Sticky side of the Web

By: Robert Sanders

No Abstract

Sanders, Robert. (1997), Sticky side of the web. *MultiMedia Schools*,[4(4), 40-43].
(ISSN: 1075-0479) (The predecessor title to Internet @Schools Magazine)

Sanders, R. L. (1997), Sticky side of the web. ***MultiMedia Schools***, September/October 4(4), 40-43. (ISSN: 1075-0479) (The predecessor title to *Internet @Schools* Magazine) Archived in NCDOCKS with permission of the publisher, Information Today, Inc.

Sticky side of the Web

Robert Sanders

Sticky problems searching the Web?

Allow me to start by stating that I use the World Wide Web every day. I have answered hundreds of reference questions, which might have gone unanswered prior to having access to the Web. As a high school media specialist, I use the Web to supplement what I have in my print collection. While we have a large and balanced collection, there are naturally certain types of information that I cannot keep up-to-date. For example, in the last week, I have used the Web to find out current weather conditions in Chicago, to determine the currency exchange rate from U.S. dollars to Mexican pesos, to find home prices on our local Multiple Listing Service database, and to look at options and pricing on several car models. Few of these reference questions could have been answered with our existing print collection.

I have witnessed students' interest in finding information go from being virtually nonexistent to seemingly insatiable when using the Web. The instantaneous accessibility of such a vast amount of information is unprecedented. Never before have we had the ability to "publish" ideas and have them accessible to such a wide audience. Several of my students put out monthly electronic newsletters complete with editorials, feature stories, jokes, and announcements.

That being said, it is the premise of this article that the Web is far from the perfect solution to finding current and reliable information as some would want you to believe. In fact, I would argue that the Web, in its current state, suffers from a surfeit of unreliable resources, poor searching mechanisms, and lack of organization.

QUALITY VS. QUANTITY

The Internet is not the magical source of information that the popular media makes it out to be. I have far too many students and even staff members who want to get on the Net to find the information they need, without questioning the validity of the source. Two years ago, the Internet consisted of approximately 6.6 million hosts. Most of the first Internet sites were designed and maintained by university professors and college students. At that time, it was somewhat manageable, and, arguably, more credible. Since then, it has nearly tripled to over 16 million hosts. The number of domains in that same amount of time multiplied from 120,000 to 828,000 (Network Wizards, 1997). Now the sites available consist of interconnected pages, ranging from these initial academic sites to advertisements to cyberporn. There is far more "information" today through which you must sift to find what you need.

The recent Sprint television ad promoting their new Internet access shows a dozen or so "computer-geek" types listening to an apocalyptic sermon from one of their own about how Sprint, by offering this new service, is ending their cyberworld as they know it. One of the young men in the audience asks, "What's to keep my Mother from getting on the Net?" The answer they shudder at hearing is, "Nothing at all." This commercial illustrates the fact that all kinds of people are now on the Web, and, in many cases, are creating their own Web pages. Some of these pages created by experts in the field may contain valuable information. Other pages may have been designed by persons less knowledgeable about the topic, and, in some cases, even by people with the intention of publishing misinformation. Sometimes, it is difficult to distinguish one from the other.

In the print publishing industry, a publisher acts as the gatekeeper between the writer and the reader. Part of the publisher's responsibility is to ensure that what is published is a quality document. This is not to say that there isn't questionable information published in book form, but a reputable publisher weeds out, if not the good from the bad, at least the mediocre from the truly useless. Not so in the online world.

In the world of Web publishing, anyone can publish anything at anytime without having to go through anyone else. Therefore, the truly useless can and often does get published. The problem with this situation is not so much access to what might be considered offensive material, but that individuals must be even more selective in choosing resources from the Web than when using print resources. Those who are not trained to question the source and reliability of information sources may have difficulties sifting through massive materials.

SEARCH ENGINES

Even when quality documents are available, people frequently have a difficult time finding them with the current array of search engines. Anyone who's taken a course in online searching knows about precision and recall--these are two concepts that usually work in reverse. The focus of precision is on the relevancy of material found. In other words, if a student did a search on a topic and found 50 documents, she would have high precision if most of these documents had to do with the topic. (There might have been 500 related articles that she did not find, but what she did find was right on target.) The focus of recall is the amount of information retrieved. (If another student did a search on the same topic and found almost all of the 500 documents related to this topic, but only 50 of which were right on target, he would have high recall.)

Web searchers will soon discover that while the search engines can boast high recall, these engines cannot even come remotely close to precision. It is not uncommon to do a search and retrieve over 100,000 documents, very few of which have anything to do with the topic being searched.

ORGANIZATION . . . OR LACK THEREOF

But, even if the search engines were improved, the way the Web is organized--or not organized--this improvement would have limited effectiveness. Web pages are "registered" with the search engine's database. Yahoo!, for example, has millions of pages in its database, as do Lycos, AltaVista, and Excite, just to name a few. So, when you search "the Web" with one of these search engines, you are only searching the pages registered with that engine, not the entire Web. Search results are limited to the scope of those search engines. Given this, the recall may be low.

For a more comprehensive search and thus higher recall, one should use several search engines. There are "Meta" search sites, which use several search engines to search multiple databases simultaneously. However, since each engines works differently (with regard to truncation and Boolean logic and nesting), it is doubtful that the individual engines being fed the query by the meta-engine translate the query in the same way. The results of the search might indicate higher recall, but the lack of uniformity between individual search engines would continue to lower precision. As a regular Web user, I tend to choose what I see as the lesser of two evils by using multiple Internet search engines. As media specialists, we need to educate students in the use of several Internet search engines.

To add to these problems, sites appear and disappear daily. It's difficult, with all the links from one site to another, to keep the whole Web connected seamlessly. It is nearly impossible for the search engines to track what's available at any given moment. In a traditional library setting, we are used to finding out that books we need are checked out. While we accept this occasional inconvenience, we would not accept the idea of these books changing in their content from day to day. On the Web, information that may be found one day may not be there the next. The page where it was found may disappear or the information on the page may change. While the addition of information or the updating of information is a good thing, this ephemeral quality to the information on the Web is disturbing to those of us who find it necessary to cite sources.

PERSONALIZED INFORMATION SERVICES

Despite all these problems, the Web is an amazing thing and is unprecedented in its delivery of text, sound, and video to millions of people worldwide. And while some may argue that television includes text, sound, and video, television is lacking in the Web's most unique and important component--interactivity. Never before have people been able to control, manipulate, customize, and create the information they receive. With the Web, we have entered an era of truly personalized information services.

But, in order for the Web to be a valuable and reliable information resource, several things must change:

1. The Web must become more organized. In its current state, the Web is a librarian's worst nightmare. It seems as if all the rules of cataloging and classification have been thrown to the wayside. I'm not advocating a centralized, bureaucratic organization to whip the Web into shape by indexing and cross-referencing every site available. However, there are currently search engines like Yahoo! that have developed ways to categorize and organize sites by doing so at the time of registration. These pioneers have demonstrated that there are ways in which organization on the Web can be improved.

2. More quality (refereed) information must be available. Recently, several companies (e.g., EBSCO, UMI, SIRS, and Grolier) have provided the option of Web access to their databases that were once available on CD-ROM, and, prior to that, paper. While these companies do charge access fees for their databases, the availability of large databases of magazine, newspaper, and encyclopedia articles will definitely be a welcome supplement to what's currently available for free on the Internet.

3. Users need to understand the nature of the Web. Our users--teachers and students--need to know the Web's good and not-so-good qualities . . . or at least recognize those qualities that are different from information as we know it in books and libraries. Second, just as people have to know how to drive a car before they get behind the wheel and actually do it, people need to know how to search (as opposed to surf) the Web and evaluate sources found. While technology will improve, I do believe that we will always have to do a little work to pose our queries, wade through the results, and determine for ourselves what resources will satisfy our needs. This is not a world of instant answers in which the Web will provide us with exactly what we're looking for merely by us typing in or speaking a few key words.

ROAD TO TECHNOPOLY?

When I think of the Web, I can't help reflecting on Neil Postman's recent book, Technopoly. Postman describes the world as a place in which we create technology for the sake of creating it, and we call it good. In fact, we do more than call it good; we worship it, believing technology to be divine. I often equate the recent advances in Internet technology with the explosion of information. I see the Internet as the epitome of the technology in which we live, wherein we create more and more information for the sake of creating information and call it good. There is much that we have invented that can be called good, and there will be much more. The Web has much that is good, and it truly has the potential of changing the way we live in the future.

My hope is that we can create a society that not only values access to information, but, more importantly, values organization, quality, and the usefulness of information. The question that needs to be addressed is, where can a person most easily access an organized collection of information in all formats--both print and electronic--and find quality information with precision? Currently, the answer to that question is still . . . the library.

REFERENCES

Network Wizards. Internet Domain Survey, January 1997 (Online). Available http://www.nw.com/, May 30, 1997.

Postman, Neil. Technopoly: The Surrender of Culture to Technology. New York: Knopf, 1992.

Communications to the author should be addressed to Robert Sanders, Library/Media Specialist, Princeton High School 11080 Chester Road Cincinnati, OH 45246, 513/552-8267, rsanders@phs.princeton. k12.oh.us