Sustainable building on the WWW

N. Maiellaro, A. Abrescia

National Research Council, Institute for Housing and Social Infrastructures, strada Crocifisso 2/B, 70125 Bari, Italy
Email: maiellaro@ba.cnr.it

Abstract

The Symposium on Engineering and Building for Sustainable Development clearly emphasized the hazards of relying too much upon the governments’ actual ability to solve or keep in check environmental problems, and the dual responsibility of designers, who are now playing a basic role either by dealing with these problems personally or by collaborating with governments in the formulation of concrete action to promote sustainable development.

In order to single out the design decisions that best contribute to achieve sustainability, the creation of a database in which to store all technical, economic and cultural solutions adopted by designers in different contexts could serve many useful purposes. In this framework the Internet plays a significant role, both as an information provider and as a huge platform for electronic publishing.

The paper first examines the problem related to finding information on the WEB; then presents ‘Sustainable Building Resource’, a site on the Internet devoted to give information in a focused way about sustainability in building.

1 The Internet as information provider

New and unforeseen technologies are changing the way we think about locating, using, and managing information. Regardless of such changes, the Internet is a rich, complex, and often-confusing environment filled with facts and fallacies, engrossing stories and tall tales, vibrant multimedia experiences. The quality and the quantity pose problems.

The WEB has a lot to offer, but not all sources are equally valuable or reliable: it is important to remember that anyone can publish on the Internet. This means that the quality of the information you find on the Web must be evaluated very
carefully. Some rules helping to validate have been suggested, e.g. Westera [7], Grassian [3].

Finding the Internet documents (Web "pages" or "sites") can be time consuming and frustrating for a non-expert. This is not surprising, given the amount of information available on the Internet (over 50 million documents) and the different ways it is stored and retrieved (the documents are not indexed in any standard vocabulary).

One way to find or discover information on the Internet is by a mailing list, a discussion group of people debating relevant topics through e-mail (first subscribing and then receiving mail from the group), or a Usenet newsgroup, accessed and viewed on the Internet. Discussion Groups can be found at http://www.listz.com and http://www.tile.net; http://www.dejanews.com is useful to find Usenets. However, the most commonly known and, for many, the default starting point to find information online, is a search tool.

1.1 Search tools

Search tools are computer programs for obtaining information; they search the contents of their database, not the World Wide Web (WWW) directly. A search tool finds information for its database by accepting listings sent in by authors wanting exposure, and by following the links on these pages to find more links to add to the database. Some of the search engines index only the title and the URL for the submitted Web page, some do full-text indexing, some retrieve the most important concepts of a Web page. Since none of these databases includes all the WWW pages in existence, there are usually different results from different search tools.

As with any other craft, productive research takes both planning and a knowledge of the right tools to use. Basically, the success of a search depends on the ability:

• to create exact matches between terms the user is searching for and terms used in the documents to find (methods to analyze a topic before beginning a search have been suggested, e.g. Barker [1]);

• to select a search tool's features for searching its database contents successfully.

The universe of the Internet search tools falls into two large categories: Subject Directories and Keyword Search Tools.

1.1.1 Subject trees - directories

A subject tree is a list of Web pages or URLs created by humans. A subject tree is useful if the items are classified in some sort of hierarchical scheme.

They are valuable for their smaller size of hand-picked, usually "good" sites; construction and maintenance of these tools is labor-intensive and has not been automated to a high degree, so they are often not updated. In spite of this, beginning with a directory for topics with vast quantities of information on the Internet often helps to sort out meritorious sites from those that may mention your topic without in-depth treatment.
Browsing large subject trees is not very efficient. The documents have to be located by trying to match the desired topic at the top layer of a subject hierarchy, then by choosing narrower sub-subject-categories in the hierarchy that seem to lead to the prefixed target. Browsing encounters the difficulty of guessing under which subject category the topic is classified. Also, the taxonomy in each subject directory differs, making browsing inconsistent from one search tool to another.

In one subject tree (http://arioch.gsfc.nasa.gov/wwwvl/engineering.html), the category "engineering" may contain documents on architectural engineering, civil engineering, and so on. In another (http://dir.yahoo.com/Science/Engineering), "engineering" may include civil engineering, environmental engineering, but not the term architectural engineering.

Subject trees more and more often offer a keyword search page.

1.1.2 Keyword search tool

Keyword searching means entering one or more keywords separated by spaces in the first white search box you encounter in any search tool. In small databases of Web documents and in subject directories, however, "simple" keyword searching (using the system's defaults), is usually the best approach: while the small size of the databases makes more complex searching unnecessary, yet it may not include documents you need. In larger databases, "simple" keyword searching usually retrieves irrelevant or too many documents; better results can be obtained using more advanced techniques. Most search engines have developed their systems of search independently. Therefore, there are no standards in nomenclature, database organisation or retrieval systems. Thus for each search engine, the best results are performed by composing the query for that particular engine[6], and it is beneficial to use different engines for different tasks (See 'Search Engine Test Results' in [5]). Using the various search tools is enhanced by knowing how they were actually designed, and especially by knowing their specific rules [4] - most engine sites include detailed documentation on how to improve searches (Search Tips, Custom Search, Advanced Searching, etc.). Best results in a search tool are performed by using synonyms, variants and related terms; entering multiple or alternate spelling when appropriate; using phrases (the search term has to be enclosed in double quotation mark); connecting two or more words with operators (like conjunction, rejection, and so on); substituting word endings with wildcard expander - the asterisk can be used in most search engines [2]. The most suitable way is related not only to the selected search tool but also to its current version: due to fierce competition between the various web search products, the use of the search tool often changes to benefit the users of the product or service. In ordinary search tools (such as Infoseek, Alta Vista, Hotbot or Excite), a different list of URLs from each search engine is obtained by using the same keyword; results from submitting very comparable searches can differ widely, but also contain some of the same sites (Figure 1-2). Some services have been put together web pages that contain a collection of search engines. More useful are the meta-search services.
Ecosystems and Sustainable Development

Figure 1: Infoseek (search engine) - Results for "sustainable building resource"

Figure 2: AltaVista (search engine) - Results for "sustainable building resource"
1.1.3 Meta-search services

Meta-search services allow querying multiple Web search engines simultaneously from a single interface. Several search engines are accessed, results are obtained and compiled. Although they do not usually return the comprehensive results produced by an in-depth search of a single search engine, they are fast and convenient. Attention must be paid to complex query, because the syntax is quite different for various search engines.

Metacrawler (http://www.metacrawler.com) utilizes AltaVista, Excite, Infoseek, LookSmart, Lycos, The Mining Co., Thunderstone, Webcrawler, Yahoo!. This meta-search service consolidates results in one large list, ranked by a "vote" score derived from the rankings of each of the search engines listing each site. It normalizes the confidence scores used by each of the services, sums the scores, and ranks them 1 to 1000.

Metafind (http://www.metafind.com) searches AltaVista, Excite, Infoseek, PlanetSearch and Webcrawler; the results are sorted into clusters for each keyword term used in the search. It is possible also to sort in alphabetical order or by domain.

Inference Find (http://www.infind.com) queries, for a user prefixed time, WebCrawler, Yahoo, Lycos, Alta Vista, InfoSeek, and Excite, merges the results, removes redundancies, and then groups the related items together in clusters (Figure 3).

A meta-search tool can save the user the trouble of visiting every search site itself, but there is the problem of sorting through too many (or not enough) hits. The solution is a search agent that enhances, augments, and helps to keep track of the searches.

1.2 Search agents

Search agents are an emerging technology that is making computer systems easier to use, by allowing people to delegate work back to computer. They help do things like find and filter information, customize views of information, and automate work. Unlike the search tools, the agents are software applications to buy, install, and run from the desktop. It is possible to enter the search query offline, and then the agent runs it through anywhere from half a dozen to 30 or more individual search engines at a time.

Instead of visiting each search site in turn, the agent queries the resources, collects the responses and presents them on a single page (a list of agents can be found at: www.searchinsider.com/agents.html).

WebCompass, the agent used in our research, simultaneously searches multiple resources on the Internet, automatically generates summaries and keywords for the documents it finds that match the search topic (Figure 4); organizes these documents into a personal hypertext index that is updated at any interval, and outputs the index as HTML files so it is possible to share the results.

The information is stored in an Access file, organized in tables (concept_master, concept_relations, concept_terms, concept_urls, search_resources, url_master).
Figure 3: Inference (meta-search) - "sustainable building resource" results

Figure 4: WebCompass (search agent) - "sustainable building resource" results
2. Sustainable Building Resource

Understanding the specifics of sustainable building and determining effective sustainable practices can be confusing. Local governments and private industry often do not have resources to perform the necessary research to assemble information on sustainable practices.

Green building initiatives, as well as sustainable development activities, offer many opportunities to local governments and communities. In this frame, Sustainable Building Resource (http://www.iris.ba.cnr.it/sustain/welcome.htm) is intended to meet the need for a comprehensive overview of sustainable building. This Web-site (Figure 5) is devoted to share information on events, researches, projects, links about sustainable building, a first step to helping architects, organizations, researchers, students, understand the wisdom of adopting sustainable principles in building.

This project aims to:

- establish and maintain a network and clearinghouse where one can find all available information on sustainable building;
- compile and maintain a network of persons who have interests in all forms of sustainable building;
- provide training, seminars, and workshops which support the acceptance of sustainable building.

Sustainable Building Resource (SBR) directs users to materials published or hosted by a large number of organisations on the Web. SBR's intention is to make good use of such materials, and to help promote easier access to materials which publishers have already decided to make available in the WWW. At present the SBR datasets includes:

- a selection of web-sites related to sustainable building ("Catalog");
- our own description of pages on the Web ("On the web reviews");
- a selection of documents related to sustainable building ("Off the web reviews");
- shortcuts to web-sites grouped and classified on the bases of their different approaches to sustainable building;
- data collected from the Web using an internet spider.

"Catalog" is the database of the web-sites related to sustainable building written by SBR; it contains URL, Title, Country and a link to the related review. "On the web reviews" are pages written for the web (using html code) by SBR, though often using the document's own words; the reviews are in the same language as the examined document. The reviews do not necessarily reflect the views of the organisation itself on the relative importance of its different products or activities.

The reviews have been catalogued in three categories (town planning, construction, materials) and eight groups (applications, codes, events, organizations, publications, resources, services, tools) to benefit the user. A small portion of each document listed in "Catalog" is stored in SBR to allow searching of this material; SBR takes every step to ensure that the user of the directory cannot be mistaken about the origins of this material.
Unlike other sites, there are a lot of navigation tools (see the function "Map"):

- insert a user-defined term or phrase ("Search")
- choose from a list of pre-selected terms ("Keyword")
- select the content to narrow the search ("Category", "Group" and "Topic")
- a query to the database of the web-sites ("Catalog")
- a shortcut to a specific key-issue ("Shortcut")

### 3. Basic functions of the Web site

The web server for the management of the site has been selected given the possibilities of connection both to DBMS (Data Base Management System) systems, or to external applications, using CGI (Common Gateway Interface) and dynamic libraries (dll). The server functions have been extended using cgi script (a cgi script is a file “outside” practicable to the web server: to any call of a script, the server makes a new process in a separated memory space).

In addition to html standard pages and pages with JavaScript code, ASP (Active Server Pages) pages have been implemented in order to realize “dynamics” web pages. In this way different pages are generated, beginning from only one document, on the basis of the query and its results.

A type of query, limited to the information available in the index-linked folder, takes place on a specify page (Figure 6) that contains:

- a textbox where the phrase or word to be found has to be inserted;
- a check-box to activate a free-text search;
- some fields to request restrictions to the size or to the date of the documents.

The free-text search makes it possible to extend the search of the introduced word even to a specific code in the html page (the content of the "META NAME" used, like: "Ms.Category", "Author", and "Keywords").

The representation of the results differs due to the several search’s features, or in relation to the type of page used (html o asp):

- the html page recalls an idq (internet data query) file in which they define the query’s parameters; the size to use for the pagination of the results is included in file of .htx size;
- the ASP Page dynamically visualizes the results processing the portion of page predisposed for the print of the results.

The information is contained in relationship database. Some asp pages make the search on database tables, using the vbscript language and ActiveX objects; in particular access to remote information managed by SQL Server has been realized by means of ADO (Active Data Object) objects.

The search’s page "On the web resources", for example, makes a search in all its fields: Item, Title, URL, Country, Category and Group. The search using these fields is made in AND, and also it’s possible to order the results on the basis of some fields (Item, Title, URL, or Country).

The htm pages provided with the tables and frames was realized with Symantec Visual Page editor; those of asp type and htx was realized with Microsoft FrontPage. For the idq file’s creation a normal editor of texts such as the WordPad was used, while the modification of the research’s page (query.asp) was executed directly by browser.
Are you searching for keywords as "sustainable building", "sustainable building design", "sustainable architecture", "sustainable city", "green building", "green materials", "ecological design", "eco village", "appropriate technology", "natural building", "alternative building materials", "energy efficient building design", "recycled building material", "renewable material"? Sustainable Building Resource is your site! We have pre-selected a number of keywords which we know from experience will provide a broad search. Shortcuts and navigation tools (see map) are available.

Sustainable Building Resource directs users to materials published or hosted by a large number of organisations on the Web and off the Web. The contents are: URL's list, data collected using an internet spider, database on the use of recycled materials.

Please help make this site a valuable resource by sending relevant links that you know of and/or helping to publicize it through a link or a banner on your websites. Reproduction in any form or medium is allowed quoting the source. Best view at 800x600 resolution 256 colours.

**Figure 5: Sustainable Building Resource - Home page**

**Figure 6: Sustainable Building Resource - Search page.**
4 Conclusions

Search tools are wonderful but far from complete. Certainly search tools can help to find something more quickly on the WEB, but sometimes people have to browse through many bad hits before finding useful information. This is a major deficiency for technicians, because designers are not interested in exploring the Internet per se but rather in obtaining information to solve problems or accomplish tasks they are dealing with. As a step towards correcting this deficiency, our WEB site 'Sustainable Building Resource', through the agent technology, checks the whole Internet for new information about sustainability in building on a regular basis, provides focused information, and promotes the creation of a 'Sustainable Building Ring', a network to link together all the institutions related to the matter: site to site links may bring in much more meaningful traffic than indexing by all the general search engines in the world.

5 Acknowledgements

This paper refers to a project supported by National Research Council of Italy. N. Maiellaro (manager), is the author of all the chapters, except "Basic functions of the Web site" whose author is A. Abrescia (software editor).

6 References


