1 The Design of Search User Interfaces

1.1. Keeping the Interface Simple

The job of the search user interface is to aid users in the expression of their information needs, in the formulation of their queries, in the understanding of their search results, and in keeping track of the progress of their information seeking efforts.

However, the typical search interface today is of the form type-keywords-in-entry-form, view-results-in-a-vertical-list. A comparison of a search results page from Google in 2007 to that of Infoseek in 1997 shows that they are nearly identical (see Figure 1.1). Why is the standard interface so simple? Some important reasons for the relative simplicity and unchanging nature of the standard Web search interface are:

- Search is a means towards some other end, rather than a goal in itself. When a person is looking for information, they are usually engaged in some larger task, and do not want their flow of thought interrupted by an intrusive interface.
- Related to the first point, search is a mentally intensive task. When a person reads text, they are focused on that task; it is not possible to read and to think about something else at the same time. Thus, the fewer distractions while reading, the more usable the interface.
- Since nearly everyone who uses the Web uses search, the interface design must be understandable and appealing to a wide variety of users of all ages, cultures and backgrounds, applied to an enormous variety of information needs.

Designers of Web search interfaces have learned that in order to be able to successfully serve their highly diverse user base, they must be very

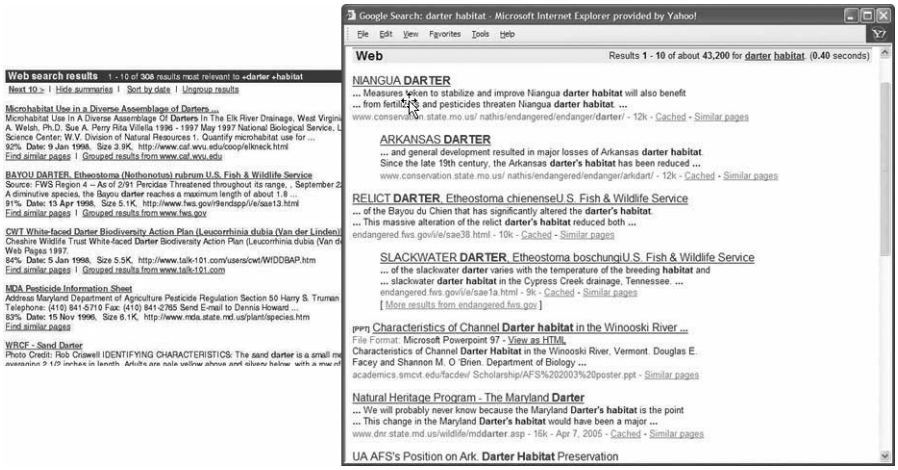

Figure 1.1. Search results listings from Infoseek in 1997 (left) and Google in 2007 (right). (Courtesy Jan Pedersen.)

careful about any complexity that they introduce. Almost any feature that a designer might think is intuitive and obvious is likely to be mystifying to a significant proportion of Web users.

To illustrate this point, despite the simplicity of the search results listings shown above, research suggests that even this spartan presentation is too complex for some people. A study of elderly users by Aula and Käki (2005) found that further simplifying the list of results reduced errors substantially. And research by Hargittai (2004) showed that some people do not understand even the very basics of keyword specification. Unlike most studies that involve university-educated participants exclusively, Hargittai obtained a random sample of 100 participants representative of the population of a county in New Jersey according to socio-economic factors. Hargittai (2004) found that, in addition to not really understanding keyword queries, many participants confused the address bar with the search entry form, and vice versa (the latter effect is common, as can be inferred from the fact that the most frequent queries for all search engines are `google` and `yahoo`). Some participants confused the syntax of the address bar with the syntax of query terms, placing spaces within URLs in the address form, as in `www.new york times.com` and `time warner.com`, or omitting all spaces from their keywords, resulting in queries like `presidentalcampaign2000`, `employmentopportunities`, and `fordescort`.

Another study by Muramatsu and Pratt (2001) with 14 participants found that most people had strong misconceptions about simple Boolean

operations. When comparing search engines that automatically applied AND versus OR to query terms, some assumed the ANDing search engine indexed a smaller collection; most had no explanation at all. When receiving empty results for the query `to be or not to be`, two thirds could not explain this phenomenon in a way that remotely resembled stopword removal. For term order variation in queries (for example, `boat fire` vs. `fire boat`), two thirds did not expect the results to differ.

Although today's standard search is a big improvement in usability over older command-line based Boolean systems, there is evidence that keyword querying is not initially intuitive. In fact, the literature suggests that people who are new to using search engines tend to start by asking a natural language question (Bilal 2000; Schacter et al. 1998). Novice searchers must *learn* to expect that a query will not yield immediately usable results, and that they must scan search results lists, navigate through Web sites and read through Web pages to try to find the information they seek. A study by Pollock and Hockley (1997) found that, for novice searchers, the notion of iterative searching was unfamiliar. Some study participants assumed that if their first attempt failed then either they were incapable of searching or the system did not contain information relevant to their interest.

Given the difficulty that some users experience in using relatively simple interface elements, it is perhaps not surprising that attempts to improve search via more complex interfaces have for the most part not been widely adopted. There are, however, some successful innovations in search interfaces which are becoming widely used; some of these are discussed in the design guidelines sections below. First though, a historical interlude explains the evolution of search interfaces over time. This is followed by a brief summary of how interface design is done in practice, and then a discussion of design guidelines for search user interfaces.

1.2. A Historical Shift in Search Interface Design

The story of search user interfaces is complicated by a radical shift that occurred after the Web became a worldwide phenomenon. Before the Web, computerized information retrieval was usually done only by members of a narrow demographic: highly educated users, such as paralegals, librarians and other search intermediaries, and journalists. These people searched over highly specialized, high-quality, information-oriented text collections such as bibliographic records for university libraries, legal

cases and opinions, and newswire articles. Often the providers of search access to these collections had monopolies on the content, and therefore did not feel the pressure of competition to provide improved interfaces for that content.

By contrast, the Internet is now accessed by 75% of the U.S. adult population, and 91% of those who use the Internet use Web search engines (Pew 2008b). The content of the Web differs from that of earlier systems in several important ways. Older systems usually did not allow search over full text; rather, the user could only search over titles and perhaps abstracts and other descriptive metadata. Search was usually used to find the name and location of a source containing this information, and then a physical paper copy would have to be obtained to see the full text. By contrast, most of what is available on the Web is the full text itself; the desired information is often immediately accessible.

The content available on the Web is vastly broader than that of older systems, and in addition to expository text, contains the equivalent of brochures and local newsletters, official information for companies and all kinds of organizations, information that can be used directly, such as guitar chords and knitting patterns, how-to information, hobbyist guides, and so on. The Web can be used to see the answers to questions, such as `what is the population of Madagascar`, directly. This was not usually possible in the older systems, which acted as gateways to more detailed information that was available only offline.

Older systems were developed before bitmapped (graphical) displays were commonplace, and so were based on command-line interfaces. These usually required complex combinations of operators – which had to be memorized – and Boolean syntax for query specification. Very few members of the lay public understand Boolean syntax and even fewer are willing to learn command languages. The lack of competitors with access to the content, plus an installed base of users who knew the old systems, probably slowed the adoption of modern user interface conventions. Another important difference between old and new search systems is that older retrieval systems often charged for use (in terms of number of queries issued, number of results returned, or amount of time used), whereas Web search has always been free of charge.

These contrasts – highly educated and trained users verses everyone as a user; high-quality, expensively edited expository text versus a huge variety and multiplicity of information types, search over document metadata (titles and abstracts) rather than over full text, TTY displays versus graphical displays, and expensive usage controlled by one provider versus free

usage provided by a multiplicity of search providers – help explain the differences seen in search user interfaces before and after the Web. These differences will be revisited throughout this book.

1.3. The Process of Search Interface Design

An important quality of a user interface (UI) is its *usability*, a term which refers to those properties of the interface that determine how easy it is to use. Shneiderman and Plaisant (2004) identify five components of usability, restated by Nielsen (2003b) as:

- *Learnability:* How easy is it for users to accomplish basic tasks the first time they encounter the interface?
- *Efficiency:* How quickly can users accomplish their tasks after they learn how to use the interface?
- *Memorability:* After a period of non-use, how long does it take users to reestablish proficiency?
- *Errors:* How many errors do users make, how severe are these errors, and how easy is it for users to recover from these errors?
- *Satisfaction:* How pleasant or satisfying is it to use the interface?

How are interfaces designed in order to attain the goals of usability? Despite the newly recognized importance of usability and user interface design, it is nonetheless surprisingly difficult to design highly usable interfaces. The field that encompasses interface design, as well as understanding how people interact with information and technology, is called *Human–Computer Interaction*, or HCI (Shneiderman and Plaisant 2004). Among many other activities, this field has led to the development of a design technique called *user-centered design* whose goal is to lead to the development of usable designs.

In user-centered design, decisions are made based on responses obtained from target users of the system. (This is in contrast with standard software practice in which the designers assume they know what users need, and so write the code first and assess it with users later.) In user-centered design, first a *needs assessment* is performed in which the designers investigate who the users are, what their goals are, and what tasks they have to complete in order to achieve those goals. The next stage is a *task analysis* in which the designers characterize which steps the users need to take to complete their tasks, decide which user goals they will attempt

to support, and then create scenarios which exemplify these tasks being executed by the target user population (Kuniavsky 2003; Mayhew 1999).

Once the target user goals and tasks have been determined, design is done in a design–evaluate–redesign cycle consisting of creating prototypes, obtaining reactions from potential users, and revising the designs based on those reactions. This sequence of activities often needs to be repeated several times before a satisfactory design emerges. Evaluation at this phase can often achieve useful results by testing with only a few participants, so the evaluation method used at this point in the design space is often referred to as "discount" usability testing (Nielsen 1989b). After a design is testing well in discount or informal studies, formal experiments comparing different designs and measuring for statistically significant differences can be conducted.

This iterative procedure is necessary because interface design is still more of a practice than a science. There are usually several good solutions within the interface design space, and the task of the designers is to navigate through the design space until reaching some "local optimum." The iterative process allows study participants to help the designers make decisions about which paths to explore in that space. Experienced designers often can begin the design near a good part of the solution space; less experienced designers need to do more exploration. Designing for an entirely novel interaction paradigm often requires more iteration and experimentation. Evaluation is part of every cycle of the user-centered design process. Because it is such an important topic, it receives a chapter of its own in this book (Chapter 2).

1.4. Design Guidelines for Search Interfaces

Researchers and practitioners in the field of Human–Computer Interaction have proposed dozens of sets of guidelines for successfully building user interfaces. Some authors have proposed guidelines for search interfaces specifically; an influential paper by Shneiderman et al. (1997) specifies eight design desiderata for search user interfaces generally (re-ordered below):

- Offer informative feedback.
- Support user control.
- Reduce short-term memory load.
- Provide shortcuts for skilled users.

- Reduce errors; offer simple error handling.
- Strive for consistency.
- Permit easy reversal of actions.
- Design for closure.

These guidelines provide good advice for search UI design. However, design guidelines can be difficult to follow, for a number of reasons. First, they are under-specified; they do not usually say *how* to achieve the guideline's goals. Second, meeting one guideline often conflicts with meeting another. For instance, in order to satisfy the consistency rule, if every results page must look identical, then an interface that shows query term suggestions in retrieval results must show a label stating "no feedback terms available" when it has no suggestions to make. This message would keep the interface consistent, but at the cost of distracting users with unnecessary information. Third, any list of guidelines is incomplete. For instance, the list above omits Nielsen's (1993) commonly stated guideline of "speak the user's language," which urges designers to adopt concepts and language familiar to users where possible. And finally, for any given interface, some guidelines will be superfluous.

Despite these drawbacks, the following sections elaborate in more detail about how some of these design guidelines should be applied to search interfaces. These guidelines and recommendations are informed by a study of the search interface literature, by cognitive considerations in search, and by a decade of experience designing such interfaces. The substance behind most of these is discussed in more detail in later chapters of this book.

It should be noted that these guidelines are specific to search interfaces; there are many other very important design guidelines for other aspects of interface design, and a number of excellent books to refer to for them (e.g., Cooper et al. 2007; Nielsen and Loranger 2006).

1.5. Offer Efficient and Informative Feedback

A bedrock principle of interface design is to provide the user with *feedback* about the status of the system and how that relates to the user's interactions with the system. A familiar example of interface feedback is the hourglass timer icon that is typically shown in a graphical operating system interface to indicate that the user has to wait while an application is launching or saving a large file.

Because the search task is so cognitively intensive, feedback about query formulation, about the reasons the particular results were retrieved, and about next steps to be taken is critically important. The subsections below describe important feedback indicators for search interfaces.

1.5.1. *Show Search Results Immediately*

Numerous studies show that an important search interface design principle is to show users some search results immediately after their initial query or navigation step (Hutchinson et al. 2006; Käki 2005a; Plaisant et al. 1997a). This information can be shown alongside other navigation aids, but at least a few initial results should be shown. This helps searchers understand if they are on the right track or not, and also provides them with suggestions of related words that they might use for query reformulation. Many experimental systems make the mistake of requiring the user to look at large amounts of helper information, such as query refinement suggestions or category labels, before viewing results directly. Information visualization interfaces that show documents as dots or icons in a two-dimensional space suffer from poor usability because the searcher cannot see the text of the titles and document surrogates (Granitzer et al. 2004; Hornbæk and Frøkjær 1999).

1.5.2. *Show Informative Document Surrogates; Highlight Query Terms*

Most search results listings today show a vertical list of results, each containing information about the document and why it was retrieved, such as the title, the URL, and a textual summary; this information is referred to as the *document surrogate*. The documents' *summaries* (also called *snippets*, *extracts*, and *abstracts*) are typically a few lines of text extracted from the retrieved documents.

An important form of feedback in search results listings is to include the terms from the query in the document surrogates in order to show how the retrieved document relates to the concepts expressed in the query. Early Web search interfaces showed the first few lines of the document in the summary, but today, summaries are designed to show the query terms in the context in which they occur in the document. Research shows that summaries are most informative if they contain the query terms shown

in their context from the document (Tombros and Sanderson 1998; White et al. 2003a).

Query term proximity information can be quite effective at improving precision of searches (Clarke et al. 1996; Hearst 1996; Tao and Zhai 2007). According to a large study by Clarke et al. (2007), when possible, all the query terms should appear in the search result surrogate, but if all of the query terms are present in the title for the hit, they need not appear in the summary, which can then include other useful relevance information. Clarke et al. (2007) also found that query terms appearing in the URL can be a useful cue, but that length and complexity of the displayed URL should be reduced where possible.

It has also been shown that visually highlighting query terms can be a useful feature for search interfaces (Aula 2004; Landauer et al. 1993; Lesk 1997; Marchionini 1995). Term highlighting refers to altering the appearance of portions of text in order to make them more visually salient, or "eye-catching." Highlighting can be done in boldface, reverse video, by displaying a colored background behind each occurrence of a query term, assigning a different color to each term. This helps draw the searcher's attention to the parts of the document most likely to be relevant to the query, and to show how closely the query terms appear to one another in the text. However, it is important not to highlight too many terms, as the positive effects of highlighting will be lost (Kickmeier and Albert 2003).

There is an inherent tradeoff between showing long, informative summaries and minimizing the screen space required by each search hit. There is also a tension between showing fragments of sentences that contain all or most of the query terms and showing coherent stretches of text containing only some of the query terms. Research is mixed about how and when chopped-off sentences are preferred and when they harm usability (Aula 2004; Rose et al. 2007). Research also shows that different results lengths are appropriate depending on the type of query and expected result type (Guan and Cutrell 2007; Kaisser et al. 2008; Lin et al. 2003), although varying the length of results has not been widely adopted in practice.

Figure 1.2 shows a screenshot from the BioText interface for searching over bioscience literature in which several kinds of document surrogate information are used (Hearst et al. 2007). Figures extracted from the articles are shown alongside each search hit, query terms are highlighted (in title) and boldfaced (in abstract and full-text excerpt), and the user can vary how much information is shown in the text excerpts by selecting or deselecting checkboxes for showing the abstract and full-text excerpts.

Figure 1.2. Search results in the BioText system (Hearst et al. 2007), in which rich document surrogate information is shown, including figures extracted from the articles, query term highlighting and boldfacing, and an option to expand or shorten extracted document summaries. From http://biosearch.berkeley.edu. (See color plate 3.)

The figure shows a case in which the second word in the query appears in the body of the article, but not in the title or abstract.

1.5.3. *Allow Sorting of Results by Various Criteria*

Another effective form of feedback in the display of search results allows for the dynamic sorting of search results according to different ranking criteria (e.g., recency, relevance, author, price, etc.). An effective interface for displaying results sortable along several dimensions at once uses a sortable columns format, as seen in email search interfaces, some product search, and some bibliographic search (see Figure 1.3). With this view, users can sorting results according to different criteria, while being able to visually compare those criteria, because the changes are directly visible (Cutrell et al. 2006a; Reiterer et al. 2000). This kind of view is typically more effective than showing choices hidden behind drop-down menus.