



SEO Book

The Basics of Search Engine Optimisation

David Burdon

Simply Clicks

June 2005

Introduction to SEO - What is SEO

Search engine optimisation – commonly abbreviated to SEO – is the process whereby a web site, or more specifically a web page or document, is constructed or amended in such a way as to improve its placement in the search engine results pages or SERPs. Search engine optimisation should not be seen as an end in itself. It is a function that should be undertaken to improve the overall commercial performance of a web site.

Good search engine optimisation will ensure that a page appears higher in the search engine results for a range of relevant, specific and valuable search terms or queries. The simple objective of SEO is to generate more valuable web site traffic. The achievement of a higher ranking against relevant search terms has commercial value for a web site because it will attract more traffic than a lower ranking. In an increasingly crowded online environment, search engine optimisation is therefore a crucial online marketing discipline.

The role of SEO is to legitimately *influence* the process of improving rankings. There are few genuine guarantees of a top placement, particularly for highly competitive search terms. Good SEO will improve a web site's ranking across a range of selected terms. However, any process whereby a search engine is illicitly manipulated in order to guarantee a high placement is referred to as spamming.

The successful execution of a search engine optimisation project requires skills in the areas of analysis, research, planning, copy writing and communication. A comprehensive search engine optimisation project is divided into four interrelated phases.

1. **Pre-site activities** – The research and planning activities undertaken before an existing or new site or page is actually touched or built.
 - Understanding your organisation's online business strategy
 - Researching your market category, customers and competitors
 - Keyword research and selection
2. **On-site activities** – The activities directly involved in the content and design of web pages.
 - Writing the title, description and keyword meta tags
 - Writing content – Body copy, titles, image tags, outbound links that reflect and enhance keywords.
 - Building internal links – Helping the search engines navigate the site
 - Site design and construction - Ensuring the web page utilises design and code that can be properly crawled and indexed by the search engines.
3. **Off-site activities** – Building a portfolio of quality inbound links to your web site.
4. **Post –site activities** – Analysing and responding to site traffic and user feedback once a web site has been optimised. Effective SEO is a continuous activity.

Changing Industry – History of SEO

The 10-year history of search engine optimisation is closely tied to the underlying growth of the internet and the development of its attendant search technologies. The three driving forces have been:

1. The growth and commercial value of the internet consumer base.
2. The rising intensity of competition in online market categories.
3. The increasing sophistication of search technologies.

With the growth in value, the rewards for success in the marketing battle have risen significantly. With an increasingly crowded internet, search has become a more important component of commercial success. Without search, how is your site found? As a result, attempts to both legitimately manage or illegitimately manipulate search results have become motivated by the greater rewards on offer.

The early days of search engine optimisation go back to mid-1990s when the internet first began to attract significant numbers of web sites and users. In those early days, emphasis was on the submission stage – getting your site placed into as many search engines as possible. The most important aspect of a search engine algorithm appeared to be entirely “on-page” based and was focused almost exclusively around meta tags and their related text.

Search algorithms could be decoded simply by analysing the results pages. During the late 1990s, ethical SEOs and spammers alike realised that search engine results could be manipulated by the simple process of adjusting a site’s meta tags to match the desired keywords. During this period there were many crude attempts by spammers to stuff meta tags with irrelevant but popular search terms. Famous spamming keyword meta tags have included “Britney Spears” on sites with nothing to do with Britney Spears. It just happened to be that Britney was one of the most searched for terms.

Google’s arrival in 1998 and the introduction of its “off-page”, link based, approach signalled the beginning of the end for the exclusively meta tag driven approach. Google was really the first engine to establish that sites carrying similar content had a propensity to be linked. Google’s strength appeared that the relevance of its results was less vulnerable to the orthodox spamming techniques of its day. Search users were attracted by its relevance to their search needs. In essence the key to success under the Google algorithm was not what your site said about itself but what the links from other sites said.

The Google spider apparently ignores keyword meta tags entirely and only the MSN spider apparently places any emphasis on them at all. Abuse of the keyword meta tag by spammers led to its downfall. Google’s subsequent rise to dominance eventually transformed the SEO industry. Google’s rise in popularity forced many competitor search engines to fall by the wayside or to be consolidated with larger parents such as Yahoo. Due to Google’s success, both Yahoo and Microsoft, through its newly independent and revised MSN search engine, have had to take on board many of the features of Google’s approach. The influence of inbound links continues to increase.

What to avoid in SEO

In recent years a number of illicit techniques have grown up to artificially manipulate a web site's ranking. These techniques are referred to as spamming or sometimes "Black hat" techniques. The "black hat" description refers to the fact that in the old western movies the bad guys always wore black hats. The core of any spamming technique is the attempt to deceive the search engine, and ultimately the site visitor, about the true nature of a web site's content.

The question is whether spamming techniques actually deliver any long term benefit. In addition, it is known that using proscribed spamming techniques can get the spammer, their client sites and organisations delisted by the major search engines. It has happened publicly in the past and the search engines particularly Google place great emphasis on their warnings. Google even has a page for reporting spamming offenders.

I have identified a list of nine types of illicit SEO or spamming techniques.

- Keyword Stuffing
- Cloaking
- Doorway Sites
- Throwing away sites
- Mirror Sites
- Hidden Text
- Tiny Text
- Link Spamming
- Comment Spam

Keyword Stuffing

Keyword stuffing is the technique of excessively using lots of keywords with express intention of influencing the search engines. Quite often this use appears in an incomprehensible or ungrammatical manner. Keyword stuffing is often used in conjunction with other spamming techniques such as cloaking, doorway sites, hidden text and tiny text.

Cloaking

Cloaking is the technique whereby the web site visible to a site visitor is entirely different from that seen by a search engine spider. The ordinary user may see one set of text and images but underneath that image, or "cloak", the site is "stuffed" with keywords. By examining the cache of a cloaked site on the Google search results, we can see that the site shows entirely different information to the spider, from that shown to the human eye.

Doorway Sites

A doorway site is a site that acts as a referring page for another site. The doorway page is highly optimised – containing hidden links and keywords that the ordinary web user never sees. The doorway site then climbs the search engine rankings but re-directs all of its traffic to the target – and perhaps poorly optimised site.

Throwaway Sites

Throwaway sites are almost always doorway sites. They are web sites built by spammers to provide a short-term and artificial boost to traffic. Once their traffic objectives are achieved they are often switched off or left to decay – hence throwaway. Throwaway sites are stuffed with links and keywords to attract and then re-direct traffic to a target web site. Typically, the spammers retain ownership of the throwaway domain. The spammers' clients initially receive large amounts of traffic. But once the throwaway site is switched off – or thrown away – the traffic comes to an abrupt halt and the clients business suffers. The clients are then effectively blackmailed into spending vast sums to retain traffic. The target web site receives no long term ranking benefits.

Mirror Sites

Mirror sites use an alternative URL to the target site but contain identical content. With automated page production, there maybe hundreds of different URLs all with the same content. This technique is sometimes referred to as domain duplication.

Hidden Text

The technique here is to fill or “stuff” a page with keywords invisible to the naked eye. This is done by using the same colour for text as for the background page. This technique is sometimes referred to as WOW, short for white on white.

Tiny Text

Tiny text is a technique of using very small text that is barely visible to the human eye. This text can be read by the engines. However, the engines will also attribute this text as spam.

Link Spamming

In many respects, due to the increasing influence of links, it was inevitable that link spamming would become an issue. Spamming of links has been a growing problem as many people have realised the importance that Google, in particular, places on links. As a significant issue it raised its head in April 2005 when Google's new release appeared to ban one of the leading SEO firms from its rankings. Few people outside of Google and the SEO firm concerned are entirely sure why this is the case. But the industry consensus is that Google are cracking down on web sites and organisations that accumulate vast numbers of irrelevant links with the sole intention of climbing the rankings.

Comment Spam

Related to link spamming is comment spam. Comment spam is where a spammer visits a publicly accessible site and deposits a comment with an anchor text link back to a designated site. Forums and blogs are typical target. This activity became identified as a major problem in January 2005 when Google took steps to prevent it from the blogs of Blogger.com. The reason was that spammers working for so called PPC (Pills, Porn and Casino) web sites were trawling legitimate blogs and posting uninvited comment advertisements with their web site's anchor text. Blogs were vulnerable because they typically possess a comment section that can be accessed without the need for passwords or even registration.

How to Optimise Your Site

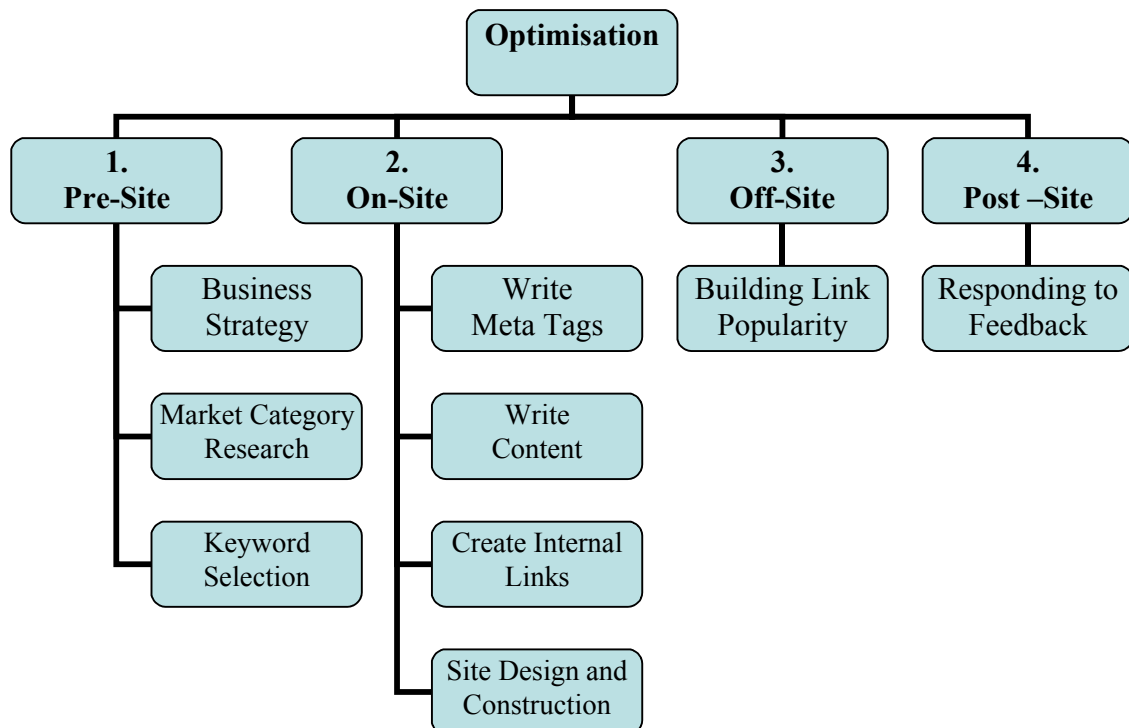
Introduction

This section describes the key processes undertaken to obtain a higher organic ranking with the major search engines.

How search engines work is part of their proprietary knowledge. The exact workings of their algorithms are closely guarded commercial secrets. However, guidance to how these algorithms (or algos) work can be found or deduced from various sources. Some general guidance is available free, directly from the search engines' own web sites. Some guidance can be found from examining the various Google and related patents. Some general guidance can be found from authoritative articles on SEO forum sites. However, real world applications of this knowledge can only be found by experimentation and trial and error.

There are some general rules. Applying them will provide a route to improved search engine visibility. The guidance in this section could be broadly applied to the three main engines – Google, Yahoo and MSN. However, given its dominance, much of the advice is derived from my interpretation of the Google “Hilltop” patent of 2001. The patent is believed by SEOs to have been the basis of the so-called Google “Florida” update of November 2003.

The Four Phases of an SEO Project



In addition to definitive information about the workings of search engines, there is much speculation, myth and rumour. There are many spurious ideas in circulation and applying them may do more harm than good. In this section, I will try to stick to tried and trusted conventions.

How Search Engines Gather Information

Search engines gather information by crawling web sites. They crawl from page to page visiting sites already known and by following the links that they find. Whilst crawling, the robots, or spiders, gather information from the source code of each site and then send back that information for indexing. The Spiders were designed to read HTML code or code related to it such as XHTML or PHP. The Spiders find it difficult to read pages written in Flash and some other popular web programmes. Spiders cannot directly read Java Script or images. They can however read the alt tags which may be provided with GIF, JPEG or PNG images.

SEO 1 - The Pre-Site Phase

Search engine optimisation is a marketing discipline. It is not a stand alone function. Before any specific optimisation activity is undertaken it is essential that two areas are non-search areas are appraised:

Understanding your Organisation's Online Business Strategy

Good SEO requires a thorough understanding of your organisation's overall business strategy. How does search fit in with activities such as advertising, e-mail and direct marketing? Is there a marketing plan? What does it say about objectives, strategy and budgets? What is the overall direction of the business and what can search contribute?

Researching your Market Category, Customers and Competitors

Good SEO also requires a thorough understanding of the market category within which the search project and web site will compete. What is the category size and how is it developing. What other channels to market are there? What information is available regarding their behaviour and attitude of customers? What role in the buying process is played by the search marketing? Who are current and likely competitors? Once the above is fully grasped you can proceed to the first real activity of SEO; Keyword selection.

Keyword Selection - Factors

Keyword selection is the first search specific discipline. Having explained that spiders read and index text, we find that some text is more important than others. That text is keywords. Valuable keywords are the words or phrases that prospective customers use when searching in your market category. Keyword selection is therefore crucial and has implications for so much else within search. I have drawn up a list of factors that should be taken into account when selecting keywords.

Category Priorities

The first thing to remember is that the number of keywords you can use on any one site or page has a finite limit. A general recommendation is that there is an overall limit of 20

individual words. In my opinion – due to other factors – the limit should be drawn much tighter than this. Rather than a limit of words, I prefer, a limit of characters – including spaces - of no more than 64. In essence, you must be sufficiently focused to sum up the key priorities of your business within this limit – typically no more than 6 to 8 words. The only way around this limit is to have an endless number of pages on an endless number of sites – all optimised, monitored and updated on a regular basis.

Search Volumes

You should use a word or phrases that have sufficient search volumes for your needs. You can find out about search volumes by checking with Word Tracker software or Yahoo's Overture keyword suggestion tool. Read more about these tools below.

Competitive Advantage

A place to look for keywords is where you enjoy some competitive advantage. How are your products or services differentiated? What are the real strengths of your business compared to your closest competitors? What proprietary advantages do you enjoy? What is it you do better that may persuade prospective purchasers to visit your site?

Competition

You may have decided on your own keyword priorities but you must also check out the competition for those keywords. Selecting a word or phrase already prioritised by a multitude of competitive sites will see you struggle for visibility. Try to find words or phrases that appear ignored or underutilised by your competitors. An alternative but higher risk approach is to see what keywords are used by competitor sites and then attempt to outmanoeuvre them by better use of links, content and meta tags.

Relevance

The keyword terms you select must be relevant, salient and part of the vocabulary used by the audience you are seeking to attract. If that audience is a consumer one it is unlikely to use jargon. The opposite may be true if you are seeking B2B prospects. My experience suggests that consumers will often use entirely different vocabulary from marketing, advertising and IT people. To avoid confusion use simpler but more specific terms.

Making your keyword choice

In essence, you must synthesise all of the above five factors in selecting and refining your keywords. Ignoring any one of the factors could create problems. Do not rush into this process. Test out your keywords by making trial searches on the major engines and see what company results you might keep. Getting it wrong may involve a large amount of reworking.

SEO 2 - The On-Site Phase

Writing Meta Tags

There is much debate about the current value of meta tags. I still find them very effective – both as an end in themselves and also as a guide to producing better and more search friendly content. Although Google apparently ignore their contents, MSN and Yahoo

both still utilise the site title and description meta tags in their search algorithms. MSN's newly launched web site still makes reference to the value of the key words meta tag.

Meta tags are so called because they sit *above* the site – in the “Head” section – and are not visible to the casual site visitor. The meta tags can be found between the <Head> and </Head> lines of HTML code, as the description suggests, at the top of the page.

Meta Tag Priorities

The Site Title Meta Tag

The site title tag is the most important meta tag. The site title meta tag is still read and indexed by all the major engines. How do we know this? Because it appears at the top of each organic search entry in the search engine results pages. However, some SEOs dispute whether it is really a meta tag at all – because the information the tag contains is clearly visible in the top left corner of the blue area surrounding the screen. The recommendation of RFC 1866, the international standard for HTML, is that the tag should contain no more than 64 characters, including spaces. There is nothing physically stopping you exceeding this limit. I have seen some major sites with 150 characters in this tag. However the typical browser can only show 70 or so characters and secondly, and with more characters, the impact of keywords within the tag is progressively diluted. From my experience the keywords in the early part of the tag carry more weight. I personally prefer a limit of 50 to 55 characters. Checking the quality of the title meta tag is the quickest way of assessing whether a site has been optimised.

A key debate, given the character limitations, is whether you should include the organisation's name in the title meta tag. Much depends on the names length and whether it includes desired keywords. My view is that with limited space, you are wasting a valuable resource if you use your organisation name here.

The Site Description Meta Tag

The site description is the second most important meta tag. It is read by the engines Yahoo and MSN and still plays a significant role in their searches. The site description should tell the engine about the nature of the web site. It is recommended that this is done in no more than 200 characters including spaces. It should be presented using good grammar and avoiding repetition. The site description should include relevant keywords.

Keywords Meta Tag

You would have read in the previous section on search history that the keywords meta tag is, today, ignored by most of the search engines. So, if the spiders do not take them that seriously, why do we still have them? In my opinion the main role they perform is one of internal guidance and discipline for the web master. If you know what keywords you are looking for it is easy to test your content to ensure there's a match.

Writing Content

Content is deemed to be increasingly important by many in the SEO field. This is apparently because with all the spamming and other optimisation techniques becoming

increasingly sophisticated, only content can give the search engines a true indication of a site's content. There are some general rules, amongst them is the more copy the better – aim for 250 words. Secondly look to use the keywords in two or three word phrases.

Content Location

Where should content with keywords be located? High up on the first page is the general rule. Certainly get keywords into the opening sentence or paragraph. The latest MSN engine picks out and uses a selection of text from the opening paragraph in their site description. This implies that the MSN algorithm is placing additional emphasis on this text. Keywords should then be spread throughout the first page and the rest of the site.

Content Relevance

Keywords and their surrounding copy should be relevant to one another. Certain words and combinations of words go together and the search engine algorithms know this. So advertising goes with marketing. Food with drink. Photographs with film. Also derivative words with different utilisation go well. So market and marketed with marketing etc.

Content Density

There is much debate about density. Too little and the keyword or phrase won't be picked up. Too much and your site may fail the spamming test. Some SEOs suggest repeating keywords no more than 7 times on any one page. Density is always measured in relative terms. A page with a lot of copy will have more word repetition than one with few words.

Titles

Text within title tags has a greater weighting than ordinary copy. Within HTML code titles are marked up <H1>, <H2>, <H3> etc. Therefore whenever a paragraph title is used it is wise to use a keyword or keyword phrase.

Alt Tags

Search engines can read the alt tags that accompany JPEG, and GIF images. Every relevant image should have an alt tag and this tag should be written to comply with your keyword objectives. The text in an alt tag is believed to be given additional weight.

Bold and Cursive Script

Both bold and cursive script are given extra emphasis by the search engines. A subtle use of bold or cursive script, when using a keyword, will enhance its presence.

Internal Links

Inbound links are important for two reasons. Firstly, their content is highlighted with a hyperlink and this is given special emphasis by the search engines and secondly it is a way of ensuring the engines can navigate and deep crawl into a site.

Outbound Links (Forward)

These are apparently growing in influence as the engines realise that inbound links (IBLs – see below) are being widely spammed. Rather link internal links, external links provide the opportunity to include keywords in the hyperlink text.

Site Map

Engines respond positively to site maps especially on larger sites with several levels. The site map is also a useful way of aiding the navigation of a spider for deep crawl purposes.

Content Change

Engines apparently respond positively to a degree of content change – this is why some blogs appear high in the rankings. Apparently, Google responds positively towards “fresh” web sites and negatively towards “stale” web sites. If content has changed between crawling cycles it signals to the spider to return again at more frequent intervals.

Business Address and Telephone Number

It is believed that engines give an additional weighting to sites that carry an address and telephone number. In many categories there are a large number of searches made using a national discriminator in the search term, so include your country in the address.

Technical Issues – Site Design and Construction

This section is about avoiding the technical mistakes or pitfalls that may hamper search engine visibility.

HTML Code

As has been explained previously, search engines were originally designed to read via HTML code or code related to it such as XHTML and PHP.

Other File Formats

As at May 2005, Google claims that it is able to read 13 different file types apart from HTML. The most common non-HTML formats are PDF and MS Office files. From my experience documents in these two formats can all rank highly. I do, however, have reservations about some of the other formats and, as mentioned above, particularly Flash.

Frames

Frames cannot be read by the major engines. So in terms of search they are an absolute no. To find out if your site is utilising frames carry out the cache test on Google.

File Size

Do not make your opening page too large. Even if an engine can read your site many internet users are still on a dial up connection.

CSS Format

CSS stands for Cascading Style Sheet. It is a way of compressing HTML code, allowing a site to load faster and, in the SEO context, improves the density and priority of keywords.

SEO 3 – The Off-Site Phase

The off-site phase deals primarily with inbound link building. Amongst the major engines Google places the greatest emphasis on links. The relevance and quality of these links has

a significant influence on the ranking of your site in all of the major engines. The search engine algorithm interprets each inbound link as a “vote” for a site. Not all links are equal and therefore the quality of the vote is important in determining the value of the vote.

Key Factors Effecting Link Quality

According to SEO convention and the information gleaned from the Google patents, there are a number of factors affecting the quality of your inbound links.

Google Page Rank

For Google ranking purposes a link from a high Page Rank site has even greater influence. A link from a PR 6+ site is extremely valuable. At the other extreme, I suggest you are prudent when exchanging links with sites of a PR of zero. The PR0 category contains a number of banned sites.

Site and Page Relevance

A link from a site and page carrying similar content would carry more influence than from a site without the similar content.

Link Density

Links from pages with fewer outbound links have more influence than from pages where there are huge numbers of links – see FFAs. Additional outbound links dilute the value of existing links on a page. My suggestion is to accept links from pages with no more than 10 to 12 links. Avoid pages with 20+ external links.

Anchor Text

Anchor text is the text that contains or sits alongside a link. This text provides additional relevance to the quality of a link. Anchor text is written in HTML. On-screen part of the text shows up as highlighted (usually coloured) or underlined type and part in normal type. The anchor text for your site could be written in HTML code as follows:

```
<a href="http://www.yoursite.com"> Your Site Title </a> - A short description of what you do. <BR>
```

Link Age

A long established link is deemed by Google to have more value than a recent link. A rapid build up in links may also be deemed spam. However Google apparently makes an allowance for a rapid build-up of links generated by news stories.

Originating Site has been Crawled and Indexed

It may sound obvious, but for search engine purposes a link is not a link if the search engines are not aware of it. The link will only exist in the records of the search engine if the page on which it is situated has been crawled and indexed whilst the link was there.

A Suggested Approach to Link Building

Obtaining a portfolio of good quality links can be a time-consuming task. There are proprietary software packages to help you in your task of tracking down web sites with

relevant content. Many organisations work on a reciprocal basis. Sites with high Page Rank may even charge for links. There are also link marketing and exchanging specialists. However, according to Google, any attempt to exchange or buy links with the explicit attempt to influence the ranking of your web site is considered link spamming. My suggestion is to start close to home, exchanging links with businesses you may have a trading relationship with. Do not rush into the link building process and do not trade links with just anyone. Poor quality links may have a negative impact on your site.

SEO 4 The Post-Site Phase

Managing and responding to search marketing feedback

Search engine optimisation is a continuous process. Having completed phases 1 to 3 your site it should start to behave positively in the rankings.

It is important that you measure the performance of your web site against the keywords you have targeted. Has the site actually risen in the rankings? If so, how far? Is ranking performance satisfactory against all targeted keywords and phrases or against just one or two? Has the improvement in rankings led to an increase in site traffic and business?

How have your competitors reacted? Do they appear aware that you have optimised your site? Has anything changed on their sites to suggest they are responding? SEO is a competitive business so you should anticipate a response to any ranking progress you achieve.

Log Files

Log files are where the records of web site activity are kept. They reveal what domains have visited the site and what pages have been accessed. The log files should be used as a check for any search activity. A regular check may reveal some useful information.

Pay Per Click

I am a particular advocate of utilising pay per click campaigns alongside organic activity. Creating and managing a pay per click campaign will generate lots of valuable data about your specific search category and the terms you use. Given the current arrangements in the pay per click market you will gather feedback about how your site and keywords perform across the Google, Yahoo and MSN networks and a few others as well.

The SEO Cycle

Concluding Phase 4 of SEO for ultimately leads you back to Phase 1. The process is essentially cyclical. The next time however, you will tackle phases 1, 2 and 3 with significantly more knowledge and experience.

How to find appropriate keywords, determine keyword saturation and the number of inbound links

Finding Appropriate Keywords

There are various methods of finding keywords. They vary from asking customers, asking colleagues or even asking friends and relatives.

A more systematic approach is to find out how many searches actually occur for each search term. Two global systems are Word Tracker and Overture keyword suggestion tool.

Word Tracker takes a large sample of searches across a number of search engine networks from around the world. Overture provides a precise picture of the volume of searches from the previous month on each of its national networks. The Overture tool is free and can be accessed directly by clicking on the hyperlink below: <http://inventory.overture.com/d/searchinventory/suggestion/>.

Word Tracker is a proprietary programme and costs money to license. However a summary of both Word Tracker and Overture search volumes can be accessed for free at www.digitalpoint.com. You need to select a word or phrase that generates sufficient search volumes but one that is relevant to your site. A two, three or four word search term is generally more useful than a single word term. However, beyond four words and the numbers of searches tends to drop dramatically.

As explained, Word Tracker is a world wide sample. And it typically includes a number of spelling variations for different English speaking countries. Ensure that the spelling variation you are checking is relevant to your geographic target.

Overture covers all the searches received on its network from whatever country you select. It differs from Word Tracker in that it tends not to distinguish between singular and plural forms or words and may also show some duplicated rather than discrete searches.

Determining the Number of Inbound Links

There are two ways of checking the presence of an inbound link. The first approach is to check the back links record held by the search engines. Go to the search box of Google, Yahoo or MSN and type "link:yoururl.com"* . A word of caution is that Google only shows a sample of the inbound links and that typically you will get much higher numbers for Yahoo and MSN links. Essentially you can use this approach to gather link information about any other web site. The results page for this query will provide the number of links and a list of sites. Some of the links will almost certainly be internal, i.e. from other pages within the same web site.

The second approach is to inspect the search engine's index of the link's originating site. This can be done for each of the major engines by using the following term "site:yoururl.com*".

- In each case enter the actual URL.

As with keyword density there are a number of SEO tools that can be accessed for free.

Google Toolbar

A particularly useful and widely used tool is the Google Toolbar. The toolbar can be accessed at:

<http://www.toolbar.google.com/>

The toolbar includes a tool displaying PageRank of any site visited.

Determining Keyword Saturation

I tend to use the SEO tools of the large SEO companies many provide free tools including keyword saturation or density.

A general purpose SEO tool

An excellent general purpose tool that I find extremely useful is called SEO Open. It is an extension to the Firefox browser. The tool can provide a complete range of on and off page data and is available to download from:

<http://www.suseroot.com/seopen/index.php> .

The SEO Open service can provide a whole range of information regarding a site. This includes Google PageRank, Inbound links recorded and the pages that have been indexed by Google, Yahoo and MSN. A keyword density checker. An internal link analyser. An HTML validator.

The Cache Information

To see what a spider can read try making a search and clicking on the cache hyperlink. Each of the major search engines provides this information. Google and MSN will even tell when the site was crawled. The Google cache allows you to see the "cached text" only. This will take you to a page without images. It shows the exact content information gathered by the Google spider including the alt text gathered from images.

The major engines have a facility whereby in response to a keyword search the target web page can be shown with the search keywords highlighted.

Another approach is to click on a text part of a website, click right and "view source". This will show the web page in HTML text format. The text shows both the meta tag and body text.

A Glossary of SEO terms.

- **Above the fold**

The section of a web page that can be read without scrolling down. The most valuable part of the search engine results pages (SERPs)

- **Alt tag**

The alternative text that accompanies an image.

- **API**

Applications Programme Interface. Used to interrogate a search engine data base.

- **Cache**

The copy of an indexed page taken when a site is crawled.

- **Conversion Rate**

The percentage of site visitors who make a position action – make an enquiry, leave their contact details, make a purchase. Also known as the “look to buy” or “look to book” ratio.

- **Cost per acquisition**

The absolute cost of gaining a new customer. Calculated by aggregating total web site costs and dividing by the total of new customers.

- **CTR**

Click Thru Rate. The percentage of search engine results pages impressions that convert to site visits.

- **Density**

The percentage of copy on a page represented by a keyword or keyword phrase.

- **DMOZ**

DirectoryMOZilla. The open directory project that supports many search engines including Google and Yahoo.

- **FFA**

Free For All. A link farm that offers unrestricted access for sites seeking inbound links.

- **GIF**

Graphics Interchange Format. An image format that when accompanied by an alt tag can be read and indexed by search engines.

- **HTML**

Hypertext Mark-up Language. The universal code of the internet.

- **IBL – Inbound Link**

A link to your web page from an external web site.

- **Index**

A search engine data base.

- **JPEG**

Joint Photographic Experts Group. An image format that when accompanied by an alt tag can be read and indexed by search engines.

- **Link Farm (FFA)**

Link farms, some times known as FFAs for Free For All are mass link sites.

Generally anybody can create a link, with few controls. The inbound links from link farms have little value.

- **OBL – Outbound Link**

A link from your site to an external web site.

- **Return of investment – ROI**

Generally calculated by dividing web site revenue by web site costs.

- **(Google) Sandbox**

A phenomenon whereby new sites are excluded from the Google SERPS, for competitive phrases, for an undetermined period of time.

- **Search query**

A word or phrase entered into a search box.

- **SERPs**

Search engine results pages. The pages shown by a search engine response to a search query.

- **Source Code**

The HTML code behind a web page

- **URL**

Uniform Resource Locator. The internet address of a web page or document.

- **W3C**

The World Wide Web Consortium. The body that coordinates and governs the rules for operating the World Wide Web.