



Best Practices: Search Engine Optimization

Detailed Recommendations

Global Navigation Text Links

Internal linking structure will greatly influence whether a search engine's spider can find other pages on your site in addition to the home page. From a Search Engine Optimization perspective, we prefer that the navigational structure consist of pure text links and with no JavaScript. In addition, the home page should contain links to every other important page within your Web site in the form of keyword rich text links. The use of text links will help push the relevancy of the target page, thereby increasing your presence in the engine's search results.

Secondary Navigation Keyword Rich Text Links

These links, like the main navigation, should be text links.

Tertiary Navigation

These links, like the main navigation, should be text links.

Content

The quantity and quality of your content is the foundation of a successful SEO campaign. The primary reason a consumer searches the Internet is to find information or solve a problem. Quality content is what allows the user to meet their goals, as well as acting as a vital component within a search engine's algorithm.

The amount of relevant content and the number of pages a site has are both important factors in determining the value or weight that a search engine's spider will assign to your site. When a search engine spider "crawls" your Web pages, it reads the content on the page, categorizes the page within its index, and ranks the page based on the engine's proprietary relevancy algorithm. Meta tags and ALT tags are not considered visible content, and therefore, are not considered as important as your copy.

Our research has shown that a site is most effective when it contains a minimum of 250 words on each page. It is not advantageous to try to cram all of your content onto one page. This makes it very challenging for a search engine to distinguish between varying subjects, causing it to skip over significant concepts. The greater the quality and quantity of content, the higher the weight allocated to a site by a search engine spider.

Use <H?> Tags for Headings and Sub Headings

Using the <H_> tags for headings and subheadings in a Web page is a key element of SEO. All of the heading and subheading tags, <H1>, <H2>, <H3>, etc. should be used according to their relevance. Many search engines give additional relevance to your heading tags, thus, the heading tag is a great place to incorporate keywords. Specific details regarding the language/wording of heading tags will be incorporated into your content review document.

Callouts and Sidebar Content

Secondary content, defined as "call outs" or secondary blocks of text, are also important to a site's search engine optimization efforts. Again, secondary content provides spiders additional content to spider.



Closing Text Links

Closing links are of the utmost importance if you do not have text links within your site's main navigation. Your text links should contain your main keywords. This practice is key to effective SEO for three reasons:

- It helps ensure that the search engine spiders have the ability to navigate through your site
- Reinforces your main keywords within your internal navigation
- Can also be used as supplemental navigation within the site

Additional Recommendations and Comments

Redirects

Redirects are used, as the name states, to redirect one page (URL) to another page. There are two main types of server redirects- temporary (302) and permanent (301). The 301 permanent redirect is the safest way to preserve your search engine rankings.

When redesigning or launching a new site, always point your old pages to corresponding pages on the new site using a 301 permanent redirect. If there is not a page that corresponds directly, use a 301 permanent redirect to send visitors to the Home Page or Site Map. Do NOT use 302 temporary redirects.

Site Map

A Site Map is incorporated into a site to ensure sure that search engine spiders are able to find and index all the pages within your Web site. To allow the search engine spiders to find all the pages on the site without the interference of graphics and image maps, a Site Map should contain straight text links to every page within your Web site, broken down by category or section, using your main key phrases. The site map should include a heading, which contains a keyword or phrase, and an introductory paragraph that includes important key phrases about your product.

Externalize your JavaScript

JavaScript and some server-side scripting can cause problems that may result in pages not being found by an engine's spider. JavaScript is code that spiders cannot read, and it must be used with caution. JavaScript is not search engine friendly, and in most cases fills up the <head> of the document that, in turn, pushes the rest of your content down to the bottom of the page. To avoid JavaScript "clutter," place your scripting code, such as mouse-overs/rollovers, drop down menus, pop-up menus and sliding menus, in a separate (.js) file, and create a single line within the <head> to call upon that external file. This will create far less code for the search engine to wade through in order to get to your main content.

404 Error Codes - "page not found errors"

404 Error Codes indicate an incorrect URL or a deleted file has been requested. Returning this code to engine requests is optimal to make sure the search engines remove old pages that no longer exist. Not returning this code can result in a site having duplicate content.

Robots.txt files

Many Web sites utilize a Robots.txt file because they may have pages within their Web site that they would not want a spider to read. Implementation of a robots.txt file specifying which pages within the Web site should not be spidered will prevent these pages from being indexed or viewed by non-authorized parties. In addition, robot.txt files can be used to prevent potential "spamming" issues within search engines such as duplicate content. A search engine's robot (spider) will look in your root domain for a



special file named "robots.txt" (<http://www.mydomain.com/robots.txt>) that tells the robot which files it may not spider (download).

Linking Resources

Link popularity is an important component of your SEO campaign. Link popularity is the number of Web sites that link to your Web site. Success with link popularity will result in top search engine rankings and an increase in traffic to your Web site. Search engines give link popularity greater weight in their algorithms because they believe it indicates quality. Google relies heavily on link popularity to rank sites. Search engines use links as a tool to help them filter out sites that aren't relevant to their users. PLEASE NOTE: There are currently linking efforts in place at the local level and we are on location #3 of all locations.

Image ALT Attributes

ALT attributes are HTML tags used describe Web site graphics by displaying a block of text when an image cannot be loaded on a page. As search engines are generally unable to view graphics or distinguish text that might be contained within them, the implementation of an ALT attribute enables search engine spiders to categorize that graphic. Associating the ALT attribute to an image is also a best practice set forth by the W3C (World Wide Web Consortium) and therefore contributes to having a site that is W3C compliant.

Anchor Title Attributes

Title attributes act in the same fashion as ALT attributes do on images, except they are used for text links and the purpose is slightly different. The title is an attribute of the anchor tag, `<a>`, and should describe the contents of the following page in which the link is directed. The title attribute will appear, just as ALT attributes appear, when the user moves their mouse over the text link.

Code example:

```
<a href="services.htm" title="Learn more about our SEM services." >Services</a>
```

Code Validation

This is the process of ensuring web pages uphold the proper structure within the intended markup language structure. The W3C or World Wide Web Consortium is the organization that defines the standards for proper markup language code structure. Writing improper code results in the spiders having to work much harder when traversing your pages. It is advisable that all pages within your site are properly coded and hence valid.

To test or validate your pages visit: <http://www.w3.org/QA/Tools/#validators>.

Canonicalization

Canonicalization is the process of converting data that has more than one possible representation into a "standard" canonical representation. Read more here. In search engine optimization, we use this to refer to the linking structure of a site. The most common case occurs when there are two different links on a site that lead to the homepage. For example, <http://www.domain.com> and <http://www.domain.com/index.htm> usually lead to the homepage. For SEO purposes, we need you to choose one or the other. In this instance however, the only option would be to go with <http://www.domain.com>, because this is how other sites will be linking to you. So in this example, all links on your site directed to your site's homepage should be linked using <http://www.domain.com>.



Additional SEO Best Practices

Externalize your Cascading Style Sheets

Similar to JavaScript, an embedded Cascading Style Sheet fills up the <head> of the HTML code and, in turn, pushes the rest of your content down to the bottom of the page. To avoid this, create an external style sheet for the entire site. This will create far less code for the search engine to wade through in order to get to your main content.

Frames

Many sites use frames for navigation, but most Search Engines have difficulty indexing framed web sites. Using frames will prevent search engines from finding pages within a web site or the search engine sends visitors to an internal page with no connections to the navigation or other areas of the site.

Flash

While spider technology has greatly improved, the search engines do not have the technology to spider flash-based pages, as well as flash-based content. As content within flash movie will not be indexed by the search engines, our optimal recommendation is to remove all flash-based content from all pages of the site and replace with plain HTML.

URL Rewrites

URL rewriting, in regards to SEO, is the process of making URLs look like static instead of dynamic via web server URL rewriting program.

Dynamically Served Content

Dynamic pages are not indexed well by the search engines and in some cases completely ignored by the spiders. The most common problem dynamic content poses is the "?" (or any ASCII symbol, i.e. equal sign, commas, etc) in the URL. Most search engines stop at the "?" symbol within the URL and as a result, try to retrieve an invalid URL (see example below).

<http://www.pittsburghregion.org/public/cfm/region/index.cfm?FUSEACTION=TheRegion>.

Although there have been some advances made in improving the ability to index dynamic content (specifically by Google), we recommend removing most (if not all) dynamic content. Replace dynamic content with static html pages, and create additional static value pages that are placed before your dynamic content (if any).

Cookies

Cookies are small text files that store user information and preferences. Cookies can cause issues with search engine optimization if they are implemented incorrectly. By using cookies on your site, the search engine spiders are not able to view your site. From an SEO perspective, cookies should not be mandatory in order for users to visit a site. Often sites require cookies to be active, usually directing users without active cookies to a default "cookies not active" warning page, which then tells the user that their cookies must be active to fully experience the site.

Search engine spiders will not be able to visit the pages of your site that make the cookies mandatory, as they are not able to visit pages implemented in this fashion. If your site is making cookies mandatory, you should change the site so the cookies are not mandatory. Under certain circumstances, changing the



User Centered Design

Elizabeth G Fagan
Principal Content Developer
HSBC Technology & Services, User-Centered Design
200 W. Adams, Suite 1400, Chicago, IL 60606
<http://connect.us.hsbc/ucd>

cookies requirements will limit certain abilities on the administrative end, but if you choose not to change them, your entire site will not be indexed—only the default “no cookie” notice page will be indexed.