

# SEO for Large Websites

There is a big difference between a Micro-Site and a Regular Website, and even bigger difference in approaching SEO for “**Large / Authority**” site. This EBook provides useful suggestions, strategic insights and real-world SEO tactics you can implement to sculpt and cultivate your site as it grows so you both can prosper.



**RULE #1:** Don't be intimidated by your website!



## LEGAL NOTICE

Before you scroll down and read anything in this guide, you need to be made fully aware of the following things...

**Rankings and Page Rank Disclaimer:** This document contains business strategies, search engine strategies and other business advice that, regardless of our results and experience, may not produce the same results (or any results) for you. We make absolutely no guarantee, expressed or implied that by following the advice below you will make any money or improve current profits, increase page rank, increase search engine rankings. There are multiple factors and variables that come into play regarding any given business and the competitive landscape. Primarily, results will depend on the nature of the product or business model, the conditions of the marketplace, the experience of the individual, and situations and elements that are beyond your control. As with any business endeavor, you assume all risk related to investment and money based on your own discretion and at your own potential expense.

**Liability Disclaimer:** By reading this document, you assume all risks associated with using the advice given below, with a full understanding that you, solely, are responsible for anything that may occur as a result of putting this information into action in any way, and regardless of your interpretation of the advice. You further agree that our company cannot be held responsible in any way for the success or failure of your business as a result of the information presented below. It is your responsibility to conduct your own due diligence regarding the safe and successful operation of your business if you intend to apply any of our information in any way to your business operations.

**In summary, you understand that we make absolutely no guarantees regarding search engine rankings as a result of applying this information, as well as the fact that you are solely responsible for the results of any action taken on your part as a result of this information.**

### Terms of Use, Personal-Usage License

**This document is FREE** – if you have paid for this Ebook, please send us the details in an email to: [abuse@seodesignsolutions.com](mailto:abuse@seodesignsolutions.com) so we can send this information to our attorneys. Furthermore you cannot use this Ebook in whole or partial as; a bonus offer, blog post, or anywhere online or offline without our expressed written consent.

We will ensure appropriate legal action shall be taken to preserve our brand, and to ensure that we preserve the exclusive nature and value of this product in the interest of our proprietary information.

Brought to you by: Jeffrey L. Smith  
<http://www.seodesignsolutions.com/>

SEO Design Solutions, Inc.

Copyright © 2010 All Rights Reserved

# SEO for Large Sites Part One

As your website grows, your grasp and understanding of [SEO](#) should also. The more content you have, the more keywords and topicality can overlap which may leave search engine algorithms in a quandary to parse and identify which pages have priority over others.



Don't be intimidated by your website!

Although [phrase based indexing and retrieval](#) is a prime method for extracting relevance and context in documents, there are multiple layers involved - ranging from site architecture, [internal links](#), external deep links and domain trust that contribute to rankings through relevance score.

This score is tallied (based on the least common denominator of multiple ranking metrics) to determine which pages are the best “*most relevant and authoritative*” match for a given query.

### **To Assist Search Engines We Suggest You Optimize By:**

- Defining Ranking Objectives.
- Planting Seedling Keywords.
- Determining Preferred Landing Pages
- Reinforcing Competitive Keyword Categories.
- Strengthening Internal Links, Consistently.
- Use Succinct Titles and Meta descriptions (as they can pull their own rankings alone).
- Managing Site Wide Duplicate Content (Custom templates, Noindex, Follow, 301 Redirects, XML Sitemaps).

**Define Ranking Objectives** - Depending on your keyword positioning and ranking objectives, you will need to select the best method to accomplish them. For example, if your focus is on the [Long-Tail of search](#) (meaning less competitive keyword combinations), then link prominence (links from other sites only to the homepage) and flat site architecture (all the pages in the root folder) will work just fine for accomplishing this type of result.

If the market were more specialized and you had less competition (like a high ticket item like MRI equipment), then you could elect for a smaller site with key nodes of relevance (like tips of an iceberg) to wrap up singular, plural and other popular keyword combinations over time.

By adding a series of keyword rich naming conventions and using pagination such as breadcrumb navigation you would reinforce 2 out of the 3 primary metrics search engines assess for relevance (url naming conventions and anchor text) within your own site to subsidize off page ranking factor.

The more work or more streamlined and granular your layers are optimized the less dependent the site becomes on getting a jump-start from other sites and the faster it develops its own authority (to rank on its own coherence/merit).

This strategy would differ from a large ecommerce site with thousands of SKU's (stock kept units) which means that the model numbers would be the keywords and there would be a large degree of duplicity across the website due to generic mfg. descriptions, etc.

Each keyword has a threshold and tipping point and term frequency within a document, how clearly the hierarchy is promoted within your own website and how large of a footprint your website and link profile make (compared to your competitors) all determines how buoyant your rankings will become.

**Plant Seedling Keywords** – With keywords, keyword research, modifiers and action words the last thing you want to do is work backwards after the fact. Keywords are the foundation of organic optimization and are the wireframes that eventually support the schema of what a strategic page acting as a piece of the puzzle can accomplish based on its position in the pyramid of relevance.

That pyramid is (1) stable site architecture (2) relevant landing pages based on keyword continuity and (3) a strong internal and external link profile. However, before you can progress to stage 2 or 3, you need to have your hit list of keywords ready to engage the market and search engines alike.

There is no need to recreate the wheel, just use search trends, the [Google Wonder Wheel](#), related searches (usually found in the footer in Google), [market and keyword research tools](#) or common sense to find a nice cross-section of relevance to use as your blueprint to frame out the site.

If you have a site that lacks coherence, it is not too late to add a new segment to consolidate the ranking factor, execute 301 redirects, add new naming conventions and implement a clear structure that implies a hierarchical method of consistency.

We often implement a an optimized blog on clients sites to concentrate ranking factors that allow us to bypass poorly implemented errors or dated site structures which omit crucial metrics for search engines (such as meta tags, internal links, a one size fits all template, etc).

By baking SEO into the template or new section, you know it is only a matter of time if you manage the areas which impact relevance over time.

A stable search engine ranking for a competitive keyword can take anywhere from 4-12 months, but you have to remain focused on the ranking objective over time.

It is not uncommon that when you finally give up and think you have done everything possible to acquire a competitive position (and just relax) that it takes a top 5 or #1 position. The main thing is not to get frustrated or abort inadvertently since you may not see the real result without patience.

So, instead of targeting a handful of keywords, target an entire set of keywords to build the relevance for that subject/topic. By adding internal links to reinforce each other, your pages can implement the buddy system to feed off each other and steamroll the competition since you are using a tight niche of keywords.

To provide a more thorough overview, here is a post from the past on [keywords and keyword modifiers](#) that will explain the importance of **alternative and supporting keywords** and what role they play. This post on [theme density](#) will also aid in getting the point across that by concentrating your topics intelligently, rankings are a by-product of great content.

I can see clearly that this topic will require more elaboration, so, if you haven't already, I encourage you to subscribe to our [RSS feed](#) to catch **a follow up video** and the rest of the topics outlined in this post above. This is a great topic and should have more than just an overview with bullet points.

In the follow up video and posts, I will break out each segment below in greater detail with examples to provide clarity:

- Determining Preferred Landing Pages
- Reinforcing Competitive Keyword Categories.
- Strengthening Internal Linking, Consistently.
- Use Succinct Titles and Meta descriptions (as they can pull their own rankings alone).
- Managing Site Wide Duplicate Content (Custom templates, Noindex, Follow, 301 Redirects, XML Sitemaps).

# SEO for Large Websites Part II

In our first post of this 3 part series named [SEO for Large Websites Part I](#), we touched on the importance of defining ranking objectives (from the onset) to allow the theme to resonate with continuity and planting the seeds of relevance that will eventually become pillars of stability for eventual rankings.

Now, in Part II, it's time to move on to the following metrics (that are equally as important) to round out the summary of what an optimal blueprint would look like for creating a robust [SEO](#) campaign for a large website (which ultimately has different needs than a smaller website).

- Define Ranking Objectives (covered in Part 1)
- Plant Seedling Keywords (covered in Part 1)
- Determine Preferred Landing Pages
- Reinforce Competitive Keyword Categories.
- Strengthen Internal Linking, Consistently.
- Use Succinct Titles and Meta descriptions (as they can pull their own rankings alone).
- Manage Site Wide Duplicate Content (Custom templates, Noindex, Follow, 301 Redirects, XML Sitemaps).

**Determining the Preferred Landing Page** – This is possibly one of the most important aspects to an organic optimization campaign. Most simply assume that [their homepage](#) is the most relevant place for conversion, yet, nothing could be farther from the truth.

You should know that instead of just entering the front door/homepage that people will enter a website from the side door, the back door, on a whim from an iframe from a social media site, a Google blog alert from Google Blog Search, an RSS feed reader or otherwise.

So, depending on your homepage excessively (like a bird who has not left the nest) **is not an option**. As your website grows and develops more authority, unless you want to leave a pile of cash on the table for your competition to devour you need to infuse purpose **“into every page”**.



Each page MUST establish **what its purpose is** and **clearly communicate a core benefit to the user** in order to facilitate conversion – so this is where your [preferred landing pages](#) come in.

If you understand that **each page can rank for specific keywords on its own merit** and understand that it is a process well within your control, then it doesn't have to be a guessing game when it comes to conversion.

Organic optimization under this light becomes more like a preferential selection of significant [landing pages](#) that are augmented by subordinate means within a site such as (a) dynamically generated php includes with topically aligned products or pages (b) [contextual linking](#) from related pages “much like Wikipedia” (c) sitemaps (d) [RSS feeds](#) or syndication formats such as Google Base or other means to provide [deep crawling](#) and indexing.

In another post named “[How to Make SEO Landing Pages That Rank Like PPC](#)“, I cover this tactic in greater detail. However, to summarize, you simply create enough topical pages or leverage the co-occurrence of keywords (from pages or posts, tags, tag clouds, footer links, etc.) to **link to the preferred landing page** with the root keyword and stemmed variations you wish to rank for.

You simply need to conduct competitive analysis on the first page you see ranking on its own merit (ranking with a specific page and not the homepage) in search engines for that keyword. Then you can check the websites' concentration ratio of frequency and co-occurrence for that [keywords saturation/percentage](#). For example in a Google search box you can use the search operator

site:competitorsdomain.com keyword

It will state “results 1-10 of X Amount”, *X-amount* will vary depending on the number of supporting pages are targeting the keyword. Targeting implies *that it appears in exact match, phrase match or broad match in the title, description or on the page or in links in some way, shape or form* in the site.

Using that [search operator](#), you can see how many **ball park / supporting pages** they used to acquire that position. Granted site age, domain authority and page strength / relevance and continuity have a great deal to do with their position; you still need comparable metrics that you can use for [gap analysis](#).



After you have a metric, you can then analyze their strongest pages as well as the other top 3 competitors to look for similarities or indications of their ranking tactics.

For example, are they employing fewer more authoritative links and dozens of supporting pages for each topic or a larger number of internal links and less external links to each page to produce their ranking factor?

Each has its advantages and disadvantages, but at least you will have a benchmark such as (100 pages and 20 links per page with varied anchor text) to topple their competitive advantage.

**Reinforcing Competitive Keyword Categories** – This is where ranking battles are fought and won daily in reinforcing multiple granular layers to send a clear signal to search engines that you are working under the premise of optimization.

For example if I wanted to rank for the keyword “**designer shoes**” or “**athletic shoes**” which is a great example of a broad category replete with brands, makes, models and features. I could then set up the parent category so that it is only a few clicks away from the root folder and then feed the relevance model by using title tags, site architecture and content to reinforce topicality.

Designer Shoes has 2.5 million competing pages in the top 1000 search results in Google. Since that is not extremely competitive, depending on how you leveraged the off page factor, I would say a timeline of 8 months to acquire that result is fair.

Looking at the threshold, the more supportive layers I add to this topic, the easier it would be to scale relevance, which means that adding multiple sub folders and using the subordinate pages to get spiders and links to those pages internally would be the first priority.

From personal experience, I would add 100 pages of **contextual content** to use to support the landing pages. The contextual articles could be a blog or articles designed to instill a footprint in search engines with an array of anchor text and relative pages “**designer shoes for sale**”, “**why you should avoid designer shoe knock offs**”, etc., to reinforce the theme of the site.

I would then structure the site in such a way that the **top level categories** are broad enough (shoes) to encompass one more layer where I can instill make, model (designer pumps/designer flats) and a unique description to set myself apart from other companies that may be employing copycat content from mfg. specs.

One example would be shoedomain.com/shoes/prada-black-highheel-pumps.html that way I could consolidate or parse each aspect/designer/make within one query string after the /shoes/ sub folder or elect to specify one additional silo to support.

The category would be where I focused the anchor text to from within the site, in addition to each specific make, model and designer after the site architecture was in place.

By [feeding the sub folder](#) “shoes” and funneling [internal and external links](#) to that “keyword-rich / relevant sub folder”, I am indirectly providing buoyancy to each page contained within that folder (making it less dependent on [off page ranking factor](#)).

If I can internally reinforce the landing pages in this capacity, then time, absolute links (using the full URL vs. relative links /page.html) will eventually create [PageRank](#) and flow upwards and back to the main categories and site theme if structured properly.

Then after I was done with that category, I would move on to the “**athletic shoes**” category and set up similar internal support mechanisms. You could then cross link the two categories with secondary navigation, footer links or the contextual content to provide a push to your deeper internal pages.

As the site grows and ages, it will be able to manifest ambient rankings from keyword stemming (queries finding near matches in the descriptions, title tags and content throughout the site) and return an array of search results from the long tail that supersede your original ranking intent.

This broadening of relevance and [keyword stemming](#) is a blessing in disguise, but it is the site architecture and the use of deep links and internal links (deep links from RSS feeds, or other sites even if it's only a few links from other websites) will differentiate your site from a competitor who has zero inbound links to deep pages.

# SEO for Large Websites Part III

Topics covered in the previous posts were:

- Defining Ranking Objectives. (Covered in Part 1)
- Plant Seedling Keywords. (Covered in Part 1)
- Determining Preferred Landing Pages. (Covered in Part 2)
- Reinforcing Competitive Keyword Categories. (Covered in Part 2)

Now, we will conclude this triage with the last few areas pertaining to optimizing large websites specifically through:

- Strengthening Internal Linking
- Using Succinct Titles and Meta Descriptions
- Managing Site Wide Duplicate Content (Custom templates, Noindex, Follow, 301 Redirects, XML Sitemaps).

**Strengthening Internal Linking** – The purpose of [internal linking](#) is to reinforce specific keywords, topics or stemmed arrays of relevant information to search engines. Each link on your page is telling search engines “**this page is important**”, so, the first thing you need to understand is, make sure you are sending the right message.

Ironically, many webmasters omit using internal links to connect [contextual content](#) or help establishing a pecking order for important content. Here are a few posts that provide more detail from a how to perspective or tactical viewpoint on [internal links](#), but for the sake of context we will stick to the premise of why it is important.

From the standpoint of a hierarchy, the page which is a destination for the most internal links from other pages within a website, earns a higher search engine position as search engines associate links with authority (both from other sites and within a website).

The larger a website becomes, the more important it is to **distinguish each page from others** so that it can (1) develop its own authority and (2) rank for a specific array of keywords. The less work you do, like share one template for example instead of create a series of topical templates to facilitate key landing pages, the less likely those pages will stand out from others.

The idea is, consider your trophy keywords; these landing pages should have the highest concentration of absolute internal links to them. To accomplish this, focus on the first occurrence of a keyword on a page within the body area of the content and use it to provide the link to the page you wish to emphasize as the new champion page for that keyword.

Do this throughout the entire site to really see the beauty of this technique or you could implement a script or [SEO plugin](#) to accomplish this tactic with ease.

It is important to seed these keywords or use keyword co-occurrence whenever possible to reinforce the internal linking effect to your top level landing pages.

Say for example you have 1000 pages in your website, and you have 10 landing pages you want to emphasize as the top level keyword and key phrases. Depending on the keyword threshold and how competitive it is, you may wish to add editorial content to support the main landing page.

This does two things (1) it helps to flag your website as a relevant match over and above say for example a site that is using a general shopping cart and just has cookie cutter descriptions to distinguish it and (2) it provides a platform for leverage for internal links and keyword stemming.

There are two solid ways to produce rankings (1) content or (2) links both have advantages but the precedence is even more prominent when you utilize both in tandem to accomplish a series of rankings for competitive keywords.

So, with 1,000 pages to work with you could potentially use 100 pages per landing page to support at the top of the hierarchy or [silo/subfolder](#) to produce buoyancy for your pages.

As a website gains more pages, the site gains more Pagerank; and not the kind in the toolbar, actual on page authority. Internal linking (linking to pages with relative keywords or key phrases) is your first line of offense and defense in layering a foundation which will provide stability and rankings as time progresses.

Each page only counts if its indexed and duplicity is not rewarded, so, using supporting pages and internal links can keep your content relevant to search engines (since they can determine the concentration of ranking factor) as well as alleviates the need to place all of your ranking factor on one page alone.

The way to leverage on page [internal links](#) is to find the other pages deemed significant and then provide a link to the target / preferred landing page potentially as close to the beginning of the body text as possible.

Not that sidebars or stand alone links in the footer don't pass value (a fractional amount as a normalization process known as a null-set), but based on block segment analysis, and [semantic connectivity](#), keywords using contextual links in the body text carry more weight for SERP (search engine result page) positioning.

Out of those 100 pages if you have one page that has 5 of the keywords from your 10 primary landing pages, then build an internal link from the content from that page to each of the 5 keyword/pages.

The idea is, concentrate as many keywords to as many pages as possible, but try to use the same target page for a given keyword or key phrase (so you don't diffuse ranking factor). Once each landing page is sufficiently reinforced, then its dependency on off page ranking factor (links from other sites) is diminished and it will (a) get or stay indexed as well as (b) be able to pass that [ranking factor](#) along by consolidating link flow to another page (by way of the links that leave it to other pages).

**Using Succinct Titles and Meta Descriptions** – Your title tag is the most important element in [search engine optimization](#). You can rank on an [“exact match shingle occurrence”](#) in a title tag alone. Then if reinforced by a supportive occurrence of the keyword, a plural or synonym variations and then reinforced in the H1 (header tag), in the links on the page and links to that page, creates a thorough degree of relevance for that page to search engines.

**Top level categories** should be succinct and descriptive. For example – ”Brand-X Designer Shoes: Shoe Collection from Company -Y” would be ideal if the company sold shoes or a collection from that designer.

The idea here is, if you have a main category page that feeds subsequent listings and then use the most competitive description in the title. The description could state... “Company-Y Provides Designer Shoes by Brand X as well as Model A, B and C”, this way the meta description reinforces the title. This type of relevant redundancy works well **as long as it is not abused**.

Never stuff titles with more than 6-8 words and keep your descriptions to 12 words or less whenever possible.

In addition, if your page has 500 words of unique text, a header (h1) tag that matches the first word in your title (your main keyword) and uses a relevant description and has internal links pointed to it, the only missing ingredient is a few deep links from other sites to cement its position from other large orphaned pages from competitors who are selling the same wares.

**Managing Site Wide Duplicate Content** – I mentioned before in Part I that search engines are not a fan of duplicity or [duplicate content](#), so, the less you replicate across a site, the more distinct the signature becomes when attempting to equate relevance.

So, is having a page that has 100 links on every page across every category the same going to really help contribute to facilitating a higher relevance or quality score? Or what about having the same meta title on every page in a category in a large e-commerce site (as if search engines will not notice)... It is possible to [SEO your CMS](#) (content management system) to fine-tune relevance and eliminate duplicity with a few custom hacks/settings.

There is one thing to be said about duplicity and large sites “**try to avoid it as much as possible**”, which means create more templates or more content and make each page unique. True, [domain authority](#) can push through a great deal of duplicity and keep pages in the index, but if rankings are the objective, then think of each page as an island that needs to have its own ecosystem to survive in the SERPs.

So, to provide methods for sustenance you can employ tactics such as custom templates, using a noindex, Follow tag in the header / meta data, 301 Redirects or XML Sitemaps to provide irrigation to pages starving from link attrition.

**Custom templates** – If you have a large site such as an ecommerce site, try to create themed categories that logically support the next.

`mysite.com/products/shoes/`

\*This is your top level category, hence anything linked from this will rank well on the merit of the page strength funneled to it from the main site. The main site/navigation would like to it with the keyword “shoes”. That way, If you offered shoes, watches, jackets, etc., each would have its own main landing page / top level category.

On the top level category this should have the least amount of links leaving it and the most inbound links from other pages internally as well as links from other sites.

Then by adding brand modifiers or descriptors in the next tier you would have [allinurl relevance](#) meaning that your URL (universal resource locator) a.k.a. web address would reflect relevant naming conventions that reinforce topically what that pages contents are.

Combined with the keyword/naming conventions used in the title and description, this is a winning combination for relevance that borders on redundancy, but really drives the point home to both users and search engines alike.

If you wanted to further stem that tier, you could add `mysite/products/shoes/black-highheel-pumps-size6.html`



As an example, or the option to truncate the .html is also an option based on [MOD rewrite](#) preferences (preferences conducted and assigned at the server level). The point is, the more information you replicate and reinforce, the easier it is for search engines to sort and assign relevance score towards.

You could also take it a step further by using a specific category based template that only refers to the top level category in the sites secondary navigation (usually text links on the left side bar) to provide additional shopping options to users.

Then instead of having the template bleed ranking factor by trying to link to everything (like the main page of the site), consolidate links to only categories that are part of the same product type.

For example the main navigation on the internal shoes pages would have primary navigation to

Shoes | Pumps | Heels | Flats

Don't forget to use [breadcrumb navigation](#) as well to aid user awareness, but then in the footer (using [footer links](#)) under each product I could have the other brands of shoes positioned there to reinforce internal links or even add links to other categories such as

Jackets | Socks | Hats – to provide a robust user experience...

The idea here is to create self contained segments that reinforce topicality and then use links from each category to push link flow deeper into the site where it can become buoyant over time. Using a series of rotating .php includes with pre assigned related links is one alternative that you can use to avoid duplicate content as well as drawing from a series of 2-3 different meta descriptions or opening paragraphs.

Nobody said getting deep pages indexed was easy and many simply rely on the brute-force tactic of links alone, yet 60% of the ranking factor comes from on page SEO and on page continuity. If you understand that as the premise, you can get pages indexed and ranking with just a few links each from the right sources and bury your competition.

**Sitemaps** - One of the best use for [sitemaps](#) is to **create a sitemap of your sitemaps** and then consolidate them to your key pages that pass the most link flow.

If you had a top level category called shoes (using the example above), then your sitemap linked from the top level category should be to a sitemap that contains all of the pages that have shoes. Each brand, subcategory, etc. should all be accessible from that sitemap (xml or html).

As a result, those pages are food for search engine spiders and as long as they (a) link back to the top level category and (b) the homepage, the cycle is complete. The idea is, not to have any page more than 4 folders or clicks away from the homepage and the more specific you are with your content the better it can rank on its own (making the home page less significant to its ranking factor).

Wikipedia applies this for hundreds of thousands of keywords and ranks in the top 2 results for virtually any keyword from any topic as a result. If you only apply 25% of their method, you can do extremely well in contrast to someone simply using a [flat site architecture](#) and relying on navigation alone to push ranking factor into thousands of “me too” shopping cart / product pages or hum drum blog posts or supporting articles.

I will have to pass on the value of using 301 redirects in the video follow up of this post since, at this point, it would have been a great topic for an e-book or webinar as each topical point leads to 10 more granular tactics that could be implemented.

I hope this has provided you with a few ideas or concepts for implementing a self supportive site architecture. For every tip provided, there are 10 more such as leveraging tags, [canonization issues](#), etc. So, this is a wrap for this 3 part post.

Granted, there are multiple intricacies and layers to every metric discussed, each in fact could be taken to the extreme from the level of minute details to a thesis level dissertation.

However, the take away here is, if you apply principles of optimization across a broad array of metrics (from server side to the [template](#)) you increase the likelihood of attaining a higher relevance score for each page and developing a broader range of authority for your domain.



Optimization is designed to increase search engine exposure and positioning, yet positioning only matters if the content is geared for the user to [take action](#). Yet, that is another topic which warrants another eBook at a later time.

## **Thanks for Reading!**

We hoped you enjoyed SEO for Large Websites. This brief was designed to share real-world SEO tactics to distinguish your website in search engines.

For additional information with hundreds of insightful posts on SEO, visit the [SEO Design Solutions](#) Blog or if you haven't already, subscribe to the [RSS feed](#).

For those of you interested in [SEO consulting](#) or [SEO Services](#), call 312.794.7883 or visit our [SEO Pricing](#) page and let us know how we can help.