

Search Engine Optimization and User Behavior

Abstract: Search Engine Optimization (SEO) is the craft of elevating websites or individual website pages to higher rankings on search engines through programming, marketing, or content acumen. This article covers the origins of SEO, strategies and tactics, history and trends, and the evolution of user behavior in online searching.

Nicholas Carroll

Hastings Research, South Lake Tahoe, Nevada, U.S.A

(Article was commissioned by UCLA Graduate School of Education and Information Studies and published in the [2010 Encyclopedia of Library and Information Sciences](#). Republished in [Understanding Information Retrieval Systems: Management, Types, and Standards](#), 2011.)

| | |
|---|----|
| INTRODUCTION | 3 |
| ORIGINS AND HISTORY OF SEO | 4 |
| PRACTITIONERS AND SOFTWARE | 5 |
| Early Search Engine Promoters and Reporters | 5 |
| The Current Search Engine Optimization Industry | 5 |
| SEO Software (Automating the Process) | 6 |
| SEO Worldwide | 6 |
| BLACK HAT VS. WHITE HAT SEO | 6 |
| SEO METHODS | 7 |
| Choosing Keywords To Target – The Core Of SEO | 9 |
| Content | 10 |
| Density | 11 |
| Placement | 11 |
| Inclusion of Related Keywords | 12 |
| Consistency | 12 |
| Technical | 13 |
| Keyword Loading, Stuffing, and Spamming | 13 |
| The Meta Keywords Tag | 13 |
| Entry Pages | 14 |
| Cloaking | 14 |
| Linking | 15 |
| Link-Weighting Fundamentals | 15 |
| Strategies for Leveraging Links | 16 |
| Website And Web Page Structure | 17 |

| | |
|--|-----------|
| Web Analytics | 17 |
| Combining Content, Technical, And Linking SEO | 18 |
| BROAD ISSUES IN SEO | 18 |
| Building Downwards vs. Outwards | 18 |
| Broad vs. Narrow Targeting and Long-Tail Terms | 19 |
| Lead Time and Longevity | 20 |
| Geographic Targeting | 20 |
| SEO Source Tracking | 21 |
| Increased Competition | 21 |
| Search Engines Only Index Words..... | 22 |
| STANDARDS AND REGULATION IN SEO | 22 |
| HTML..... | 22 |
| Certification | 22 |
| SEO Industry | 22 |
| Search Engines..... | 22 |
| Government Regulation | 23 |
| TRENDS | 23 |
| Fading Trends..... | 23 |
| Current Trends | 24 |
| Trends Affecting SEO | 25 |
| New Search Engine Presentation Methods..... | 25 |
| Specialized Searches | 26 |
| Alternate Search Channels | 27 |
| USER BEHAVIORS IN SEARCH..... | 28 |
| Basic Behaviors..... | 28 |
| Searching Popular Subjects | 29 |
| Use of Boolean Syntax | 29 |
| Search Engine Loyalty | 29 |
| Hardened User Behaviors | 29 |
| User Behavior Worldwide | 30 |
| CONCLUSION..... | 30 |
| ACKNOWLEDGEMENTS | 30 |
| REFERENCES..... | 30 |
| General References..... | 31 |
| Citations..... | 31 |
| Trends In SEO | 33 |
| Trends Affecting SEO | 33 |
| Users | 34 |
| Conclusion | 35 |
| Tables and Illustrations..... | 35 |

INTRODUCTION

Search Engine Optimization (SEO) is the craft of elevating websites or individual website pages to higher rankings on World Wide Web (WWW) search engines through programming, marketing, or content acumen. The definition often includes specifications for increased traffic to a given website, improved quality of traffic, increased profits, or brand awareness.

“Higher rankings” in the popular press or lay discussion generally equates with the goal of having a web page appear in the first 10 or 20 search records for a particular search term, as it is broadly established by tracking that few search engine users will click through to any links beyond the 20th record. SEO professionals usually discuss rankings in terms of SERP (search engine results page) position, example below:

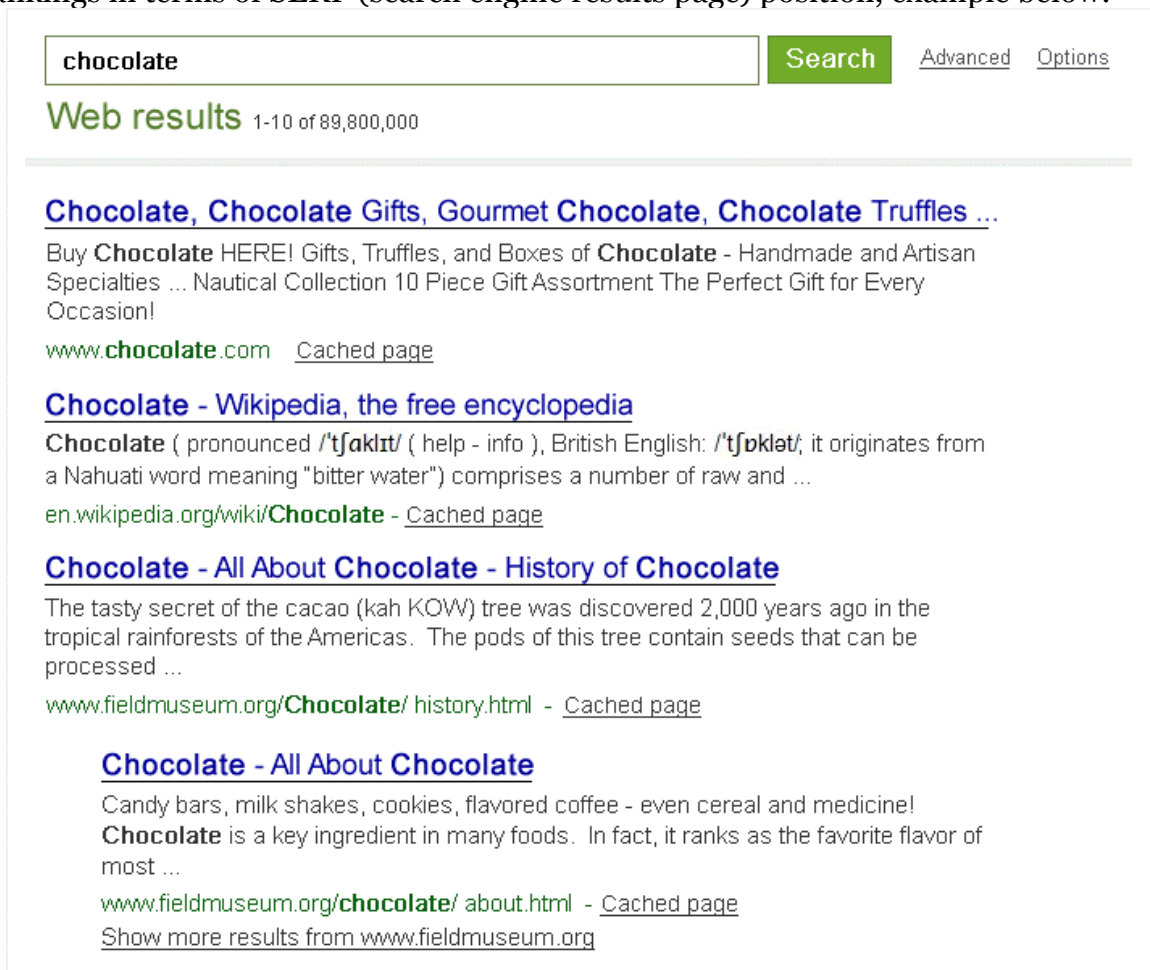


Fig. 1. Typical SERP in search for chocolate, showing the first four results.

Before search engines accepted paid advertisements, SEO was considered a unique form of promotion, radically different from all traditional forms of advertising. It is now more often considered a subset of search engine marketing (SEM), and is

sometimes referred to as "organic search" or "natural SEM," as opposed to paid advertisements placed on the pages of search engines or their affiliates.

Online commerce was the originating force behind search engine promotion and remains the primary driving force behind SEO. Nonprofits and government bodies apply some SEO methods but tend to rely on their unique identity to assure them a prominent SERP position, e.g., the Red Cross, Amnesty International, the Vatican, the New York Department of Motor Vehicles, or the Peoria Public Library.

All the methods described have been or continue to be effective SEO to some degree. Changes in search engine indexing protocols as well as their methods for displaying search results mean that no strict description of the best method or methods can remain entirely accurate indefinitely.

Because the discipline originated in the mid-1990s, terminology is still in flux. General references at the end of this article point to the more authoritative websites and definitions.

This section covers the history of SEO, strategies and tactics, trends, and the evolution of user behavior in online searching. It does not cover pay-per-click, other forms of online advertising, or the resale of website traffic or links.

ORIGINS AND HISTORY OF SEO

The term search engine optimization came into popular use in 1997-98 and is frequently attributed to Danny Sullivan, then operating "Search Engine Watch," though Sullivan states he is uncertain who coined it.(1) Previous terms included search engine placement, search engine ranking, search engine positioning, and search engine promotion, the latter attributed to Jim Rhodes, author of "The Art of Search Engine Promotion."(2) Predecessors such as Jim Heath in his 1995 article "Pointers on how to create business websites that work" did not have a formal name for SEO.(3)

The period from the mid-1990s to about 2000 was characterized by broad experimentation on the part of both search engines seeking a business model and website creators intent on promoting themselves. Search engines were relatively under-powered and minimally staffed; their primary focus was on keeping pace with the growth in new websites. SEO quickly became part of the American "wild, wild Web" metaphor, with more active website owners engaging in a huge variety of methods to gain higher rankings, as described in "SEO Methods." In a period when many multi-national organizations did not have websites at all, smaller and more nimble organizations and individuals aggressively practiced SEO to establish a beachhead on the WWW.

The early days of search engines were in some ways a struggle against pornography. Visitors who started using search engines after 2000 have little conception of how

pervasive the online sex industry was at one time, with their records appearing among search results for cooking, art, quilting, travel, and other innocuous subjects.

As search engines have grown and become more sophisticated, and the number of websites has increased more than ten-fold, aggressive maneuvering to rank well in the SERPs has to some extent given way to a focus on ranking well in particular niches, and executing well on fundamentals rather than exercising brilliance in manipulating search engines.(4)

PRACTITIONERS AND SOFTWARE

Early Search Engine Promoters and Reporters

Early SEO innovation was an individualistic endeavor, primarily developed by small-to-midsize businesses ranging from small hotels to makers of custom sports equipment. Several such niche and popular products were financially successful, demonstrating that SEO was a path to profit.

Websites with forums for exchange of SEO strategies began to appear by 1996, including still-existing websites such as www.virtualpromote.com, www.searchengineforums.com, and the archived www.deadlock.com/promote.

Websites reporting on search engines and optimization began about the same time, including www.wilsonweb.com, www.searchenginewatch.com, and www.webmasterworld.com. More recent additions include www.seobook.com, www.searchengineland.com, and www.toprankblog.com/search-marketing-blogs.

The Current Search Engine Optimization Industry

Whether offered as a subset of SEM services or sold alone as organic search, SEO has become a niche industry with its own sales forces and conventions. The service may be sold as a one-time audit or on an ongoing basis with monthly billing and performance reporting.

Pure SEO consulting firms now number in the thousands, primarily concentrated in North America and the U.K. Tens of thousands more web designers and developers offer the service as ancillary to building websites, and an unknown number of webmasters and website owners apply SEO methods to their own websites.

Gross expenditures on SEO in 2006 were estimated at over USD one billion, with steady annual growth anticipated. This figure describes organic search and does not include paid advertising or most in-house work in smaller companies.(5)

SEO Software (Automating the Process)

Automated web page submission tools came into being in the mid-1990s soon after search engines began indexing the WWW. For some years they were a vital part of SEO strategy for large websites, as manually entering thousands of URLs into the submission pages of a dozen search engines was extremely time-consuming. Submission tools fell into disuse as search engines became more adept at deep crawling (following links into the lower levels of websites' file hierarchies). The ones currently in use typically offer additional features such as help in choosing keywords and automated reporting on SERP positions.

Web analytics tools move beyond automated reporting to features such as tracking visitors' paths through websites, and integrating the data with financials.(6)

Most content management systems (CMS), originally focused on intranets, now include management features for websites, and many are programmed to allow a significant modification of variables important to SEO.(7) Likewise most blogging software includes features that encourage users to add keywords to the title tags, the web page body, and the filename of a given web page.

SEO Worldwide

While search engines are now used worldwide, most SEO reporting is about the U.S.A., Canada, and the U.K. Activity in other countries must be inferred. Australia, Ireland, and Russia in particular have significant indications of entrepreneurial SEO, including blogs on the subject.(8)

BLACK HAT VS. WHITE HAT SEO

"White Hat" SEO typically refers to strategies and tactics that are in concordance with the policies of online search engines, in a loose tacit agreement to provide web surfers with "relevant content." "Black Hat" SEO describes tactics that ignore generally accepted conventions of ethical WWW behavior to advance an agenda or commercial interest. The focus on search engine policies before other interests flows from the pervasive position of search engines, at the mid-point in the process of aggregating web page data and delivering it to users. The tilted point of view is buttressed by the oligopoly of major search engines, with a small number of them processing the vast majority of searches, while millions of websites vie for high rankings.

Technically sophisticated tactics are often equated with Black Hat SEO, and a focus on high-quality content with White Hat SEO. However, there is substantial overlap; few if any tactics can be inherently classified as good or evil. From the perspective of a user, the main criteria of legitimate SEO is whether a given search return is relevant to their interests, regardless of how it achieved its ranking. Most experienced SEO practitioners consider intent to be the defining factor.

Typically, methods considered Black Hat SEO develop high page rankings faster, while those known as White Hat SEO tend to create longer-lasting rankings. A preference for one approach or the other is not entirely a matter of the SEO practitioner's personal preferences; it also depends on the business model. Websites or pages promoting products and services with short lifecycles are suited to Black Hat methods, as they do not suffer from being burned, meaning they have drawn the attention of search engine administrators and been banned from that search engine's index entirely. (This has also happened to sites of long-view organizations such as auto manufacturers when SEO subcontractors acted with an excess of zeal.)

White Hat SEO is better suited to websites that offer products with long life cycles. It is also suited to academic or government websites, where there is likely to be a consistent focus over decades – such websites often gain high and enduring SERP rankings simply by publishing high quality content about a particular subject.

SEO METHODS

In describing SEO it is necessary to distinguish between automated search engines and manually created directories such as Yahoo! Directory (<http://dir.yahoo.com>) or the Open Directory Project (www.dmoz.org), where listings are added by human editors rather than automated protocols.

SEO describes strategies and tactics for influencing page rank on search engines that use robots (www.robotstxt.org/wc/faq.html) or spiders to crawl web pages, traveling from page to page through hyperlinks, and indexing those pages by algorithms and protocols. While directories such as the original Yahoo! Directory can be searched from within Yahoo!, and the ODP pages can frequently be found in SERPs, it is the manner of creating the index that differentiates search engines and directories.

Elements of a Web Page To Be Optimized

The following example of HTML (Hypertext Markup Language) is referred to throughout this article. The markup language and page structure used are explained in the table below the example. (Web page structure can be viewed in most web browsers by selecting the "View > Source" or "View > Page Source" option.)(9) To see how the HTML below actually renders on your browser, click [this link](#).

```
<html>
```

```
<head>
```

```
<title>ELIS – Encyclopedia of Library and Information Sciences 2010</title>
```

```
<meta name="description" content="Search Engine Optimization and User Behavior.">
```

```
<meta name="keywords" content="encyclopaedia, optimise, marcia, bates, nicholas carroll">
```

```
</head>
```

```
<body>
```

```
<h1>ELIS – Encyclopedia of Library and Information Sciences</h1>
```

```

```

```
<p>Body content written for SEO is rich in keywords – and also readable. Proper nouns such as "encyclopedia" are used instead of pronouns. Initializations such as "SEO" are spelled out as "search engine optimization," both to include alternate search terms and to increase keyword density in the web page. To target both sophisticated and lay searchers, common terms such as "acronym" are used in addition to precise terms such as "initialization". Concepts are described by all likely variants, such as "SERP", "search engine position", or "ranking".</p>
```

```
<h2>Further Encyclopedia Resources</h2>
```

```
<p>See the <a href="http://informationr.net/ir/12-4/colis/colis29.html">encyclopedia description</a> for further information on ELIS.</p>
```

```
<p><font size="-1">Keywords: encyclopaedia, optimise, marcia, bates, nicholas carroll</font></p>
```

```
</body>
```

```
</html>
```

| Tags and Text in a Simple Web Page, as seen through an HTML editing tool | Explanation of HTML Tags and Structure |
|---|---|
| <pre><html></pre> | Flag to user's browser: begin parsing for HyperText Markup Language. |
| <pre><head></pre> | Begin the head content (which is not visible in a browser). |
| <pre><title>ELIS – Encyclopedia of Library and Information Sciences</title></pre> | The page title is the single most important part of web page text for search engine ranking. While not shown in the body text in a web browser, it can be seen in the title bar at the top, and is shown as the title of the web page in almost all search engine listings. |
| <pre><meta name="description" content="Search Engine Optimization and User</pre> | Most search engines will use the meta description as the default description in SERPs. SEO practitioners consider this tag a minor influence on |

| | |
|---|---|
| Behavior."> | SERP position, and usually orient the text towards readers rather than search engines. |
| <meta name="keywords" content="encyclopaedia, optimise, behaviour, marcia, bates, nicholas carroll"> | The meta keywords tag, originally intended as a field for general metadata, was widely abused by SEO practitioners and is now a minor or irrelevant element of SEO. |
| </head> | End of the Head section. |
| <body> | Begin the text visible in a browser. |
| <h1>ELIS – Encyclopedia of Library and Information Sciences</h1> | Headings tags, from H1 down to H6, are the equivalent of chapter and section headings in books, and are usually considered influential in SEO. |
| | IMG tags define what image to display as part of a web page. The "alt" attribute is used to store a brief description, typically keywords. |
| <p>Body content written for SEO is rich in keywords ... [etc.]</p> | The paragraph tag is the fundamental "container" for body content. |
| <h2>Resources</h2> | The second-level heading tag may have influence on SERP position, as well as significance for metadata and document readability. |
| <p>See the encyclopedia description for further information on ELIS.</p> | An outbound link where the anchor text contains the relevant keyword "encyclopedia". |
| <p>Keywords: encyclopaedia, optimise, marcia, bates, nicholas carroll</p> | This is an unsightly but often necessary addition to the body content, displaying related keywords in the body text, where the search engines will read and index them. This practice evolved when search engines stopped indexing meta keywords. |
| </body> | End of the body element. |
| </html> | Flag to user's browser: end of HTML file. |

Table 1. Further basic information on HTML is available at www.w3.org/MarkUp/Guide

Choosing Keywords To Target – The Core Of SEO

Regardless of preferred methods, White Hat or Black Hat, content or technical, the most critical part of effective SEO is deciding which keywords to target for high rankings. While search engines continually work to improve their ranking algorithms, they are not clairvoyant; targeting the right keywords is still the foundation of reaching the right audience.(10)

A self-centric viewpoint will usually lead an organization to aim at high rankings for its own name, services, agendas, or products. For example, the Raffles Hotel or the band U2 might focus on targeting searches for their own name, in a reasonable belief that most potential customers are specifically seeking them, rather than hotels in Singapore or rock music in general.

While this may be a successful strategy for broadly known organizations or people, it is considered poor quality SEO when an unknown product, service, or agenda is being advanced. Products and agendas without name recognition are better served by targeting a generic search term such as stainless steel ball bearings than a term like Smith ball bearing company. A multi-national company such as General Electric or Hitachi – that manufactures a huge variety of products – might likewise target searches for the products rather than its own company name.

Regardless of whether the focus is on the organization or the products, SEO can fail when names are chosen without forethought to online search, particularly when they compete with long-established names. Two examples:

- Any product named "Guardian" must compete in the SERPs with dozens of well-established newspapers throughout the English-speaking world.
- Organizations using initializations that might be unique in a local telephone directory frequently face obscurity when competing for recognition in a global medium – "ABC" is not only an initialization for the American Broadcasting Company, but hundreds or thousands of other organizations throughout nations that use the Roman alphabet.

This leads to a fundamental truth about the limits of SEO as a promotional avenue: it can only succeed if people are searching for relevant keywords. If an idea or product is beyond the public's conception, it cannot be promoted through search engines – in contrast to promotion through traditional untargeted media such as radio, TV, or print. This usually makes SEO a poor method for promoting radically new ideas. For example, anti-gravity belts is a search only used by ten-year-old boys.

When possible, experienced SEO practitioners perform keyword analysis before a website is built, expanded, or redesigned, using paid or free online tools such as www.keyworddiscovery.com, www.wordtracker.com, or <https://adwords.google.com/select/KeywordToolExternal> to view aggregated searches conducted on search engines.

Content

At its simplest, the content-focused approach to SEO is to tell the story thoroughly and precisely, in hopes that the words (content) will find a fit with search engine indexing algorithms, and rise in the rankings for particular search terms. The presumption in this strategy is that the writer's words will also find harmony with the terms used by searchers, and if the writer does have the same interests – and frame of

reference – as the target audience, this by itself can result in highly successful SEO. (Conversely, web pages written in ignorance of either subject matter or audience may achieve high SERP positions, but not necessarily for the keywords that draw the desired audience.)(11)

A highly content-oriented strategy is often seen in websites that were conceived and designed by an individual or small group. Content-oriented websites or sub-sites may also come into being in an organization with no SEO strategy at all, such as a department of a university, where the authors are discussing the same or related subjects.

Density

Keyword density receives attention because it is easy to understand and simple to calculate through online density-checkers such as <http://www.keyworddensity.com>. (Such tools are easily found through search engines with the term "keyword density analyzer".)

Though the effectiveness of intentionally repeating keywords throughout a web page is debated by SEO practitioners, a search for highly-promoted products such as loans, chocolate, or discount shoes will return numerous high-ranking web pages with the core keyword repeated dozens of times, and keyword densities reaching up to 7% or higher of the page's content. (However, exceptionally high-density pages can suddenly plummet in the SERPs or be dropped from the index entirely if search engines classify them as examples of keyword spamming.)

Conversely, using a set of keywords only once in a web page usually results in SERP obscurity, even if the set is extremely precise. For example, even dodo bird nesting returns tens of thousands of web pages, and of the first few dozen search engine records, almost all will have the three keywords used several times.

Many SEO practitioners and analysts favor well-written content over keyword density, as search engines become better at differentiating between conventional prose and deliberately enhanced web pages. This school of thought says that content should be primarily aimed at humans, not search engines, and that good writing not only entices users to read and act, but is better search engine bait in the first place. The belief is often expressed as "Content is king."

Placement

The placement of keywords in a web page is a significant factor in SEO, since search engine algorithms accord more weight to certain positions, whether visible or within tags (see Table 1). Notable areas for placement include:

Visible in web page

Inside heading tags <h1, h2, h3....>

Placed at the beginning (top) of a web page.

Within outbound links, i.e., "anchor text" (example on in Table 1).

Placed NEAR other keywords (generally within 10 words).

Visible in SERPs and browser title bar
Inside the HTML <title> tag.

As part of the URL

Competition over search terms and SERP positions has expanded the use of keywords to the point that they are now routinely made part of file names (e.g., www.mysite.com/used-software-for-sale.html), or may even form the domain name itself, e.g. www.usedsoftwareforsale.com.

The title tag is almost universally agreed to be highly important by SEO practitioners. Most other placements are debated. The <meta="keywords"> tag is largely considered of no importance. The alt content in image tags is considered significant in image SERPs, less so in the overall ranking of a web page.(12)

Inclusion of Related Keywords

Due to the unpredictability of user search behavior, many SEO practitioners include synonyms, cognates, tangential terms, and misspellings in web pages. In the mid-1990s, SEO practitioners placed such related keywords in the meta keywords field, but massive abuse and misuse of that field led virtually all search engines to stop indexing meta keywords. As a result, SEO practitioners reluctantly turned to placing them as a visible footer on the web page, though this compromises the esthetics of page design. Some examples of keyword (or "tags") line listings:

Synonyms, cognates, and closely related terms for a medical page about trigeminal neuralgia, formerly called tic douloureux:

"Keywords: tic douloureux, neuropathy, facial pain, face pain"

Tangential terms (statistical outliers) for a site on vegetarian diet:

"Keywords: low sodium, low fat, organic"

Variants on the author's name for a web page written by Jon Smythe:

"Keywords: John, Smith, Smyth"

Common misspellings for a travel website about Colombia:

"Keywords: accomodation, Columbia"

(However, when a keyword tag is used to repeat words already in the body content, it simply becomes keyword stuffing.)

Consistency

Building a website with a consistent pattern of clearly related keywords is commonly known as creating a "theme."(14) A website with a clear theme stays on a single topic or closely related topics throughout its pages (barring such standard pages as "Contact Us"). This strategy builds the overall ranking of the website and, as a consequence, the SERP positions of the individual web pages.

The concept "clearly related" restricts thematic rank-building to keyword relationships that search engines are capable of recognizing. Where ship and marina might begin to build a theme relating to recreational sailing, search engines would be less likely to recognize a theme in a web page with the separate keywords ship and space. Once combined into the more specific keyword spaceship, the page becomes identifiable as related to science fiction or space exploration, and if the website also contains the words galaxy and parsec, a theme begins to build.

The interest in themes peaked around 2004, alarmists suggesting that websites without a keyword theme would plummet in the SERPs.(15) However, there are exceptions to the rule: newspapers, encyclopedias, and many news blogs will never have a clear theme, yet they still can reach high SERP positions for a variety of keywords, based on meeting search engine criteria other than consistency of subject matter.

Technical

A technically oriented SEO strategy emphasizes programming skill or ingenuity over command of language, familiarity with the target audience, or interweaving related content. Tactics vary from the simple, which can be executed by anyone familiar with HTML, to the sophisticated, which require knowledge of programming, website servers, or WWW and Internet protocols.

Keyword Loading, Stuffing, and Spamming

Keyword density can be taken to an extreme. Efforts to increase keyword density or variety that involve excessive or awkward repetition are known as keyword loading. An example would be replacing every instance of "it" in a page with the noun in question.

When this repetition reaches the level of incoherence, with the same word or words used dozens of times in the body, title, or tags of a web page, it is usually called keyword stuffing (a.k.a. cramming), spamming the index ("index" referring to the search engines' databases of web pages), or spamdexing. One of the earliest SEO tactics, it began with simply repeating keywords hundreds of times, generally at the bottom of the page, and frequently with the font color the same as the background color, thus rendering the text invisible to humans. As search engine algorithms began to discount this tactic, keyword spamming evolved into a more precise metering of keyword density. This obsolete technique periodically sees a resurgence as search engine administrators let down their guard.

The Meta Keywords Tag

Located in the head section of a web page's HTML (see Table 1) and invisible to users viewing a web page in a browser, the meta keywords tag was created purely for metadata.

In the mid to late 1990s meta keywords were highly popular as a quick path to higher rankings, and indeed stuffing the meta keywords field showed some success when only a few million pages were being indexed. By the time the tactic became broadly known – with "keyword-jacking" lawsuits over copyright and trademark infringement – search engines were on their way to down-ranking the meta keywords tag contents, and usually not indexing the keywords in the field at all.

Entry Pages

Entry page is the broadest descriptor for the strategy of creating particular pages to rank well with search engines, sometimes finely tuned to rank well with a particular search engine. Doorway page and gateway page imply that a page contains some relevant content that has been tuned for high SERP position. Bridge page or jump page imply that a page may be little more than an uninformative landing page that either urges the visitor to click through to the rest of the website, automatically redirects them to the rest of the website, or redirects them to an entirely different website.⁽¹⁶⁾ Most media coverage of such pages classifies them as Black Hat SEO. However, there is no clear agreement on the definition of these pages or their relative level of deception. Many SEO practitioners observe that while bridge and jump pages may be technical and possibly Black Hat as well, all pages visible to search engines are in a real sense entry pages.

Cloaking

Cloaking is used to describe a broad range of tactics. The common element is that the human visitor and the search engine spiders "see" different content when visiting a web page. In the simplest form, the SEO practitioner creates a page tuned for high SERP position, and replaces it with a human-readable page after the page has been indexed by search engines (of course the artificially high ranking only lasts until search engine spiders visit the page again).

Two of the more technical methods, which attempt to deceive search engines on a continuing basis:

- User-agent-specific page delivery, in which the web page server "sniffs" the incoming page request, extracts the data that identifies what browser the visitor is running (e.g., "googlebot" or "msnbot"), and delivers a special web page tuned to gain higher SERP on that SE.
- IP-specific page delivery, in which the web page server delivers a special web page based on the visitor's Internet Protocol address (e.g., "127.0.0.1").

Cloaking is broadly considered one of the most aggressive and sophisticated forms of Black Hat SEO, but search engine analysts point out that even cloaking has legitimate purposes, as when a website is migrating to a new domain name; the owners may want to keep the old website available to the public until the new one becomes established in SERPs.⁽¹⁷⁾

Linking

Links have become such an important determinant of SERP position that under some circumstances they may override all other elements of ranking, and catapult a web page of little or no relevance to a top SERP position.

Link-Weighting Fundamentals

Since the mid-1990s, search engines have routinely tracked the links from (and to) websites, often assigning a relative value based on the links' source or destination. Links pointing to a website are generally called inbound links, and links pointing to other websites outbound links. (Inbound links are frequently called "backlinks," though the term has other technical meanings in computing.)(18)

Link-weighting is frequently described as a "link popularity" or even "popularity" measurement by popular media. More accurately, search engine link analysis algorithms attempt to infer the value of a web page or entire website based on four factors:

1. The number of inbound links.
2. The "quality" of the inbound links, based on the quality of the originating website. Assessing the quality of websites is where link weighting moves into more sophisticated and sometimes arcane mathematics.(19, 20) In simple terms, an inbound link from a major university has more value than a link from an obscure small business, and has a greater beneficial effect on SERP position.
3. The "relevance" of the inbound link. A website about a baseball team benefits more by inbound links from other baseball-oriented websites than by links from the personal pages of fans.
4. The keywords contained within inbound links (the linked text visible in a browser), known as anchor text. The most famous example was the 2003 link manipulation of Google SERPs (known as "Google bombing") in which hundreds or more website owners inserted the link `Miserable Failure` into their pages, causing the U.S. White House biography of George W. Bush to rise to the top ranking for the term miserable failure.(21 (The gwbbio.html page rose to #1 ranking on Google for that term even though it did not contain either of the keywords in question.)

The cumulative importance of these factors in SERP position has led to naming the collective effect of inbound links "link juice."(22).

Websites with many relevant or high-quality links pointing to them (inbound links) are known as authority websites; those with many links pointing to other relevant websites (outbound links) as expert or hub websites. The presumption built into the search engine algorithms is that a website with many inbound links from high-quality web pages is an authoritative source, and that pages with many outbound links to authority websites serve as a WWW resource. (The mathematics can become circular

and even self-reinforcing, as when www.wikipedia.org web pages briefly started to dominate #1 SERPs positions for thousands of subjects.)

Strategies for Leveraging Links

Link-weighting was implemented to in effect create a weighted voting scheme and take some of the control of rankings away from a website's designer. Since weighting means that not all inbound links are of equal value to SERP position, the methodology also opened the door to SEO opportunities of bewildering complexity, in which websites buy, sell, and trade links to and from their websites in order to improve their SERP positions.

Always a factor in SEO, link building is now broadly considered a core element of strategy, and in some schools of thought the most important element.(23) Increasing media attention brought linking as a SERP ranking factor to the attention of website owners and the quickly growing number of SEO practitioners, and today organizations spend substantial time and effort on link building aimed at high rankings.

Non-reciprocal links (a.k.a. generosity links) were the first evolution, in the mid-1990s. At that time website creators linked to almost any remotely related website, including their competitors. By the late 1990s reciprocal links had become the standard, though the exchange was offered with a view to a direct increase in traffic rather than SERP position.

Soliciting inbound links, at first done casually, now often means assigning employees or subcontractors to solicit links from high-ranking websites. The return on investment is questionable – high-ranking websites have little to gain by giving an outbound link, and the site owners may be concerned that linking to a low-quality website will harm their own rankings. (Soliciting links should not be equated with link building, a broader term that covers all strategies for gaining inbound links.)(24)

The practice of posting links in discussion groups also originated with the goal of direct traffic rather than influencing SERP positions. FFAs (Free For All websites), link farms, and link rings were crude early link-exchange schemes, most variants indiscriminately exchanging thousands of links without regard to relevance. While such sites are still active, their value in SEO has largely been eliminated by increasing sophistication of search engine ranking protocols. Most SEO practitioners today consider participating in them to be useless at best, and at worst, possibly injurious to a website's SERP positions.

Paid links – inbound links that have been purchased, sometimes masking as editorial recommendations – existed before the WWW on Internet bulletin boards, always with the intent of publicity or direct financial gain. With the broadening awareness of SEO, using paid links to influence SERP position has become both a business strategy and a source of contention between website owners and search engines.

Good content, the oldest link-building strategy of all, is a somewhat indirect way to build links. Coupled with even a modest amount of self-promotion beyond good SEO – or occasionally just with good content that perfectly targets a popular search term – competent writing on a particular subject can generate hundreds or thousands of inbound links to a given web page, often without any communication at all with the websites that are giving the links. Because content-inspired linking may produce results slowly (in months or years), and is often difficult to quantify, few organizations devote serious effort to the method.

Website And Web Page Structure

The directory (folder) structure of an SE-friendly website looks similar to a clearly and logically organized hard drive on a personal computer, with the additional proviso that every document is directly or indirectly linked from the home page or some other prominent web page. Ideally the structure is "shallow" (three or fewer sub-directory levels), to make it easier for search engines to spider, though that has become less important as all major search engines now perform deep crawling.

Actual page structure of HTML pages is in theory dictated by an adherence to W3C standards. In practice, websites use almost any markup code that can be rendered by a web browser and leave the difficulties of indexing to the search engines.

There are drawbacks to unorthodoxy where search engine rankings are concerned. Use of highly irregular website structure, page structure, or file naming conventions can seriously harm SERP position. In extreme cases, search engines simply do not add a website to their indexes; poor website structure can be as destructive to website rankings as the most extreme Black Hat tactics, and site maps generated specifically to aid search engines in indexing a website are not a substitute for logical website structure.(25)

Despite a broad disregard for standards among website owners, most SEO practitioners consider disciplined site and page structure fundamental good practice, though these are seen as a foundation for SEO rather than an SEO strategy in themselves.

Web Analytics

The most thorough SEO practitioners use web analytics to analyze website traffic for patterns that can lead to enhanced SEO. While the term embraces areas more concerned with usability and user behavior while on a website, analytics sweeps in SEO functions such as keyword analysis and SERP position monitoring.

Data can be collected by tagging individual web pages with Javascript and other programming, or by placing cookies on visitors' browsers.

Keyword analysis is often initiated at the server level through log analysis. Because WWW communications protocols usually pass the full URL of the previously visited page to the destination website, the headers can be processed for search terms, and those terms then organized and further analyzed.

In either case, data is then analyzed manually, or with in-house programming, or with one of the many commercial web analytics programs.

The data below is a sample of some of the information that can be extracted from website server logs: visitor's ISP and specific IP (Internet Protocol) address; number of pages visited; time and date; visitor's browser, operating system, and language setting; the website the visitor came from; the search terms they used; and the landing page.

A visitor from dynamic.dsl.com (88.104.65.000) was logged twice, starting at 12:46:36 on Sunday, October 14, 2007. The initial browser was Firefox/2.0.0.7. (Windows XP; en)

This visitor first arrived from www.google.co.uk while searching "ecommerce business models 1-10" and visited digitalenterprise.org/models/models.html

There has been a general shift in web analytics from log analysis to page tagging, in part due to the limits of log information compared with information gathered by tagging, and in part driven by vendors of analytics software.

Combining Content, Technical, And Linking SEO

Strategies that combine all three forms of SEO may be the most effective in gaining high SERP positions. Combined strategies are uncommon because websites are normally designed without regard to SEO (although the most effective SEO is begun at the conceptual level), and because website owners and managers rarely allow significant changes to the structure or content of their websites after they are built and online.

An alternate way of viewing SEO is "on page" or "off page" SEO, the former being methods applied directly to websites and pages, the latter focusing on methods external to a website, such as link strategies.(26)

BROAD ISSUES IN SEO

A number of strategic issues that may affect a website go beyond fundamental SEO tactics or current "best practices."

Building Downwards vs. Outwards

Large or growing organizations face the question of whether to house all their web pages under a single WWW domain, or to establish separate websites for different

services, product lines, or agendas. An established organization with a strong brand will usually lean keyword targeting towards its own brand name in SERP positions, and thus benefit from housing all its web pages under a single domain, whereas a newer organization with separate divisions would be more likely to establish separate websites for each product line or goal.

This is not strictly a business issue; a university establishing a new campus in a different city would probably create a new website for that campus, just as a business that sells to both architects and game designers might choose to divide its product lines into two separate websites. The university would certainly link its two websites for the link-weighting benefits; the company selling unrelated products might not link its websites at all.

Broad vs. Narrow Targeting and Long-Tail Terms

The potential benefits in targeting highly specific searches were known to SEO practitioners by 1995-96. Bed and breakfasts were one of the earliest and most evident examples, showing how highly specific keywords (e.g., bed breakfast Ireland Kilkenny) could reach precisely the right audience.

Since the web was still sparsely populated, the more skilled practitioners were able to gain high SERP rankings for broad searches at the same time as they targeted narrow niches. For example, in 1997 it was possible for an Irish bed and breakfast's website to gain a high SERP position for both its own locale and Irish B&Bs in general. That grew more difficult with increasing competition, and today a search for bed breakfast Ireland will generally return a SERP dominated by bed and breakfast directories and associations.

The changing situation became somewhat better understood in 2003, when Zipfian distributions were mentioned in an article by Clay Shirky about "power laws" as applied to blogs.(27) Power laws were later popularized as "The Long Tail" by an article in Wired magazine.(28) With use of the term growing, many clients and SEO practitioners now refer to any three- to four-word term as a "long tail term"; others use the description more correctly to describe an uncommon search term.

There are now indications that the pendulum of interest has swung too far towards uncommon terms, and that organizations are targeting long-tail terms without a clear view towards long-term benefits such as memberships or profits.(29)

Balancing Targeting and Serendipity

While some organizations have only one easily described product or agenda, for most organizations SEO can be too successful when it targets particular search terms so tightly that their web pages can be found by little else. When virtually all traffic to a website comes from a small number of terms, analyzing incoming search terms becomes an exercise in analyzing what one already knows. Ideally, improvements in SEO create an increase in targeted traffic along with an increase in unanticipated but relevant searches. While this can be analyzed in terms of Zipfian distributions, in

proactive marketing it is better represented by the APUPA (Alien-Penumbral-Umbral-Penumbral-Alien) bell curve.

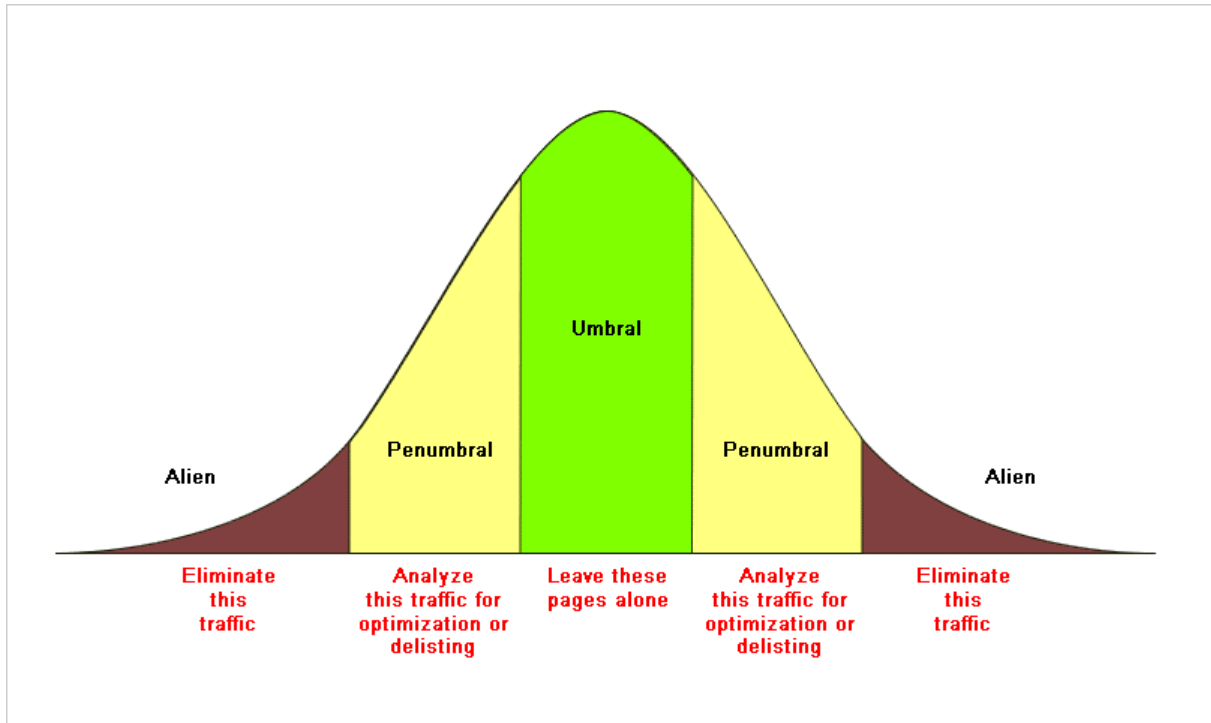


Fig. 2. APUPA curve as applied to SEO.

SEO can also be too successful when a web page captures a high SERP position on a major search engine for an extremely broad term like health. Where this might be satisfactory to a large organization like the U.S. National Institutes of Health, an organization focusing on a particular niche of health could be swamped with masses of unwanted website visitors and email.⁽³⁰⁾ Websites deluged by unwanted traffic sometimes convert a liability to an asset by "reselling" the traffic or the entire website.

Lead Time and Longevity

Search engines rarely assign top rankings to newly indexed web pages. A website can take months or even years to reach its "natural" position in SERPs. This is sometimes called the sandbox, referring to the period a new website may wait for a good SERP position or even to be listed at all by a search engine.

Equally, search engines assign value to longevity, and older well-ranked websites dating from the 1990s can be notoriously difficult to dislodge from their SERP positions by new competitors.

Geographic Targeting

Geographic targeting through SEO had limited success through the mid-00s. Few early website designers made a diligent attempt to anticipate all the geographic keywords searchers might use, and those who did target searches with geographic

keywords often used so many keywords (e.g., London, Pimlico, Knightsbridge, Belgravia, Dulwich, Stratford, Paddington, Fleet Street – and every other village and neighboring town) that the pages were treated as index spam by search engines, and down-ranked to obscurity. Other designers created dozens or even thousands of mirror pages (pages with substantially similar content), each targeting a particular locale, with the result that the pages were down-ranked for duplicate content.

In a cause-and-effect loop, users learned that searching by locale was nearly useless, and abandoned the effort (with the exception of specific travel destinations), leading most SEO practitioners to abandon their efforts at geographic targeting.

With the now-growing success of geographic targeting efforts by search engines – typically displaying maps showing physical locations – users have again started searching geographically, and most SEO practitioners advise making some effort to target searches containing geographic keywords, even though geographic searches may be defined by commercial databases rather than SEO efforts.

SEO Source Tracking

The wealth of data that accompanied visitors to websites led many ecommerce pioneers to think the perfection of marketing analysis had arrived. While this sometimes bordered on true for new businesses that were entirely WWW-based, the opposite turned out to be the case for large, established organizations already being marketed through advertising, mail, public relations, and retail outlets; the WWW added another layer of complexity to their market tracking. Due to its 24-hour availability, the WWW also creates a "smoothing" effect on response to promotional efforts, which can hinder analysis of event-driven spikes in visitor traffic.

Refinements in web analytics have not entirely compensated for these factors, and in many ways website traffic source analysis remains less precise than that of traditional promotion and advertising.

Increased Competition

The continuing growth in the number of websites and web pages means newly created websites have increasing difficulty gaining rankings for common keywords, so their approach is often to use a fanciful domain name and drive web traffic via social networking or publicity instead of SEO.

Aside from the natural growth in websites, huge numbers of "made for ads" (MFA) websites have been created for no purpose other than to make money by hosting online advertisements – they have no products, services, or agendas of their own. Because advertising revenue indirectly comes from organizations that pay for online ads, MFA websites often specifically target keywords used by existing organizations. Nonprofit informational websites such as Wikipedia (<http://www.wikipedia.org/>) also compete directly for many search terms, often gaining very high search rankings.

Search Engines Only Index Words

While search engines are often technically capable of distilling words when they are represented in images, and in theory from audio tracks, in practice they index words from readable text documents, file names, or metadata attached to other forms of electronic files. There is a persistent failure to understand this aspect of SEO, with many website owners believing that if the human eye can read the words or grasp the meaning, the search engines will correctly index the files.

Further restricting SEO, search engines by choice index only certain types of text documents or metadata. At one time that meant pure text or HTML only. Most search engines are now willing to crawl and index file formats such as PDF (Portable Document Format), Microsoft Word and Excel, and in some cases Flash.

STANDARDS AND REGULATION IN SEO

The practice of SEO is marked by the absence of standards, other than those imposed by the willingness of major search engines to index web pages, their level of effort in returning the most relevant web pages, and punitive actions (down-ranking or de-listing) taken against what search engine administrators consider unacceptable manipulation of SERP positions.

HTML

Standard web pages (HTML) are in theory defined by W3C standards – which are largely ignored by both web developers and amateurs. Other formats now being indexed by search engines – including PDF, Microsoft Word and Excel, and still images and videos – follow even fewer standards.

Certification

Certification is in beginning stages. Some businesses have begun offering SEO credentials, and accredited colleges are now beginning to offer courses in SEO.

SEO Industry

The SEO industry may be loosely divided into SEO professionals of varying ability; web designers and programmers who know the rudiments; and a much more visible group of low-cost SEO companies that operate through advertising and email marketing.(31)

While there are no formal standards or regulation, informal standards are slowly evolving through consensus of trade associations and websites reporting on SEO. (See General References at the end of this article.)

Search Engines

Because it is difficult to mathematically determine whether a given web page's content accurately reflects the page's relevance to visitors (much less the creator's true agenda), search engines have faced and continue to face an impossible task in down-ranking deceptive pages without penalizing legitimate ones. (In the classic if

now simplistic example, a medical website can unwittingly trigger an algorithm aimed at filtering out pornography pages.)

Attempts to "game" (manipulate) search engine rankings have been so relentless, from the first significant appearance of online search engines, that many experts consider search engine administration to be equally a process of excluding Black Hat pages and elevating relevant pages(32). While penalizing irrelevant content faces the same problem as returning relevant content – a struggle to develop artificial intelligence – most search engines strongly downgrade websites that are found to be using the more technically complex SEO tactics such as IP-specific page delivery.

This issue of controlling Black Hat SEO without penalizing White Hat SEO is a continued source of tension between search engine administrators and SEO practitioners. Search engines selling advertising space on their own SERPs has exacerbated the tension by raising the issue that pursues traditional media, of whether there is a true division between editorial and advertising departments (or search relevance and advertising, in the case of search engines).

Like the evolution of military tactics, regulation is a game of innovation and counter-measure. The counter-measures may make a given tactic so ineffective that it is completely forgotten by both attackers and defenders, at which point it may be reintroduced by the attackers. (A medical analogy would be the mutation of microorganisms in reaction to new antibiotics, accompanied by the resurgence of forgotten diseases when society no longer guards against them.)

Government Regulation

At present there is no hint of SEO regulation. As subcontractors, SEO consultancies are in a position similar to advertising agencies – with no actual control over organizations or distribution channels – so any government regulation is likely to flow indirectly from regulation of websites themselves (trade practices) or of search engines (particularly in the U.S., where anti-trust law may affect conglomeration of major search engines with software, hardware, or media companies).

TRENDS

Fading Trends

Meta keyword tags were the first major chimera in SEO. They have been followed by other imaginary fast-tracks to high SERP position, all beginning with a grain of truth then blown out of proportion.

Google PageRank has been one of the most persistent areas of focus in SEO. Because of media attention and the easy access to the publicly visible PageRank via the Google Toolbar, the scale has commanded a great deal of interest from website owners, many of whom consider the published PageRank a practical scale for performance-based SEO contracts. Most SEO professionals now consider public PageRank in itself a

minor factor in a website's actual SERP positions, more an effect than a cause. The latter opinion is to some extent corroborated by Google, Inc.(33) Over the course of 2007, countless websites saw their Google public PageRank drop significantly, which may further diminish interest in the measurement. ("PageRank" in popular usage should not be confused with "pagerank", an internal Google term.)

Website traffic – whether measured in hits (requests for individual files, whether pages or images), page views, unique visitors, or the dot-com measure of "more eyeballs" – is slowly losing its popular connection to rankings, as organizations focus on conversion rate of users' visits to desired actions. (In extreme cases, traffic volume will have greater influence on SERP positions than relevance; however, this is uncommon and often results from a situation such as a major news event, where the boost in rankings may be due to a proliferation of inbound links, rather than traffic volume itself.)

Current Trends

- Link building
- Pursuing "long tail" searches.
- Bringing SEO in-house. Major companies in North America and Europe are increasingly bringing SEO in-house as their online sales grow to (USD) millions or hundreds of millions. Companies are hiring in consultants, training current employees, or both.(34)
- Using best practices. Clients and employers are now commonly asking that practitioners follow "best practices." For SEO, the term describes such basic practices as prioritizing meta titles and heading tags, including keywords, and avoiding discredited (Black Hat) tactics.(35)
- Tuning metadata for search engine interfaces. Increasingly, practitioners structure web pages so that titles, descriptions, and filenames are presented appealingly on SERPs. Eye-tracking studies, where eye motion is represented by printed "heat maps," are currently the basis for most decisions in tuning metadata.(36)
- Tagging web pages – adding keywords or allowing visitors to add keywords to the visible text – is common on many social networking websites and some ecommerce and media websites, but has shown mixed results in improving SERP positions, possibly due to the indiscriminately chosen keywords selected by lay users, website owners, and bloggers. Tagging may benefit SERP position most for websites that encourage commentators to use controlled vocabularies.(37) It may also boost SERP positions for "long tail" search terms.

- Web 2.0 designs. Web 2.0 methods have been criticized as being detrimental to SEO. This criticism has arisen with each evolution in website development, from dynamic page delivery (assembling web pages from databases "on the fly") through Cascading Style Sheets (CSS) and PDF files. In most previous instances, organizations' design preferences have won out over SEO considerations, and eventually the major search engines have adapted their protocols to classify new web page structures and document formats.

Since search engines do not immediately retool for new forms of web page markup and structure, websites that employ Web 2.0 methods such as AJAX (Asynchronous Javascript and XML) or that stream together content from different sources ("mashups") may drop in SERP positions until search engines decide how to interpret these new document structures.

Trends Affecting SEO

New Search Engine Presentation Methods

Clusters of search records such as those displayed by <http://www.clusty.com> may improve in relevance; if they increase in popularity, each cluster link will constitute a second-tier SERP, and SEO may begin targeting particular clusters.

Drop-down contextual search suggestions are now offered at <http://www.ask.com> and <http://www.yahoo.com>. Because users see contextual suggestions as they type, even before seeing the first SERP, "contextual position" may become a sought-after goal of SEO. Since contextual suggestions are all displayed on the first screen of a search engine's website, without the need for scrolling down, position is not likely to be as critical as SERP position, and "contextual inclusion" may prove to be the desired goal. (Contextual search flows from a given search engine's database of "most likely" requests; it should not be confused with the URL suggestions that Web browsers make by accessing a user's own search history.)

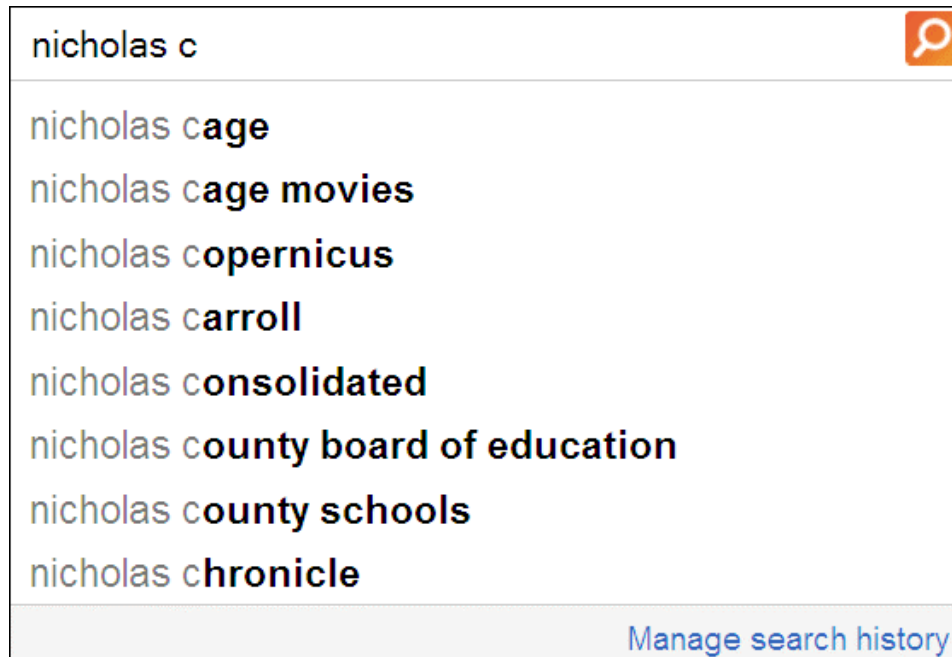


Fig. 3. Example of drop-down contextual search suggestions.

With exceptions, such as the European search engine Kartoo (<http://www.kartoo.com>), interfaces such as topic maps and link maps show little sign of entering the mainstream of WWW search, and thus do not affect SEO.

Blended search – the blending into the primary WWW SERPs of news, images, maps, videos, and other types of records once considered niche searches – has now been instituted by major search engines.(38). This presents opportunities to the more aggressive SEO practitioners, some of whom are attempting to dominate blended SERPs with a mix of varied data formats. (Blended search has also been called universal search.)

Specialized Searches

Specialized searches have been offered for decades through pre-Web online information providers. They are often called vertical searches (i.e., niche search). Specialized search could disrupt current SEO strategies, whether it is offered by major search engines or niche suppliers. First, if niche searches or tools proliferate in the GYM group (Google, Yahoo, Microsoft) – each with its own classification protocols – they may dilute the value of SEO methods that address the GYM standard algorithms. Second, niche search engines such as www.vivisimo.com have entered the search race, and others such as www.northernlight.com (founded 1996) are reentering public search; these specialized search engines are usually less susceptible to technical methods, as they tend to use proportionately more human editing than broad reach search engines. Either trend could diminish the importance of technical- or link-based SEO.

Personal search has been a holy grail of the Internet industry since the mid-1990s. Early attempts at personal search included Yahoo! personalization features and numerous failures in push technology.(39) The current trend is towards analyzing individual users' search requests in the context of their previous searches (remembered by the server) and delivering customized results. Taken to an extreme, this could theoretically filter out a majority of available web pages regardless of the SEO efforts invested, as search engines "learn" to focus on individual users' core interests.

Local search (geographic search) has become a serious goal for search engines. However other than travel destinations, few small retail businesses have websites, so search engines cannot easily find data for them by crawling the WWW. The primary solution to date has been to buy or barter for data such as traditional telephone book databases. Other new local search companies are aggregating local data for search engines, often crawling online local directories for source data, and search engines themselves are encouraging data entry from local organizations. Some trades with guilds such as law or medicine have niche websites with search functions; as yet few are comprehensive or definitive. Trades without guild organizations such as auto repair or beauty services may depend entirely on the success of aggregation if they are to be located through broad search. In few of these cases do current SEO methods provide a clear path into SERPs, and paid listings may vie with SEO in creating online exposure for local businesses.(40)

Alternate Search Channels

Alternate search channels could have an even more dramatic effect on SEO practices than blended or niche search services:

SEO for mobile communications devices – Analysis of users' WWW search behaviors on mobile devices has shown an emphasis on local retail search followed by entertainment. Coupling mobile search with GPS navigation feedback opens the prospect of delivering content or ads raises the prospect of directing mobile users to the nearest restaurant or a movie showing, thus both of these searches are potentially lucrative advertising venues – search engines are pursuing them, and SEO will follow. It is not yet certain whether search engines will convert conventional web pages to mobile-friendly formats, or whether organizations will have to create new pages specifically targeting mobile devices.(41, 42)

SMO (Social Media Optimization) – It is arguable how much SMO involves "search" in the sense of users searching by keywords, because the spread of information on social networking websites is largely viral (self-promotion coupled with word-of-mouth). SMO is currently practiced by creating profiles on popular social media sharing and news web sites, building a large base of "friends" and contributing unique promotional content with the option for other community members to vote in favor or against. Popular content that goes "hot" is placed on the high traffic home pages of the social media sharing and news sites, sometimes generating tremendous amounts of exposure, direct traffic and secondary effect inbound links from bloggers that post about the content. SMO also involves using software to make posts

automatically, this latter method being simply spamming. There may be a growing synergy between SEO and SMO, and SMO may become a parallel profession lumped under the umbrella of search engine marketing.(43, 44)

USER BEHAVIORS IN SEARCH

User behavior while searching online shows more consistency than change in the period from 1995 to 2008. This section will briefly address the consistent behaviors and available research on the course of changes.

Users' search behavior has a strong effect on SEO, since many users click the first listing in a SERP. On the other hand, gaining a #1 SERP position for a particular search term does not guarantee that a user will click that link; they might click the #2 link or the #10 link if those page titles, descriptions, or URLs are more compelling.

In the same vein, there is no clear evidence that users who click the first record will take further action; many SEO practitioners believe the more motivated users will scan an entire SERP before deciding which link to click. Regardless, user behavior while visiting SERPs affects SEO decisions as well as the search engines' goal of relevance.

Research into specific behaviors has revealed a great deal about what users do in the specific environment of a search engine interface when tested in a laboratory setting. It leaves unanswered many questions about why users search the way they do – in part because most research has focused on what users do when looking at a SERP, without inquiring how or why they arrived at that particular SERP, or what actions they take after clicking a particular link. More research is required to develop consistent conclusions. Substantial research from information science has yet to be incorporated in SEO; when it is, it may transform strategies and tactics.

Variables that affect user search behavior:

- Users' level of subject knowledge
- Improvements in results returned by search engines
- Users' confidence in the quality of search results
- Degree of users' sophistication in search
- Presentation style of individual search results (title, description, URL, etc.)
- Motivation/curiosity/laziness
- Habituation among users (hardening behaviors)
- Search engine interface design
- Number of search engines available

Basic Behaviors

Few users methodically click search results in sequential order, from the first result to the bottom of the page. Typically they skip over unappealing titles or URLs, and may bounce back and forth between organic search listings and paid listings. (Depending on SERP design, users may not always know the difference between paid and organic

listings.) If a relevant web page is not found quickly, users may change search engines, change their search term, or migrate to a general information website such as a dictionary, encyclopedia, or user-fed Q&A website.

Searching Popular Subjects

Patterns of search seen on Wordtracker, Yahoo! search tool, or KeywordDiscovery continue to show a preponderance of searches for celebrities and popular news, as described since 1999 in the weekly "Lycos 50" listing of the most popular topical searches (<http://50.lycos.com/archives.asp>) with written analysis and opinion. Google Trends (<http://www.google.com/trends>), which shows the number of searches for chosen terms in chart format, is well-suited to display spikes in interest, relative interest in different subjects, or seasonal patterns in popular searches; like the Lycos 50, it demonstrates that a huge proportion of web searches are for popular subjects rather than niche information.

Use of Boolean Syntax

In the 1990s, WWW search engines typically limited the use of search syntax to Boolean AND, OR, and NOT, with OR as the default. (With AND or NOT typically applied by +/- symbols.) Initially these options were only available by typing them into the search term entry box. HotBot search engine was one of the earliest to offer Boolean logic to untrained users, with the addition of a drop-down box offering AND as well as "exact phrase" options, but the options were rarely used.

In the late 1990s, search engines moved towards a default (or forced) Boolean AND; by early 2003, it was the default on all major search engines search engines.(45) Although this change narrowed search results, it also hugely increased the relevance of results, and, at the same time, reduced the average user's motivation to learn Boolean syntax. Some data samples suggest that Boolean search skills are known to a smaller percentage of users today than in 1997, though it is uncertain whether the decrease is due to a loss of interest in search syntax or an influx of less sophisticated users.(46)

Search Engine Loyalty

Available statistics strongly indicate that users are now more likely to refine or expand their search term than switch to another search engine.(47) The disinclination to switch is often called "loyalty" in a broad sense; motivating factors could be the generally improved quality of search returns over the last few years, or users' increased recognition of the value of using precise search terms. In terms of user behavior, this could be described as a slight shift from berrypicking (wandering about and collecting bits of information) to more linear search behaviors.(48)

Hardened User Behaviors

Users' search behavior search engines may have become "hardened," or fixed, by their expectations, in particular in the U.S. and Canada, because the majority of popular search engines have presented returns as linear lists since the mid-1990s. Current SEO is predicated on the idea that users type keywords into a box, click a "Search" button, and scan a list from top to bottom.(49) Because they represent a

large conceptual shift, interfaces that present information in clusters, topic maps, or relationship maps face an entry barrier.

User Behavior Worldwide

Patterns of user behavior outside the U.S. and Canada do not show identical evolution. While less analysis has been done, in regions as disparate as England and China users are apparently more willing to examine entire SERP pages, as well as second and third pages, and less likely to click on the first or first few SERP positions.

CONCLUSION

Beginning in the mid-1990s, search engine optimization evolved from placing random keywords in all possible parts of web pages, to more focused doorway page strategies targeting particular keywords. By 2000, improvements in search engines were rendering most of such tactics obsolete, and SEO practitioners moved on to the integration of overall website structure, and then on to establishing a website's relationship with the World Wide Web as a whole through linking.⁽⁵⁰⁾ For most, the reluctant and secondary focus of SEO was on users or quality of content. For others, a focus on user experience as well as SEO provided long term, sustainable results throughout the evolution.

SEO now faces broad changes in search and user behavior. The growth of online information is outpacing the indexing rates of all search engines. Search engines are struggling to deal with that overload, and in the process changing their strategies about what information to present and how to present it. Hardware – notably mobile communications devices – is redefining the technical limits of information presentation and also creating new niches in user search behaviors.

###

ACKNOWLEDGEMENTS

To saddle partners Jim Rhodes and Jim Heath who were working the range before I arrived; Lee Odden for copious, charitable, and incisive editing; Guy Shalev for catching lingering errors as the article went to press; Danny Sullivan and Gary Price for their frequent feedback; Kelly Bryan and Paula Sheil for tenaciously copyediting both grammar and meaning.

REFERENCES

General References

A timeline of SEO: <http://blog.searchenginewatch.com/blog/060417-130526> (accessed Oct. 2008)

<http://www.searchengineland.com> (accessed Oct. 2008)

<http://www.searchenginewatch.com> (accessed Oct. 2008)

<http://www.seopros.org> (accessed Oct. 2008)

http://www.sempo.org/learning_center (accessed Oct. 2008)

http://www.sempo.org/learning_center/sem_glossary (accessed Oct. 2008)

Citations

1. Sullivan, Danny. <http://forums.searchenginewatch.com/showthread.php?t=78> (accessed Oct. 2008)

2. Rhodes, Jim. *Art of Search Engine Promotion*. 1997.

<http://www.deadlock.com/promote/search-engines> (accessed Oct. 2008)

3. Heath, Jim. "Pointers on how to create business websites that work." Published online. 1995.

Earliest available version at: <http://www.viacorp.com/pointers1995.html> (accessed Oct. 2008)

4. "November 2006 Web Server Survey." Netcraft. 2006.

http://news.netcraft.com/archives/2006/11/01/november_2006_web_server_survey.html (accessed Oct. 2008)

5. *The State of Search Engine Marketing 2006: Survey of Advertisers and Agencies*.

p.5. Search Engine Marketing Professional Organization (SEMPO). 2007.

6. Web Analytics Association. <http://www.webanalyticsassociation.org> (accessed Oct. 2008)

7. CMSWatch. <http://www.cmswatch.com> (accessed Oct. 2008)

8. *Global Search Report 2007*. Wilsdon, Nick, editor. E3internet. 2007.

<http://www.multilingual-search.com/global-search-report-2007-released-free-pdf/15/10/2007> (accessed Oct. 2008)

9. "The global structure of an HTML document."

<http://www.w3.org/TR/html401/struct/global.html> (accessed Oct. 2008)

10. *Search Engine Optimization, An Hour A Day*, p. 46-47, 101. Grappone, Jennifer. and Couzin, Gradiva. Sybex, June 2006.

<http://www.yourseoplan.com/book-inside-contents.html> (accessed Oct. 2008)

11. Whalen, Jill. "Avoiding Clueless-Is As Clueless-Does SEO." <http://searchengineland.com/070927-071933.php> (accessed Oct. 2008)
12. Stamoulis, Nick "The Top 8 SEO Techniques (A Dispute)". <http://www.searchengineoptimizationjournal.com/2007/09/17/the-top-8-seo-techniques-a-dispute> (accessed Oct. 2008)
14. How to Win Sales & Influence Spiders, p. 5. Seda, Catherine. New Riders, February 2007.
15. Bradley, Steven. "Using Keyword Themes To Structure Your Site Content." 2006. <http://www.yellowhousehosting.com/resources/2006/09/19/using-keyword-themes-to-structure-your-site-content>
16. Sullivan, Danny. "What Are Doorway Pages?" Published online, 2007. <http://www.searchenginewatch.com/showPage.html?page=2167831> (accessed Oct. 2008)
17. Sherman, Chris. "In Defense of Search Engine Cloaking." Published online, 2001. <http://www.searchenginewatch.com/showPage.html?page=2157261> (accessed Oct. 2008)
18. Literary Machines. Nelson, T. H [Theodor Holm Nelson, a.k.a. Ted Nelson]. Mindful Press, 1982. (This is the earliest specification of the term hyperlinking, ca. 1965.)
19. Google PageRank and Beyond: The Science of Search Engine Rankings, Langville, Amy N. and Meyer, Carl D. Princeton University Press, 2006.
20. Kleinberg, Jon. Authoritative sources in a hyperlinked environment, Proceedings of the 9th ACM-SIAM Symposium on Discrete Algorithms, 1998. Extended version in Journal of the ACM 46(1999). Also appears as IBM Research Report RJ 10076, May 1997. <http://www.cs.cornell.edu/home/kleinber/auth.pdf> (accessed Oct. 2008)
21. Sullivan, Danny. "Google Kills Bush's Miserable Failure Search & Other Google Bombs." 2007. <http://searchengineland.com/070125-230048.php> (accessed Oct. 2008)
22. American slang definition of "juice" : influence, clout (<http://www.m-w.com>)
23. The SEObook, p.197-199. Wall, Aaron. 2007. <http://www.seobook.com>
24. Odden, Lee. "Making Sense of Linking and Site Promotion." 2007. <http://www.toprankblog.com/2007/09/linking-site-promotion/> (accessed Oct. 2008)
25. Thurow, Shari. "The Right Way To Think About Site Maps." 2007.

<http://searchengineland.com/070809-091501.php> (accessed Oct. 2008)

26. "Off Page Optimisation VS On Page Optimisation."
http://www.directtraffic.org/on_page_optimisation.htm (accessed Oct. 2008)

27. Shirky, Clay. "Power Laws, Weblogs, and Inequality." 2003.
http://www.shirky.com/writings/powerlaw_weblog.html (accessed Oct. 2008)

28. Anderson, Chris. "The Long Tail." Wired magazine. 2004.
<http://www.wired.com/wired/archive/12.10/tail.html> (accessed Oct. 2008)

29. Brynjolfsson, Erik; Hu, Yu "Jeffrey"; Smith, Michael D. "From Niches to Riches: The Anatomy of the Long Tail." Sloan Management Review. 2006.

30. Carroll, Nicholas. "The Anti-Thesaurus: A Proposal For Improving Internet Search While Reducing Unnecessary Traffic Loads." 2001.
<http://www.hastingsresearch.com/net/06-anti-thesaurus.shtml>. (accessed Oct. 2008)

31. How to Win Sales & Influence Spiders, p. 16-22. Seda, Catherine. New Riders, 2007.

32. Price, Gary. [personal communication about gaming search engines]
<http://www.resourceshelf.com>

Trends In SEO

33. <http://www.google.com/corporate/tech.html> (accessed Oct. 2008)

34. Search Marketing Benchmark Guide 2008, p. 43-50. 2008.
<http://www.sherpastore.com/Search-Marketing-Benchmark-2008.html> (accessed Oct. 2008)

35. Fusco, P.J. "SEO Best Practices: 20 Questions." 2007.
<http://www.clickz.com/showPage.html?page=3624560> (accessed Oct. 2008)

36. <http://www.enquiroresearch.com> (accessed Oct. 2008) Whitepapers and videos on user eye tracking.

37. Tagging advisory from DailyKos.com political blog:
http://www.dkosopedia.com/wiki/Standard_Tags_-_alpha_list (accessed Oct. 2008)

Trends Affecting SEO

38. Sullivan, Danny. "Search 3.0: The Blended & Vertical Search Revolution"
<http://searchengineland.com/071127-091128.php> (accessed Oct. 2008)

39. push technology: a means of automatically delivering information via the Internet to subscribers based on their choices for customized news, etc. Webster's New Millennium Dictionary of English. 2007.
Originally described by H.P. Luhn [Hans Peter Luhn] and Phyllis Baxendale as "SDI" (Selective Dissemination of Information).
<http://www.ischool.utexas.edu/~ssoy/organizing/l391d2c.htm> (accessed Oct. 2008)
40. "Guide On How To Get Your Business Listed On Major Local Search Engines, Yellow Pages Sites and Social Local Networks."
<http://www.locallytype.com/pages/submit.htm> (accessed Oct. 2008)
41. Carroll, N.; McGraw, M.; Brahms, S.; Rodgers, D. "Wireless Usability 2001-2002 Report."
<http://www.hastingsresearch.com/reports/62-wap-2001.shtml> (accessed Oct. 2008)
42. Holahan, Catherine. "The Battle For Mobile Search." BusinessWeek. 2007.
http://www.businessweek.com/technology/content/feb2007/tc20070220_828216.htm (accessed Oct. 2008)
43. Bhargava, Rohit. "5 Rules of Social Media Optimization (SMO)." 2006.
http://rohitbhargava.typepad.com/weblog/2006/08/5_rules_of_soci.html (accessed Oct. 2008)
44. How to Win Sales & Influence Spiders, p. 58-74. Seda, Catherine. New Riders, 2007.

Users

45. Sullivan, Danny. "Search Features Chart." 2001.
<http://searchenginewatch.com/showPage.html?page=2155981> (accessed Oct. 2008)
46. Hastings Research databases of real-time searches and web server logs, 1995-2008.
47. Search Marketing Benchmark Guide 2008, p. 135. 2007.
<http://www.sherpastore.com/Search-Marketing-Benchmark-2008.html>
48. Bates, Marcia .J. "The Design Of Browsing And Berrypicking Techniques For The Online Search Interface." 1989.
<http://www.gseis.ucla.edu/faculty/bates/berrypicking.html> (accessed Oct. 2008)
49. Nielsen, Jakob. "Mental Models For Search Are Getting Firmer" 2005.
<http://www.useit.com/alertbox/20050509.html> (accessed Oct. 2008)

General References to the USERS section

Mezei, Cristian. "Website And Search Engine User Behavior Analysis." 2006.
<http://www.searchnewz.com/latestsearch/senews/sn-4-20061101WebsiteandSearchEngineUserBehaviorAnalysis.html> (accessed Oct. 2008)

Search Engine User Behavior Study. iProspect. 2006.
http://www.iprospect.com/about/whitepaper_seuserbehavior_apro6.htm (accessed Oct. 2008)

Conclusion

50. Rhodes, Jim. [personal communication on early SEO practices]

Tables and Illustrations

Table 1: Tags and Text in a Simple Web Page.

Fig. 1: Generic SERP for the search term chocolate.

Fig. 2: APUPA chart expressed in SEO terms, © Hastings Research, 2005.

Fig. 3. Example of drop-down contextual search suggestions.

Copyright © 2010, Nicholas Carroll and Taylor & Francis. All rights reserved.
Please contact Taylor & Francis at www.taylorandfrancisgroup.com for reprint rights.
The author cannot give permissions for reprints or excerpts.