

## Introduction to SEO Theory

SEO Theory is the study of techniques and methodologies intended to affect or improve the visibility of Web documents in search engine results. An informal discipline, SEO Theory is a natural outgrowth of the widespread interest in and practice of Search engine optimization. "Search engine optimization" has been defined in many ways (Cf. <http://www.google.com/search?hl=en&safe=off&q=define%3A%22search+engine+optimization%22>).

As the scope of search engine optimization is very broad, for the sake of this discussion, SEO will be treated as "the art of designing or modifying Web pages to rank well in search engines" (Source: <http://www.seomoz.org/blog/the-martinez-dictionary-of-seo-and-spam-terminology-2006-edition>).

### Fundamental Principles of SEO Theory

SEO Theory is founded upon the assumption that, through the review and analysis of search results, technical papers, patent applications, search engine Webmaster guidelines, and other authoritative sources, a search engine's algorithm can be in whole or in part reverse engineered for the effective production of Web pages that will rank well in search results.

The value of reverse engineering search engine algorithms is questioned or challenged by many advocates of best practices SEO including Doug Heil and Jill Whalen. Commonly referred to as *algorithm chasing*, search algorithm reverse engineering is often associated with *black hat SEO* or the practice of gaming or deceiving search engines.

In addition to Web document design, and because of the importance placed upon link analysis by several important search engines (including Ask, Google, Live Search, and Yahoo), SEO Theory has evolved to include the study and analysis of linking practices, patterns, and placement.

Although its history may be more strongly identified with the darker side of search engine optimization, SEO Theory today helps to guide the implementation of best practices in Web search marketing.

### History of SEO Theory

SEO Theory was born in the popular *Virtual Promote Gazette* Enewsletter and associated Search Engine Forums founded by Jim Wilson (whose Web properties have been handed on to successors after his death in May 2003). Wilson brought together the first organized community of Web site developers, marketers, and promoters with the purpose of sharing information about how Web search works, how it can be used to benefit the Internet community, and how to build successful online business resources.

Wilson's forums and newsletters often documented new tools and promotional ideas and techniques. But the foundation of SEO Theory lay in the often free-ranging discussions and sharing

of page design techniques in his forum community. However, Wilson openly disavowed association with so-called "search engine spammers" and within a few years the SEO community he built began to branch out into younger forum and Enewsletter ventures.

As the spammers built their own communities, they began to share ideas more openly about how they were able to successfully place doorway pages in search engine results. The spammers developed new techniques and tactics to counteract search engine anti-spam methodologies. Only a few of the spammer tactics have ever been adopted or co-developed by the best practices community.

SEO Theory shifted its emphasis away from on-page elements toward off-page elements with the ascendancy of Google and the discovery of the effectiveness of link bombing in 2003. Although best practices advocates discourage the use of link bombs, they acknowledge the necessity of obtaining links from other documents.

Today, SEO Theory is less emphasized by the best practices community, who generally rely upon well-established Website marketing and search engine Webmaster guidelines to structure their Web promotion campaigns. Nonetheless, SEO Theory has made significant contributions to widely accepted and advocated Website marketing methodologies.

## **How SEO Theorists Work**

SEO Theory requires the testing and/or analysis of search engine behavior with respect to Web document design and organization. Much SEO Theory is simply based upon the study of authoritative documents that reveal significant or potentially significant details about search engine design. Bill Slawski is the most well-known search engine patent analyzer. He shares detailed analysis about patent applications obtained by search engines on his blog (Cf. <http://www.seobythesea.com/>) and on popular SEO news sites.

Many SEO FAQs and SEO tutorials have also analyzed a few technical papers published by Larry Page and Sergey Brin about their Google search engine design and PageRank algorithm.

Some people engage public testing of optimization ideas by sharing carefully designed documents on their blogs. These documents attempt to measure the impact of various on-page and off-page optimization techniques on search engine rankings.

Public SEO tests are not historically developed according to scientific principles and their results provide minimal value. One reason for the insufficiency of SEO test results (and the conclusions their authors reach) is that they fail to approximate real search results. That is, SEO tests are usually built around idiosyncratic terminology that is so rare as to be unique to the tests.

Most SEO testing appears to be done in secret, usually by corporations seeking to develop strategically competitive advantages for use in online marketing campaigns. SEO testing has not been documented or studied even by the SEO community. The practice is usually conducted in a "don't ask, don't tell" environment and test results are carefully hoarded and shared only among confidential associates.

It is impossible to gauge the quality or value of the vast majority of SEO testing methods and standards because of the level of secrecy associated with the tests.

Some members of the SEO community have argued that SEO contests contribute to SEO Theory, but the value of such contests has not been well articulated. SEO contests are usually acknowledged with a mixture of reservation and enthusiasm [<http://blog.searchenginewatch.com/blog/060116-162523>]. It is customary for SEO contest participants to rely mostly upon link networks to influence search results, so at best SEO contests only repeatedly show that search results can be manipulated by link networks.

## Examples of SEO Theory Concepts

*Link building* seems to be the most widely practiced aspect of SEO Theory today. Because search engines now rely almost exclusively upon Web crawling to find the most relevant documents to satisfy user queries, direct submission to search engines has fallen off in popularity among the SEO community. The practice of link building has been defined in several ways (Source: <http://www.google.com/search?hl=en&safe=off&q=define%3A%22link+building%22>) but can best be summed up as “acquiring links that point toward a designated Web page, usually with specific Anchor text”.

So-called Black Hat SEOs often employ aggressive link building tactics, including the use of software to seek out and “drop links” on Web documents (such as guest books, blogs, and forums). The White Hat or best practices SEOs use less aggressive, generally non-invasive techniques. Best practices SEOs may contact a few Webmasters whose sites have similar content and ask for links to new sites.

In-between the White Hats and Black Hats are a community of people who create and manage linking relationships and users of such services. These “link brokers” may facilitate the paid placement of links or the formal reciprocation of links between Websites. Reciprocal links have long been recognized by the SEO community as a valid and useful means of helping pages build visibility and prominence. However, excessive reciprocal linking has brought an unintended stigma to the practice.

*Link farms* are an outgrowth of reciprocal linking. Considered to be an unethical or Black Hat approach to search engine optimization, link farms are penalized or ignored by search engines when found. Best practices SEOs strenuously advise clients and students to avoid participating in link farms while often skirting the issue of reciprocation.

*On-page optimization* remains an important skill for many SEOs, but its continued value is disputed particularly by advocates of strong SEO linking strategies. Some on-page optimization techniques clearly violate search engine guidelines and are considered to be spam or unethical. Other on-page optimization techniques are considered to be a normal function of Web page design but can be misused.

On-page optimization advocates generally favor the creation of user-friendly content that is compelling and interesting. A full on-page optimization methodology may look at title tags, header tags, use of **bold** and *italics*, use of alt= text for images, design of anchor text for outbound links, and other factors.

*Website optimization* looks at the structure and layout of a collection of Web documents. Some SEOs refer to this practice as “theming a Website”, but the “theming” label has fallen into disrepute. Website optimization includes the design of internal or interpage navigation, styles applied to all pages, URL formatting, and page redirection as well as the contents of the Robots.txt file. Other elements of on-page and on-site management and organization are also considered.

*Link baiting* has become popular among some SEOs because it borrows the best techniques from other areas of SEO Theory. First coined by Nick Wilson (founder of the popular SEO forum ThreadWatch), “link bait” is the practice of creating content that is so unique, compelling, and interesting that people will naturally want to link to it on their own sites.

The goal of link baiting is to attract enough attention to earn natural 1-way links from a large number of Websites. Link baiting was developed as a means of overcoming the Google Sandbox Effect, which was first identified by the SEO community in 2004. The “sandbox effect” prevented new Websites from ranking well in search results on Google for up to a year. After it became apparent that Google was measuring the value and quality of links to new sites in a different way, SEOs began seeking out methods of acquiring the right links.

Link baiting has recently come under fire as being too dependent upon Social media and Social bookmarking resources. Some critics of link baiting also point to its increasingly formulaic appearance as SEOs experiment with titles and article formats. Nonetheless, link baiting may help swing the pendulum back toward greater emphasis on on-site optimization, as well-structured articles tend to draw more attention than rambling, poorly organized articles.

*Bait optimization* has only just begun to emerge as a set of techniques and practices by content creators toward the end of 2006 and in early 2007. A growing number of SEO blogs and news sites now publish articles about the most effective link baiting practices. Search engines including Google have encouraged the practice of link baiting either through tentative endorsements or through the creation of tools to help Webmasters add value to their content.

## SEO Theory Terminology

*Algorithm chasing* is the practice of reverse engineering search engine algorithms. Best practices advocates discourage the practice of algorithm chasing on the grounds that it requires constant monitoring of search results and adjustment of page contents and links. Some best practices advocates also associate algorithm chasing with so-called Black Hat SEO.

*Cloaking* is the practice of showing one content page to a search engine and another content page to actual people. “Cloaking” is achieved through a variety of means and some people argue that there are legitimate reasons for cloaking. Best practices advocates usually advise people not to cloak. Google has been criticized for allegedly permitting a select group of sites (such as academic paper archives, newspapers, and at least one SEO forum where Google employees have participated in discussions) to engage in cloaking or cloaking-like behavior while penalizing and/or banning other sites for using similar technology.

*Doorway page, landing page, presell page, gateway page* all refer to specially designed, minimal content pages that are created solely for the purpose of receiving traffic from a search engine and

sending it to another destination, either through static links or redirection. Traditional doorway pages were used to acquire traffic from search engines for thousands of query expressions, but pay-per-click advertising uses similar pages as “landing pages” to receive click-through traffic from PPC advertising campaigns.

*Crawlability* refers to the effectiveness of a Website's internal navigation. Some internal linking practices, such as use of Javascript and Flash for interpage menuing structures, inhibit search engine crawling of pages. The easier it is for search engines to find and retrieve pages from a Website, the more crawlable the site is said to be. The best links for crawlability are static HTML links embedded directly in page content. These links are usually found in margin space (top, bottom, left, or right).

*Link baiting* is the practice of creating content that naturally attracts unsolicited links.

*Link building* is the practice of acquiring links to Web content through active solicitation, construction of linking pages, reciprocation, or purchase of linking services.

*Link dropping* is the practice of visiting a forum, guest book, or blog for the sole purpose of leaving a link in a usually vacuous and insincere comment. Most link drops are usually deleted on sight in actively managed forums and blogs.

*Link optimization* is the practice of designing or selecting links on the basis of the presumed quality of a Website, and/or to pass or convey anchor text in a very specific format. “Link optimization” also refers to the managed distribution of artificial links in an attempt to avoid detection by search engine analysts.

*Link spam* refers to the indiscriminant or excessive inclusion of links on a Web document. Link spam is often created by third-parties or spammers who send software to drop unwanted links on Web pages. However, link spam can also be intentionally created by Webmasters.

*Meta optimization* refers exclusively to the design of content for meta tags. The two most useful meta tags today are the “description” meta tag and the “robots” meta tag. The “description” meta tag is used to control or influence the snippets that search engines display in their search results. The “robots” meta tag is used to direct search engine crawlers to ignore, index, follow, or handle pages in specific ways.

*Optimization* refers to how a Web document may be designed or modified to rank well in search engines. Search engines return their results based on relevance and on-page optimization can be more effective for many queries than off-page optimization. However, most people do not know much about optimization and they only look at title tags and meta tags.

*Organic SEO, organic listings, organic search* all refer to unpaid search results. Paid search results are usually called *PPC* or *pay-per-click* to distinguish them from organic results. Organic search engine optimization requires more time and strategic planning than PPC marketing (also called *search engine marketing*) but successful organic SEO incurs diminishing costs, where PPC campaigns may be continuous or may spiral into unprofitable cost ranges.

*Scraper* (also called "scraper site" or "scraper page") is a Web document consisting of content taken from other sources. Search engine news and blog alerts are popular sources of scraped content for blogs, which can be configured to accept content from email. Scraper pages may be used to mask other content pages, or they may be used to generate revenues by providing context for Javascript-driven ads, or they may be used to mask link spam.

*Sitemap* refers to a document that provides a list of URLs for pages on the same site. Traditional sitemaps evolved from "hallway pages" or "crawl pages" that were used by spammers to help search engines find their doorway pages. Best practices SEOs realized that large content Websites could benefit from simple, functional, less ornately designed pages. These types of sitemaps are now often called "HTML Sitemaps" to distinguish them from the "XML Sitemaps" that Google and other search engines allow Webmasters to upload to help crawling.

*Value* refers to several concepts, such as the value that links provide in terms of visibility, PageRank (or PageRank-like value), trust, and anchor text. "Value" may also refer to how interesting or compelling Website copy is to real human visitors.

*Visibility* refers to how easy it is to find a Website on the Internet. A Website is said to have no visibility if it is not indexed by search engines, has no inbound links, and is not otherwise being promoted (as through advertising, word-of-mouth, etc.). *Search visibility* refers to how visible a Website is in search results. If a Web document can be found for any query at all, it is considered to be "search visible".

## **The Future of SEO Theory**

SEO Theory will continue to evolve as search engines find new ways to index and promote Web-based content. SEO Theory will also continue to drive Black Hat SEO practices as they react to the constantly changing criteria for inclusion in search engine databases. But SEO Theory should also remain a viable part of White Hat or best practices SEO because it embraces the holistic approach that White Hats take to Web design and promotion.

# # #