Constantine & Lockwood, Ltd.

Insight on In-Site Searching

Larry L. Constantine
University of Technology, Sydney
Constantine & Lockwood, Ltd.

In Search of an Answer

Searching is such a common and integral activity on the Web that it has come to occupy center stage. Indeed, it would seem that many things would be virtually impossible to accomplish on the vast and diversified Web were it not for the wide-angle access that generic, Web-wide search engines like AltaVista or Yahoo. On the other hand, searching within a site can actually create problems for users and lower the overall usability.

Feedback from a visitor to our Web site (forUse.com) prompted this note. The comment read, in part:

> I came to your site hoping I might be able to search for "extreme programming." Lo and behold, no search function on the home page! After awhile I stumbled upon it on the site index page, which I have to admit makes a certain amount of sense. On the other hand, if it was in on the home page as usual, I would have found it much more quickly. I steadfastly resisted going to the site index page because I didn't want to go browsing through a large list--I knew I wanted to search from the beginning. [Norman Azadian, used with permission.]

As some of you may have already suspected, there was indeed a reason why the search facility was not directly available on the home page. Of course, designers and developers always have their reasons, but having a reason, even a good reason, is no excuse for abusing your users. So what is the issue here? What are the usability issues in searching and how do you design good search facilities. Considering that we are in the business of improving product usability, it might seem perverse of us not to put a search box on the home page or at least to have a direct link to a search page.

In truth, as we said in reply to our visitor quoted above, the issue of site-based searching is actually fairly complex. Some users, especially more sophisticated, technically-oriented ones like our visitor, may insist on searching and, by force of habit, may start a session with a search even before skimming the home page contents. Searching, however, is a means to an end, not a goal in itself.

Finding Intentions

When Web designers and developers are asked about their site visitors' needs and intentions, many will mention searching. Ordinary users do not, however, head to a site thinking, "Oh, I really want to use a search box." Users undertake searches because they want to find something, to buy something, to learn what something costs, or to answer some other question. If the desired product or tidbit of information were displayed on the home page, they would have no reason to search. In fact, many users turn to searching within a site only when all other means fail. The search function is the Web-based equivalent of going to the Information Booth

The author may be contacted at Constantine & Lockwood, Ltd., 58 Kathleen Circle, Rowley, MA 01969; tel: 1 (978) 948 5012; fax: 1 (978) 948 5036; email: larry@foruse.com

in a store or at a park. If you can't find something on your own or with the help of the posted map or guide, you might go ask Information.

Of course, not everyone behaves like this. Some drivers would rather wander for hours than ask stop to ask directions at a service station. In contrast, the queries window near my office in the Computer Sciences building has no shortage of students who show up to ask questions that could have been more readily answered by a glance at an information sheet or a flip through the student guide. Similarly, there are probably those Web users whogo straight for the search box—if they can find it—without a pause for thought or forethought.

It Doesn't Work

The real problem with site search functions is that they do not actually work. By that I do not mean that the code is buggy or the algorithms are incorrect, but rather that, in combination with users, they do not, on average, enable the user to accomplish their intentions in an efficient manner.

In the classes we teach, we often hear developers say that a search is good because it allows users to go directly to whatever they want rather than having to follow links, but this efficiency is a programmer's fantasy seldom obtained in practice. In reality, the odds are incredibly low that the first search attempt will yield a results list with the desired page at or near the top. More likely, the page will be buried far down in the list or will not show up at all until the user's second or third attempt, or worse, will be effectively unreachable.

Indeed, the research on searching is both clear and consistent. If a visitor uses a site-based search engine, their chances of finding what they are seeking, even given that it is on the site, are drastically reduced. Jared Spool has found that using the search box can cut a visitor's chance of success in half. In other words, if ,instead of searching, visitors stay with browsing and following links, they are twice as likely to find what they seek. The implications of this for the design of e-business sites are enormous.

Causes of Search Failure

There are many different ways that within-site searches can fail. To give one example, when we learned from our publisher that our book, ***Software for Use***, had been nominated for a Jolt Product Excellence Award as the best book of 1999 (It won, by the way!), we went straight to <http://www.sdmagazine.com/> to confirm the news and to see who else had been nominated. Searching for "Jolt" or "Jolt award" returned stacks of links, none of them about the current nominations. As it turned out, there was a page about this year's nominees, but it could only be reached via a graphic link that was buried down "below the fold" on the home page. For unknown reasons, however, the target page was unreachable through the site search engine and was also invisible to external search engines as well.

Another problem with searching is that frequently the user's terms do not match the ones by which the site can be searched. A consumer going to http://www.llbean.com/ and searching for "children's sweaters" would get a message stating:

> No matches were found for "children sweaters" in the Shop section.

Visitors who are unsure whether L. L. Bean sells children's sweaters or who are not willing to persist doggedly in the pursuit are likely to turn to an alternative eTailer. What the reasonable but naïve would-be customer does not know is that the search only recognizes the sought-after item as "kid's sweaters."

Boolean Illogic

Although as designers we may often seem to forget it, we all know that people do not think like computers. What seems perfectly clear and logical to human beings can be contradictory nonsense to a computer, and what makes sense in software terms is often unfathomable to ordinary folk. The vast majority of everyday site visitors have only limited abilities when it comes to forming effective search queries. Formulating search criteria is a subtle and unnatural art that often requires people to use language in highly specialized ways that go against the grain of how they think. For example, most people would find it reasonable to enter "Vermont and New Hampshire" in looking for retail outlets in Northern New England, but, of course, their query turns up nothing because the software "knows" with perversely logical certainty that no retail outlet is simultaneously in both states.

Because of difficulties in formulating well-formed queries, most users will inadvertently either limit searches too narrowly or frame them too broadly. The result is either that the search returns nothing or returns so many hits that users cannot find what they are looking for anyway, even if it is actually present among the 17,426 matches being offered. Formulating a good search strategy, a series of successive limitations and expansions that not only yield at least one usable answer but also give access to the full range of relevant results, is an art that most people who are not professional researchers are unlikely ever to master.

In the absence of advanced skills that cannot be assumed to be widely distributed in the general population, searching is risky for both the user and the provider. If a query comes up empty, the visitor is likely to conclude that what is sought is not available and may leave. If the query returns too much, visitors are likely to be overwhelmed. Even if they eventually find what they are after, accumulated frustration and irritation make them more likely to abandon the site. If more than a few search attempts are needed, the visitor's quick departure into cyberspace is all but assured.

Effective searching requires users not only to learn an artificial way of thinking about their interests but also to learn a sort of programming language with a special syntax and semantics in which, for example, uppercase means uppercase but lowercase means either lowercase or upper case and in which common words and punctuation marks take on special meanings and have uncommon uses. So much for "easy" and "intuitive."

Natural-language queries based on so-called artificial intelligence are not the answer. The problem with these techniques is that they are more artificial than intelligent. For the most part, natural-language query engines do not actually understand the query entered but only pick up on key words or match phrases out of context. The resulting interpretations are often ludicrous and the search results so irrelevant that many more experienced users prefer simple keyword techniques, however ineffective these may be.

Advancing the Cause

Separating out "search" and "advanced search" capabilities is a common technique intended to shield the average user from the bewildering realities of sophisticated searching, but this, too, does not work particularly well in practice. If a high percentage of queries need the "advanced" facilities, then most users will be punished by an extra click and page load. Moreover, the very term "advanced" itself puts a bias on use, tending to scare off some users while teasing others. If the "simple search" query box actually accepts the same queries, then advanced users have been misled and beginners risk having their queries misinterpreted.

The Eye is Faster

Search boxes, as well as lists of links, drop-downs, and site maps, are just alternative solutions to a common set of problems. It is the designer's job to organize for the site's users the best solution to these problems. Within-site searching is one example of the conundrum of whether designers should give users what they want or what they need. We have long come down squarely on the side of meeting user needs. Your users may say they want a search box, but if you meet their real needs simply and efficiently, the majority will be pleased with or without it.

In the case of the visitor quoted earlier, the real need was not to search but to find what we had to offer on extreme programming, which was, in truth, nothing. Thus, no matter what means he might have used to scour our site, he would have come up empty-handed. What he needed in this case was to find out that there was nothing to be had. The kindest thing to do would be to make it as painless as possible to reach the conclusion that our site was devoid of material on extreme programming but full of interesting bits on usability, use cases, and user interface design.

For meeting this objective, indexes and site maps have substantial advantages over searches: They are far better at communicating to the user what is on the site as well as what is not on the site. An empty text box next to a "submit" button says absolutely nothing about what is to be found on a site. An empty search result is ambiguous, conveying either that we do not have what you are looking for or that you did not ask for it in the right way.

One of the fastest, most flexible, and effective searching mechanisms known is visual scanning. The eye-brain combination is a powerful search tool, particularly for aligned vertical lists of items. The greatest advantage of visual scanning is that it works even when users do not know exactly what they are seeking. Life is full of queries of this sort. The customer says, "I don't know, it was one of those Italian brands. I'll know it when I see it." It turns out to be Gato, a Spanish brand. A visual scan of a list of brands picks it out in an instant, while a search for "Italian" would come up without it. No search algorithm of today can cope with this sort of query, but a quick visual scan of an index may often do the job handily. In fact, an index more readily supports browsing for related or unanticipated material.

Both/And

The root of the problem is that searching may not be in the best interests of users, even when they want to do it. From a usage-centered perspective, the aim of the designer should be to offer searching to the user but to make it subtly easier to use other means first. That is what we had in mind when we put the search feature on the site index page. However, our visitor pointed out some problems with this solution.

> We can hop into a rental car and drive off without delay because we know approximately where most things are. If you now suddenly go re-locating something as basic as the search function to a "better" place, then you shouldn't be surprised when the uninitiated are not pleased. In fact, I would argue that the search function is the ONE thing that MUST be on the home page. A possible compromise would be to re-label your site map page (link) to the "finder" or "locater" page or some such name which would at least hint at the possibility of finding a search function there.

That, as recent visitors to our site may have noticed, is precisely what we did, calling the link "Find It!" with a pop-up <title> tag that reads, "Site Content Index and Search." Of course, all design is a matter of tradeoffs; there is no one right answer. User interface designers are always balancing competing factors.

Structured Searching

In designing a search facility, the aim should not only be for ease and simplicity of use but also to maximize the chance that the user will find what is sought while minimizing potential information overload from "result glut." The most effective search facilities are designed to fit with the structure of the information being searched and of the tasks of interest to the users. In other words, they are tailored to particular applications. The worst results are achieved with generic search engines and simple-minded designs. (We apologize for the standard Microsoft FrontPage search facility that still graces our Web site; it will be replaced.)

To fit the tasks, you must understand what kinds of searching users are most likely to attempt, then make those the easiest. For example, you can allow users to qualify or constrain searches through check boxes or drop-downs, with the most likely or safest selection already defaulted. For example, a book club site might include checkboxes of this ilk:

Search in: [X] title [X] keywords [ ] text

As the lesson of llbean.com highlights, indexes and search facilities should incorporate synonyms and equivalents. For large sites, you should establish an internal glossary of terms and a thesaurus that maps equivalents.

Extreme Programming Revisited

As it turns out, at least one of our papers is on modeling shortcuts for crunch-mode projects <http://foruse.com/ApplicationNotes/AppNote4.htm> and could be construed as related to extreme programming, although we would never have thought of it in those terms. You would find it indexed under "crunch-mode modeling," "modeling shortcuts," and "shortcuts in modeling," which illustrates another important point in structuring information access: rotate your terms so phrases can be found in multiple ways. (If only the writers of software help files would take this to heart!)

Now, of course, our earnest visitor would actually find something because a search on "extreme programming" will lead to this paper, which refers to the term numerous times but really has nothing to do with it. Unfortunately, search functions do not understand this sort of thing and will often point you to things for the wrong reasons. Such is the nature of the beast and yet another reason to keep the beast under control.