SEO—short for Search Engine Optimization—is the art, craft, and science of driving web traffic to web sites.

Web traffic is food, drink, and oxygen—in short, life itself—to any web-based business.

Whether your web site depends on broad, general traffic, or high-quality, targeting traffic, this PDF has the tools and information you need to draw more traffic to your site, and build your bottom line. You’ll learn how to effectively use PageRank (and Google itself—effective use of SEO means understanding how Google works), site design and analysis tools, SEO best practices, and much more.

When you approach SEO, you should take some time to understand the characteristics of the traffic that you need to drive your business. Then go out and use the techniques explained in this PDF to grab some traffic—and bring life to your business.
Search Engine Optimization

SEO—short for Search Engine Optimization—is the art, craft, and science of driving web traffic to web sites.

Web traffic is food, drink, and oxygen—in short, life itself—to any web-based business.

Some web sites depend on broad, general traffic. These businesses need hundreds of thousands or millions of hits per day to prosper and thrive. Other web businesses are looking for high-quality, targeting traffic. This traffic is essentially like a pre-qualified sales prospect: already interested and able to buy your product.

This PDF has the tools and information you need to draw more traffic to your site, and build your bottom line. You’ll learn how to effectively use PageRank and Google itself—effective use of SEO means understanding how Google works: how to boost placement in Google search results, how not to offend Google, and how best to use paid Google programs. You’ll also learn how to best organize your web pages and web sites, apply SEO analysis tools, establish effective SEO best practices, and much more.

When you approach SEO, take some time to understand the characteristics of the traffic that you need to drive your business. Then go out and use the techniques explained in this PDF to grab some traffic—and bring life to your business.

SEO’s Evolution

Originally fairly narrowly conceived as a set of techniques for rising to the top of search engine listings, search engine optimization, or SEO, has conceptually expanded to include all possible ways of promoting web traffic.

Learning how to construct web sites and pages to improve—and not harm—the search engine placement of those web sites and web pages has become a key component in the evolution of SEO. This central goal of SEO is sometimes called core SEO (as opposed to broader, non-core, web traffic campaigns, which may include paid advertisements).

Search engine placement means where a web page appears in an ordered list of search query results—it’s obviously better for pages to appear higher up and toward the beginning of the list returned by the search engine in response to a user’s query.

Not all queries are created equal, so part of effective SEO is to understand which queries matter to a specific web site. It’s relatively easy to be the first search result returned for a query that nobody else cares about.
Clearly, driving traffic to a web site can make the difference between commercial success and failure. So SEO experts have come to look at search engine placement as only one of their tools—and to look at the broader context of web technologies and business mechanisms that help to create and drive traffic.

SEO is rapidly evolving into an advertising discipline that must be measured using the metrics of cost-effectiveness that are applied to all advertising techniques.

**SEO at its Core**
The core practices of good SEO are fairly simple:

- Understand how your pages are viewed by search engine software
- Take common sense steps to make sure your pages are optimized from the viewpoint of these search engines
  
  Fortunately, this essentially means practicing good design, which makes your sites easy to use for human visitors as well.
- Avoid certain over-aggressive SEO practices, which can get your sites blacklisted by the search engines

From a broader viewpoint, good SEO involves creating an effective business campaign: understanding your sales proposition and benefits, creating a strategy for drawing and converting prospects, and being better at what you do than your competition.

**Understanding Search Engines**

Search engines, such as Google, are highly complex implementations of software technology that have evolved into mega-businesses, and certainly Google is a colossus when it comes to providing access to the information you can find on the Internet.

<table>
<thead>
<tr>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>Google is far and away the most important search engine. So SEO experts tend to focus on Google. But most other search engines work basically the same way that Google does. By better understanding Google’s mechanisms, you’ll also be able to improve your placement with other search engines.</td>
</tr>
</tbody>
</table>

Effective SEO requires a basic understanding of how the pieces of search engine technology fit together.

A search engine, such as Google, implements four basic mechanisms:
• Discovery, meaning finding web sites. This is accomplished using software that travels down web links, which is sometimes called a *bot, webbot, or robot.*

• Storage of links, page summaries, and related information. Google calls the systems used for this purpose its *index servers.*

• Ranking, used to order stored pages by how important they are. Google uses a complex mechanism called PageRank to accomplish this task.

• Return of results, used to organize the display of search results, based on ranking, in response to a specific user search query.

**Discovery, Storage, Ranking, and Return (DSRR)** are all important to SEO. In particular, you’ll need to have a basic grasp of Discovery and Ranking in order to be effective with SEO implementation—so these mechanisms are explained in greater detail later in this article (see “Using PageRank”).

**Natural and Paid Listings**

Natural search result listings are listings that appear as search results without the payment of a special fee to the search engine provider. The goal of core SEO is to appear high in natural listings.

Paid listings take several forms. Some search engines (but not Google) accept *pay for play* search fees. Some consumer advocates frown upon these niche search engines because the appearance of a natural search result that is actually a paid listing can be confusing.

Google—and other first tier search engines—take the high road. If you purchase an advertisement keyed to search terms using Google AdWords, when your ad appears in response to a search query, it will be separate from the natural listings and clearly marked as sponsored.

As previously mentioned, the goal of core SEO is to obtain high natural listings. As the world of search has grown, however, the SEO discipline has also expanded. There’s no longer any stigma associated with paid listings, particularly when using a program like AdWords that appropriately labels content. If paid listings help drive the traffic that you need in a cost effective fashion, they should be considered a valuable part of your extended SEO campaign management. You can find more information about effectively using the AdWords program in my book, *Google Advertising Tools* (http://www.oreilly.com/catalog/buildinggoogle/index.html), from O’Reilly.
The SEO Advantage

If you understand SEO, you have an edge. It pays to nurture this understanding—whether you are coding your web pages yourself, working with in-house developers, or outsourcing your web design and implementation.

True, some web sites do just fine without consciously considering SEO. But by consciously developing a plan that incorporates SEO into your web sites and web pages, your web properties will outrank others that do not.

Just as success begets success in the brick and mortar world, online traffic begets traffic. (What you plan to do with the traffic, and how you plan to monetize it, are other issues.)

One way to look at this is that sites that use core SEO have an incremental higher ranking in search results. These sites don’t make gauche mistakes that cost them "points" in search engine ranking. They use tried-and-true SEO techniques to gain "points" for each web page.

Page by page, these increments give you an edge.

This edge is your SEO advantage.

What SEO Can (and Cannot) Do

SEO can drive more traffic to your web site. If you plan carefully, you can impact the kind of traffic driven to your site. This means that you need to consider SEO as part of your general market research and business plan. Sure, most businesses want traffic. But not just any traffic. Just as the goal of a brick-and-mortar business is to have qualified customers—ones with money in their pocket who are ready to buy when they walk in the door, an online business ideally wants qualified traffic.

Qualified traffic is not just any traffic. It is made up of people who are genuinely interested in your offering, who are ready to buy it, and have the means to buy it. This implies that to successfully create a SEO campaign you need to plan: this means understanding your ideal prospect, their habits and who they are, and creating a step-by-step scheme to “lure” this prospect to your site where he or she can be converted to a customer.

In addition, SEO cannot spin gold from straw, or make a purse out of a sow’s ear. Garbage sites—or sites that exist as scams—will not draw huge amounts of traffic. Or if they do, these sites won’t draw traffic for long. Google and other search engines will pull the plug as soon as they see what is going on.
As time goes by, SEO needs to be regarded as an adjunct to the first law of the web: good content draws heavy traffic. There is no substitute for content people really want to find.

While best practices SEO should always be observed, there needs to be a sense of proportion in how SEO is used. It may not make sense to create a "Potemkin Village" using SEO to draw traffic to a site if the fundamental site doesn’t yield high returns. In other words, SEO that is costly to implement is becoming regarded as one more aspect of advertising campaign management—and subject to the same discipline of cost-benefit analysis applied to all other well-managed advertising campaigns.

Driving Traffic to a Site

The first goal of SEO is to draw traffic to a site. But traffic is drawn to sites that are good. So let’s take a look at two perspectives on what kind of content draws traffic.

First, a web site worth practicing SEO upon should be a worthy beneficiary: a site with content that at least theoretically has the ability to draw traffic.

Second, one SEO technique—in the absence of this worthy content—is to create it from scratch. So bear in mind that creating content to draw traffic is one of the most effective—and simplest—SEO techniques. As such, it’s worth having a look at content that draws traffic.

Note

In case you haven’t noticed, there’s a web content arms race going on. For Google—and other search engines—to stay in business, the results they deliver have to be meaningful to users. This means that there’s only so much content spam that Google will put up with.

Just as email spammers keep trying to outwit email filters, and filters keep getting better in response, content spammers and Google are involved in an arms race.

Your goal with SEO should not involve content spam, sometimes referred to as black hat SEO. Instead, be one of the good guys, a white hat SEO practitioner.

SEO skills are SEO skills whether the hat is black or white. But use your white hat SEO skills to draw even more traffic to sites that are already good and useful. In other words, SEO should be used in an ethical and legitimate fashion as a branch of market research and business promotion for products and services that are genuine—and not as part of a scam to rip people off.
**Tip**
Most sites need to present breadth—meaning many pages that are organized cohesively. A site that presents a single web page is unlikely to present sufficient depth of content to justify extensive SEO.

The SEO trick is to draw traffic with desirable content, then to "seduce" the traffic into portions of the site that may not directly have anything to do with the content—this is the ultimate goal of the SEO campaign.

**Content That Draws Traffic**
There’s a great deal of variation in good—successful—content web sites. The purpose of these sites varies from humor to practical to editorial opinions and beyond. It’s hard to generalize. But successful content sites typically do tend to fall into at least one (maybe more than one) of the following categories:

- The site is humorous, and makes visitors laugh. While humor is a matter of personal taste, and varies tremendously depending on demographics of the target, humorous features can draw a great deal of traffic to a site. One downside is that humor tends to get stale quickly. An example of a humorous site that is popular is Googlefight, http://www.googlefight.com, a site that compares the Google rankings of two terms such as “God” and “Satan.”

- The site provides a useful free service. Web services that are free (and desirable) can draw astounding levels of traffic. A good example is TinyURL, http://tinyurl.com, which provides a practical and very useful (but simple) service: it allows you to convert long, unwieldy URLs—for example, those you often see from Amazon.com when you select an inventory item—to short, convenient URLs that are easy to use in HTML code (and easy to enter in a browser). TinyURL gets hundreds of millions of hits per month, and is able to convert some of this traffic for other purposes (such as making money from advertising).

- The site is an online magazine, newspaper, or blog. Newsworthy content or interesting opinion pieces can draw considerable traffic. But the fact that anyone can put up a blog means that you’ll have to go to considerable lengths to distinguish your content.
The site provides practical information. Many people turn to the Web as their first line of approach for finding information: about technology, relationships, travel destinations, health, and much more. Adding useful articles is often one of the first steps taken by SEO consultancies to beef up traffic to a destination site. Obviously, the more closely related the articles are to the kinds of traffic you want to target, the better this will work from a SEO perspective.

The site services a community, and provides communication tools for that community. The Web is largely about community, and involving community in your site serves several SEO purposes. Not only will your community help keep your content fresh, it’s also self-selecting: if you enable your community with care it can help to tell you which topics are of interest, and what content areas will draw traffic.

**Tip**
From a SEO viewpoint, adding a blog to a site provides one very desirable feature—it helps make sure the site has fresh content. Freshness of content—preferably on a daily basis—is one of the factors that search engines use when evaluating how to rank sites.

**Enabling Community on a Site**
You might not want to program an application that enables community functionality from the ground up, but your web host may provide this software for free, versions may be available from the open source community that are also free, or you may be able to inexpensively outsource applications that facilitate community.

You can build community on a site include by providing:

- Message boards
- Chat rooms
- Calendars with information about events in a specific field
- Instant messaging applications
- Reader reviews
- Blog comments, pings, and trackbacks
Web Site Metrics and Measuring Traffic

An important component of SEO is getting a handle on web site metrics and measuring traffic. This is not as easy as it sounds, because a great many competing terms are used, and data is not always reliable.

From a SEO perspective, you need to establish a plan for measuring traffic so that you can find out objectively which SEO measures have succeeded.

How much traffic do you aspire to? Another important question, because SEO approaches will differ depending on whether you want to generate tons of general broad traffic, or if you are targeting a narrow (but significant) niche.

A good (and reasonably objective) source of information and metrics about general and high-trafficked sites is Alexa.

Alexa


Note

I mention Amazon’s ownership of Alexa because ownership can be destiny, and you should bear Amazon’s relationship with Alexa in mind when you read the opinion pieces on the site.

For example, when Alexa indicates that Google’s traffic is perhaps peaking, and that MySpace’s continues to climb, and puts the graph of each together—as they did recently—you might want to consider whether this is unbiased information or speaks to a rivalry with Google.

On the Alexa site, you can click the Traffic rankings tab to see an ordered list of the top 500 Sites updated daily. The Movers and Shakers list is also interesting. It is a snapshot of the “right here and now” on the Web, and is useful for aligning your SEO efforts with Web-wide trends in real time.

It is worth spending time learning about popularity on the Web if you want to build successful sites. Alexa provides the tools you can use to see for yourself what is trafficked, and what is gaining or losing among top-ranked sites.

You can also use Alexa to see traffic statistics for sites that are not in the top 500. For almost any site that has been around a while, Alexa will give you an idea of traffic statistics, and whether it is gaining or losing traffic.

Alexa lets you enter descriptive information about your web site, which others can see if they check your site traffic using Alexa. You can also make sure that Alexa provides a snapshot of your home page along with its statistics. Since this service
is free, it is certainly worth entering a site description and monitoring your Alexa-garnered statistics.

Alexa works by collating results from users throughout the Web who have installed the special Alexa Toolbar. (If you’d like, you can install the Alexa Toolbar and help with popularity statistics.) There’s some question about the statistical validity of Alexa for less-trafficked sites because of this method of gathering data—Alexa’s results are probably skewed towards users who are already web savvy and heavy users.

Most likely, Alexa’s results are not very meaningful for sites that are ranked below 100,000 in popularity (very roughly, with fewer than 10,000 visitors per week).

**Measuring Traffic**

The metrics of web site traffic is a huge topic just by itself, and goes way beyond Alexa. There are a number of books just about web metrics—if you are interested in this topic, you might want to check out Jim Sterne’s *Web Metrics: Proven Methods for Measuring Web Site Success* ([http://www.wiley.com/WileyCDA/WileyTitle/productCd-0471450502.html](http://www.wiley.com/WileyCDA/WileyTitle/productCd-0471450502.html)), which is comprehensive and excellent, if a little dated. You can also take a look at the “Tracking and Logging” thread on WebMasterWorld, [http://www.webmasterworld.com](http://www.webmasterworld.com).

There is also, of course, quite a bit of software designed simply to help webmasters gather and understand the metrics of their sites. Also, your web server’s logs contain a great deal of traffic information that can help provide you with useful metrics.

**Note**

You can use your web server logs to find out which search queries most commonly land visitors on your pages.

Measuring traffic is a very important topic: to optimize your site you need to have baseline information, as well as feedback, so you can understand whether changes improve site traffic—and to also see which elements in your site draw traffic.
Using PageRank

Google uses the PageRank algorithm to order the results returned by specific search queries. As such, understanding PageRank is crucial to core SEO efforts to improve natural search results.

Depending on who you ask, PageRank is named after its inventor, Lawrence Page, Google’s co-founder—or because it is a mechanism for ranking pages.

When a user enters a query, also called a search, into Google, the results are returned in the order of their PageRank.

Originally fairly simple in concept, PageRank now reportedly processes more than 100 variables. Since the exact nature of this "secret sauce" is, well, secret, the best thing you can do from a SEO perspective is more or less stick to the original concept.

The underlying idea behind PageRank is an old one that has been used by librarians in the pre-Web past to provide an objective method of scoring the relative importance of scholarly documents. The more citations other documents make to a particular document, the more “important” the document is, the higher its rank in the system, and the more likely it is to be retrieved first.

Let me break it down for you:

Each web page is assigned a number depending upon the number of other pages that link to the page.

The crucial element that makes PageRank work is the nature of the Web itself, which depends almost solely on the use of hyperlinking between pages and sites. In the system that makes Google’s PageRank algorithm work, links are a web popularity contest: Webmaster A thinks Webmaster B’s site has good information (or is cool, or looks good, or is funny), so Webmaster A may decide to add a link to Webmaster B’s site. In turn, Webmaster B might return the favor.

Links from Web site A to Web site B are called outbound (from A) and inbound links (to B)

The more inbound links a page has (references from other sites), the more likely it is to have a higher PageRank. However, not all inbound links are of equal weight when it comes to how they contribute to PageRank—nor should they be. A web
Search Engine Optimization

page gets a higher PageRank if another significant source (by significant source I mean a source that also receives a lot of inbound links, and thus has a higher PageRank) links to it than if a trivial site without traffic provides the inbound link.

Note
PageRank is essentially a recursive algorithm; a given page’s PageRank is the sum of the PageRanks of the pages that link to it (weighted by the total number of links, of course). In this scheme, a link from a high PageRank page clearly counts for more than a link from a low-ranking page.

The more sophisticated version of the PageRank algorithm currently used by Google involves more than simply crunching the number of links to a page and the PageRank of each page that provides an inbound link. While Google’s exact method of calculating PageRank is shrouded in proprietary mystery, PageRank does try to exclude links from so-called link farms, pages that contain only links, and mutual linking (which are individual two-way links put up for the sole purpose of boosting PageRanks).

Tip
The easiest way to see the comparative PageRank for your web pages is to install the Google Toolbar. With a web page open, the PageRank is shown in the Toolbar on a scale of 0 to 10. These PageRanks are really between 0 and 1, so although the 0 to 10 scale is useful for comparison purposes, it does not represent an actual PageRank number.

Note that you may have to specifically turn on the feature that displays PageRanks in the Google Toolbar; in some installations this feature is not enabled by default.

From the viewpoint of SEO, it’s easy to understand some of the implications of PageRank. If you want your site to have a high PageRank, then you need to get as many high-ranked sites as possible to link to you. Paradoxically, outbound links reduce the PageRank of the linking site because they reduce overall traffic on the linking site (users are more likely to leave the original site if they have several links they can click).

However, useful outbound links draw traffic to the linking site and encourage other sites to return the favor because they respect the quality of the links the original site provides. So for SEO there’s a delicate balancing act with outbound linking: some quality outbound links add merit to a site, but too many outbound links decrease desirability. Trial and error is probably the only way to get this one right.
Getting Technical: PageRank and Random Surfers

The PageRank formula can be thought of as a model of user behavior of “random surfers.” Such a random surfer visits a random web page, keeps clicking links randomly, never clicking the back button, and eventually gets bored enough to visit a new random page by typing in the web address into the browser. The probability that the random surfer visits a particular page is its PageRank. The probability at each page that the random surfer will get bored and request a new random page is called the damping factor, represented by \( d \) in the formula.

Put this way, the PageRank for a specific web page can clearly be calculated by going through all the inbound links to a page, calculating the PageRanks of all these pages, backing up to calculate the inbound links in turn to the new set of pages, and so on, all the way back until there are no more inbound links. A little more technically, a web page’s PageRank can be calculated by iterating recursively through all of its inbound linked pages. This is the fundamental method behind Google’s search engine, although in the real world (as you likely know if you’ve read this far in this sidebar) there are usually non-recursive techniques that calculate results more quickly than the corresponding recursive algorithm.

The original formula for PageRank with further explanation is contained in the Brin and Page page at Stanford University (http://www-db.stanford.edu/~backrub/google.html). Here it is (PR stands for PageRank; \( A \) stands for a random page, identified as Page A; \( T_1 \ldots T_n \) signifies all the pages that link to Page A; \( C(A) \) represents the number of Page A’s outbound links):

\[
PR(A) = (1-d) + d\left(\frac{PR(T_1)}{C(T_1)} + \ldots + \frac{PR(T_n)}{C(T_n)}\right)
\]

PageRanks form a probability distribution over web pages, so the sum of all web pages’ PageRanks is one.

The formula for PageRank has, of course, evolved since this formulation—and, as I’ve mentioned, now involves more than 100 variables—and its exact nature is part of Google’s proprietary technology. It’s still the case that the best insight for SEO purposes into how Google works come from this early academic formulation.

Getting Listed

Understanding PageRank doesn’t get you very far if you are not indexed in the first place.

Search Engines

If your page (or site) has inbound links from sites in a search index, then Google (or any other broad search engine) will most likely find you pretty quickly. However, it’s peculiar but true: different search engines index different portions of
the Web. Also, at any given time, it is impossible for any search engine index to include the entire Web!

To avoid being left out, it makes sense to manually add your URLs to search engines. (In early times there might be more of a delay before your sites were found, and it really made sense to make sure you were listed!)

To manually add your URL to Google, go to http://www.google.com/addurl/. You can get listed at Yahoo! at http://submit.search.yahoo.com/free/request.

**Tip**
Be careful to create pages that are friendly to webbots—which constitute the advance scouts for search engines—as I explain later in this article in “Working with the Bot.”

Getting listed is, of course, only the beginning of the battle. One of the key goals of core SEO is to get both highly ranked and listed in reference to some specific search terms. Achieving this goal requires taking affirmative control of the information that the search index may use to make ranking decisions—and that it will use to understand the quality of your pages.
Submission Tools
You may also want to use an automated site submission tool that submits your site to multiple search engines in one fell swoop.

It’s quite likely that your web host provides a utility with this functionality that you can use to submit the URLs for your hosted domains to a group of search engines. It’s in your web host’s interests to help you generate traffic, and most are pleased to provide this service.

Before using a site submission tool, you should prepare a short list of keywords and a one or two sentence summary of your site. Alternatively, you can use the keywords and description used as meta information for your site for search engine submissions.

If you Google a phrase like “Search Engine Submit,” you’ll find many free services that submit to a group of search sites for you. Typically, these free submission sites try to up-sell or cross-sell you on a product or service, but since you don’t have to buy anything, why not take advantage of the free service? The two best-known examples of this kind of site are Submit Express, http://www.submitexpress.com, which will submit your URL to 40 sites for free (just be sure you pass on the various offers you’ll find on the site) and NetMechanic, http://www.netmechanic.com, which is another search engine submission site along the same lines.

Taxonomies
Taxonomies are categorized directories of information. A directory differs from the index used by a search engine because a directory uses a structured way to categorize sites. An often-used analogy is that a directory is like the table of contents in a book, while the search engine index is like the, well, index to the book. In other words, the directory tells you how the book is organized, while the index allows you to search within the book for specific keywords.

It’s a not-so-well-kept secret that one of the best approaches for getting good placements in the search engine listings is to enter through a back door, by working with the two most important structured directories: the Open Directory Project (ODP) and the Yahoo! Directory.

Getting Open Directory Project (ODP) Listings
The Open Directory Project (ODP), http://dmoz.org, is the most important taxonomic directory on the Web. Formally hosted and administered by the Netscape division of AOL, the ODP is run along the lines of an open source project with the belief that “humans do it better.”
The ODP believes that web-automated search is ineffective, and getting worse, and that the small contingent of paid editors at commercial web search engine companies cannot keep up with the staggering rate of change on the Web—decaying stagnant sites, link rot, new sites, sites intended as search spam, and so on.

The ODP is run and maintained by a vast army of volunteer editors. These editors follow internal checks and balances to keep the integrity of the directory. See http://dmoz.org/guidelines/ for more information about the ODP review process and guidelines for site inclusion.

You, too, can become an ODP editor in an area of your interest and expertise. See http://dmoz.org/help/become.html for more information about becoming an ODP editor. This is a partially facetious suggestion, but one of the most effective ways to use SEO to promote your sites is to follow the patterns and practices of the ODP to get your sites included. You’ll find a FAQ about how to add your site at http://www.dmoz.org/add.html (this FAQ is also available via a link from the ODP home page).

The ODP taxonomy (categorization system) and the sites included in the categories are freely available as data for use by anyone who wants to run their own search engine as long as the terms of the ODP’s free-use license is complied with.

Google, and most of the major search engines, do use information derived from the ODP, but each major search engine uses it in their own way. With Google in particular, information from the ODP is used to form the Google Directory, http://directory.google.com.

Most significant, inclusion within an ODP category means that your site is very likely to be included within the Google Web index (as well as the Google Directory, and in other major web indices) as a high-ranking result for searches within that category.

So it’s worth submitting your site to the ODP, if only because it’s the best way to get indexed by search engines, including Google—and, to a significant extent, to manage how your site is categorized.

**Getting Yahoo! Directory Listings**

The Yahoo! Directory, a somewhat lesser known part of Yahoo!, works in pretty much the same way as the ODP, except that it is privately maintained Sites added to the Yahoo! Directory tend to end up in the Yahoo! index, as well as other important search indices, often with high ranking in response to searches related to the Yahoo! Directory category for the site.

You can also find the Yahoo! Directory by opening the main Yahoo! home page, selecting Directory as your search category, and searching for a term. The search results you will be presented with are from the Yahoo! Directory (not the Yahoo! Web index), and the display will show where in the taxonomy you are, so you can browse through related categories.

**Note**

Just as the Google Directory is not the Google search engine, the Yahoo! Directory is not Yahoo! itself.

**Wikis**

Wikis—and particularly the Wikipedia, found at [http://www.wikipedia.org/](http://www.wikipedia.org/)—are community-based knowledge systems. The Wikipedia, and other select wikis, turn out to be excellent grist for the SEO link placement mill. Anyone can add content to a wiki, and the content tends to be authoritative. A link to your site strategically placed in a Wikipedia article may generate considerable traffic, and is likely to boost your site’s standings with Google and other search engines.

The note of caution here is that placements should be relevant. Irrelevant spam links that don’t have anything to do with a topic in a wiki are likely to be deleted by the community quickly. In addition, SEO has grown up from the early days, and practitioners realize that what goes around comes around, and that spam is evil. In other words, craft wiki text and links with care and make it valuable content for the wiki—and, in doing so, better serve your SEO goals.

**Getting Links**

Sometimes it seems like all of life has the same themes as high school: what’s important is being popular. A significant measure of popularity on the Web is how many inbound links—links from other sites to your site—you have.

Inbound links are an important component of Google’s PageRank system—which is a way to order the sites returned from a search.

Obtaining inbound links is not rocket science, but it is labor intensive and does require some thought. The best way to get another site to link to your site is to ask for it, as obvious as that may seem.

Link farms—sites that exist for the sole purpose of providing inbound links to better a page’s search ranking—will not help your site become more popular, and are even likely to damage your standing with Google and other search engines if they are noticed.
It makes sense for sites to link to your site when they have similar or related content. This is a reasonable thing for a webmaster in charge of the other site to do because it adds value for his or her site’s visitors. (If your site is not adding value, you might want to rethink its premise.)

**The Best Inbound Links**

The best—meaning most likely to drive traffic—inbound links come from:

- Sites that publish content that is complementary and related to the content on your site.

- Hub sites that are a central repository, discussion area, and community site for a particular interest group. For example, a mention on SlashDot ([http://www.slashdot.org](http://www.slashdot.org)) can drive huge amounts of traffic to sites related to technology—so much so that the phenomenon of a sudden up-tick in traffic due to inbound links has become known as the “Slashdot Effect”.

**Finding Sites to Make a Link Request**

To find sites that are appropriate for an inbound link request, you should:

- Consider the sites you find useful, entertaining, and informative.

- Use the Web’s taxonomic directories to find sites in your category and in related categories

- Use specialized searching syntax to find the universe of sites that search engines such as Google regard as “related” to yours. For example, the search `related:www.digitalfieldguide.com` produces a list of sites that Google thinks are similar to `www.digitalfieldguide.com`.

**Making the Link Request**

The very best way to get someone to link to you is to link to them! If your readers find their content useful, their readers will likely find yours useful as well. And because many sites pay attention to the sites that link to them, you may often get a reciprocal link without any further action. However, if that doesn’t work, you may want to contact them by email.

You should be aware that a blatant request for a link is likely to be perceived as spam. Such a request is likely less effective than an email that lets the site owner know about your site, and why you think that your content might be of interest to his or her readers.
Finding email addresses

The first step in writing an email requesting an inbound link is to find the email address for the webmaster you want to contact. This can take quite a bit of poking around, but it is amazing how often you can uncover the right email address with a bit of persistence if you just look at all the pages on a web site.

If a web site has a contact form but no explicit email address, you can often find an address by viewing the HTML source code for the contact form’s page and looking for a submission address. Another place to look for email addresses is within a syndication feed. If the site provides an RSS or Atom feed, the creator’s email address is often included as part of the feed.

As you may know, you can use the Whois service of Internet domain registrars to find contact information for site owners, although with multiple domain registrars this information is more fragmentary than it used to be. In addition, some sites intentionally do not publish information about the real domain owners when they register domains, for example, by putting the domain in the name of the web host.

A good first stop if you want to try using a Whois service to get email contact information is Network Solutions (http://www.networksolutions.com), the “classic” Internet domain name registrar. Next, try Whois.net, http://www.whois.net/, which has one of the largest databases for Whois information.

If these two sources fail, do not give up. Go to Internic, http://www.internic.net/whois.html. The Internic service will not give you contact information, but it probably will tell you the specific domain registrar who registered a given site, and the address of the domain registrar’s Whois server. You can then go to the Whois server maintained by the appropriate registrar, and usually find email contact information there.

If this sounds time consuming, well, it is. To justify the time, any sites that you contact should indeed be related to your site.

Emails should not spam

Generally, you should not send email that reads like spam. Don’t send mass emailings to request links (it will probably get intercepted and marked for deletion by anti-spam filters, anyhow). Personalize each email with each recipient’s name, something about their site, and information about why they should link with you.

It’s OK to offer a reciprocal link in exchange for your inbound link. But the better approach is already to have a link to any site you approach. You can set aside a resource page for this purpose. Why bother with trying to get an inbound link from a site that isn’t worth linking to? If it is worth linking to, then go ahead and do it on your own without requiring payback. You’ll be surprised at how often the other party decides to reciprocate.
Comments, Trackbacks, and Discussion Threads

The easiest way to get inbound links is for you to post them, using a mechanism such as a blog comment, a blog trackback ping, or a discussion thread. These links do not have the permanence or credibility of a link from a stable site, but can draw considerable short-term traffic if posted on a popular site.

There’s nothing wrong with adding a link to a comment on a blog, or in a discussion thread, or using a trackback mechanism, provided you have a valid hook for hanging your URL. In other words, it’s OK to enter a discussion if you really have something to say, and it’s also OK to link back to relevant material on your site—but don’t come completely from left field. It will undermine the credibility that you are trying to build. You should also be aware that many popular blogs use the "nofollow" tag on any links inserted into comments. This tag tells search engines to disregard those links. This means that they will help humans to find your site, but will not actually help with search engine optimization.

Keeping Content Fresh

Search engines look for fresh content, and for content that is regularly freshened. Observation indicates that bots visit sites with new content daily more often than content that is not updated—and vice versa.

The moral here is that you will need to find a mechanism that works for you to keep your content fresh. As I mentioned earlier, if constantly adding pages sounds like too much work, then you might want to consider prominently featuring a blog on your site.

Using Syndication

As you may know, syndication is a simple XML-based mechanism for publishing content. Syndication feeds come in two predominant flavors: RSS and Atom. From the viewpoint of publicizing your web site, you don’t need to worry about the differences between them. (If you’d like to learn more about how to use syndication feeds, check out O’Reilly’s What Are Syndication Feeds PDF, available at http://www.oreilly.com/catalog/syndicationfeeds/cover.html.)

You syndicate content by encoding it within an RSS or Atom feed. This feed can—and usually does—include links to the site originating the content.

Subscribers can view syndication feeds in all different kinds of software, including web browsers, email clients, standalone programs, and on HTML web pages. There’s no mechanism built into syndication to pay for subscriptions, but once you are subscribed, your feed display is automatically updated when a new item is added to the feed. It’s up to the syndication viewing software to decide how to
render feeds, but software that can display web pages often shows the underlying pages that the feed links to.

There’s some controversy about how publishers can best use syndication feeds, since it’s not obvious how to make money from them. (Google has introduced a program allowing publishers to insert contextual ads within syndication feeds, but this is a controversial step.)

However, syndication feeds work well as a device for driving traffic to a site because:

- Feed content is under the control of the publisher
- Most feeds contain items that are thematically linked (and can be related to a site)
- Feed items provide content along with links back to more content on a publisher’s site
- It’s easy to distribute a syndication feed

In other words, many savvy web publishers use syndication feeds as a kind of teaser for their real web content. In addition, syndication aggregation engines drive traffic to a site, and increase a site’s PageRank (by sending the site inbound links from the aggregator).

Creating Feeds

If you maintain a blog, it’s likely that you are already publishing a syndication feed (even though you may not be aware of it). Check your blog templates to see if there is a template for an index.xml, index.rdf, or atom.xml file. If so, have a look at the root directory for your blog. Voila! You’ll probably find a syndication feed. You may want to tweak the template tags to make sure that you are syndicating the content you want, and only the content you want.

If you don’t have a blog feed, or want to publish content other than the entries of your blog, it’s easy to construct an XML syndication feed by hand using a text editor. There are also a great many tools available to help you construct your own feeds. Some tools use a Wizard interface, so you don’t need to know anything about coding in XML to create a syndication feed.

You’ll find links to these tools and to syndication resources in general on the RSS Compendium, online at http://allrss.com/, and on O’Reilly’s Xml.com site, http://www.xml.com/pub/rg/RSS_Software. General information about syndication is available in Shelly Power’s Digital Digest What Are Syndication Feeds?
Telling the World about Your Feed

Once you have your syndication feed, the key to getting some bang out of it is to get it distributed. As with a web site, in the long run, this requires constant addition of fresh content. You probably should not try to distribute a syndication feed until you have a minimum of a dozen entry items, and can reasonably expect to add at least an item a week.

You can (and should) mark your web site with a graphic that is linked to your syndication feed. To create the graphic, you can create a button using FeedForAll’s free RSS Graphics Tool, http://www.feedforall.com/public/rss-graphic-tool.htm, or you can grab a pre-made button from RSS Specifications, http://www.rss-specifications.com/rss-graphics.htm.

You also need to add code into the head section of your HTML pages to let syndication viewers and aggregators automatically know about your feed. For example, if you include this code in a page, someone visiting your site using a web browser that is capable of displaying syndication (such as Firefox) is automatically offered a subscription to the feed.

The general form of the code to be added is:

```html
<link rel="alternate" type="application/rss+xml" title="RSS" href="http://www.yourdomain.com/rss.xml" />
```

Obviously, you need to specify the actual location of your own feed when you add this code to the head section of your HTML page. For example, I maintain a syndication feed for the Googleplex Blog at http://www.braintique.com/research/mt/index.xml. The link code on my page looks like this:

```html
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">
<html xmlns="http://www.w3.org/1999/xhtml">
<head>
...  
<title>The Googleplex Blog</title>
...
<link rel="alternate" type="application/rss+xml" title="RSS" href="http://www.braintique.com/research/mt/index.xml" />
...
</head>
...
```

Submitting Feeds

The next step is to submit your syndication feed to syndication aggregators and search engines. The RSS Compendium provides a great list of sites for submitting syndication feeds for inclusion at http://allrss.com/rsssubmission.html. The RSS Specifications site also has an extensive list of sites that maintain syndication feed databases at http://www.rss-specifications.com/rss-submission.htm.
It’s a good idea to continue to submit your feeds as you add content items. If you are publishing multiple feeds, however, this can become an unpleasant chore, so you may want to use a tool that automates the process. RSS Submit is one such program—it’s available for download at http://www.dummysoftware.com/rsssubmit.html in an evaluation version or (with free updates) for $35.

The updates to RSS Submit add new syndication indexes as they come online, and make sure the submission pages for older feeds stay accurate.

**Tools of the Trade**

Driving traffic to your site is macro-level SEO strategy. The other important tools of the trade have to do with crafting your web sites—and web pages—so that they "put their best faces forward" in search engine ranking.

**Working with the Bot**

To state the obvious, before your site can be indexed by a search engine, it has to be found by the search engine. Search engines find web sites and web pages using software that follows links to crawl the web. This kind of software is variously called a crawler, a spider, a search bot, or simply a bot (bot is a diminutive for “robot”).

To be found quickly by search engine bot, it helps to have inbound links to your site. More important, the links within your site should work properly. If a bot encounters a broken link, it cannot reach—or index—the page pointed to by the broken link.

**Note**

As I explained earlier, you don’t have to wait for the bot to find your site on its own if you list your site—manually or using a submission tool—with the important search engines. However, it’s still important, even if you list your site, that it be "bot friendly."

**Images**

Pictures don’t mean anything to a search bot. The only information a bot can gather about pictures come from the alt attribute used within a picture’s `<img>` tag, and from text surrounding the picture. Therefore, always take care to provide description information via the `<img>` tag’s alt attribute along with your images, and to provide at least one text-only (e.g., outside of an image map) link to all pages on your site.
Links

Certain kinds of links to pages (and sites) simply cannot be traversed by a search engine bot. The most significant issue is that a bot cannot login to your site. (This is probably a very good thing, or we’d all be in big trouble!)

So if a site or page requires a user name and a password for access, then it probably will not be included in a search index.

Don’t be fooled by seamless page navigation using such techniques as cookies or session identifiers. If an initial login was required, then these pages can probably not be accessed by a bot.

Complex URLs that involve a script can also confuse the bot (although only the most complex dynamic URLs are absolutely non-navigable). You can generally recognize this kind of URL because a ? is included following the script name. Pages reached with this kind of URL are dynamic, meaning that the content of the page varies depending upon the values of the parameters passed to the page generating the script (the name of the script, often code written in PHP, comes before the ? in the URL). You can try this example by comparing the two URLs to see for yourself the difference a changed parameter makes!

Dynamic pages opened using scripts that are passed values are too useful to avoid. Most search engine bots can traverse dynamic URLs provided they are not too complicated. But you should be aware of dynamic URLs as a potential issue with some search engine bots, and try to keep these URLs as simple—using as few parameters—as possible.

File Formats

Most search engines, and search engine bots, are capable of parsing and indexing many different kinds of file formats. For example, Google indexes file types including: pdf, asp, jsp, html, shtml, xml, cfm, php, doc, xls, ppt, rtf, wks, lwp, wri, and swf.

However, simple is often better. To get the best search engine placement, you are well advised to keep your web pages, as they are actually opened in a browser, to straight HTML.
Google puts the “simple is best” idea this way: “If fancy features such as JavaScript, cookies, session IDs, frames, DHTML, or Flash keep you from seeing all of your site in a text browser, then search engine spiders may have trouble crawling your site.” The only way to know for sure whether a bot will be unable to crawl your site is to check your site using an all-text browser.

**Viewing Your Site with an All-Text Browser**

Improvement implies a feedback loop: you can’t know how well you are doing without a mechanism for examining your current status. The feedback mechanism that helps you improve your site from a SEO perspective is to view your site as the bot sees it. This means viewing the site using a text-only browser. A text-only browser, just like the search engine bot, will ignore images and graphics, and only process the text on a page.

The best-known text-only web browser is Lynx. You can find more information about Lynx at [http://lynx.isc.org/](http://lynx.isc.org/). Generally, the process of installing Lynx involves downloading source code and compiling it.

The Lynx site also provides links to a variety of pre-compiled Lynx builds you can download.

Don’t want to get into compiled source code, or figuring out which idiosyncratic Lynx build to download? There is a simple Lynx Viewer available on the Web at [http://www.delorie.com/web/lynxview.html](http://www.delorie.com/web/lynxview.html). You’ll need to follow directions carefully to use it. Essentially, these instructions involve adding a file to your web site to prove you own the site. The host of the Lynx Viewer is offering a free service, and doesn’t want to be deluged, but it is not hard to comply.

Using Lynx Viewer, it’s easy to see the text that the search bot sees when you are not distracted by the “eye candy” of the full image version.

---

**Note**

Even though a file opens in straight HTML in a browser, it can be generated using a server-side script. *How* a file was created is essentially irrelevant to the search engine, which cares only about the actual file that is browsed.

So check the source file as shown in a browser rather than the script file used to generate a dynamic page to see what the search engine will index.
Meta Information

Meta information—sometimes called meta tags for short—is a mechanism you can use to provide information about a web page.

The term derives from the Greek word “meta,” which means "behind" or "hidden". "Meta" refers to the aspect of something that is not immediately visible, perhaps because it is in the background, but which is there nonetheless and has an impact.

The most common meta tags provide a description and keywords for telling a search engine what your web site and pages are all about. Each meta tag begins with a name attribute that indicates what the meta tag represents. The meta tag

```
<meta name="description" ...></meta>
```

means that this tag will provide descriptive information. The meta tag

```
<meta name="keywords" ...></meta>
```

means that the tag will provide keywords.

The description and keywords go within a content attribute in the meta tag. For example, here’s a meta description tag (often simply called the meta description):

```
<meta name="description" content="Quality information, articles about a variety of topics ranging from Photoshop, programming to business, and investing."/>
```

Keywords are provided in a comma-delimited list. For example:

```
<meta name="keywords" content="Photoshop, Wi-Fi, wireless networking, programming, C#, business, investing, writing, digital photography, eBay, pregnancy, information">"
```

Meta Information Tactics

In the early days of SEO, intelligent use of meta information was a crucial element of SEO. But since anyone can add any meta keyword and description information they’d like to a site or page, the feature has been widely abused—and search engines discount meta data in favor of automated page content analysis using PageRank and other variables.

That said, the price of adding meta information—it’s free—is right, so you should add it to each page as a matter of standard SEO practice. Try to provide targeted

Tip

Users, web designers, and advertisers are often very fond of fancy graphics and graphic effects on the Web. But bear in mind that search engines are almost exclusively concerned with words (meaning plain text rather than pictures). Therefore, your SEO efforts should be focused on text, not the "look and feel" of a site. When practicing SEO stick to word-craft, and don’t get sidetracked by irrelevant visual issues.
meta keyword lists and descriptions (for example, the samples I provided just above are way too general to be helpful as SEO).

Meta keywords should be limited to a dozen or so terms. Don’t load up the proverbial kitchen sink. Think hard about the keywords that you’d like to lead to your site when visitors search.

Note
Meta tags can contain a lot more than just descriptions and keywords, including (but not limited to) a technical description of the kind of content on a page, and even the character encoding used:

```
<meta http-equiv="Content-Type" content="text/html; charset=iso-8859-1" />
```

One kind of page that really does need a meta description and keyword list—even today—is the page that primarily consists of images. Recall that I’ve suggested that as a matter of SEO you avoid such pages. But if you can’t help yourself, or have bowed to unspeakable pressure applied by your web designer, you should know that Google and other search engines won’t have a clue what is on your page—unless you describe it for the search engine using meta information.

Tip
If you want to include a phrase containing more than one term in your keyword list, quote it. For example: “digital photography”. However, there is not much point in including a compound term if the words in the phrase (“digital” and “photography”) are already included as keywords.

For the keywords that are really significant to your site, you should include both single and plural forms, as well as any variants. For example, a site about photography might well want to include both “photograph” and “photography” as meta tags.
**Site and Page Design**

*Keyword density* means the ratio of keywords that you are trying to target for SEO purposes to the other text on your pages. Getting keyword density right—enough so that your SEO goals are achieved, not so much that the search engines are "offended"—is a key goal of core SEO practice. Search engines do look for keywords, but they take away points for excessive and inappropriate keyword “stuffing.”

Even from the point of view of your site visitors, you want a nice density of keywords in your pages—but you don’t want so many keywords that the content of your pages is diminished from the viewpoint of visitors.

**Site Design Principles**

Here are some design and information architecture guidelines you should apply to your site to optimize it for search engines:

**Use text wherever possible**

For most sites, the fancy graphics do not matter. If you are looking for search engine placement, it is the words that count. Always use text instead of or in addition to images to display important names, content, and links.

---

**Meta Includes**

An include file is code referenced, and opened when it is referenced, in another file. It’s easy to use includes on the Web in many file formats, including .shtml (like HTML but with server-side includes) and .php.

As a matter of site architecture, meta information should be put in include files referenced by each web page. You can have multiple meta includes, with each include referenced by pages in a different portion of your site.

The advantage of doing this is that it is easy to modify meta information across a site (or across portions of a site with related content).
Navigability
Pages within your site should be structured with a clear hierarchy. Several alternative site navigation mechanisms should be supplied, including at least one that is text-only. The major parts of your site should be easy to access using a site map. If your site map has more than about 100 links, you should divide the site map into separate pages.

Provide static text links
Every page in your site should be accessible using a static text link.

Words and Keyword Density
By now, you probably understand that the most important thing you can do on the SEO front involves the words on your pages. There are three issues you need to consider when placing keywords on a page:

- How many words should be on a page?
- Which words belong on what page?
- Where should these be placed on the page?

Page Size
Ideally, pages should be between 100 and 250 words. Shorter than 100 words, and Google and other search engines will tend to discount the page as a lightweight. In addition, you want to include as many keywords as you can without throwing the content off-kilter. With less than 100 words, any significant inclusion of keywords is going to look like keyword stuffing—and get "points" taken off your pages. Pages that are longer than 250 words are not terrible, but do tend to diminish traffic—both actual, and measured as a per page statistic. From the viewpoint of advertising, lengthy pages waste content; 250 words is about as many as will fit on a single monitor screen, so your visitors will have to scroll down to finish reading the rest of the page if you publish longer pages. You might as well provide navigation to additional pages for the content beyond the 250 words—and gain the benefit of having extra pages to host advertising.

Tip
Do make sure you provide accurate alt attribute text for any images that are on your pages.
Choosing Keywords

Beyond the mechanics of crafting sites and pages that are search engine friendly lies another issue: what search queries does your site answer? You need to understand this to find the keywords to emphasize in your site construction—a very important part of search engine optimization.

Keyword Placement

Keywords are emphasized by their placement within a page. For example, important keywords should go in a page’s HTML `<title>` and in `<h1>` headers.

You can use some of the same keywords on your page as part of a page meta information, but meta information is not as important to search engines as the actual content of the page.

There’s no magic bullet for coming up with the right keywords to place in a page. A good starting place is the “elevator pitch” story, and related keywords, that you’ll need to develop as part of a SEO campaign.

It’s likely that you’ll want to vary keywords used in a page depending on the page content, rather than trying to stuff a one-size-fits-all approach across all the pages on your site.

If the answer is X, for example, what is the question? This is the right way to consider keyword choice. X is your web site or web page. What did someone type into Google to get there?

As you come up with keywords and phrases, try them out. Search Google based on the keywords and phrases. Ask yourself if the results returned by Google are where you would like to see your site. If not, tweak, modify, wait for Google to re-index your site (this won’t take too long once you’ve been initially indexed), and try your search again.

Ultimately, the best way to measure success is relative. It’s easy to see how changes impact your search result ranking: just keep searching (as often as once a day) for a standard set of half a dozen keywords or phrases that you’ve decided to target. If you are moving up in the search rankings, then you are doing the right thing. If you ranking doesn’t improve, then reverse the changes. If you get search results to where you want them (usually within the top thirty or even top ten results returned), then start optimizing for additional keywords.

You should also realize that the success that is possible for a given keyword search depends upon the keyword. It’s highly unlikely that you will be able to position a site into the top ten results for, say, “Google” or “Microsoft”—but trivial to get to
the top for keywords phrases with no results (such as “nigritude ultramarine” or “loquine glupe” two phrases that became the fodder for SEO contests).

The trade-off here is that it is a great deal harder to do well with keywords that are valuable—so you need to find a sweet spot: keywords where you stand a chance, but that also will drive significant site-related traffic.

**Tip**

Since feedback is ultimately determined by financial incentive, an interesting approach to keyword selection is to see what words cost the most to advertisers. If you are registered with Google AdWords, you can use the AdWords tools to do just that—and get valid cost estimates for keywords and phrases.

**Keyword Placement**

The text on your web page should include the most important keywords you have developed in as unforced a way as possible. Try to string keywords together to make coherent sentences.

Not all text on a page is equal in importance. First of all, order does count: keywords higher up in a given page get more recognition from search engines than the same keywords further down a page.

Roughly speaking, besides the body of the page itself and in meta information, you should try to place your keywords in the following elements—presented roughly in order of descending importance:

- **Title**: putting relevant keywords in the HTML title tag for your page is probably the most important single thing you can do in terms of SEO
- **Headers**: keyword placement within HTML header styles, particularly `<h1>` headers towards the top of a page, is extremely important
- **Links**: use your keywords as much as possible in the text that is enclosed by `<a href="">...</a>` hyperlink tags on your site in outbound and cross bound link. Ask webmasters who provide inbound linking to your site to use your keywords whenever possible
- **Images**: include your keywords in the alt attribute of your HTML image `<img>` tags
- **Text in bold**: if there is any reasonable excuse for doing so, include your keywords within HTML bold (`<b>`...`</b>`) tags
Inbound, Outbound, and Cross Links

The links on your site constitute a very important part of how Google and other search engines will examine, analyze, and rank your pages. Links can be categorized into inbound links, outbound links, and cross links:

- Inbound links point to a page on your web site from an external site somewhere else on the Web
- Outbound links point from a page on your site to an external site somewhere else on the Web
- Cross links point between the pages on your site

Broken Links

Broken links are links that do not work, either because they have been miscoded or because the pages they point to don’t exist (perhaps it has been moved).

It’s quite important to a search engine that none of the links on your site are broken. It shouldn’t be that big a problem to go through your site and check to make sure each link works manually. Doing this will also give you a chance to review your site systematically, and understand the navigation flow from the viewpoint of a bot.

Even though you’ve checked your links manually, you should also use an automated link-checking tool. Quite a few are available. A good choice is the simple (and free) link checker provided by the World Wide Web Consortium (W3C) at http://validator.w3.org/checklink. All you need to do is enter the domain you want checked and watch the results as the links in your site are crawled.

Inbound Links

You want as many inbound links as possible, provided these links are not from link farms or link exchanges. With this caveat about inbound linking from “naughty neighborhoods” understood, you cannot have too many inbound links. The more popular (and the higher the ranking) of the sites providing the inbound links to your site, the better.

Inbound links are considered a shorthand way of determining the value of your web site, because other sites have decided your site has content worth linking to. An inbound link from a site that is itself highly valued is worth more than an inbound link from a low-value site, for obvious reasons.

Outbound Links

The “everything in moderation” slogan is really apt when it comes to outbound links. You could also say that the “outbound link giveth and the outbound link
taketh.” Here’s why: you want some respectable outbound links to establish the credibility of your site and pages, and to provide a useful service for visitors. After all, part of the point of the Web is that it is a mechanism for linking information, and it is truly useless to pretend that all good information is on your site. So on-topic outbound links are themselves valuable content.

However, every time your site provides an outbound link, there is a probability that visitors to your site will use it to surf off your site. As a matter of statistics, this probability diminishes the popularity of your site, and Google will subtract points from your ranking if you have too many outbound links. In particular, pages that are essentially lists of outbound links are severely penalized.

If you follow the word-per-page guideline I make in "Words and Keyword Density"—a little less than 250 words per page—you’ll get best results if you try to provide at least two or three outbound links on every page and no more than ten or fifteen per page.

**Cross Links**

Cross links—links within your site—are important to visitors as a way to find useful, related content. For example, if you have a page explaining the concept of class inheritance in an object-oriented programming language, a cross link to an explanation of the related concept of the class interface might help some visitors. From a navigability viewpoint, the idea is that is should be easy to move through all information that is topically related.

From a SEO perspective, your site should provide as many cross links as possible (without stretching the relevance of the links to the breaking point). There’s no downside to providing reasonable cross links, and several reasons for providing them. For example, effective cross-linking keeps visitors on your site longer (as opposed to heading off-site because they can’t find what they need on your site).

One reason for cross linking is that ideally you want to have dispersal through your site. You may have established metrics in which one page that gets 100,000 visitors is doing well. In this case, 100 pages that each get 10,000 visitors should be considered a really great success story. The aim of effective cross-linking is to disperse traffic throughout the pages of relevant content on your site.

**SEO Analysis Tools**

You’ll find a wide range of SEO analysis tools available to help you optimize your web pages and sites. From the viewpoint of your pocketbook, there are two kinds of SEO analysis tool:

- Free tools: these free tools generally tackle a single piece of the SEO puzzle, such as generating good keywords, understanding how Google operates on
specific sites and keywords, checking who links to your sites, and displaying rankings in multiple search engines at once

- Commercial SEO analysis software that must be licensed for a fee.

**Free Tools**

There are a myriad of free SEO tools available, and even many sites that list these free tools (the compendium sites are generally supported by advertising, and must therefore practice good SEO themselves to be successful!).


Some of the most useful (and free!) single-purpose SEO tools are:


These tools can definitely be time savers, particularly if you have a large amount of content you need to optimize. The price is certainly right!

Individual tools also can serve as a reality check: by running your pages through one of these tools you can get a pretty good feeling for how well you have optimized a page. However, you should bear in mind that there is nothing that one of these tools can do for you that cannot also be done by hand given the knowledge you have learned from this article.

Individually, SEO analysis tools available on the Web can help you with your SEO tasks. However, to get the most from these tools you need to understand underlying SEO concepts, as explained in this article, before you use these tools.

Over time, as you progress with SEO, you will probably accumulate your own favorite SEO analysis toolkit.

**Commercial SEO Analysis Software**

Commercially-licensed SEO analysis software is unlikely to prove worth its cost unless you are in the business of performing SEO for numerous web sites and a great deal of content, or for a major enterprise.
If the advertising for this kind of software claims too much, beware!

Essentially, the closer a commercial SEO product comes to true enterprise web analytic capability, like that provided by Google Analytics or WebTrends, the more likely it is to provide value—but only to larger enterprises. So it’s fairly rare for software labeled specifically for SEO purposes to be worth the licensing fee—you can either do the job yourself manually, or a free tool is available.

If your need for SEO analysis falls into the enterprise category, I’d suggest you start with Google Analytics, http://www.google.com/analytics/, or WebTrends, http://www.webtrends.com/, even though these products are not specifically aimed at SEO.

If you do want to look into licensed SEO software, some of the better known commercial SEO analysis products are:

- Keyword Elite (keyword analysis) and SEO Elite (automated SEO analysis) available under commercial license from Top Software, http://www.topsoftwaredownloads.co.uk/SEO.htm.

- Clicktracks: a variety of commercial analytic packages that show how visitors react to your site, and provides SEO analysis for those sites, http://www.clicktracks.com/.


**Avoiding SEO Pitfalls**

Google—and other major search engines—urge you to avoid overly aggressive SEO practices when you build your site. Google has actually taken the trouble to spell out SEO practices it regards as naughty (the list can be found at http://www.google.com/webmasters). You should pay attention to this list as if it speaks the mind of every major search engine, not just Google. Google’s position that building sites that get highly ranked is simply a matter of providing useful content isn’t totally off-the-wall, although it assumes a world where everything always works.

According to Google, good search-engine-citizen web sites do not:

*Employ hidden text or links*

For example, users cannot read white text on a white background (and will never even know it is there). But the search engine will still parse this text. This rule comes down to making sure that the search engine sees the same thing that users view.

*Cloak pages*
Also called stealth, this is a technique that involves serving different pages to the search engine than to the user.

*Use redirects in a deceptive way*

It’s easy to redirect the user’s browser to another page. If this is done for deceptive purposes—for example, to make a user think they are on a page associated with a well-known brand when in fact they are on a web spammer’s page—it’s frowned upon.

*Attempt to improve your PageRank with dubious schemes*

Linking to web spammers or bad neighborhoods on the Web may actually hurt your own PageRank (or search ranking), even if doing so provides inbound links to your site.

Bad neighborhoods are primarily link farms or link exchanges—sites that exist solely for the purpose of boosting a site’s inbound links without other content. Web spammers are sites that disguise themselves with pseudo descriptions and fake keywords—the descriptions and keywords do not truly represent what the site contains.

*Practice keyword loading*

This is the practice, beloved by SEO “experts,” of adding irrelevant words to pages (the page can then be served as the search result based on a query for the irrelevant words that actually don’t have anything to do with the page content). Proper keyword usage highlights your legitimate content but does keyword load.

*Create multiple similar pages*

Google frowns on the creation of pages, domains, and subdomains that duplicate content—although obviously there are places to legitimately duplicate content.

*Present “doorway” pages*

Pages created just for search engines are sometimes called doorway pages. (The term doorway page covers a variety of techniques that are used to substitute one page for another—either by redirection or actual substitution of pages on the web server—when the first page is optimized for specific keyword searches, and the page to which the user is actually sent has little or nothing to do with that search.)

*Pages that lack content*

Google frowns on pages that lack original content, such as a page that exists simply to present affiliate links.

*Create domains with the intention of confusing users*

Likely you’ve landed on a site with a domain name that’s confusing because it’s sharing the same name with a different domain suffix (for example,
Best Practices: SEO That Works

SEO in the broadest sense works when you carefully consider and plan your campaign.

This means having a good understanding of:

- Your offering
- Your unique selling proposition
- Your competition
- The resources that can be used for your campaign
- Best SEO practices

This section will help you achieve clarity on these points, and provide a best practices SEO checklist.

Create an Elevator Pitch

A good way to begin planning your SEO campaign is to create an elevator pitch. An elevator pitch is a one-or-two sentence "story" that makes your offering and target audience crystal clear. So named for something one could deliver between stops on an elevator, one of the most important aspects of the elevator pitch is its succinctness.

A good elevator pitch is dramatic, tells your story, and can be used for sales purposes. Most important, if your elevator pitch is right, you should be able to use it to understand the most important keywords to use from a SEO perspective.

Note

You can use your elevator pitch to generate meta information for your site.

Know Your Competition

Learn from your competition. Compile a provisional list of keywords and phrases you want to target with SEO. Then Google, Google, Google these keywords like crazy.
Examine the results of these searches carefully. On the one hand, the search results from your targeted keywords should be related to your offering. If there’s no connection—if the sites you come up with belong to completely different realms—then something is wrong and you should reexamine your SEO campaign premises.

On the other hand, your site and pages need to be at least as good as the competition’s offerings. Otherwise, SEO or no SEO, you won’t pull ahead of the competition in natural listings.

What can you learn from the competition? How can you improve upon it?

**Use AdWords as a Search Engine Optimization Tool**

As core SEO becomes simply one of the tools used to advertise destinations on the Web, it makes sense to supplement pure SEO with targeted CPC—Cost per Click—programs, such as that provided by Google AdWords.

AdWords should be used as an auxiliary to SEO programs, not as the primary focus of a SEO campaign. Attention should be paid to distinguishing within AdWords between content and search ads.

**Note**

Search ads, which appear on Google search results pages, are generally worth paying more for per click than content ads, which appear on web pages enrolled in the Google AdSense program.

As an auxiliary SEO tool, the best approach with AdWords is oblique: don’t target your primary SEO keywords (you can rely on SEO for this) as much as conceptually related topics that might draw visitors who might otherwise overlook your site. For example, a relationship and dating site might use AdWords to target people interested in Meg Ryan movies such as *Sleepless in Seattle*.

**Tip**

A benefit of using AdWords as part of your SEO campaign is that it is very easy to track results because you know when your targets get clicked.

**Work with Consultants**

SEO consultancies have become big businesses, and many of them are reputable outfits that will help you implement the practices outlined in this article.
But beware: this is also an arena that tempts shady operators. Review SEO success claims with some degree of skepticism, and make sure to verify claims whenever possible. Remember, if something sounds too good to be true, it probably is.

**Best SEO Practices**

No matter what the size of your organization, if you are mounting a SEO campaign, you should consider implementing some best practices.

The matrix below presents the most significant best SEO practices and indicates whether you should consider implementing the practice if you are a:

- A: Small business
- B: Entrepreneurial web-based business
- C: Large enterprise

Obviously, different businesses will have differing priorities, and which aspects of SEO are most important to you will vary. But the matrix in Table 1 gives you a starting place to evaluate what steps you may wish to undertake.

**Table 1: Best SEO Practices Matrix for Different Size Organizations**

<table>
<thead>
<tr>
<th>Best SEO Practice</th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Review traffic to your site using web server logs and third-party software</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Understand the content and queries that draw your traffic</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Establish base-line metrics for site traffic, and criteria for success</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Regularly monitor site traffic</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Create content specifically for SEO purposes</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Work to improve PageRank</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Submit to search engines, manually or with a submission tool</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>List your site in taxonomies</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>List your site in wiki articles</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Request inbound links</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Post comments and pings containing links</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use syndication for SEO purposes</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Create a bot-friendly site</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Review your site in a text-only web browser</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>
Add meta information to your site  X X X
Use text rather than graphics when possible  X X X
Add descriptive alt attributes to your images  X X X
Create an easily navigable site  X X X
Keep pages between 100 and 250 words  X X X
Keep keywords dense without stuffing  X
Add keywords to <title> and <h1> tags  X
Add moderate outbound linking  X X
Make sure to include cross links  X
Assemble toolkit of free SEO analysis tools  X X X
Deploy web analytics platform (Google Analytics or WebTrends)  X
Avoid dubious SEO practices  X X X
Create a clear and brief campaign plan  X X
Use AdWords in Conjunction with Core SEO  X X X

Summary
SEO is an important aspect of advertising campaign management on the Web. Core SEO impacts almost every aspect of web site and page design. Careful and responsible use of SEO tools and techniques can greatly increase your traffic to web pages, and improve your cost-benefit analysis for maintaining a web presence.

About the Author
Harold Davis is a strategic technology consultant, hands-on programmer, and the author of many well-known books, including Google Advertising Tools (http://www.oreilly.com/catalog/buildinggoogle/index.html), from O’Reilly Media. He is a popular speaker at trade shows and conventions, giving presentations on topics ranging from digital photography, wireless networking, web services, and programming methodologies.

Harold has served as a technology consultant for many important businesses, including investment funds, technology companies, and Fortune 500 corporations. In recent years, he has been Vice President of Strategic Development at YellowGiant Corporation, a company providing infrastructure for Internet marketing, Chief Technology Officer at a CRM analytics startup, a Technical Director at Vignette Corporation, a leader in customer-centric content
management, and a Principal in the enterprise consulting practice at Informix Software.

He has earned a Bachelors Degree in Computer Science and Mathematics from New York University and a Juris Doctorate from Rutgers Law School, where he was a member of the law review.

Copyright

Search Engine Optimization: Building Traffic and Making Money with SEO, by Harold Davis

Copyright © 2006 O’Reilly Media, Inc. All rights reserved.

Not for redistribution without permission from O’Reilly Media, Inc.

ISBN: 0596527861