory to explore and explain, so many new ideas about how to do so, and so much potential for having an impact on innovative engineering and imaginative design.

Two exciting opportunities await science and engineering in this field. The first moves us from prescientific conceptual frameworks about information seeking to more rigorous scientific theories and predictive models. Progress in cognitive science and human-computer interaction is moving toward a coherent set of theories and models to address the complexity of modern-day information seeking at a range of scales, from the long-term social to individual moment-by-moment interaction.

The second opportunity expands the kinds of things we study and develop. Information seeking in the current world involves much more than isolated solitary users working with a single tool to retrieve some document or fact. The information environment has become a place to explore and learn over longer periods. It has become much more social. People use many tools and systems fluidly for many purposes. Information is no longer just passive text, but includes rich media that often seek users as much as users seek media.



**Figure 1.** Classic information retrieval model. This conceptual framework for information retrieval prompts a user with an information need to reformulate it as a system query, which in turn retrieves a document whose system representation best matches the query.

## CLASSIC CONCEPTUAL FRAMEWORKS

Researchers have a long tradition of developing ISSS conceptual frameworks (often called conceptual models) that provide a general pool of concepts, processes, assumptions, methods, and heuristics that orient researchers to a particular way of viewing and describing the world. For example, Figure 1 depicts a classic conceptual framework for information retrieval.<sup>1</sup> A user has an information need that must be reformulated as a query to a system, which in turn retrieves a document whose system representation best matches the query. This classic information-retrieval model motivated a wave of research that crested with the search engine technology now pervasive on the Web, which addresses a large class of everyday information needs for hundreds of millions of users.

However, it also became clear throughout the 1980s and 1990s that information-seeking systems and behavior must include more than just query-based search. For example, the model depicted in Figure 1 did not capture browsing in hypermedia systems, such as the Web. The classic model did not address more complex sense-making activities in which large amounts of information about a situation or topic are collected and deliberated on to form an understanding that becomes the basis for problem solving and action.

Researchers studying real user behavior and systems began to expand on the classic conceptual framework in Figure 1. For example, analysis of real-world information seeking<sup>2</sup> led to the identification of an expanded set of general processes and factors involved in information-seeking behavior. Conceptually, an information seeker is seen as situated in an environmental context with an embedding task or goal that triggers a need for more knowledge. Information systems provide interactive access to information (that must be shaped into knowledge) to meet those needs and achieve those tasks.

The system provides information that enhances the ability of searchers to find the right information for their purposes. This expanded conceptual framework leads to the identification of many factors that can shape information seeking, including those associated with the following:

- *information seekers*, including prior knowledge, skills, and other individual differences;
- *tasks*, which can vary greatly in terms of complexity, structure, and constraints; drive and shape information-seeking and problem-solving behaviors; and provide criteria by which information is evaluated;
- *domains*, which are the general bodies of knowledge that can vary in complexity, kinds of content, rate of growth, change, and organization, among other things;
- *contexts*, which can include physical, social, and temporal factors and constraints; and
- *search systems*, which can vary enormously in how domain information is represented and how they present and interact with users.

Information seeking as berrypicking<sup>3</sup> also became an influential metaphor and conceptual framework. Users often start with some vague information need and iteratively seek and select bits of information that cause the data needs and behavior to evolve over time; there is no straight line of behavior to a single best query and retrieval set.

In real libraries, users employ a variety of information navigation strategies, such as footnote chasing, citation chaining, reviewing a journal series, browsing entire areas at different generality levels, and browsing and summarizing an author's works. Today, all these techniques seem mundane, a testament to how specific system features and user interface designs provide better support for real-world information-seeking strategies.

## THEORIES AND MODELS

Conceptual frameworks have proven useful for providing conceptual tools to analyze and describe observed behavior, which in turn can suggest new functions for information-seeking systems. They also provide a common ground for summarizing findings and accumulating results, formulating hypotheses and analyses, and contrasting and debating different ways of characterizing ISSS systems.

However, conceptual models are not scientific theories or models that provide a basis for making predictions about the design effects and engineering decisions for information-seeking support systems. Researchers construct scientific theories within frameworks by providing additional assumptions that let them make predictions they can test. Typically, researchers achieve this by specifying a model that makes precise predictions they can fit to observation and measurement for a specific situation or class of situations.

Two exciting challenges involve developing truly predictive and explanatory scientific theories and models, then applying them to address the full complexity of information-seeking behavior at multiple time scales. Among other applications, this permits the prediction of how minute changes at the microscale of individual-user interaction can percolate upward to emergent macroscale phenomena such as the evolution of wikis and tagging systems.

Predictive models can provide a basis for understanding and control over ISSSs and the behavior they support. In practical terms, it means that designers and engineers can explore and explain the effects of different ISSS design decisions before undertaking the heavy investment of resources for implementation and testing. With additional research, this design space exploration will become more efficient and innovative.

## HUMAN-INFORMATION INTERACTION

To get a sense of what might be possible, we can consider the hierarchical organization of human behavior and the phenomena that emerge at different analysis time scales, as Table 1 shows.

Allen Newell and Stuart Card<sup>4</sup> argued that human behavior—including information seeking—can be viewed as the result of a hierarchically organized set of systems rooted in physics and biology at one end of the spectrum, and large-scale social and cultural phenomena at the other. Table 1 shows this framework, which has proven useful in cognitive science<sup>5</sup> and human-computer interaction research.

The basic time scale of operation for each system level in this hierarchy increases by a factor of approximately 10 when moving up the hierarchy in Table 1. Some behaviors, such as seeking information on the Web, can be modeled at multiple time scales. However, the most exciting work on developing models that map out this territory for different kinds of ISSSs and phenomena has yet to begin. Within the timeframes in Table 1, there are layered bands differentiated by the kinds of factors that shape behavior:

- *Biological band phenomena*—spanning approximately milliseconds to tens of milliseconds—are mainly determined by biochemical, biophysical, and especially neural processes, such as the time it takes for a neuron to fire.
- *Psychological band activity*—spanning approximately hundreds of milliseconds to tens of seconds—is where the elementary psychological machinery for perception, cognition, and action play a major part in shaping behavior. This has traditionally been the domain of cognitive psychology.
- *Rational band phenomena*—spanning approximately tens of seconds to minutes to days—are where the

**Table 1. Human behavioral time scales: Different bands show different phenomenological worlds.**

Scale (seconds)	Time Unit	Band
$10^7$	Months	Social
$10^6$	Weeks	
$10^5$	Days	
$10^4$	Hours	Rational
$10^3$	10 minutes	
$10^2$	Minutes	
$10^1$	10 seconds	Psychological
$10^0$	1 second	
$10^{-1}$	100 ms	
$10^{-2}$	1 ms	Biological

task structure and other environmental and contextual constraints come to dominate the linking of actions to goals. Longer-term goals are typically realized by task structures hierarchically composed of many shorter-term goals. Individuals tend to approximate a principle of rationality: Based on their perceptions of the ongoing situation and their current knowledge, they prefer actions that will move them toward their goals.

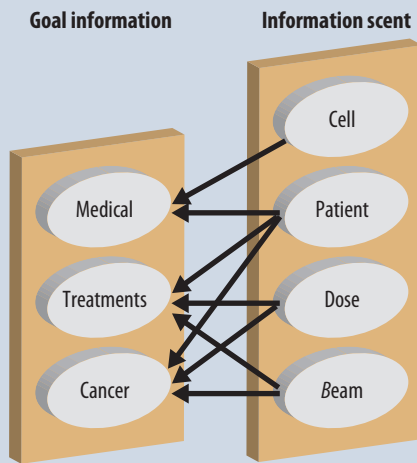
- *The social band*—spanning approximately days to weeks to months and beyond—addresses social systems that involve many individuals in communication and interaction. At this level, factors such as communication, coordination, cooperation, trust, reputation, and others play a role, as do the structure and dynamics of social networks.

## WEB NAVIGATION CHOICES

Consider this simple example of information-seeking on the Web, where accurate models<sup>6</sup> at the rational, psychological, and social bands have been developed for simple tasks such as finding products, event dates, and particular documents.

These have been incorporated into Bloodhound, an automated Web usability evaluator.<sup>7</sup> Analysis at the rational band of Web information seeking involves rational analysis,<sup>8</sup> which focuses on the task environment's performance, the information environment that structures access to valuable knowledge, and the adaptive fit of the human-information interaction system to these environments' demands. This method progresses through the following steps:

1. Specify the user's information-seeking goals.
2. Develop a formal model of the task and information environments—the Web's information architecture, for example.



**Figure 2.** A cognitive structure in which cognitive chunks representing an information goal are associated with chunks representing information scent cues from a Web link.

3. Make minimal assumptions about cognitive costs.
4. Derive the user's rational behavior, considering steps 1 through 3.
5. Test the rational predictions against data.
6. Iterate.

In the case of the Web tasks, one focus of rational analysis concerns how users choose the most cost-effective and useful browsing actions to take, based on the relation of a user's information need to the perceived cues associated with Web links. Such user interface cues have been called information scent because users follow these cues on trails through an information architecture. A random utility model can predict link choices to an approximation.<sup>9</sup> This model asserts the probability that a user will choose a particular link  $L$ , having a perceived utility  $V_L$ , from a set of links  $C$  on a webpage, given a user information goal  $G$ :

$$Pr(L | G, C) = \frac{e^{\mu V_L}}{\sum_{k \in C} e^{\mu V_k}} \quad (1)$$

where  $\mu$  is a scaling parameter.

A spreading activation model of information scent can model the details of how users judge the utility of Web links such as  $V_L$  at the finer-grained psychological band for text links. This model assumes that the user's cognitive system represents information scent cues and information goals in cognitive structures called *chunks*, as Figure 2 shows.

These chunks can be thought of as representing mental concepts in human memory. Figure 2 assumes that a user seeks information about "medical treatments for cancer"

and encounters a Web link labeled with text that includes "cell," "patient," "dose," and "beam." We assume that when users focus on Web links, their attention to information scent cues activates corresponding cognitive chunks into conscious awareness. Activation spreads from those attended chunks along associations to related chunks. The amount of activation accumulating on the representation of a user's information goal provides an indicator of the utility—for example,  $V_L$ —of that link. When incorporated into a computational simulation of users, the rational and psychological band models can predict up to 90 percent of the variance in users' link choice behavior.<sup>6</sup>

These Web link choice models can be mapped to a social band model that simulates the flow of users through a website. This provides the core algorithm used by the Bloodhound Web usability systems.<sup>7</sup> A flow network can model the aggregate flow of users grouped for a particular task, where each node in the network represents a webpage, and each edge in the network represents the flow of users from one page to another. A simple version employs discrete steps corresponding to users clicking to move from one page to another. The number of users arriving at webpage node  $i$  at step  $t + 1$  from the  $k$  nodes that can reach it are modeled as follows:

$$N_{i,t+1} = f_t \sum_k S_{i,k} N_{k,t} \quad (2)$$

where  $f_t$  is the proportion of users who continue to surf after  $t$  time steps, and  $S_{ji}$  is the proportion of users who decide to move from page  $j$  to page  $i$ , which can be determined by Equation 1. Tests of this algorithm, when incorporated into the Bloodhound usability evaluator, showed that the predicted pattern of visits demonstrated moderate to strong correlations with observed patterns in 29 out of 32 tasks conducted across four different websites.

### THREE THESES

This simple illustration about modeling the Web also demonstrates three theses<sup>5</sup> that promise a future science of ISSSs.

#### Decomposition

This thesis states that complex IS behavior occurring at large time scales—ranging from minutes through hours and days to months—can be decomposed into smaller elements and events. For example, the complex behavior of an individual interacting with a Web browser can be decomposed into individual elements of perception, judgment, decision, and action selection. Decomposition of the Web information-seeking task—of which link choice provides just one subtask—is required to develop rich user simulations. Specific features of ISSSs can improve or degrade those elements in a way that affects the full shape of IS behavior in complex situations.

## Relevance

This thesis holds that the microstructure of behavior is relevant to macrostructure phenomena. For example, small perturbations in the quality of information scent can cause qualitative shifts in the search cost of Web navigation. There is also evidence<sup>10</sup> that changes in the time cost of fine-grained user interaction in information rating and social bookmarking systems have reliable effects at the social band on contribution rates in user populations. Tweaking a user interface's fine-grained structure can have effects on the phenomena that emerge at the level of large social groups.

## Modeling

This thesis claims that predictive models can be developed to specify precisely how changes at the microstructure of individuals interacting with and through ISSSs can percolate upward to affect longer-term complex information seeking. In some cases, a single unified model can go from the lower bands to higher, as in the SNIF-ACT model.<sup>6</sup> More likely, however, there will be layers of models at different bands that capture essential features of models at the lower bands, just as statistical mechanics models particles at a level that only captures crucial features of individual particles' behavior. The graph flow models in Bloodhound, for example, only capture the average or asymptotic behavior of many individual users interacting with the Web.

## ENRICHING THE IS TASKS CLASS

As the Internet and communication technologies become ever more pervasive, we see an astounding number of new ISSSs and behaviors. As a result, we must continually expand the kinds of phenomena that conceptual frameworks, theories, and models must address. A few interrelated conceptual frameworks can illustrate this expanding territory.


## Sensemaking

Many information search tasks form part of a broader class of tasks called *sensemaking*. Such tasks involve finding and collecting information from large information collections, organizing and understanding that information, and producing some product, such as a briefing or actionable decision. Examples of such tasks include understanding a health problem to make a medical decision, forecasting the weather, or deciding which laptop to buy. In general, these tasks include subtasks that involve information search, but they also involve structuring content into a form that can be used effectively and efficiently in some task.

As an extreme example, intelligence analysts perform this sort of sensemaking as a profession, working to gather and sift through vast amounts of incoming information to write briefings that shape decisions af-

fecting national security. Cognitive task analyses of intelligence analysis<sup>9</sup> suggest that the overall process can be organized into two major activity loops: an information foraging loop that involves processes aimed at seeking information, searching and filtering it, and reading and extracting information, and a sensemaking loop that involves iterative development of a mental model—a conceptualization—that best fits the evidence.

Information processing is both driven by bottom-up (from data to theory) and top-down (from theory to data) processes. The foraging loop essentially trades off among three process types: information exploration, enrichment, and exploitation. Typically, analysts cannot explore this entire space and must forego coverage to actually enrich and exploit the information.



**Exploratory search involves three major kinds of activities: lookup, learn, and investigate.**

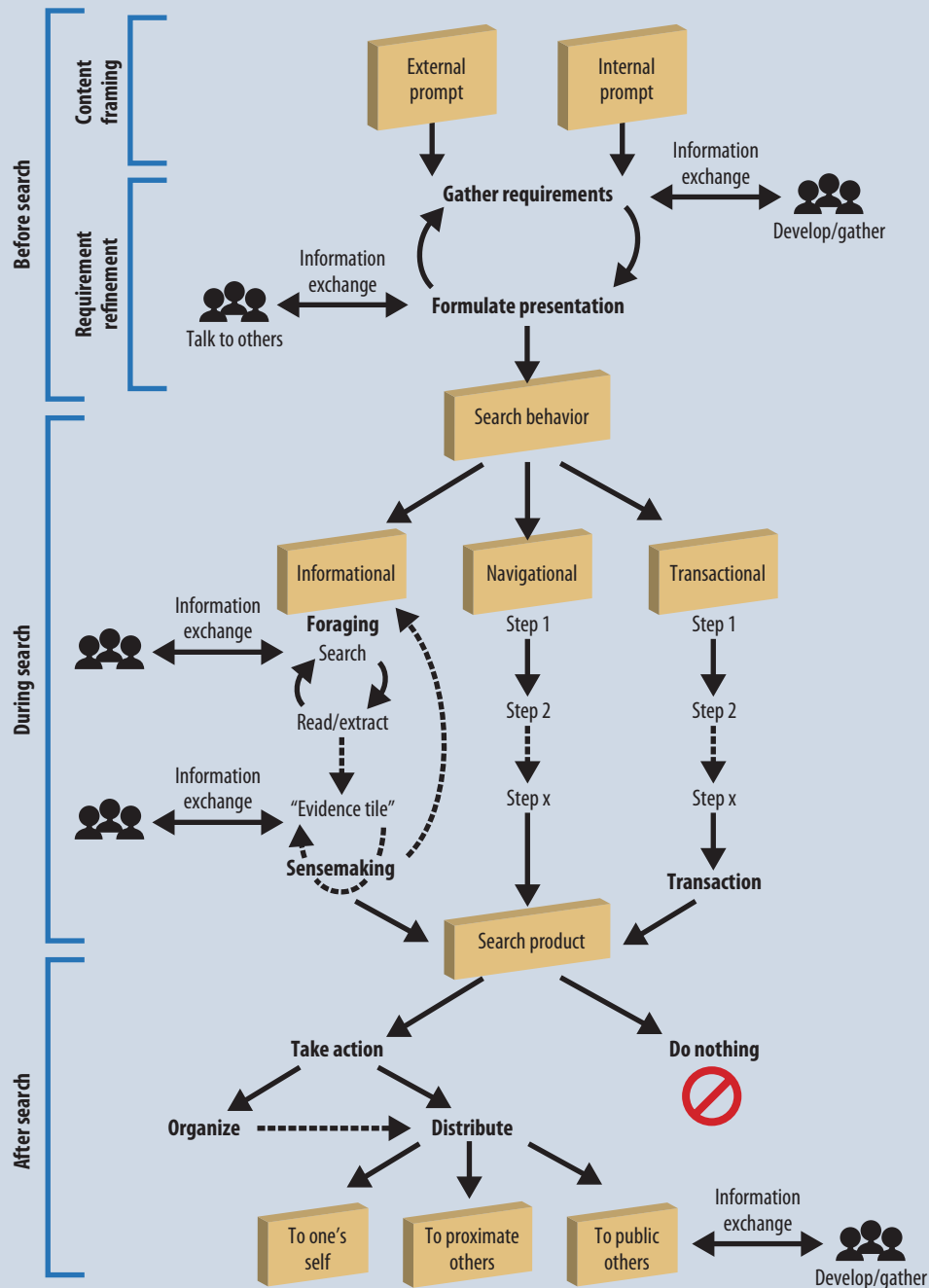
The sensemaking loop involves substantial problem structuring (the generation, exploration, and management of hypotheses), evidentiary reasoning (marshaling evidence to support or disconfirm hypotheses), and decision making (choosing a prediction or course of action from the set of alternatives). Many well-known cognitive limitations and biases affect these processes.

## Exploratory search

This undertaking provides another rich domain,<sup>11</sup> one that includes activities involving information lookup, learning, and investigation, which can overlap in time. For example, looking for health-related information is one of the Web's most prevalent information-seeking activities. Typically, this involves a prolonged engagement in which individuals iteratively look up and learn new concepts and facts. These activities in turn dynamically change as searchers refine their information-seeking goals and ask better questions. The process can be viewed as a subcomponent of sensemaking.

Exploratory search involves three major kinds of activities: lookup, learn, and investigate. Whereas lookup activities have been the traditional focus of ISSSs, as Figure 1 shows, exploratory search emphasizes learning and investigation activities.

Searching to learn includes activities involved in making decisions, such as purchases, up through professional and lifelong learning. It also includes social searches to find communities of interest—via social network systems, for example. Investigative activities include analysis, synthesis, and evaluation.



**Figure 3.** Conceptual model of social search. Social exchanges of information can occur before, during, and after an actual information search.

### Social search

People frequently turn to their social networks to find information, thus social search has become a recent focus of study.<sup>12</sup> Figure 3 shows some of the complexity that arises in social search analysis. Social information exchanges can occur before, during, and after an actual search for information. Social transactions can influence the gathering of requirements and formulation of a prob-

lem representation before a search, the gathering and extraction of information during search, and the distribution of results after search.

Even more generally, new forms of sociotechnical information systems—such as social bookmarking or rating sites—let users participate with low effort and contribute their unique nuggets of knowledge—such as data about navigation, organization, and recommendations—in a



## PRECISION AND RECALL

Daniel Tunkelang, *Endeca*

Information retrieval (IR) research today emphasizes precision at the expense of recall. Precision is the number of relevant documents a search retrieves divided by the total number of documents retrieved, while recall is the number of relevant documents retrieved divided by the total number of existing relevant documents that should have been retrieved.

These measures were originally intended for set retrieval, but most current research assumes a ranked retrieval model, in which the search returns results in order of their estimated likelihood of relevance to a search query. Popular measures like mean average precision (MAP) and normalized discounted cumulative gain (NDCG)<sup>1</sup> mostly reflect precision for the highest-ranked results.

For the most difficult and valuable information-seeking problems, however, recall is at least as important as precision. In particular, for tasks that involve exploration or progressive elaboration of the user's needs, a user's progress depends on understanding the breadth and organization of available content related to those needs. Techniques designed for interactive retrieval, particularly those that support iterative query refinement, rely on communicating to the user the properties of large sets of documents and thus benefit from a retrieval approach with a high degree of recall.<sup>2</sup>

The extreme case for the importance of recall is the problem of information availability, where the seeker faces uncertainty as to whether the information of interest is available at all. Instances of this problem include some of the highest-value information tasks, such as those facing national security and legal/patent professionals, who might spend hours or days searching to determine whether the desired information exists.

The IR community would do well to develop benchmarks for systems that consider recall at least as important as precision. Perhaps researchers should revive the set retrieval models and measures such as the F1 score, which is the harmonic mean of precision and recall.

Meanwhile, information scientists could use information availability problems as realistic tests for user studies of exploratory search systems, or interactive retrieval approaches in general. The effectiveness of such systems would be measured in terms of the correctness of the outcome (does the user correctly conclude whether the information of interest is available?); user confidence in the outcome, which admittedly may be hard to quantify; and efficiency—the user's time or labor expenditure.

Precision will always be an important performance measure, particularly for tasks like known-item search and navigational search. For more challenging information-seeking tasks, however, recall is as or more important, and it is critical that the evaluation of information-seeking support systems take recall into account.

### References

1. K. Järvelin and J. Kekäläinen, "Cumulated Gain-Based Evaluation of IR Techniques," *ACM Trans. Information Systems*, Oct. 2002, pp. 422-446.
2. R. Rao et al., "Rich Interaction in the Digital Library," *Comm. ACM*, Apr. 1995, pp. 29-39.


*Daniel Tunkelang is chief scientist of Endeca, a provider of enterprise search, information access, and guided navigation solutions based in Cambridge, Mass. Contact him at dt@endeca.com.*

highly independent and modular way. These contributions in turn improve the system's performance as a whole. The emergence of such Web 2.0 technologies shows how systems that can efficiently let users make contributions, and that have architectures that harness those contributions in a highly efficient way, can win big in the public and commercial world.

**O**ur conception of ISSSs has expanded from simple information search engines. Similarly, the field is moving from conceptual frameworks to scientific models, expanding the range of systems and phenomena studied. As the "Precision and Recall" sidebar explains, as the field moves away from a focus on simple precision and recall metrics to a more comprehensive understanding of information's utility to the range of human goals in modern-day society, the need arises to better understand user experience and evaluations of credibility, trust, and reputation.

The field is realizing that we need to understand why people search, explore, annotate, and decide to participate

and share the efforts of their knowledge work. People no longer work with a single ISSS, but even for simple tasks will effortlessly move among an array of tools. This is a complex territory to map with scientific models. It spans 10 powers of 10 of time scale, from tens of milliseconds to years, with enormous complexity at multiple bands of phenomena, from the psychological to the social.

A great opportunity beckons as these phenomena increasingly happen online in large living laboratories. This great attractor for scientific minds in diverse areas will range across fields as varied as behavioral economics, incentive mechanisms, network theory, cognitive science, and human-computer interaction. 

### References

1. B.M. Evans and E.H. Chi, "Toward a Model of Understanding Social Search," *Proc. ACM 2008 Conf. Computer-Supported Cooperative Work*, ACM Press, 2008, pp. 485-494.
2. G. Marchionini, *Information Seeking in Electronic Environments*, Cambridge Univ. Press, 1995.
3. M.J. Bates, "The Design of Browsing and Berrypicking Techniques for the Online Search Interface," *Online Rev.*, vol. 13, 1989, pp. 407-424.

4. A. Newell and S.K. Card, "The Prospects for a Psychological Science in Human-Computer Interactions," *Human-Computer Interaction*, vol. 1, no. 3, 1985, pp. 251-267.
5. J.R. Anderson, "Spanning Seven Orders of Magnitude: A Challenge for Cognitive Modeling," *Cognitive Science*, vol. 26, no. 1, 2002, pp. 85-112.
6. W. Fu and P. Pirolli, "SNIF-ACT: A Model of User Navigation on the World Wide Web," *Human-Computer Interaction*, vol. 22, no. 4, 2007, pp. 355-412.
7. E.H. Chi et al., "The Bloodhound Project: Automating Discovery of Web Usability Issues Using the InfoScent Simulator," *Proc. ACM Conf. Human Factors in Computing Systems, CHI Letters*, vol. 5, no. 1, 2003, pp. 505-512.
8. J.R. Anderson, *The Adaptive Character of Thought*, Lawrence Erlbaum Assoc., 1990.
9. P. Pirolli, *Information Foraging: A Theory of Adaptive Interaction with Information*, Oxford Univ. Press, 2007.
10. D.M. Wilkinson, "Strong Regularities in Online Peer Production," *Proc. 9th ACM Conf. Electronic Commerce*, ACM Press, 2008, pp. 302-309.
11. S.E. Robertson, "Theories and Models in Information Retrieval," *J. Documentation*, vol. 33, no. 2, 1977, pp. 126-148.
12. G. Marchionini, "Exploratory Search: From Finding to Understanding," *Comm. ACM*, vol. 49, no. 4, 2006, pp. 41-46.

*Peter Pirolli is a research fellow at Palo Alto Research Center's Augmented Social Cognition Area. His areas of research are human-information interaction, cognitive science, information foraging theory, information scent, and sensemaking. Pirolli received a PhD in cognitive psychology from Carnegie Mellon University. Contact him at pirolli@parc.com.*

## Unsure about the path forward?

IEEE Security & Privacy magazine can help with useful recommendations, authoritative analysis, and practical insights for securing your networks, applications, and systems.

Pick up on best practices and new developments in our unique mix of peer-reviewed articles and lively columns on the security and privacy issues that matter most.



Subscribe for \$32 a year!  
[www.computer.org/security](http://www.computer.org/security)

IEEE  
**SECURITY & PRIVACY**

Building dependability, reliability, and trust

