15
Conceptual Models for Search Engines

D.G. Hendry and E.N. Efthimiadis

Summary Search engines have entered popular culture. They touch people in diverse private and public settings and thus heighten the importance of such important social matters as information privacy and control, censorship, and equitable access. To fully benefit from search engines and to participate in debate about their merits, people necessarily appeal to their understandings for how they function. In this chapter we examine the conceptual understandings that people have of search engines by performing a content analysis on the sketches that 200 undergraduate and graduate students drew when asked to draw a sketch of how a search engine works. Analysis of the sketches reveals a diverse range of conceptual approaches, metaphors, representations, and misconceptions. On the whole, the conceptual models articulated by these students are simplistic. However, students with higher levels of academic achievement sketched more complete models. This research calls attention to the importance of improving students' technical knowledge of how search engines work so they can be better equipped to develop and advocate policies for how search engines should be embedded in, and restricted from, various private and public information settings.

15.1 Introduction

Search engines are remarkable for their mediating power: Every day, millions of people speak through their writing, while millions of others search for this "speech" with their queries. Popular quantitative and demographic measures (Lenhart et al. 2004; Media Metrix 2004) show that search engines are an important cultural phenomenon, matching searchers' queries with producers' content. The popular press, over the last several years, has created an impressive groundswell of public interest in search engines – how they work and the cultural phenomena surrounding them. Search – surprisingly given its dusty, technical roots – has become fashionable. In turn, search has shifted interest in such important civic issues as universal access, privacy rights, informed consent, and one's autonomy to pursue one's own interests to a new space – the Internet. Perhaps the most significant long-term implication of search engines is how they have raised these issues, which have been dormant, and how they prompt society to address them.

A. Spink and M. Zimmer (eds.), *Web Search, Springer Series in Information Science and Knowledge Management 14.*

The networked infrastructure that enables information services like Google is an *artificial world* (Simon 1996), which presents people with a menagerie of new concepts, intricately interrelated. To list just a few: Web pages, keywords, meta tags, hyperlinks, caches, Web servers, robots.txt, file permissions, search engines, rankings, URLs, spiders, users, content providers, advertisers, spam, spammers, search-engine optimizers, tags, log files, and PageRank ™. While human-made, this is not a neat world. Indeed, many important relationships between elements are hidden and the intricacy of the overall system is largely due to localized technological improvement. The protocol for Web cookies is a classic example that illustrates how a seemingly straightforward technical protocol can have significant, unanticipated consequences on public policy in such important areas as privacy and informed consent (Friedman et al. 2002). To discuss the merits of such a technical protocol on privacy and similar values, one must draw upon technical knowledge for the protocol. Nevertheless, like the natural world, we engage this artificial world without complete understanding or even being aware of its underlying complexity.

However, when we encounter a phenomenon that triggers our interest or when we encounter a barrier that prevents us from obtaining a goal, we may ask a question that can only be answered by investigating the intricacies of this artificial world. Consider, for example, this barrier: "When I type my name into Google, why does my Web page not appear within the results on the first page?" To answer this question, we might follow a process of deductive thinking and draw on established concepts and principles to propose an explanation. From this explanation, we might then pursue a course of action to overcome the barrier. Alternately, in order to address the problem, we might seek the advice of experts and consider their explanations in light of our current understanding. Finally, we might follow a more inductive process and gather data related to the phenomenon and attempt to identify a general pattern. Of course, the rigor associated with each of these modes of inquiry will vary. Often, the process will be quick and ad hoc and sometimes it will be based on incorrect or only partially correct facts. Nevertheless, like a scientist seeking to understand the natural world, a person who seeks to understand the artificial world of search engines will appeal to his or her existing technical knowledge.

The question we address in this chapter is: What is the nature of this technical knowledge held by students of information science? We assert that knowledge of basic technical concepts for search engines is an important kind of scientific literacy. This assertion follows from the position that a healthy democracy requires a scientifically literate public where people understand basic scientific constructs such as "The Earth revolves around the Sun once each year", which can be assessed by closed and open questions in telephone surveys (Miller 1998). Certainly, technical knowledge about how a search engine works is needed in order to both search effectively, as well as to teach others how to search. This technical knowledge is also necessary to participate in higher level debates, such as participating in civic dispute about search engines, as well as advocating for their proper use. At the same time, it is important to acknowledge the social constructivist position, in which people learn by creating interpretations that are based on their past experiences and their current interactions with the world. In the context of public policy disputes concerning the environmental health of a river basin, Roth and Lee (2002), for example, show how scientific literacy can be constituted in a social setting of intense dialog between people of various backgrounds. Analogously, we expect that serious public dialog about search engines, involving people of varied backgrounds, would enable people to express knowledge that is not available to them when completing a survey over the telephone. We take the view, in short, that the ability for a single person to generate explicit facts about how search engines work is the only one kind of knowledge about them. Nevertheless, in this work we focus on just this form of knowledge. As educators, our goal is to take measure of students' knowledge of search engines so that we can provide better instruction and be more effective teachers.

In the next section, we develop the argument that the public discussion of search engines centers at the fuzzy junction of culture and technology. Indeed, we show that the popular press serves an important role for educating people about how search engines work and for identifying social consequences of their operation. Then, we review the literature on mental models for search engines, showing that the literature has focused on users' understandings for particular kinds of search systems. Not addressed to date are people's understandings for search at the cultural level; yet, this is clearly needed as search engines have moved from well-bounded settings, such as a library's catalog, to an information network that pervades home, work and play. Next, we report the results of an exploratory experiment where we ask students to draw sketches of how search engines work. A content analysis of the sketches reveals a tremendous diversity of approaches for conceptualizing search engines, and yet, on the whole, students have relatively weak models for how search engines actually work. Finally, we discuss the implications of this data for educators in information science.

15.2 Background

15.2.1 Everyday Reasoning about Internet Search-Engines

We begin by considering the popular activity of *Googling people*. In an episode of the popular and edgy HBO series *Sex and the City* we hear:

> Unidentified Woman #1: … ridiculous. And according to my new best friend, Google.com …
> Unidentified Woman #2: You Googled him!
> Unidentified Woman #1: … the man has dated every woman in New York from 19 … (Edwards, 2004, April 13).

Taking up the ethics of Googling people, *the Ethicist*, a weekly column in the New York Times Magazine, begins with a reader's question: "My friend went on a date last week and ‘Googled’ the man when she got home …. What do you think about using Google to check up on another person?” (Cohen 2002, December 15). And, continues:

> I’m for it … Had your friend labored all afternoon at the courthouse checking equally public information on her date, she’d have crossed the border between casual curiosity and stalking. Her Googling, however, was akin to asking her friends about this fellow – offhand, sociable and benign. ... By calling an act “checking up” on someone, you make typing someone’s name into a search engine sound devious and sinister. But that is less a consequence of malevolence than of its novelty … As more and more people routinely Google their blind dates, nobody will feel uneasy doing so.

On the other side, some people seek as many Google hits as possible to demonstrate their social standing: “Guys all over town are on the phone saying ‘I bet I can get more Google hits than you’ … It’s become this ridiculous new power game” (Hochman 2004, March 14). With these two quotations, we see in uncommonly compact form how search engines can lead to important ethical questions and, what’s more, influence cultural values at a remarkable pace.

At the same time, these and other newspaper pieces on *Googling people* beg many questions about the underlying operation of search engines: Why use Google and not some other engine? Who can you find through Google? How is information about people collected by Google? How reliable is the information? What responsibility does Google have for its *credibility*? How is it shared? How are queries about people processed? Does Google track searchers’ interests in people? Answers to such questions are important because they often inform conversations about information access, dissemination, and privacy.

An illustrative case is the phenomenon known as *Google bombing*, or more generally as *link bombing*, where arbitrary mappings between precise phrases and targeted Web pages are manufactured by a coordinated group of pranksters. For example, a politically motivated link bomb was created for the phrase *miserable failure*, which was linked to President George W. Bush’s official biography by approximately twenty bloggers. This small citation network was enough to boost the weight of the ranking to first place. Of course, the phrase *miserable failure* is nowhere to be found on the page itself. How, then, is this connection possible? Only with a fairly sophisticated understanding of how search engines work, can we arrive at an understanding of this quandary.

In a series of articles, the popular press attempted to explain the Google bomb phenomenon, assuring readers that this was not a political statement by Google itself (Hansell, 2003, December 8; McNichol 2004, January 22). The important role of these articles played has been to provide people with accurate conceptual models for how search engines work, including the algorithm that causes Google bombs, known as PageRank (Brin and Page 1998). These articles cover to some degree such topics as fetching content over the network, document parsing, term frequency analysis, citation analysis, and so on. In short, search engines raise important social, political, and commercial concerns that can often best be addressed, at least in part, by invoking and reasoning with technical abstractions.

Our claim, then, is that everyday questions concerning search engines lead to technical questions about their underlying computational processes and data structures. To further this claim, consider the following scenarios, drawn from articles in the popular press, and reflective of the general cultural conversation regarding search engines.

Example 1: Consider a mother who publishes stories and photographs about family outings on a 'hidden' page on their Internet Service Provider Website. While she has not been able to find her family's page by searching Google with her family's name and other (common) words and phrases found on her site, she nevertheless wonders if Google, to anthropomorphize, knows about the page and if there is anything she can do to make sure that Google does not find it. On the other hand, the popular press has reported that *Googledorks*, also known as Google hackers, seek out supposedly private documents by discovering holes in *digital gatekeepers* (Noguchi 2004, February 9). These hackers, taking advantage of Google's exhaustive crawling and extensive index of sites, develop knowledge for terms, file types, and other features that turn up putatively private documents. While an owner of a document can request that it be removed from Google's index, it is likely that he or she won't think of exercising this option until after the privacy of the document has been compromised, at which point it is often too late. However, for the mother to fully understand her question about the privacy of her family's Website, she must in turn understand such technical minutia as spiders, directory permissions, robots.txt files, the notion of 'informal technical protocols', and so forth.

Example 2: A landscape architect, who knows that potential clients often 'Google her name', in order to look for information about her past projects. Thus, she would like the link to her home page to appear on the first page of results. A knowledgeable friend has told her that the *keywords meta-tag*, a protocol for associating keywords with pages, is an ineffective technique, but she doesn't understand why. To explain why this is, we must begin by modeling the relationship between information providers and search engines, which is adversarial. Then, we must examine how keywords are extracted from Web pages, how words are normalized, how weights that indicate the importance of keywords are calculated, and so on (Belew 2000; Liddy 2001). The adversarial stance that is generally taken between the producers of content and search engines is needed in order to appreciate why these various techniques are needed and thus why associating keywords with meta-tags is usually ineffective. A collaborative stance, where content producer and a search engine cooperate in the spirit of fairness, leads to a different set of implications.

Example 3: A high school teacher suspects a student of plagiarism and attempts to verify that a passage from the essay is not original by typing a couple of suspicious word choices from the passage into Google. She is not impressed with her search results and wonders if there are better approaches to searching for plagiarized text. In order for the teacher to devise a better search strategy, she must have, at the very least, some understanding of the probability of matching word phrases, stop words, exact match queries, and so on. Indeed, responding to this need, new companies have recently formed to commercialize specialized approaches for detecting plagiarism (e.g., www.turnitin.com).

Example 4: A business analyst notices that the following queries generate unexpected hit counts: *water* (97,800,000 hits), *skiing* (7,160,000) *water skiing* (2,440,000), *water OR skiing* (14,100,000), and *skiing OR water* (13,900,000).[153] He wonders about the logic underlying this simple experiment: Shouldn't the expression *skiing OR water* yield more results than *water* alone, and shouldn't *water OR skiing* and *skiing OR water* yield identical hit counts? Perplexed that the *OR* operator does not work as expected (i.e., the commutative property of the disjunction operator does not hold) and that the sizes of the result sets are illogical, he questions his understanding of Boolean logic and wonders what rules Google follows. As this example illustrates, even experts, without proprietary information, cannot answer certain kinds of operational questions that emerge from the ordinary use of search engines.

Example 5: A marketing manager is dismayed when her company's Web site ceases to appear on the first page of Google. She has heard that the *Google dance* has reduced the relevance of her site. That is, Google has computed new relevance information that has caused changes in how results are ranked. Further, she has heard that nothing can be done except to buy keywords from Google. Companies that sell search engine optimization services, meanwhile, have promised her that their techniques can improve the relevancy of her site to particular queries. But, the practices followed for such companies, such as *link farms*, can run afoul of Google's guidelines, leading to genuine confusion in the minds of information providers over the fairness of various publishing and linking practices (Totty and Mangalindan 2003).

Example 6: An article in the New York Times reports that before submitting a pair of *chandaleer earrings* to eBay, the owner checked the spelling of *chandaleer* on Google (Schemo 2004, January 28). She found 85 hits and assumed the spelling was correct and submitted the item. The article reports that "She never guessed … that results like that meant she was groping in the spelling wilderness. Chandelier, spelled right, turns up 715,000 times." On the other hand, others troll eBay listings, looking for items that are spelled incorrectly because items that are misspelled have lower bidding activity and therefore they generally have lower prices. Indeed, it is remarkable that lexical errors and simple word choices can have such significant commercial consequences (Gleick 2004, March 21). Perhaps, greater awareness of how words are harvested and processed by Google would have enabled this person to detect her lexical error.

Each of these scenarios demonstrates how interaction with a search engine can be facilitated with a little technical understanding. Sometimes the necessary technical knowledge is in the public domain. For example, while the *robots.txt* file can be used to communicate areas of a site that should be visited, it does not guarantee that spiders will respect this informal protocol. In other cases, the technical knowledge is closely held, proprietary information and without it, it is virtually impossible to develop an accurate model for what is going on. For example, the

[153] In January 2004 these hit counts were produced by Google in response to the queries.

unexpected result set sizes for the queries concerning ‘water skiing’ appear to be caused by probabilistic methods for estimating result set sizes. Even this is speculation. Search engines do not publish information about their algorithms in order to keep themselves competitive. Perhaps it is nothing more than a temporary error – who can tell?

Of course, this lack of technical knowledge does not prevent people from hypothesizing about the operational mechanisms of search engines that lead to particular phenomena. On the one hand, people show great resourcefulness in trying to predict how a search engine functions, as can be readily observed in many online discussions. For example, the newsgroup, google.public.support.general, which is located at www.google.com, is filled with questions and answers, sometimes speculative and sometimes plain wrong, about how search works. At www.googlewhack.com, search fanatics share and discuss queries that return one and only one result. By studying these special-case queries, these searchers claim that it is possible to reverse engineer some of the methods Google employs to filter results. This knowledge, if accurate and durable, is commercially valuable because it can lead to approaches for defeating the filters and promoting a given Web page’s rank. Consultants at firms that promise *search-engine optimization* (i.e., creating Web pages that appear high on Google search results) draft intricate models of Google’s ranking process and test them by running empirical studies, tracking patent applications, job postings, and so on (e.g., see www.webworkshop.net/florida-update.html). It seems likely that this cycle of escalating competitive intelligence will continue for some time. On the other hand, it is in the search engine’s best interest to not disclose information that leads to practices that artificially improve the ranking of pages or that divulge information that might be exploited by competitors. Indeed, it is in the search engines’ best interest to present a biased conceptual model for its operations, leading people to perform behaviors that favor the search engine. The relationship between these two positions is hence adversarial: Outside stakeholders seek a full understanding of a search engines’ operation, yet to protect its intellectual property, and to satisfy its operational goals, a search engine must be highly selective in what it reveals about itself.

15.2.2 Metaphors and Mental Models for Search

Consider these neologisms from the above scenarios: Google hits, Google bombs, Google dance, search engines, link farms, spiders. From this list, we see evidence of explanatory metaphors being used to conceptualize search, as well as to prompt discussion about search engines in a given cultural milieu. Lakoff and Johnson (1980) show that metaphors are pervasive in everyday speech in order to support reasoning by using a source domain (*flies like an arrow*) to explain a target domain (*time*); indeed, they argue that metaphors are a fundamental tool to how we structure and conceptualize the world and our lives within it. While the above neologisms suggest dramatic technical mechanisms, alone they do not always tell the whole story. While *spider* is suggestive of an entity that creeps across a Web of pages, and suggest the presence of pests that owners might want to be rid of, other metaphors make sense only when you understand the underlying technical functional operation.

Consider, for example, the more complex concept *link bomb* (example given in the Introduction), which relies on the concrete domain of planted, physical bombs to explain the abstract domain of link bombs. Just as a bomb must be manufactured, packed with explosives, and set, so too must a citation network be constructed by linking a set of pages with a keyword trigger that ultimately point to the target page). Just as a bomb has a time-delay fuse which is triggered by some event, so too is time required for a search engine to process the citation network and be triggered by a keyword. Just as a bomb needs to be hidden to have its intended sudden impact, so too must the citation network be hidden. Just as persistent detective work is often marshaled to find hidden bombs, so too must search engines actively seek to detect manufactured citation networks. As with all metaphors, however, "bomb" is an imperfect mapping between a relatively more concrete source domain and a more abstract target domain (e.g., mapping the concepts of a *physical bomb* to the concepts of a *link bomb* on the Internet). For one, *link bombs* seem to be generally benign (no one dies or gets injured because of them). Indeed, they are by and large unnoticeable, except in the most publicized examples (as in the case of Mr. Bush's biography). Yet, pernicious effects can occur.[154] In sum, this metaphor encapsulates a significant amount of technical detail, but the metaphor in itself does not present a rigorous technical analogue that enables a person to understand the relationship between a source domain (*bomb*) and its target domain (*impacts of manufactured citation networks*).

This discussion leads to an obvious set of questions: What understandings and implications do people draw out of such metaphors related to search engines? How do these understandings initially develop and how do they then evolve over months and years? How are these understandings used to reason about individual and social consequences of search engines? How can technologies and educators best intervene to clarify the information issues surrounding search-engines? One approach for addressing such questions is to draw upon the theoretical notion of mental models (Gentner and Stevens 1983).

In the literature on Human-Computer Interaction, the term "mental model" is often used informally and without consistency; therefore, this construct can appear to lack analytic usefulness (Payne 2003). The term, which originated in psychology in the 1940s (Johnson-Laird 1983), appeals to the observation that over time, people develop understandings for the behaviors of other people, natural systems, and

[154] An example is that the query Jew returns anti-Semitic material. According to Google the term Jew brings up anti-Semitic material because, in general, anti-Semitic sites frequently employ the word Jew and not other words such as Judaism, Jewish, or Jewish people. After explaining the technical subtleties, an explanatory note reads: "The only sites we omit are those we are legally compelled to remove or those maliciously attempting to manipulate our results" (Google, April 30, 2004).

human-made artifacts. People, in short, learn. Then, when necessary this knowledge is used to anticipate future events to some probability and actions are selected that are believed to result in desired outcomes, to explain the reasons for the occurrence of observed phenomena, and so on. In addition, the term "model" entails the idea that one's knowledge about a given system is in some sense formal, that is, accurate and complete, thus allowing a person to identify the initial parameters of their model, simulate it in their heads, and calculate a set of consequences. For example, the operation of an elevator might be represented as a set of location states (above-floor, below-floor, and on-floor) and movements (moving-down, moving-up, stopped). With this understanding of an elevator and the starting condition (above-floor-and-moving-down), people, assuming they are waiting in a lobby and that the elevator is operating correctly, can anticipate when the elevator will arrive. Thus, in the most basic sense, a mental model allows a person to predict future events on the basis of an initial set of parameters.

Norman (1983) introduced some distinctions concerning mental models. He observed that to understand how a person interacts with a target system, called t, it is necessary to have a description of the system. He called this description a "conceptual model", labeled *C(t)*. The mental model of the system, labeled *M(t)*, is the long-term knowledge of the system. He noted that an analyst's conceptualization of a person's mental model, *C(M(t))*, will only be an approximation of *M(t)*. Thus, the manner in which an analyst elicits a person's mental model and, indeed, the manner used to describe users' models is an important consideration. Finally, Norman (1983) introduced the term "system image" to refer to the outer surface of the system, the displays, controls, help documents, and so on that inform users about the system and help users develop mental models. Ideally, a system image supports the development of a user's mental model that is congruent with the designer's conceptual model for the system. But, of course, this ideal is often not reached and, as we shall see, people typically hold only rudimentary approximations of the designer's conceptual model.

In a separate line of research, Johnson-Laird (1983) used the term "mental model" to label a cognitive architecture that enables people to perform deductive reasoning. Unlike the conceptualization of "mental models" found in Gentner and Stevens (1983), which focus on the long-term knowledge for how things work, Johnson-Laird's conceptualization hypothesizes a specific mechanism of working memory which enables people to infer valid conclusions. With deductive reasoning tasks, people are presented with a set of facts and are required to deduce a correct conclusion. The classic example is a syllogism, which takes one of a small number of forms. The simplest of the forms is:

All people like search engines
X is a search engine
Therefore, all people like X.

Johnston-Laird's theory describes how deductive reasoning tasks, such as the above *modus ponens* (if p then q, p therefore q) and *modus tollens* (if p then q, not q, therefore not p), are performed by people. The theory explains, for example, why *modus tollens* is more difficult and takes longer to perform as well as why it produces more erroneous deductions, than *modus ponens*. A general conclusion of this and other research in psychology is that such mental logic is universally difficult for people to perform because of how the human mind works. In sum, these two conceptualizations of mental models – Norman's knowledge-oriented perspective versus Johnston-Laird's short-term memory mechanism perspective – address different levels of analysis (Payne 2003 for careful analysis of the claims made of mental models). Both types have advanced our understanding for how people understand and use information retrieval systems. Next, this literature is briefly reviewed.

Borgman (1985, 1986) was the first to inquire into people's mental models – as conceptualized by Norman (1983) – for information retrieval systems. (Work preceding Borgman's seminal studies took a strongly cognitive perspective to understanding the nature of search and to derive insights for how systems could better support; for example, see Belkin et al. 1982; Ingwersen 1996) The systems investigated by Borgman were library catalogs that allowed people to enter Boolean expressions that formally specified information needs. As part of the study, she prompted undergraduate student participants to explain how these electronic catalogs worked. She found that participants had very weak models for how an electronic catalog worked even for participants who were given an explicit model of an electronic catalog and Boolean search expressions in pre-study training. In addition, she found that some participants from the undergraduate student population of the study had great difficulty writing simple Boolean expressions involving just one operator. She conjectured that the differences were due to differences in individual cognitive factors. In support of this conjecture, Greene et al. (1990) showed that higher scores on tests measuring the ability to reason correlated with a higher percentage of correct Boolean expressions. The search tasks were very similar to Borgman's study. The difference between the best and worst performers was very large at approximately 10% versus 90% correct solutions. The authors also showed, however, that this difference could be eliminated, enabling all participants to score at the 90% level, by replacing the query language with a query -by-example dialog, which enabled users to select exemplars of desired results. Thus, this study showed that the difficulties associated with generating correct Boolean expressions could be predicted by differences in individual cognitive factors but, more importantly, could be significantly reduced by changing the "system image" (Norman 1983) for querying. Other work has also sought to represent Boolean query languages through visualizations and guided user-interface dialogs that are intended to reduce the cognitive difficulties associated with Boolean expressions (Spoerri 1993; Topi and Lucas 2005; Young and Shneiderman 1993).

Taking a different approach, Internet search engines have largely supplanted Boolean searching by deploying complex algorithms for best-match keyword search. Boolean queries are typically available in advanced mode if at all (and even when offered, as seen in the example given previously in this chapter, they may not work as you would expect them to). In general, Internet search engines, with their short input fields and one-button operation, make the value proposition:

> You enter some words. Your words will be analyzed and matched against billions of documents. Only the best documents will be returned. Amazing – isn't it?

Under this oracle-like system image, the complexity of the system is hidden behind a vague description of the most straightforward pattern of interaction. With Web-based search engines, the vexing problem that plagues typical interfaces to library resources, which is thoroughly reviewed and analyzed by Borgman (1996), are addressed with a radical simplification of the query and results. When considering the external forces that act upon Web search – such as the complexity of the Internet's infrastructure, the diversity of the target audience and their information needs, the diverse motivations of the content providers, and a competitive landscape where the costs assumed by users to switch between engines is very low or entirely absent – this vagueness of operation is actually a virtue. Yet, it does beg the question: Does presenting a richer conceptual model of the underlying matching process improve the ability of searchers to find documents and, if so, for what kinds of information needs?

Koenemann and Belkin (1996) sought to answer this question by varying the degree of visibility and control of an underlying best-match retrieval engine, which also offered relevance feedback. They report that the interface with the greatest degree of visibility and control enabled users to achieve better retrieval effectiveness, and participants reported stronger positive feelings for these interfaces, in terms of usability and trust. These findings, at least for the specialized system and document collection used in this study, illustrate that by improving the visibility of the matching and retrieval process, participants could develop more accurate mental models of the system, and thus use it to a higher degree of effectiveness. Muramatsu and Pratt (2001) examined peoples' understandings for how popular Web search engines transform and match queries against documents. They observed that search engines process queries in quite different ways and that, for optimal results, one must formulate queries differently for each search engine used. For example, some search engines treat two word queries with an implied AND while others assume an OR. Some engines remove stop words while others do not. Some engines are sensitive to term order while others are not, and so on. Muramatsu and Pratt (2001) asked the question: Do users understand these operational differences? In order to answer this question, they presented 14 participants (profiles not reported) with representative query transformations and probed participants for their understandings of the search results. For example, they asked participants to explain why the query "to be or not to be" returned zero results for a particular search engine. Only two of the 14 participants were able to invoke some approximation of the notion of stop words, which explain this phenomenon. In general, they found that participants have weak mental models for query transformation. They, in turn, conjecture that users' mental models could be improved with an interface that makes the transformation visible; however, they also carefully note that they have no evidence that by improving the visibility of how queries are processed the overall search process is improved.

Other work has elicited understandings for Web search in naturalistic environments. Fidel et al. (1999) studied the information-seeking behavior of high school students, and reported that students had strikingly naïve understandings of Web-based search. One student, for example, said: "There's like a master program or something and everyone just puts information in, and it can be sent out to all the computer systems that hook up to it" (p. 27). They also report that the high school students of their study had expectations that everything is available on the Web. Slone (2002) interviewed library users at the library aiming, in part, to describe the mental models that people new to the Web employ when searching and browsing. Her data shows that while people had largely positive impressions of the Web, expressing ideas such as "everything is available" and "magical abilities." These participants also had vague understandings of search and employed naïve metaphors, and simplistic technical descriptions.

In all, these studies are fully consistent with the literature on mental models for devices, even simple devices: People have rudimentary, incomplete understandings for their functions. Second, logic-based query languages present a significant barrier in the information-seeking process and innovations in search interfaces have not been able to significantly lower this barrier. Third, while it seems that improvements in the visibility of the matching process might lead to better mental models, and in turn, improved searching, no framework for the specific factors concerning what to make visible and how has been proposed. Fourth, the mental models' orientation has not directly led to significant improvements in search interface design. Nevertheless, as argued in the previous section, knowledge of the operation of search engines can be important for understanding possibilities for expressing queries and understanding results. Thus, seeking to uncover how users' concepts of search engines lead to the expression and reformulation of queries is an important level of analysis.

Yet, broader levels of analysis also seem important. Search is no longer restricted to specialized systems for experts or to systems used by non-experts in well-defined settings (e.g., library catalogs). Rather, as we have seen, Web search engines have entered the everyday infrastructure of the general public. Thus, it is important to inquire into how people currently conceptualize how search engines work, and, even more, to inquire into how these homegrown mental models affect policy debates concerning search engines, as well as policy on the use of the Internet. Search engines, in short, are at the intersection of renewed civic-technological disputes, and they present new demands on the public's understanding of science and technology (Miller 1998).

Insofar as we know, no one has investigated the "folk theories" for how search engines work. This term signals that one's mental models, which as we have seen, consist of a set of associated abstractions that enable explanation and prediction, have been shaped to a significant extent by social factors – friends, colleagues, and communities (Holland and Quinn 1987). Consider, for example, a study of mental models, where the investigator prompted participants for explanations of how their home thermostats work (Kempton 1987). Participants were found to understand how thermostats work via either the feedback theory (i.e., a thermostat is used to set a target temperature and the heating system turns itself on and off in order to hit that temperature) or the valve theory (i.e., a thermostat is like a gas peddle that regulates how much heat flows into the room). While participants that used the feedback theory to understand the thermostat, rarely adjusted it, those who used the valve theory tended to adjust the thermostat more frequently throughout the day. This work has been applied to the design of thermostats so that they match a given mental model and save energy. Now, turning to a domain more closely related to search engines, Payne (1991) asked people to explain how automatic bank machines functioned. In individual sessions participants were probed for their understanding of these machines by means of what-if questions such as: What happens to the card during the transaction? Why does it stay inside the machine? When analyzing the verbal protocols, he found a great diversity of explanations concerning the how the computational processes were decomposed and related and the roles of various storage devices (e.g., bank card, local teller machine, and centralized data bank).

The participants in studies of mental models are often non-specialists. Comparing their understanding of devices against expert models provides a method for exploring how information is imparted through specific devices or cultural sources. In turn, by examining the difference between people's understandings and the original conceptual model, one can seek to change the system image in order to clarify the conceptual model and hence improve the usability of the system. Moreover, the models that specialists hold are also worthy of investigation especially when specialists from different backgrounds need to communicate across disciplinary or institutional boundaries. An interesting example of this kind of a conceptual model for Web search has been created by Matt Leacock, a visual designer (Brown 2001). This conceptual model represents the search process with approximately 60 concepts and 100 relationships between these concepts. The model is divided into five conceptual zones and the concepts and relationships are very carefully laid out. To see the complete model in its entirety requires that it be printed on a 36 in. by 36 in. poster. An elided version, consisting of 20 concepts has also been published (Wurman 2001: 158). The aim of these complex models was to externalize a complete map of how a complex, enterprise-critical search system functioned. To produce the model, Leacock interviewed individual members of product groups and developed a composite model of how people understood the search system. This model was posted in public locations along with a red pen to encourage annotations and revisions. He found that no single person understood how the system operated but that by developing a complete model and placing it in public forums he was able to make the complexity of the system visible. This enabled people to communicate better, despite shifting teams and priorities, as well as differences in technical perspective (Brown 2001). Thus, the manner in which people tell stories about search and externalize their knowledge of search is an interesting type of technical communication.

15.3 Exploratory Study

To examine how people conceptualize Web search we decided to prompt students to draw sketches of how they thought search engines work. Then, we performed a content analysis of the resulting body of material. In Norman's terms (see previous section) this method elicits conceptual models, $C(M(t))$, from non-experts. We make no claims concerning how these models are put to use when reasoning about search engines in specific problem-solving or conversional contexts; in fact, for most participants this is likely the first time that they expressed their understanding for search-engines in any form. Furthermore, it is important to note that participants varied in their level of ability and comfort to draw sketches in a short period of time. The task, in short, was quite demanding. We decided on this form of expression because sketching is an expressive, open-ended form of communication, allowing people to stress what is important to them through both drawings and words.

Participants in this work were students at various levels of academic achievement in Information Science, ranging from freshman with undeclared majors to Ph.D. students in Information Science. This participant group is an interesting population to study for two reasons. First, as a group we can expect a diversity of experiences with Web search engines. Some students in Information Science, especially at the graduate level, will have had opportunities to develop their knowledge for search and to explain search to other people. Other students will have limited or no formal training in search but can be expected to have a high level of exposure to and interest in search engines. Thus, these students provide a population of users with a broad range of experience of search. Certainly, we expected graduate students to reflect the upper bound of knowledge. In any case, because of these students' level of educational accomplishment, generally high use of the Internet and search, and specific area of interest (Information Science), one would expect that this sample would have a relatively high-level knowledge. Second, as instructors of classes on Database and Information Retrieval systems, we were extremely interested in both the technical and folk knowledge that our students held for search systems. Thus, collecting this data, analyzing it, and reflecting upon it have also served a very practical need: to enable lively classroom discussions about Web search and to orient us to our students' understanding of how search works.

This exploratory study, in sum, addresses four research questions: 1) What concepts do people include and emphasize in their conceptual models; 2) What misconceptions are found in these models? 3) What visual forms do people use to express their understanding of search engines? 4) What metaphors and technical terms are used? Following the existing literature, we hypothesized that the models would reflect only a rudimentary understanding of search engines and that participants with greater levels of academic accomplishment in Information Science would produce more nuanced conceptual models with more correct concepts. Preliminary findings of this research were presented in Hendry and Efthimiadis (2004) and Efthimiadis and Hendry (2005).

15.4 Method

15.4.1 Instructions

At the top of a blank 8 × 11 in. paper sheet, undergraduate and graduate students at the University of Washington were instructed to draw and label a sketch explaining how a search engine works. Students were given approximately 10 minutes to complete the task at the beginning of a regularly scheduled class. The exact instructions and time available to complete the task varied because different moderators collected data in different classes. A sample of 232 sketches was collected in the spring and autumn of 2003.

15.4.2 Participants

The student participants (N = 232) were from the following academic levels: 1) Freshman taking their first college-level course; 2) Juniors and Seniors pursuing an undergraduate degree in Information Science; 3) Fulltime students pursuing a master's degree in Library and Information Science; 4) Working professionals pursuing a two-year executive degree in Information Management; and 5) Fulltime students pursing a doctorial degree in Information Science. For this analysis, student participants were assigned to the following three groups: 1) Undergraduate-freshman (n = 53); 2) Undergraduate-informatics (n = 95); and 3) Graduate-information-science (n = 84). While these categories represent three general levels of academic achievement, the demographic profiles for the participants within these groups are heterogeneous, especially for the second two categories, with broad ranges in ages, work experiences, and educational achievement.

15.4.3 Reference Model of Internet Search Engines

In order to analyze the sketches, a conceptual model for search was chosen as a reference point. This model drew upon standard textbook components of search engines (Belew 2000; Liddy 2001) and identified the major conceptual components of any generic search engine. The model divides search into three phases, indexing, searching, matching, each of which contains its own processing components, as follows:

A. INDEXING: Processing documents so they can be retrieved later

1. *Content:* The search engine accesses documents, such as Web pages.
2. *Spidering/Crawling:* The search engine fetches Web pages
3. *Parsing:* Words from Web pages are extracted and analyzed in some fashion
4. *Inverted-index-creation:* An index that maps words to Web pages is created
5. *Link-analysis:* The search engine analyzes the linking structure among Web pages
6. *Storage:* Web pages and indexes are stored at the search engine

B. SEARCHING: Users formulate a query and inspect results

7. *User:* A person interacts with the search engine
8. *User-need:* A 'need' triggers a user to perform a search
9. *Query:* An interface is used to submit a query to a search engine
10. *Results:* The output from a search-engine are a list of Web pages

C. MATCHING: Queries are matched against Web pages

11. *Query processing:* Keywords and operators are extracted from the query
12. *Matching:* Words from the query are matched against words in Web pages
13. *Accessing-inverted-file:* Keywords are used to access the inverted file
14. *Ranking:* A ranked ordering of Web pages is created

In the analysis below, this model is used as a baseline instrument to assess the completeness of the participants' conceptual models.

15.5 Results

Figures 15.1–15.7 show seven sketches () that are representative of the full sample of 232 sketches. Notably, these sketches – and the full sample – reveal a tremendous diversity of approaches for explaining the operation of search engines. Figure 15.1 is noteworthy for employing multiple metaphors while maintaining compositional coherence and Fig. 15.2 is noteworthy for employing both symbolic and representational elements while also maintaining compositional coherence. Sometimes, metaphoric imagery or idiomatic symbols are used; for example, a cloud is often used to depict the Internet and a cylinder is often used to depict an information store (e.g., see Figs. 15.1 and 15.2).

Figure 15.3, one of the most detailed and complete sketches in the sample, is an extreme example where, in a reversal of typical roles, the visual language

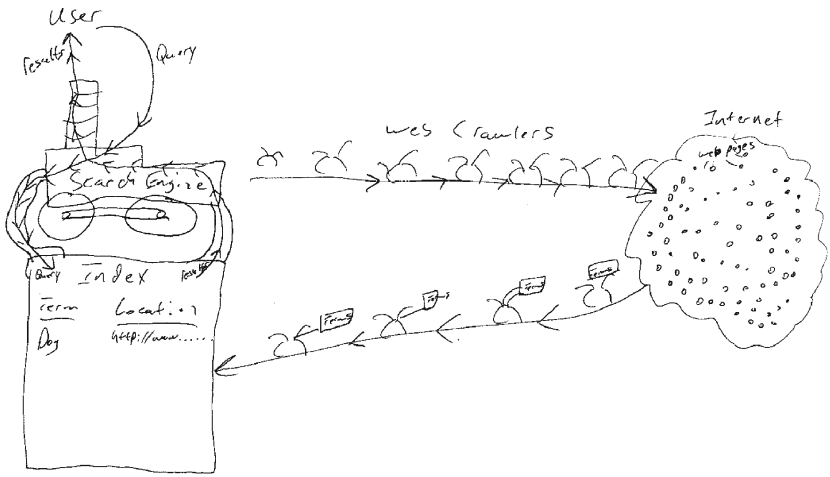

Fig. 15.1 Sketch of search engine illustrating the use of various metaphors, including a mechanical engine, complete with drive-train between wheels and a smoke stack, that performs the matching process, a cloud of particles indicating Websites on the Internet, and spiders that leave the search engine empty-handed and return with terms. In addition, the inverted file, user, query, and results are depicted

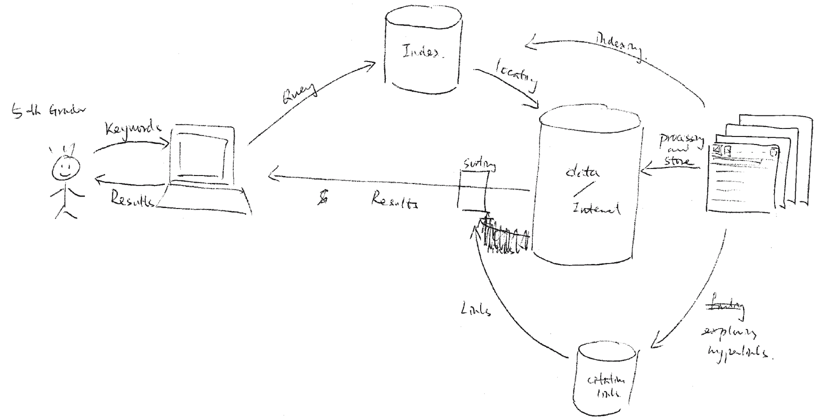

Fig. 15.2 Sketch of search engine illustrating the use of idiomatic symbols including cylinders for data stores, stick figure for users, computer monitor and keyboard client computer, and Graphical User Interface window for content. The processing steps are depicted with labeled lines between data stores and system inputs and outputs. This sketch illustrates an uncommon degree of coherence

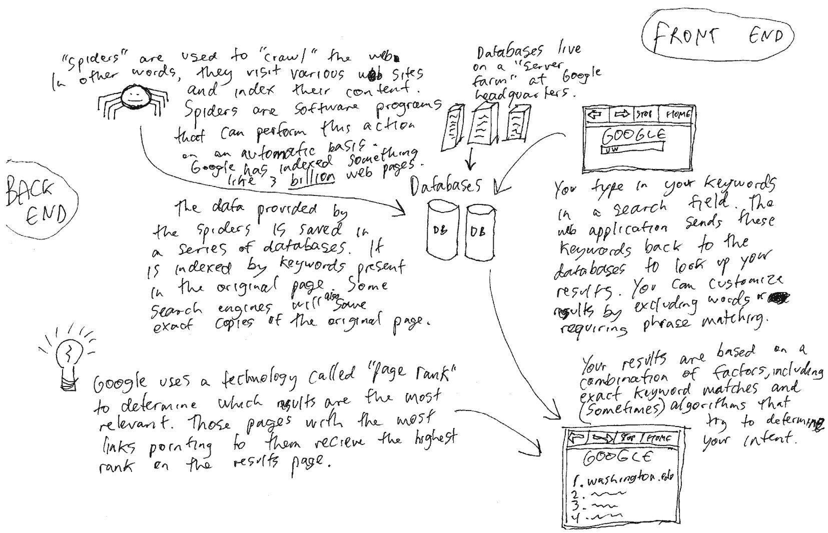

Fig. 15.3 Sketch of search engine that reveals a significant technical maturity, including an explanation of PageRank, approximate size of the WWW, and the complexity of determining a ranking of pages. The sketch segments the process into the front-end and back-end components. The use of visual symbols and user interface representations is noteworthy because to a large degree this visual language supports the written annotations – the reverse of many sketches. Finally, the light bulb, suggesting innovation and intelligence, draws attention to PageRank, a distinguishing characteristic of Google's matching algorithm

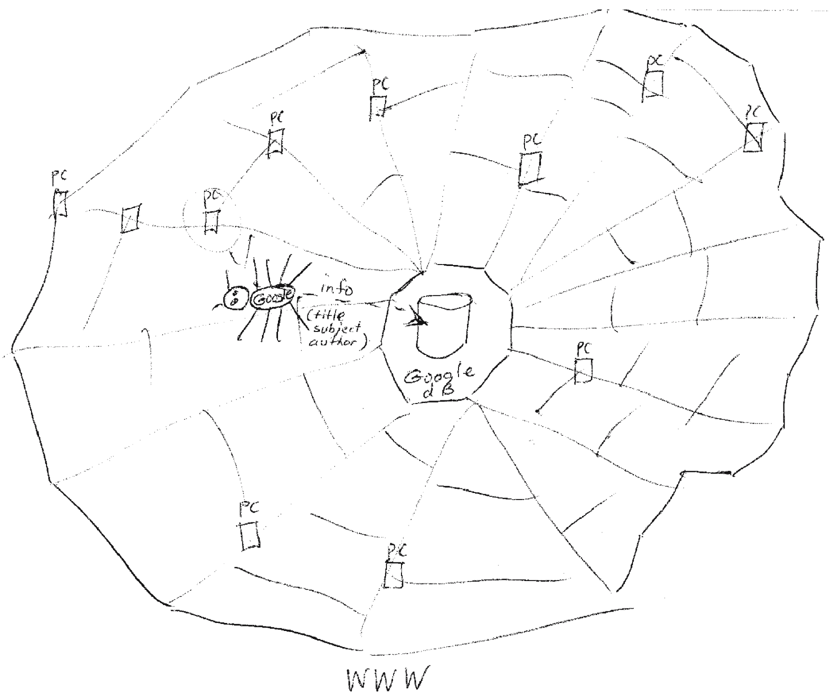

Fig. 15.4 Sketch of search engine illustrating the centrality of search with the Google DB at the center of a neatly organized Web of connections between PC computers. The Google spider crawls the Web, sending back information in the form of title, subject, and author

clarifies the narrative text. Some of the sketches are largely representational, and in such cases metaphors are depicted in a relatively simple manner or, for example, a query dialog and results display is sketched and the underlying machinery is not depicted (e.g., see Figs. 15.5 and 15.7). Other sketches are more general where box-and-line symbols are used to identify information types and communication pathways, such as those between client and server computers (e.g., see Fig. 15.2). None of the 232 sketches, however, employed a formal notation for representing systems, such as an Entity-Relationship modeling. Finally, unlike Figs. 15.1 and 15.2, many of the sketches depicted only a few concepts and relationships (e.g., see Figs. 15.5 and 15.6). The following sections summarize the information found in the sketches.

15.5.1 Concept Analysis

To assess the overall presence of search concepts in the sketches, each of the sketches was coded for concepts in the normative model presented above. As can be seen in even the small sample of eight sketches, these concepts manifest themselves

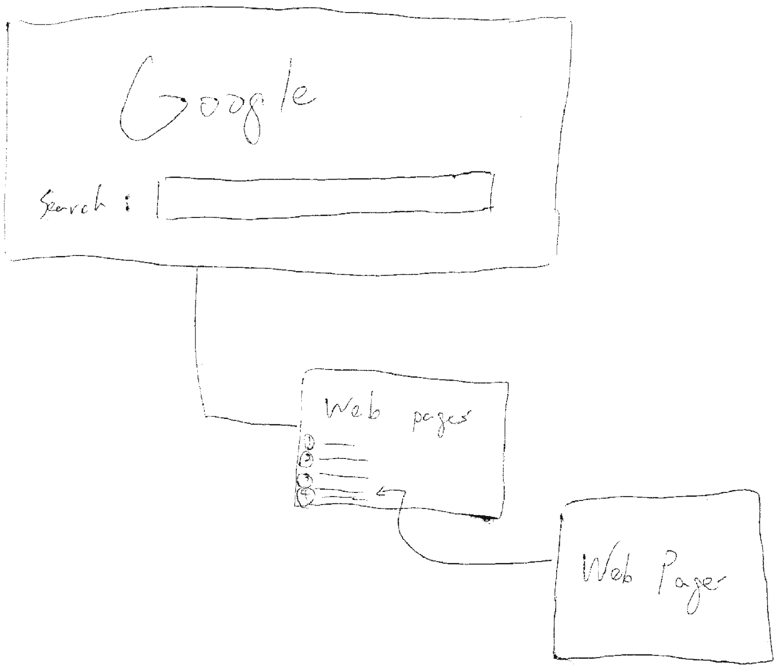

Fig. 15.5 Sketch of a search engine that illustrates the user interface. The first screen is recognizable as the Google input form for its use of whitespace and results pages shows a ranked list of Web pages

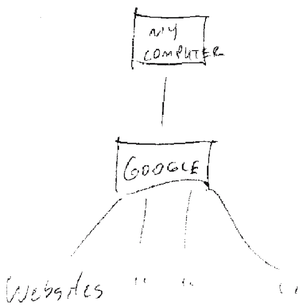

Fig. 15.6 Sketch of search engine that distinguishes between the client and server components and that indicates Google links to web sites

in numerous and different ways. For example, a query concept might be depicted as a box labeled 'query', as an input field and submit button, or as an annotation such as 'enter your keywords here'. Figures 15.1 and 15.5 each depict a query but in different styles. In this analysis, each of these manifestations of the concept would be counted.

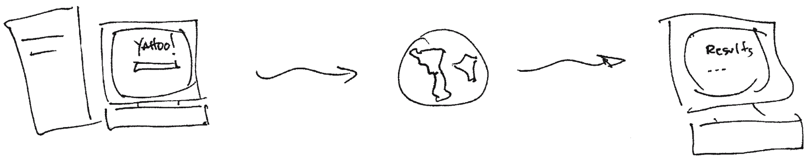

Fig. 15.7 Sketch of search engine that illustrates that a client machine communicating with the world and returning results

The process for coding the sketches followed these steps:

1. The normative model was documented and a group of four coders, including the authors of this chapter, discussed this model and developed a common understanding for its concepts.
2. Working independently, the coders coded a sample of four sketches by inspecting each sketch and making a judgment for the presence or absence of each of the 14 concepts. Below, we call these binary judgments "votes".
3. The coders met to review each others' votes and discuss any differences in judgment. After three rounds of independent voting followed by group discussion, it was decided that the sketches were being coded in a sufficiently consistent fashion that the whole sample could be analyzed.
4. Working independently, each coder inspected each of the 232 sketches for the 14 concepts. This resulted in 12,992 votes for the presence or absence of concepts (4 coders × 14 concepts × 232 sketches).

The votes were analyzed for intercoder reliability by computing the percentage of agreement between each pairwise combination of the four coders for all 12,992 votes ($M = 0.84$, $N = 6$, $SD = 0.02$). At first glance, this may suggest a relatively high degree of agreement. But, in fact, these numbers overestimate the intercoder reliability because percentage agreement does not correct for cases where there is agreement by chance. This is especially important in this analysis because, as we shall see, the likelihood that a concept will be absent from a sketch is much higher than the likelihood that it will be present. Cohen's kappa statistic corrects for chance and is used extensively in the evaluation of intercoder reliability in medicine and content analysis. Unlike percentage agreement, which is rather liberal, Cohen's kappa is a rather conservative measure. This is because kappa accounts for the differences in the distribution of values across the categories for different coders and only gives credit for agreement beyond the distributions of values in the marginals (Lombard et al. 2002: 592). Cohen's kappa was calculated for each pair of coders (kappa = 0.57, $N = 6$, $SD = 0.04$). In general, this level of agreement is considered as moderate level of agreement beyond chance (Landis and Koch 1977: 165). Consensus on calculating, reporting, and interpreting intercoder reliability is lacking in the literature on content analysis, an especially important method of analysis in studies of media use and human-to-human communication (Lombard et al.

2002). Nevertheless, given the complex nature of the data and its overall pattern, we believe that a sufficient level of reliability is obtained when the following cut-offs are made: 1) if 3 or 4 votes inclusive, concept present; and 2) if 0–2 votes, concept absent. Using these cut-offs, the votes were counted to determine the presence-or-absence status of each concept in each sketch. This transformed data is used in the analysis below. It is also important to note that the intercoder agreement vary across concepts. For example, the coders could more reliably identify the presence or absence of the concept *query* than they could for the concept *accessing-the-inverted-file* because *query* is a simpler concept.

Figure 15.8 presents the frequency distribution of concepts across all sketches, showing that a sketch contains on average about 4.5 concepts ($SD = 3.0$) with a low of 0 concepts ($n = 25$) and a high of 13 concepts ($n = 2$). Examples of sketches with 0 concepts are written notes such as "I don't know" and "Magic" and uninterruptible sketches such as one depicting an octopus, a stickman exchanging documents, or sketches of cartoon characters that seem to be processing information generally

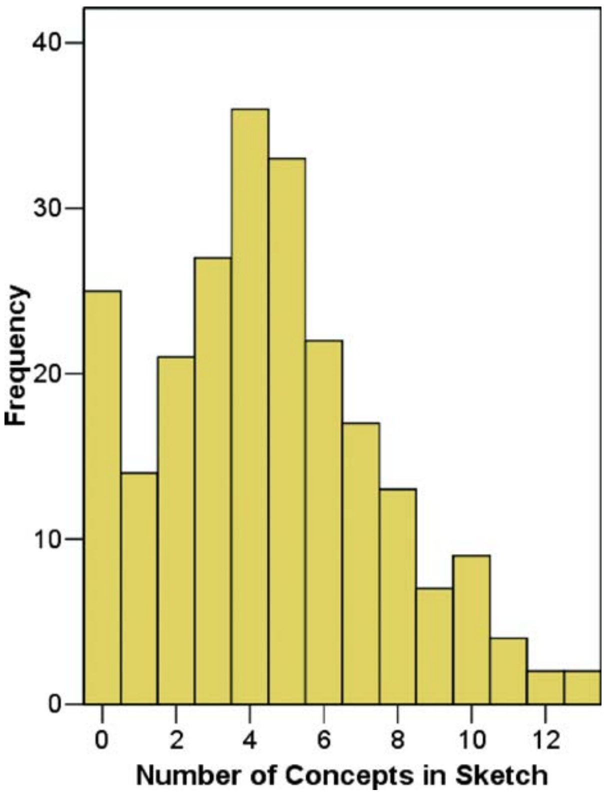

Fig. 15.8 The frequency distribution of number of concepts depicted in sketches ($N = 232$). On average, 4.5 concepts ($SD = 3.0$) are depicted in each sketch with a low of no concepts ($n = 17$) and a high of 13 concepts ($n = 1$)

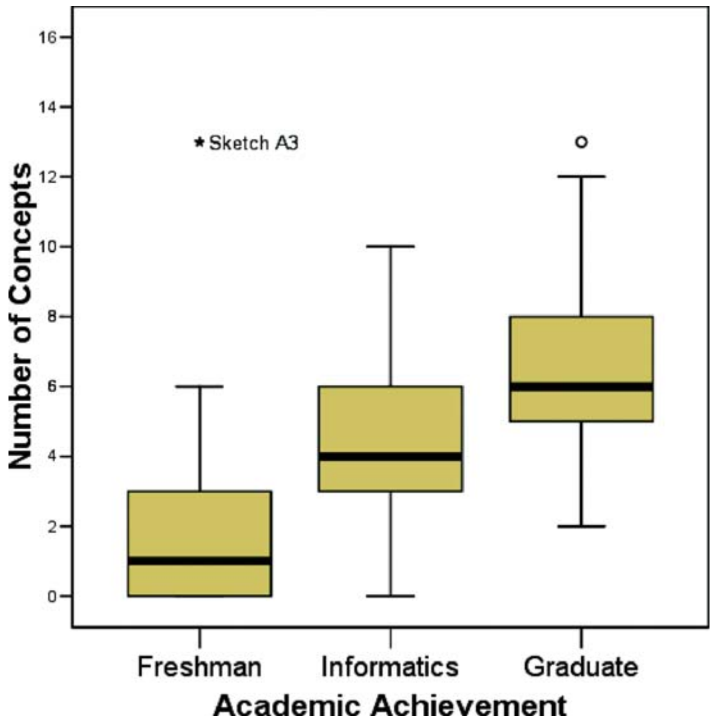

Fig. 15.9 Summary of concepts depicted in sketches by participant category, undergraduate-freshman (*Mdn* = 1.0, *SD* = 2.3, *n* = 53), under-graduate-informatics (*Mdn* = 4.0, *SD* = 2.4, *n* = 95), and graduate students (*Mdn* = 6.0, *SD* = 2.5, *n* = 84)

but lacked any identifiable explanations. Figure 15.9 presents the data collected by student group, showing, as might be expected, that graduate students in Information Science are able to depict more concepts than undergraduate freshmen or other undergraduate students in Information Science. Turning to the concepts depicted in the sketches, Fig. 15.10 presents the distribution of concepts found in the sketches with *query, results*, *content* and *user* being the four most frequently occurring concepts and *user need*, *link analysis*, *inverted-file-access*, and *query processing* being the four least frequently occurring.

15.5.2 Use of Metaphor, Notation, and User-Interface Imagery

Many of the sketches employ one or more metaphors to explain how search engines work, with, for example, Fig. 15.1 making a visual play on the metaphor *engine*. Figures 15.1, 15.4, and 15.7 are typical of the metaphors found in the sketches.

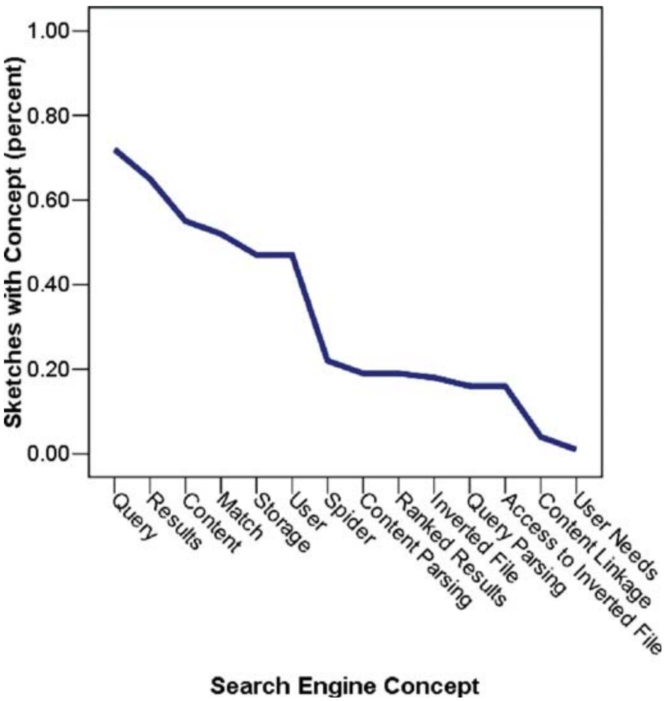

Fig. 15.10 Summary of the 14 concepts depicted in all sketches ($N = 232$)

Images of *clouds* (e.g., see Fig. 15.1) and the *earth* (e.g., see Fig. 15.7) were commonly used to suggest the vast, undifferentiated yet ultimately connected and contained nature of the Internet. *Spiders*, *crawlers*, and *Webs* were used to illustrate the process of discovering and fetching content. *Books*, *bookshelves*, *store rooms*, and *card catalogs* were used to represent information stores or to indicate a degree of information organization. *Computers* were often given *arms*, *faces*, *smiles* and other anthropomorphic features to indicate such notions as agency and intelligence. *Gnomes*, *bots*, *robots*, *brains*, *stick-figure dogs* and other agents were used to suggest autonomous action and intelligence. *Eye glasses*, *magnifying glasses*, and *eyes* were used to suggest that the search process is about looking. *Radio towers*, an *orbiting satellite*, and a *bridge* were used to indicate all encompassing communication. A *message in a bottle* was used to suggest the challenge of finding relevant information. *Stick-figure people* with raised arms or *scoring a goal* with a foot were used to suggest successful searches. See Hendry (2006) for a detailed qualitative analysis of the conceptual metaphors that were employed in the sub-sample of sketches depicting algorithmic processes.

Turning to notation, many of the sketches contain symbols that represent a type of information and process. The symbol *cylinder* is frequently used to represent the storage of data. Figure 15.2, for example, depicts Web pages as a neat pile of documents that is in turn transformed into three different types of data: *indexes*, *data/Internet*, and *citation data*. *Monitors* and *keyboards*, as shown in Fig. 15.7, are often used to represent computers. It was more common, however, for participants to draw box-and-line diagrams with labeled inputs, outputs, processes, and data stores. For example, the concepts *user*, *search-engine*, and *database* are represented with labeled boxes and connected with directed lines. Many sketches employ subtle differences in notation to signal differences. For example, solid and dotted lines are sometimes used to signal a firm versus tenuous relationship between two artifacts, and symbols, such as circles and squares, are sometimes used to signal different types of entities. While such subtle differences occur frequently in the data, rarely is the meaning of the notational differences explicitly stated or used consistently. Finally, it was very common for participants to employ a mix of notation, metaphoric imagery, and representation within the same sketch. For example, Figure 15.3 uses *cylinders* and *towers* to represent data and server farms, representational *boxes* to represent Web pages for query and result, and the image of a *spider* to represent the computational concept of "spidering" and "crawling." In Fig. 15.7, an image of the *world* is used to associate a query with a result that are both depicted as computers. As a metaphor, the image of the *world* was also used to refer to multiple ideas, including the geographic spread of the Internet as well as a repository of information at global proportions.

15.5.3 Misconceptions

The sketches also reveal a variety of misconceptions. Regarding information structure and organization, some sketches depict an Internet where a full list of Websites can be readily enumerated or an Internet that is an organized collection of Websites. Some participants gave search engines a privileged position to information: Google is often depicted at the center of the Web, and sometimes Google is even shown to, or at least implied to, directly link to Websites (see Fig. 15.6). Some participants depicted information as residing inside a search engine with, for example, Web pages arranged on bookshelves or pre-computed search results waiting to be retrieved and presented. Some sketches suggested an automatic categorization process where items found by a spider, for example, are sorted into categories by a computational process; other times, participants depicted human intervention, where people make selections based on editorial and legal standards during the indexing process. Meta tags were often denoted as a source for the indexing process, although search engines treat these terms with great care. Concerning the search process, some participants suggested collaboration amongst Web search engines: one engine asking for results from another engine, or a hierarchy or engines with Google at the center and other commercial services subsumed by it on a secondary tier. Some participants depicted de-centralized algorithms where Google initiates a search by asking a second tier of computers, which, in turn, ask a third tier. Concerning the matching process, participants often illustrated naïve sequential letter-by-letter matching algorithms (akin to regular expressions for matching) or vaguely expressed notions of indexed-based lookup.

15.6 Discussion

The main finding of this exploratory study is that students in this sample produced mostly rudimentary conceptual models for how search engines work. Even Graduate students in Information Science were, in general, only able to describe a few concepts within their sketches, and these were often the most obvious concepts (e.g., query and results). Undergraduates and freshman in Information Science produced sketches with still fewer concepts. A second finding is that the sketches reveal a great diversity of approaches for expressing a conceptual model. Some sketches proposed algorithms, illustrating successive transformations of data. Others were highly representational, showing iconographic depictions of such things as results, queries, and communication networks. Still others relied on the metaphoric language available, such as spiders and Webs. In sum, students seem to know relatively little about how search engines work and they describe what they know in very different ways.

Thus, this study follows the pattern of much of the literature on mental models. As Norman puts it: "… most people's understanding of devices they interact with is surprisingly meager, imprecisely specific, and full of inconsistencies, gaps, and idiosyncratic quirks" (1983: 8). Indeed, as we saw earlier, this is the main conclusion of previous studies of people's understandings of search. This study reproduces these findings in the current technical milieu. The instrument used in this study – drawing a conceptual model in a short period of time for a very complex system – is admittedly demanding, and the results likely underestimate students' knowledge, which would otherwise be expressed more robustly in situated or diagnostic settings. Nevertheless, in general, we believe that students' performance on this task should be much higher if the conceptual knowledge for how search engines work was a basic component of technical literacy. Without this knowledge students are ill-equipped to engage in topics associated with search engines and, indeed, to teach others about search engines – an activity that many students of information science programs will engage in during their careers. This argument for knowing the central concepts of search engines, moreover, applies to non-student populations as well, including everyday users of the Internet who, as recounted above in the stories from the popular press, often understand search as a perplexing phenomenon.

Assessing people's knowledge for search engines can be seen as a special case of the general problem of *civic scientific literacy* (Miller 1998), that is, having sufficient competence with science to understand public policy debates that center on science and technology. The argument is that a healthy democracy requires a scientifically literate citizenry; otherwise, citizens will be poorly equipped to influence public policy in such matters as nuclear power, reproductive technologies, global warming, and so on. Thus, an important goal is to measure the scientific literacy of a population for the purposes of benchmarking and garnering support for public education in science and technology, for making cross-cultural comparisons, and so on. Survey instruments of open- and closed-end questions that measure a person's knowledge for the standard of scientific inquiry and the knowledge of scientific vocabulary have been developed. These instruments, with the appropriate sampling procedures and statistical analysis, are claimed to produce a durable, meaningful measure of populations' scientific literacy (Miller 1998). This approach, it is important to note, privileges knowledge in the head. And, as a result, it has been attacked for not accounting for the situated, collective development and application of scientific knowledge, especially when technological issues play out in social contexts (Roth and Lee 2002). For now, we put this dispute aside and simply note that both positions have merit.

Next, we turn to the question of how best to intervene to improve the public's knowledge of search engines. One observation is that it is important to equip people with conceptual knowledge for search engines that can be put to use in different problem situations. Of course, the application of this conceptual knowledge may require other forms of knowledge that are specific to the problem setting (Borgman 1996). A second observation is that the conceptual knowledge of search is not localized to a well-bounded setting or system. Rather, it is distributed amongst a diverse number of sub-systems that make up the artificial world of the Internet, including Web servers, browsers, Internet protocols, search engine operations, and so on. Thus, approaches to explaining Web search engines will have to take into account the full complexity of the Internet, networking, fiber optics, etc. Below, we organize approaches of intervention into three categories:

1. Models and simulations of search engines;
2. Forums for discussing search engines;
3. Contextually relevant explanations.

15.6.1 Models and Simulations of Search Engines

Halttunen (2003) and Halttunen and Jarvelin (2005) seek to teach students about search engines by developing a constructive learning environment, called the Information Retrieval Game, which allows students to develop skills and conceptual knowledge for how search works. With this learning tool, students perform searches against a test collection and are given specific feedback on the quality of their searches. Thus, this approach helps students to develop specialized skills in searching. In contrast, to this pedagogically-centered approach are specialized tools, largely designed for programmers, for visualizing search processes. The Luke tool (Luke 2004), for example, allows programmers to inspect the search indices, query processing, and matching process for the Lucene search engine; indeed, in our teaching experience, it has proven to be quite effective for helping novice programmers learn the Lucene Application Programming Interface. For students in a library and information science program, Efthimiadis (2003) has developed the IR Toolbox, an experiential teaching tool for learning about information retrieval systems (Efthimiadis & Freier, 2007). Through hands on interaction, the IR Toolbox helps students develop their conceptual model of search engines by exploring, visualizing, and understanding IR processes and algorithms without needing to program. In a sequential fashion, the IR Toolbox presents the following processing steps: a) Document analysis (e.g., tokenizers, stemmers, stop lists), b) Indexing (e.g., ability to browse inverted file and extract statistics), c) Searching (e.g., ability to enter queries and select weighing algorithms such as IDF, TF-IDF, OKAPI), d) Evaluation (e.g., evaluate results using the TREC evaluation software and associated collections, presenting recall-precision tables and graphs). The IR-Toolbox uses Lucene as its underlining search engine. Students can interact with the IR Toolbox at different levels of complexity on individual or group exercises that help them understand the different IR processes and build a more detailed conceptual model of search engines.

For a more general audience, a viable approach would be to develop specialized simulations of the operation of Web search engines. These simulations would present conceptual models of Web search, and allow people, especially non-specialists, to visualize search engine processes, focusing particularly on the issues of Internet search. This approach would be an elaboration of the models often presented in the popular press – perhaps; the best analogue would be an interactive science-center museum exhibit for explaining a complex process. A further step would be to give people the ability to construct their own search engines though an end-user programming environment which allowed them to visualize and refine their work (Fischer et al. 2004; Hendry 2006; Hendry and Harper 1997).

15.6.2 Forums for Discussing Search Engines

A second, complementary approach would be to develop a forum for discussing search engines. The root concept would be to create an open, constructive place that supports learning about how search engines work for everyone. From this root concept, we propose the following three general requirements. First, the forum should be run by a neutral organization that does not give preference to any particular search engine. This is important because, as we have seen, it can be in the search engine's interest to misinform its audience so that people tend to behave in ways that are commercially advantageous. This requirement is derived from the relationship between search engines and content producers, which, as discussed previously, is fundamentally adversarial. Second, experts in search need to participate in the forum. They need to help guide the conversations as good teachers do, correct misinformation, add nuance to conjectures and speculations, propose "experiments" that clarify how search engines work, and explain when and why firm conclusions cannot be drawn. Third, and perhaps most of all, the forum needs to track and clarify the public policy disputes related to the use and development of search engines. It seems inevitable that as search engines undergo technological advancement, value-oriented issues centered on fair access to information, autonomy to pursue one's own interests, information credibility, and others will become important to the public (Friedman and Kahn 2003). The forum we have in mind would seek to educate the public through collective participation, allowing experts and non-experts to engage in serious dialog. In short, the forum would enable inquiry into the science and technology of search to be socially grounded. Insofar as we know, a forum with these aims does not exist, but however utopian this may sound, it would be of great benefit to the public if it did exist!

15.6.3 Contextually Relevant Help

The final approach for helping people to develop a robust conceptual model for how search engines work is to enable people to probe the operation of a search engine in a highly situated fashion. This, of course, is easier said than done as the lack of meaningful help messages, in general, and of context sensitive help, in particular, has dogged retrieval systems since the seventies. During the eighties there was an effort to include context sensitive help in front-end systems and expert intermediary systems with varied levels of success (see Efthimiadis, 1990, for a detailed literature review). The explosion of end-user search, on CD-ROM products and the Web, during the nineties shifted attention to other issues with no satisfactory solution to the problem. Research in this area includes work by Gauch and Smith (1993), Oakes and Taylor (1998), and more recently by Jansen (2005).

Our design ideas differ from the implicit suggestions that search engines make to users. These query refinements are not consistently correct and, in addition, require that the user could recognize them as well as distinguish them from sponsored results.

Triggered by some kind of breakdown, we therefore envision users being able to engage in meaningful interaction with the search engine, either by receiving system prompted context sensitive help, or by entering a diagnostic mode where they could ask questions about the problematic interaction or the problematic operation of the search engine. For example, if a person's home page does not appear on the first page, the user could ask a search engine to explain why this happened in the context of a particular query and set of results. In such a situation, the influence of query keywords, keywords on links, and page-to-page citation patterns could, in principle, be presented to users. Given the interests of the search engine, however, the searcher would do well to be skeptical. Obviously, such functionality would be used only rarely by those who are trying to understand the inner working of search engines. Nevertheless, being able to systematically explore and diagnose in the context of actual searching a particular query and set of results could provide a strong learning environment if the search engine were willing to disclose key information.

15.7 Conclusion

Search and culture are entwined in a dynamic dance. It is clear that people develop conceptual models for how search engines work, and, as this and previous studies have shown, these models are relatively weak. Less clear, however, is how educators, reformers, and activists can intervene effectively to improve the public's understanding of search. Yet, it seems clear that as search becomes even more embedded in our lives, value-oriented questions about the responsible and fair use of search will become more and more important. The adversarial relationship between content providers and search-engines is a transformative change in search that will be reckoned with for many years to come. In summary, the problem of search is one aspect of a larger question regarding the public's understandings of science and technology, of civic scientific literacy. Miller (1998: 220) says: "It is important to learn more about the magnitude and dynamics of [informal learning resources and processes] and about adults' selection of and trust in various kinds of communications [such as libraries, newspapers, magazines, television shows, and museums]". Quite right.

Acknowledgements We would like to thank Kreg Hasegawa for a very thorough review of a draft of this manuscript, Peyina Lin, Kristene Unsworth, and Hui P. Yang for help in conducting a content analysis of the sketches, and John LaMont who helped track down citations.

References

Belew RK (2000) Finding out about: a cognitive perspective on search engine technology and the WWW. Cambridge University Press, Cambridge

Belkin NJ, Oddy RN, Brooks HM (1982) ASK for information-retrieval: 1. background and theory. Journal of Documentation 38: 61–71

Borgman CL (1985) The user's mental model of an information retrieval system. Paper presented at the Proceedings of the 8th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (pp 268–273, June 5–7). Montreal, Quebec, Canada

Borgman CL (1986) The user's mental model of an information retrieval system: an experiment on a prototype online catalog. International Journal of Man-Machine Studies 24: 47–64

Borgman CL (1996) Why are online catalogs hard to use? Lessons learned from information-retrieval studies. Journal of the American Society for Information Science 37: 387–400

Brin S, Page L (1998) The anatomy of a large-scale hypertextual search engine. Paper presented at the 7th International World Wide Web Conference. Retrieved July 1, 2004, from http://wwwdb.stanford.edu/~backrub/google.html

Brown, DR (2001) SuperModeler: hugh dubberly. Gain: AIGA Journal of Design for the Network Economy, 1: 1–8. Retrieved, June 1, 2004, from http://gain1.aiga.org/pdf/profile.pdf

Cohen R (2002, December 15) Is googling okay? New York Times Magazine, p 50

Edwards B (2004, April 13) Morning edition: search engine wars, Part II. (Radio Broadcast). Seattle, National Public Radio, KUOW

Efthimiadis EN (1990) Progress in documentation: online searching aids: a review of front-ends, gateways and other interfaces. Journal of Documentation 46: 218–262

Efthimiadis EN (2003–2007) The IR toolbox. Available at http://irtoolbox. ischool.washington.edu

Efthimiadis EN, Freier NG (2007) IR-Toolbox: an experiential learning tool for teaching IR. In: Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (Amsterdam, The Netherlands, July 23–27, 2007. SIGIR'07, ACM Press, New York

Efthimiadis EN, Hendry DG (2005) Search engines and how students think they work. In: Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (Salvador, Brazil, August 15–19, 2005). SIGIR '05. ACM Press, New York, NY, pp 595–596. http://doi.acm.org/10.1145/1076034.1076145

Fidel R, Davies RK, Douglas MH, Holder JK, Hopkins CJ, Kushner EJ, Miyagishima BK, Toney CD (1999) A visit to the information mall: web searching behavior of high school students. Journal of American Society for Information Science 50: 24–37

Fischer G, Giaccardi E, Ye Y, Sutcliffe AG, Mehandjiev N (2004) Meta-design: a manifesto for end-user development. Communications of the ACM 47: 33–37

Friedman B, Kahn PH Jr (2003) Human values, ethics, and design. In: Jacko JA and Sears A (eds), The human–computer interaction handbook (pp 1177–1201). Mahwah, NJ: Lawrence Erlbaum Associates

Friedman B, Howe D, and Felten E (2002) Informed consent in the Mozilla browser: implementing value sensitive design. In: Proceedings of the 35th Annual Hawaii International Conference on System Sciences (HICSS '02) Vol 8 (January 7–10, 2002). Washington, DC: IEEE Computer Society. Retrieved July 15, 2005, from www.hicss.hawaii.edu/HICSS_35/HICSSpapers/PDFdocuments/OSPEI01.pdf

Gauch S, Smith J (1993) An expert system for automatic query reformulation. Journal of the American Society for Information Science 44: 124–136

Gentner D, Stevens AL (eds) (1983) Mental models. Hillsdale, NJ: Erlbaum

Gleick J (2004, March 21) Get out of my namespace. New York Times Magazine: 44–49

Greene SL, Devlin SJ, Cannata PE, Gomez LM (1990) No IFs, ANDs, ORs: a study of database querying. International Journal of Man–Machine Studies 32: 303–326

Halttunen K (2003) Students' conceptions of information retrieval: implications for the design of learning environments. Library and Information Science Research 25: 307–332

Halttunen K, Jarvelin K (2005) Assessing learning outcomes in two information retrieval learning environments. Information Processing and Management 41: 949–972

Hansell S (2003, December 8) Foes of bush enlist google in group prank. New York Times, p C.8

Hendry DG (2006) Sketching with conceptual metaphors to explain computational processes. In: Proceedings of IEEE Symposium on Visual Languages/Human-Centric Computing 2006, September 4–7, 2006, Brighton, UK (pp 95–102). IEEE Computer Society Press

Hendry DG, Efthimiadis EN (2004) Students' mental models of information retrieval systems. In: Proceedings of the American Society for Information Science and Technology, ASIST'04, Providence, Rhode Island, November 13–18, 2004. 41: 580–581. http://dx.doi.org/10.1002/meet.1450410186

Hendry DG, Harper DJ (1997) An informal information-seeking environment. Journal of American Society of Information Science (Special Issue on Human–Computer Interaction) 48: 1036–1048

Hochman D (2004, March 14) In searching we trust. New York Times, p 9.1

Holland D, Quinn N (eds) (1987) Cultural models in language and thought. Cambridge University Press, New York

Ingwersen P (1996) Cognitive perspectives of information retrieval interaction: elements of cognitive IR theory. Journal of Documentation 52: 3–50

Jansen BJ (2005) Seeking and implementing automated assistance during the search process. Information Processing and Management 41: 909–928

Johnson-Laird PN (1983) Mental models: towards a cognitive science of language, inference, and consciousness. Harvard University Press, Cambridge, MA

Kempton W (1987) Two theories of home heat control. In: D. Holland and N. Quinn (eds) Cultural models in language and thought (pp 223–242) Cambridge University Press, New York

Koenemann J, Belkin NJ (1996) A case for interaction: a study of interactive information retrieval behavior and effectiveness. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Common Ground (pp 205–212). ACM Press

Lakoff G, Johnson M (1980) Metaphors we live by. The University of Chicago Press, Chicago

Landis J, Koch G (1977) The measurement of observer agreement for categorical data. Biometrics 33: 159–174

Lenhart A, Horrigan J, Fallows D (2004) Content creation online: Pew Internet and American Life Project. Retrieved 1 June 2004 from http://www.pewInternet.org/pdfs/PIP_Content_Creation_Report.pdf

Liddy E (2001) How a search engine works. Searcher 9(5). (Also Retrieved 1 June 2004 from http://www.infotoday.com/searcher/may01/liddy.htm)

Lombard MJ, Snyder-Duch J, Bracken CC (2002). Content analysis in mass communication: assessment and reporting of intercoder reliability. Human Communication Research 28: 587–604

Luke (2004, 1 June) Lucene Index Toolbox. Retrieved 1 June 2004 from http://www.getopt.org/luke/

McNichol T (2004, January 22) Your message here. New York Times, p G.1

Media Metrix (2004) Press Release, RESTON, Va. March 19, 2004. Retrieved 1 June 2004 from http://www.comscore.com/press/release.asp?press=443

Miller JD (1998) The measurement of civic scientific literacy. Public Understanding of Science 7: 203–223

Muramatsu J, Pratt W (2001) Transparent queries: investigating users' mental models of search engines. Paper presented at the Proceedings of the 24th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, New Orleans, LO

Noguchi Y (2004, February 9) Online search engines help lift cover of privacy. Washington Post, p A01

Norman DA (1983) Some observations on mental models. In: Gentner D, Stevens A (eds) Mental models (pp 7–14). Hillsdale, NJ: Lawrence Erlbaum Associates

Oakes MP, Taylor MJ (1998) Automated assistance in the formulation of search statements for bibliographic databases. Information Processing and Management 34: 645–668

Payne S (1991) A descriptive study of mental models. Behaviour & Information Technology 10: 3–21

Payne S (2003) Users' mental models: the very ideas. In: Carroll JM (ed) HCI models, theories, and frameworks (pp. 135–156). Morgan Kaufmann, New York

Roth W-M, Lee S. (2002). Scientific literacy as collective praxis. Public Understanding of Science 11: 33–56

Schemo J (2004, January 28) Online auctions, misspelling in ads often spells cash. New York Times, p A1

Simon H (1996) The Sciences of the artificial (3rd ed). MIT, Cambridge, MA

Slone D (2002) The influence of mental models and goals on search patterns during web interaction. Journal of the American Society for Information Science and Technology 53(13):1152–1169

Spoerri A (1993) InfoCrystal: a visual tool for information retrieval & management. Paper presented at the Proceedings of the Second International Conference on Information and Knowledge Management (pp 11–20). New York: ACM Press

Topi H, Lucas W (2005) Mix and match: combining terms and operators for successful web searches. Information Processing and Management 41: 801–817

Totty M, Mangalindan M (2003, February 26) As google becomes web's gatekeeper, sites fight to get in. Wall Street Journal, p A11.1

Wurman RS (2001) Information anxiety 2. Que, Indianapolis, IN

Young D, Shneiderman B (1993) A graphical filter/flow representation of boolean queries: a prototype implementation and evaluation. Journal of American Society for Information Science 44: 327–339