

## Search Engines Optimization

### SEARCH ENGINE OPTIMAZATION

Search engine optimization (SEO) is the process of improving the volume and quality of traffic to a web site from search engines via "natural" ("organic" or "algorithmic") search results for targeted keywords. Usually, the earlier a site is presented in the search results or the higher it "ranks", the more searchers will visit that site. SEO can also target different kinds of search, including image search, local search, and industry-specific vertical search engines.

As a marketing strategy for increasing a site's relevance, SEO considers how search algorithms work and what people search for. SEO efforts may involve a site's coding, presentation, and structure, as well as fixing problems that could prevent search engine indexing programs from fully spidering a site. Other, more noticeable efforts may include adding unique content to a site, ensuring that content is easily indexed by search engine robots, and making the site more appealing to users. Another class of techniques, known as black hat SEO or spamdexing use methods such as link farms and keyword stuffing that tend to harm search engine user experience. Search engines look for sites that employ these techniques and may remove them from their indices.

The initialism "SEO" can also refer to "search engine optimizers", terms adopted by an industry of consultants who carry out optimization projects on behalf of clients, and by employees who perform SEO services in-house. Search engine optimizers may offer SEO as a stand-alone service or as a part of a broader marketing campaign. Because effective SEO may require changes to the HTML source code of a site, SEO tactics may be incorporated into web site development and design. The term "search engine friendly" may be used to describe web site designs, menus, content management systems, URLs, and shopping carts that are easy to optimize.

### History

Webmasters and content providers began optimizing sites for search engines in the mid-1990s, as the first search engines were cataloging the early Web. Initially, all a webmaster needed to do was submit a page, or URL, to the various engines which would send a spider to "crawl" that page, extract links to other pages from it, and return information found on the page to be indexed. The process involves a search engine spider downloading a page and storing it on the search engine's own server, where a second program, known as an indexer, extracts various information about the page, such as the words it contains and where these are located, as well as any weight for specific words and all links the page contains, which are then placed into a scheduler for crawling at a later date.

Site owners started to recognize the value of having their sites highly ranked and visible in search engine results. According to industry analyst Danny Sullivan, the earliest known use of the phrase "search engine optimization" was a spam message posted on Usenet on July 26, 1997.

Early versions of search algorithms relied on webmaster-provided information such as the keyword meta tag, or index files in engines like ALIWEB. Meta tags provided a guide to each page's content. But using meta data to index pages was found to be less than reliable because the webmaster's account of keywords in the meta tag were not truly relevant to the site's actual keywords. Inaccurate, incomplete, and inconsistent data in Meta tags caused pages to rank for irrelevant searches. Web content providers also manipulated a number of attributes within the HTML source of a page in an attempt to rank well in search engines.

By relying so much on factors exclusively within a webmaster's control, early search engines suffered from abuse and ranking manipulation. To provide better results to their users, search engines had to adapt to ensure their results pages showed the most relevant search results, rather than unrelated pages stuffed with numerous keywords by unscrupulous webmasters. Since the success and popularity of a search engine is determined by its ability to produce the most relevant results to any given search allowing those results to be false would turn users to find other search sources. Search engines responded by developing more complex ranking algorithms, taking into account additional factors that were more difficult for webmasters to manipulate.

## Search Engines Optimization

Graduate students at Stanford University, Larry Page and Sergey Brin developed "backrub", a search engine that relied on a mathematical algorithm to rate the prominence of web pages. The number calculated by the algorithm, Page Rank, is a function of the quantity and strength of inbound links. Page Rank estimates the likelihood that a given page will be reached by a web user who randomly surfs the web, and follows links from one page to another. In effect, this means that some links are stronger than others, as a higher Page Rank page is more likely to be reached by the random surfer.

Page and Brin founded Google in 1998. Google attracted a loyal following among the growing number of Internet users, who liked its simple design. Off-page factors such as Page Rank and hyperlink analysis were considered, as well as on-page factors, to enable Google to avoid the kind of manipulation seen in search engines that only considered on-page factors for their rankings. Although Page Rank was more difficult to game, webmasters had already developed link building tools and schemes to influence the Inktomi search engine, and these methods proved similarly applicable to gaining Page Rank. Many sites focused on exchanging, buying, and selling links, often on a massive scale. Some of these schemes, or link farms, involved the creation of thousands of sites for the sole purpose of link spamming.

To reduce the impact of link schemes, as of 2007, search engines consider a wide range of undisclosed factors for their ranking algorithms. Google says it ranks sites using more than 200 different signals. The three leading search engines, Google, Yahoo and Microsoft's Live Search, do not disclose the algorithms they use to rank pages. Notable SEOs, such as Rand Fishkin, Barry Schwartz, Aaron Wall and Jill Whalen, have studied different approaches to search engine optimization, and have published their opinions in online forums and blogs. SEO practitioners may also study patents held by various search engines to gain insight into the algorithms.

### Webmasters and search engines

By 1997 search engines recognized that some webmasters were making efforts to rank well in their search engines, and even manipulating the page rankings in search results. Early search engines, such as Infoseek, adjusted their algorithms to prevent webmasters from manipulating rankings by stuffing pages with excessive or irrelevant keywords.

Due to the high marketing value of targeted search results, there is potential for an adversarial relationship between search engines and SEOs. In 2005, an annual conference, AIRWeb, Adversarial Information Retrieval on the Web, was created to discuss and minimize the damaging effects of aggressive web content providers.

SEO companies that employ overly aggressive techniques can get their client websites banned from the search results. In 2005, the Wall Street Journal profiled a company, Traffic Power, which allegedly used high-risk techniques and failed to disclose those risks to its clients. Wired magazine reported that the same company sued blogger Aaron Wall for writing about the ban. Google's Matt Cutts later confirmed that Google did in fact ban Traffic Power and some of its clients.

Some search engines have also reached out to the SEO industry, and are frequent sponsors and guests at SEO conferences, chats, and seminars. In fact, with the advent of paid inclusion, some search engines now have a vested interest in the health of the optimization community. Major search engines provide information and guidelines to help with site optimization. Google has a Sitemaps program to help webmasters learn if Google is having any problems indexing their website and also provides data on Google traffic to the website. Yahoo! Site Explorer provides a way for webmasters to submit URLs, determine how many pages are in the Yahoo! index and view link information.

### Getting indexed

The leading search engines, Google, Yahoo! and Microsoft, use crawlers to find pages for their algorithmic search results. Pages that are linked from other search engine indexed pages do not need to be submitted because they are found automatically. Some search engines, notably Yahoo!, operate a paid submission service that guarantee crawling for either a set fee or cost per click. Such programs usually guarantee inclusion in the database, but do not guarantee specific ranking within the search results. Yahoo's paid inclusion program has drawn criticism from advertisers and competitors. Two major directories, the Yahoo Directory and the Open Directory Project both require manual submission and human

## Search Engines Optimization

editorial review. Google offers Google Webmaster Tools, for which an XML Sitemap feed can be created and submitted for free to ensure that all pages are found, especially pages that aren't discoverable by automatically following links.

Search engine crawlers may look at a number of different factors when crawl a site. Not every page is indexed by the search engines. Distance of pages from the root directory of a site may also be a factor in whether or not pages get crawled.

### Preventing indexing

To avoid undesirable content in the search indexes, webmasters can instruct spiders not to crawl certain files or directories through the standard robots.txt file in the root directory of the domain. Additionally, a page can be explicitly excluded from a search engine's database by using a meta tag specific to robots. When a search engine visits a site, the robots.txt located in the root directory is the first file crawled. The robots.txt file is then parsed, and will instruct the robot as to which pages are not to be crawled. As a search engine crawler may keep a cached copy of this file, it may on occasion crawling pages a webmaster does not wish crawled. Pages typically prevented from being crawled include login specific pages such as shopping carts and user-specific content such as search results from internal searches. In March 2007, Google warned webmasters that they should prevent indexing of internal search results because those pages are considered search spam.

### White hat versus black hat

SEO techniques are classified by some into two broad categories: techniques that search engines recommend as part of good design and those techniques that search engines do not approve of and attempt to minimize the effect of, referred to as spamdexing. Some industry commentators classify these methods, and the practitioners who employ them, as either white hat SEO, or black hat SEO. White hats tend to produce results that last a long time, whereas black hats anticipate that their sites may eventually be banned either temporarily or permanently once the search engines discover what they are doing.

An SEO technique is considered white hat if it conforms to the search engines' guidelines and involves no deception. As the search engine guidelines are not written as a series of rules or commandments, this is an important distinction to note. White hat SEO is not just about following guidelines, but is about ensuring that the content a search engine indexes and subsequently ranks is the same content a user will see.

White hat advice is generally summed up as creating content for users, not for search engines, and then making that content easily accessible to the spiders, rather than attempting to trick the algorithm from its intended purpose. White hat SEO is in many ways similar to web development that promotes accessibility, although the two are not identical.

Black hat SEO attempts to improve rankings in ways that are disapproved of by the search engines, or involve deception. One black hat technique uses text that is hidden, either as text colored similar to the background, in an invisible div, or positioned off screen. Another method gives a different page depending on whether the page is being requested by a human visitor or a search engine, a technique known as cloaking.

Search engines may penalize sites they discover using black hat methods, either by reducing their rankings or eliminating their listings from their databases altogether. Such penalties can be applied either automatically by the search engines' algorithms, or by a manual site review.

One infamous example was the February 2006 Google removal of both BMW Germany and Ricoh Germany for use of deceptive practices. Both companies, however, quickly apologized, fixed the offending pages, and were restored to Google's list.

### As a marketing strategy

## Search Engines Optimization

Eye tracking studies have shown that searchers scan a search results page from top to bottom and left to right (for left to right languages), looking for a relevant result. Placement at or near the top of the rankings therefore increases the number of searchers who will visit a site. However, more search engine referrals does not guarantee more sales. SEO is not necessarily an appropriate strategy for every website, and other Internet marketing strategies can be much more effective, depending on the site operator's goals. A successful Internet marketing campaign may drive organic traffic to web pages, but it also may involve the use of paid advertising on search engines and other pages, building high quality web pages to engage and persuade, addressing technical issues that may keep search engines from crawling and indexing those sites, setting up analytics programs to enable site owners to measure their successes, and improving a site's conversion rate.

SEO may generate a return on investment. However, search engines are not paid for organic search traffic, their algorithms change, and there are no guarantees of continued referrals. Due to this lack of guarantees and certainty, a business that relies heavily on search engine traffic can suffer major losses if the search engines stop sending visitors. It is considered wise business practice for website operators to liberate themselves from dependence on search engine traffic. A top ranked SEO blog Seomoz.org has reported, "Search marketers, in a twist of irony, receive a very small share of their traffic from search engines." Instead, their main sources of traffic are links from other websites.

### International markets

A Baidu search results page  
A Baidu search results page

The search engines' market shares vary from market to market, as does competition. In 2003, Danny Sullivan stated that Google represented about 75% of all searches. In markets outside the United States, Google's share is often larger, and Google remains the dominant search engine worldwide as of 2007. As of 2006, Google held about 40% of the market in the United States, but Google had an 85-90% market share in Germany. While there were hundreds of SEO firms in the US at that time, there were only about five in Germany.

In Russia the situation is reversed. Local search engine Yandex controls 50% of the paid advertising revenue, while Google has less than 9%. In China, Baidu continues to lead in market share, although Google has been gaining share as of 2007.

Successful search optimization for international markets may require professional translation of web pages, registration of a domain name with a top level domain in the target market, and web hosting that provides a local IP address. Otherwise, the fundamental elements of search optimization are essentially the same, regardless of language.

### Legal precedents

On October 17, 2002, SearchKing filed suit in the United States District Court, Western District of Oklahoma, against the search engine Google. SearchKing's claim was that Google's tactics to prevent spamdexing constituted a tortious interference with contractual relations. On May 27, 2003, the court granted Google's motion to dismiss the complaint because SearchKing "failed to state a claim upon which relief may be granted."

In March 2006, KinderStart.com, LLC filed a First Amendment complaint against Google and also attempted to include potential members of the class of plaintiffs in a class action. The plaintiff's web site was removed from Google's index prior to the lawsuit and the amount of traffic to the site plummeted. On March 16, 2007 the United States District Court for the Northern District of California (San Jose Division) dismissed KinderStart's complaint without leave to amend, and partially granted Google's motion for Rule 11 sanctions against KinderStart's attorney, requiring him to pay part of Google's legal expenses.