

The importance of Information Architecture in Search Engine Optimization.

Implementing the right semantic hierarchy is often overlooked, but an important component to On-Site Factors to consider when trying to optimize for search engines. It is amazing to me how even the simplest On-Site Factors are ignored by web designers. The additional effort to pay attention to some tried and true page construction guidelines will definitely pay off in the long run. The value of Social Media Optimization is not dependent on good baseline Search Engine Optimization, but the impacts of a Strategic Social Media Marketing campaign can be exponentially better if the base website is correctly optimized.

In the hopes that this information will be read and adopted by web designers serious about giving their clients value in the SEO arena, I am laying out the Information Architecture, or semantic hierarchy I've proven over the years to be most effective. By semantic hierarchy I mean the search engine program (spider, bot, crawler) is looking for the following items, in this order to determine a page's content and a corresponding ranking in the search engine:

- Domain name
- Domain folder naming convention
- Domain file naming convention (including image naming convention), which is also the Page's URL structure (excluding root domain)
- Page's Title
- Page's Headline
- Page's Subheadline
- Page's Sub-subheadline
- Page's content
- Bolded and italicized words within the content
- Links within your page, linking to internal pages (breadcrumb trails, left navigation and right article list)
- Links within your page, linking to external websites

I'll go through each section with some tips on how to improve it and explain the logic for search engines and people, where the change will effect what people see.

Domain name

If someone owns "world-class-apples-are-delicious.com" and you write an article called "world-class apple trees are delicious" they might get higher in search engine results for "world class apple trees are delicious" if anyone links to them, because domain names are the most important element in most search engines (since they are not editable, unlike html). Having a URL with at least one keyword phrase contained within the domain name is imperative to ranking well.

Domain folder and file naming convention

Utilizing a file folder naming convention that utilizes the desired keyword phrase is an under-utilized strategy that does have an impact. People may dismiss this as a small impact, but I've found it to have more impact than most SEO consultants are willing to give credit for, so I

always recommend this strategy. If you have www.worldclasswidgets.com as your URL, then you're going to want to create a subfolder naming convention that supports the selected keywords for the target pages. For example, if you have three services, 1) worldclass widget design, 2) worldclass widget prototypes, and 3) worldclass widget marketing, you will want to create 3 folders with names that support each of the keyword phrases. What previously might have been <http://www.worldclasswidgets.com/service1.html>, should look like <http://www.worldclasswidgets.com/worldclass-widget-design/worldclass-widget-design.html>.

Since search engines give priority consideration to keywords within the root domain (since it cannot be edited, manipulated or further optimized for SEO), the same logic applies to a lesser degree for folder structure. Although folder structure can change, it is not typically manipulated as much as the HTML on individual pages, therefore is calculated differently from a competitive site not having the available naming convention.

Page's Title

Always include your site's name (like *Worth1000.com*) and the actual headline for the page you are at. (*Worth1000.com* | Out of Bounds Contest) in your title tags. Search engines give this a lot of emphasis.

Page's URL structure

If all of your pages have dynamic extensions (like .asp or .php) and use querystrings to point you to different dynamic articles you're making a mistake.

An example of a bad URL would be: *http://www.worth1000.com/contest.asp?contest_id=1163*

This is bad for two reasons:

1. PageRank effects articles at the document level, which means querystrings are excluded (so you won't develop separate pageranks for different articles... all will share the same rank).
2. Some search engines and stat logs cut off the querystring when indexing your content. Having a link in someone's stat logs or in a non-Google engine that point to "*http://www.worth1000.com/contests.asp*" is useless and is a wasted opportunity to get more visitors and pagerank.

Structure your URLs like this instead:

http://www.worth1000.com/contests/photography/1163-Title-of-the-page.html

All that's important in there is the number 1163, but you're already getting some important extra keywords in there that will help with SEO.

I don't even need to change my code. I can simply use a program called modrewrite which will redirect that URL to my actual page: *http://www.worth1000.com/contests.asp?contest_id=1163*

(the rest of the stuff is ignored by modrewrite, so I can use anything there). Search engines will see it as a static page, even though it's exactly the same as it is now.

This also has the added bonus of users who link to this URL in forums that automatically parse URLs will now have some keywords as a part of the URL itself. This kind of clean url will definitely not annoy users either.

Page's Heading and Subheadings

Don't use fonts and CSS to make a headline appear big. Wrap it in <H1> tags instead.

A note to SEO folks: This tag is not as deprecated as you might think. <h1> is still semantically telling the search engine "this is the most important headline on this page" <h2> are less important for subheadlines, but if you have subheadlines in an article, using them can only help the spider understand how your content is structured and that is a good thing.

So long as you style the <h1> in the CSS the same way you have been making larger headlines in the past, users won't visually see a difference at all.

Best of all about this tip is that very few sites make use of it.

Page's content

This is your keyword text. Just make sure you have alt and title tags for all images, so search engines can read them. I'd use <h3> for all captions under the images, since images (and the captions that explain them are usually important to the article. Search engines place minor emphasis on words that are put in between bold and italic tags in your content. Very minor emphasis, but just worth noting that if an article is about a certain celebrity for instance, and their name appears multiple times, bold each appearance of the same and it will help the spider think that celebrity is the focus of the page. Don't keyword stock though. Search engines will notice if the same words are appearing too often (called keyword density) and it will annoy your users in any case, so don't try to stuff your page with redundant keywords. Always write your articles for your users.

Links within your page

You need to make your navigation load before the page's content so that spiders can access the rest of your site. You can use CSS to make this not effect your positioning of the navigation. This will help your links get indexed better. When a spider visits your site it sees the page as one huge block of html code, no matter how it's laid out, so where in this block certain elements are placed is very important. Remember that using CSS you can have an element display visually to users *first* on the page, even if it's the *last* thing to load in the html text block for spiders. If those links are all at the bottom of the text it interprets it as not important, and in some cases may not even download enough of the page to see them (some search engines only download the first 50,000 bytes of html).

Link to your sitemap prominently on your homepage if you have one. If you don't then make one. This is really for the search engines more than the people, but it should even be helpful for those people who do see it.

All links in your page should include title tags. `blog`.

Links outside of your page

When linking to any content that isn't related to your site (let's say you're plugging a blog article about tech you enjoyed, but your site is about catfood), always add "rel=nofollow" to the hyperlink or else Google follows it and gives some of your outgoing page rank to them and also may assume that your sites are similar (since the logic is that sites only link to other sites that will be interesting to their visitors).

Here is what I mean: `great read `.

If the site is something you think is related or you want to pass pagerank leave rel=nofollow out.

Fundamentals:

There have been some recent changes on Google, and just like the previous Hilltop Algorithm, there is a new algorithm in the works. It is titled Google Caffeine and its causing uncertainty and unpredictability right now. Regardless of how and when this latest algorithm tweak settles down, ranking well on Google boils down to some very fundamental things which most people are not disciplined enough to truly implement. If you are serious about ranking well on Google, once you do the above, here are the simplest ways to stick to the basics:

- Write interesting and keyword phrase rich content
- Seek links from related, quality sites
- Link out to related, quality sites
- Use relevant keyword rich **anchor text** on internal links
- Use keywords in **naming files and folders** (and domain name if you can)
- Be sure to have a unique **title tag** and **description tag** on each page

.. however, under Caffeine, there appears to be a few other points worth observing that may get you extra Google love either directly or indirectly through their use; and they are:

- Keep adding fresh (quality) relevant content; don't let your site go stale
- **Optimize your page load speed**