

Understanding Internet Search Engines
May 2006
Bahá'í Internet Agency

Introduction

As more Bahá'í communities create web sites, a question often asked is how a site can secure a prominent listing on Google or one of the other Internet search engines. This document will provide an overview of Internet search engines, how they work and what your webmaster can do to improve your site's relative ranking.

What is a search engine?

A search engine allows users to search among the millions of public web sites based on specific key words or other criteria. For a search engine to suggest relevant web sites to the user, it has to scan vast numbers of public web sites, index the words it finds, and then determine which sites are most closely related to possible search words or phrases. While there are many search engines, the most important search engines are: Google (google.com), Yahoo (yahoo.com), and MSN (search.msn.com).

How do search engines work?

There are two basic models for building a search engine. One is based on a computer program that weights the importance of a web site based on a long list of characteristics. The second method is to create a directory built by human content experts. Google is an excellent example of the first method and Yahoo's directory is an example of the second system. Most search sites offer some form of both systems.

The human-based systems usually require that the owners of the web site wanting to be listed submit their site for consideration. It is important to make sure that the site be submitted to the directory using relevant category terms. So, for example, if you have a site or a page that relates to Bahá'í history, you would want to be listed under "religion/Baha'i Faith/history." If you apply for a category that does not fit, your site may not be listed at all.

Google's system uses a vast system of web site scanning, called spidering, to analyze the contents of millions of public web sites. Just collecting information about all the various web sites is not enough. For example, searching in Google using the term "Baha'i Faith" returns almost 1,400,000 hits. To be useful, the search engine has to make sure the site you are looking for is in the first ten listings and not the last ten listings.

How does a search engine know what is important?

Taking a million and a half possible listings and ordering them in terms of relevancy is not an easy task. The task is made all the more difficult by the fact that many of the web

sites' creators are attempting to make sure their site displays on the first page. The hard reality of search engine listings is that any site that is not on the first few pages of results is essentially invisible. In fact, anything not on the first page is at a severe disadvantage.

Search engines use a number of factors to assess web sites. Some of these factors are well known. However, the degree to which specific factors impact actual web site rankings is a closely guarded trade secret of each search engine. This is so because unscrupulous webmasters will abuse the system to get a higher listing than their site would otherwise merit.

The basic factors a search engine is looking for are:

1. **Relevance:** Is your site about the search term being used?
2. **Authority:** Is your site a trusted and authoritative source of information on the subject?
3. **Functionality:** Is your site well designed and functional?

Relevance

There are two basic ingredients to determining relevance. The first is use of the term that was searched for in the web site. If a person searches for "Baha'i Faith" and your site does not use this term or phrase, your site will not appear high up on the list of results. In addition to using a term, the term has to be used in prominent ways to get the most attention. For example, a site with the domain name of "Bahai.org" is considered more relevant to the search term of "Baha'i" than would the domain name of "BlessedIsTheSpot.org." To a Bahá'í user, both domain names are clearly Bahá'í, but a search engine is not able to understand the allusions and other forms of references. It is of great advantage to use the most frequently searched for terms or phrases on a Web site or in a domain name. But such terms or phrases must be incorporated in a legitimate manner in the site's content.

Because key terms can be manipulated by webmasters, search engines also look to see what other public web sites are linking to a site. A web master may be able to use the term "Baha'i" all he wants to make it appear that the site is about the Bahá'í Faith, but if no other Bahá'í web site links to it, the search engine may suspect that it is being fooled. The search engine people know that it is difficult to trick someone into linking to a site about something when it is really not about the declared topic.

Naturally, webmasters have an incentive to trick the system using links from sites that they create just for the purpose of creating links back to their primary site. That worked for a short time, but now search engines look at the linked sites to see how long those sites have been on the Internet and how many well-listed sites link to them.

This cat-and-mouse game makes for a very complicated system. Fortunately, if you follow basic guidelines, your site should be listed fairly.

Authority

Computers cannot understand the content they process. They have no way to distinguish a brilliant paper from a hoax. They look at a variety of criteria that usually characterizes authoritative information. For example, has the information been on the Internet for long? This is a measure of commitment. Does the content site's content change? Too much change may be bad, but no change may indicate an abandoned site.

Search engines also look at information not obviously related to the web site's content. For example, a search engine looks at the domain name's registration. If the site is listed to a known spammer or the registration information is missing, the search engine may suspect a fly-by-night operation.

Functionality

One can tell a great deal about a person by how well they take care of their appearance and maintain their home. Sites that are broken or have large numbers of "under construction" pages don't inspire confidence. Search engines are inclined to keep users away from sites that provide a disappointing experience. Additionally, the search engine assumes that if your web site is broken, your attention to the quality of your content may be equally off-putting.

Even if the site is functional, it may be harming itself with the search engines. A search engine cannot index a web page it cannot find. This may sound obvious, but many web sites use technologies that searching engine spiders cannot handle. The normal method for a search engine spider to operate is for it to visit the site's home page and then it follows every link found there. It does the same to every page it subsequently visits. Some forms of technology, such as frames and certain uses of Flash can have the effect of cutting off parts of the web site because the links are invisible to the search spider.

Spamming search engines

As already mentioned, webmasters have a strong incentive to boost their results artificially. This is referred to as spamming. Search engines will look for these techniques and will either penalize the site or remove it completely.

While attention to how well your site will be listed is very important, it is important not to compromise your site by designing the site around search engine performance. Remember, search engines look for how many trusted web sites link to your site. A site can be so distorted by its attempts to obtain maximum listings that it no longer appeals to the visitor. The final result will be a site no other web site wants to link to.

If you find sites attacking the Bahá'í Faith that are using spamming techniques or seem unreasonably highly placed in the search results, please contact the Bahá'í Internet Agency so that we may investigate.

Basic steps for optimal listings

Here are some of the steps that are generally agreed to be important to the ranking of web sites. These steps should be used where possible but not to the point that the site's quality is diminished. The examples below will use the example of a hypothetical site on Bahá'í history.

1. **Pick the right terms.** Make sure you are attempting to be found using terms that truly represent your site. Think about what your site is about and how the prospective visitor would most likely search for you. If your site is for the non-Bahá'í public, don't use Bahá'í jargon.
2. **Use key terms in important places.** Where and how you use a term on a page is more important than the total number of times you use a term. Here are some high-value locations listed in their general order of importance.
 - a. Domain name: www.bahaihistory.org
 - b. Web URL: www.bahaihistory.org/earlyhistory.html
 - c. Page title: Baha'i History: Early Years
 - d. Headings: <H1>Baha'i History: The Early Years </H1>
 - e. Use soon and often: The sooner you use your term in the text, the better. Likewise, using a term more often is better unless it starts to harm the content.
 - f. Meta information: There are meta-tags and alt tags where information can be provided. For example, in the meta-tag for keywords, you might want to put in "Baha'i, Bahai, Bahá'í" as possible terms the user may type in for your site. Unfortunately for us, search engines don't always recognize that all the various ways users type in "Baha'i" should go to the same site.
3. **Have a properly functioning site.** Make sure your site can be scanned by the search engines. Check for broken links or other evidence that the site is not well taken care of.
4. **Get external links to your site.** A few links from well-established sites in your subject area can help a great deal. For example, a site on religious history would be a great place to have a link from. Search engines would take this as a significant sign that your site was respected. Numerous links from sites that have nothing to do with the Bahá'í Faith or history could actually hurt your site's rankings.
5. **Size of the site.** Search engines treat larger sites as more important. That means having at least 25 pages on your web site. The more, the better.
6. **Freshness of the content.** Adding content or updating current pages makes your site look more interesting and relevant.
7. **Individualize.** Optimize every page for the content on the page. Having every page with the title of "Baha'i History" does not help if the page is not about Bahá'í history. If you have an "About Us" page, label as "Baha'i History: About Us" or "About Baha'i History".
8. **Uniqueness of content.** Unique content can get your site to the top of the list if the user is looking for that content. If you have a page on the history of the Bahá'í Faith in New Zealand, make that clear. It will come up for anyone looking

for this information. Likewise, don't use the exact same content on multiple pages or on multiple sites. The search engines penalize this.

9. **Submit your site to the search engines.** When you register a new domain name, there are invariably a slew of email messages from commercial entities asking you to hire them to list your site on hundreds of web search engines. This is not cost effective. Initially, your site only needs to be listed on Google, Yahoo, and MSN and perhaps a few other search engine sites with a particular language or thematic focus.

Patience

Search engines scan public web sites on a continuous basis. How often they visit your site will depend on a wide variety of factors. The most important factor is the freshness of your content. The more often your content changes, the more often the search engines will return. If you want to see the last time your site has been indexed by Google, search for your site and look at the cached version. The date of the last scan will be there. Any changes you've made after that time will not be reflected until the next scan.

Even if your site has been recently scanned, it may take some time for the search engines to recalculate your site listing. This re-indexing happens about once a month. This means that it takes months to see the full impact of all your actions.

Helping other Bahá'í sites

The Bahá'í International Community's Office of Public Information (OPI) will make sure there is a link from bahai.org to your National Assembly's public site. The ranking of Bahá'í International Community sites on search engines can also be enhanced if your site links to them (e.g., bahai.org, news.bahai.org, reference.bahai.org, www.onecountry.org). Such cross-linking in turn helps your site because the value to their links goes up. Similarly, the same is true between your national site and the official sites for Local Spiritual Assemblies in your country. Have them put in a link for your web site and consider doing the same.

Further help

This information is basic in nature. The Bahá'í Internet Agency has access to search engine optimization experts for situations that require the application of advanced techniques. If you have a site, apply as many of the basic solutions as you can. If you are still unsatisfied with your listings, we can provide suggestions to raise your site's rankings. If you are designing a new site, it may be advantageous for us to review your prototype.