

Mihai's SEO Cheat Sheet

mihaiغهza.wordpress.com

SCOUTING

- **KeyWord Research**

Use [WordTracker](#) for an estimate on daily searches and [Overture](#) for an approximate report on KW usage in the last month. Use Wikipedia, dictionaries and other resources for KW inspiration to try and **stretch your long tail**. Remember, this is the cornerstone of your future actions and it may determine what domain name, titles, URLs, and content you chose.

- **Competition Spy**

Every gosu player knows that **scouting early in the game is crucial**. Find out who your top 3 competitors are and determine their marketing strategy, before you even buy your domain name or write a single line of code!

Check for backlinks using the "link: mycompetitor.com" query on Google. Don't stop at the directories they submit to. See if they have a Wikipedia page or reference; see if they have any links on social bookmarking websites, like del.icio.us, and what do people commonly associate them with; check for blog posts on Technorati; check for buzz on digg and other prosumer websites.

See what they did best and think of ways you can do better and differentiate.

HEAD TAGS

- **Title Tag**

Probably the most important of them all. Use relevant KeyWords, yet do Link Baiting; remember that **this tag will show up in SERPs**. Use a different title on each page to avoid duplicate content penalty. But keep them standardized by branding them with your website name. Trust a separator character. (eg: ACME | Buy Red Bicycles on Discount)

- **Meta KeyWords Tag**

Try to stick to a maximum of 20 KW combinations; stuffing is no good. Use **different KW tags on each page** to avoid duplicate content penalty.

- **Meta Description Tag**

Describe the page content as accurate as possible, while using relevant KW. Remember that this should be no longer than 250 chars and that **the first 150 chars may show up on SERPs**. Use a different tag on each page to avoid duplicate content penalty.

- **Meta Foo Tags**

Don't forget about the http-equiv, resource-type, classification, rating, distribution & other meta tags out there. Try teaching the spider as much data about the page as you can.

ARCHITECTURE

- **URLs**

Describe your content and keep it short. Separate words with hyphens, as they work best. Watch out for too many subfolders, as the further away from the home page you are, the lower the Page Rank will be. Careful with subdomains, they will be treated almost as separate websites. **KeyWords won't hurt**, you know!

- **Dynamic Pages**

Use a URL rewriting engine, like mod_rewrite on Apache, and turn your string of parameters into meaningful folders. Take the time to understand this technology and you will harvest the goods. **End your URLs in a .html**. A good starting point would be reading [Professional SEO with PHP](#).

- **Session IDs**

If you serve a crawler the same page, with different URLs, for a thousand times, it will eventually get angry and do nasty things to your ranking. Turn those ugly numbers off and implement cookies if you really have to. Or, if you feel like living dangerously, use cloaking for spiders.

- **Canonicalization**

www.website.com, website.com, www.website.com/index.html, website.com/index.html, will be treated as separate pages. **Permanently redirect (301)** all of the above combinations to a single address, preferably www.site.com.

- **Frames**

It ain't '97 anymore. So just forget about frames, as weird thing will happen to your indexing if you use them. Yes, the same with iframes!

- **Forms**

Crawlers are smart, but not smart enough to input text in search boxes and then index the resulting pages. Content available only by forms won't be indexed. If your website is designed in such a manner, engineer a workaround with HTML sitemaps.

- **Navigation**

If your website is structured on different levels and you need a drop-down menu, don't use JavaScript or Flash (with cheesy background music). **Use a SEO friendly, list menu**. If crossbrowsing issues drive you crazy and you must go with DHTML, find a script that hides the list div instead of generating it on-the-fly.

MARKUP

- **Semantic Code**

Emphasize your keywords and essential info with font weight, size and color attributes. Always use heading and subheading tags (hx) and apply whatever style you like to them. Use li tags when displaying a set of items; use tables when displaying a grid; use blockquote when quoting. Remember what the semantic web is all about: data about the data, and give meaning to your markup.

- **Anchor Text & Title**

Use KW combinations that clearly describe the page you are linking to. *Click here* is not an option. Remember this is very powerful as it has been used for link bombing. Stay away from JavaScript links as chances are they will not be crawled.

- **Valid Code**

Don't hurt the web and do your best to stick to standards. Write valid XHTML/CSS2. There are hundreds of communities that use valid code and design superb websites, like [csszengarden](#). Find your support group and [Reboot](#) your website.

CONTENT

- **Copy writing**

Write your own, original content. Surf for editorial tips, like [CopyBlogger](#) and pay much attention to scannable text. Study [Jacob Nielsen's articles on readability](#). Remember to use an inverted pyramid structure, with the essential info at the top. Organize your copy into sections, short paragraphs and lists.

- **Freshness**

Spiders love fresh and juicy content. And so do your visitors. If you sell stuff, get a *Latest Products* section; if you write stuff, get a *Latest Articles* section; if you're a boring company, at least get a *Latest News* section. To make things even more attractive, generate an RSS feed and advertise it on your home page.

- **Crawlable Content**

Although bots have achieved great technical progress, Ajax, Flash, JavaScript generated content, text in images are all blind spots on your page. Put your important info in plain HTML. If you must really show your fancy font, use sIFR as text will still be indexed. Most bots behave as a text browser, so use a tool like [Lynx](#) to analyze your website.

- **Image Name & Alt Text**

Think of crawlers as visually impaired beings, learning to comprehend images. The text will come naturally to you. If you have multiple sets of images, like in a shopping cart, group your images into folders named appropriately.

SERVER INTERACTION

- **robots.txt**

Make use of the robots.txt file on your web server. This file tells crawlers which directories can or cannot be crawled. Make sure it's current for your site so that you don't accidentally block Googlebot . Visit www.robotstxt.org to learn how to instruct robots when they visit your site.

- **sitemap.xml**

This file is now a standard practice. Visit www.sitemaps.org to get the protocol specifications. There are plenty of free tools for automatic generation of the XML. **Submit it to all major search engines:** Google, Yahoo!, MSN.

- **Server Config**

Make sure your web server supports the If-Modified-Since HTTP header. This feature allows your web server to tell the bot whether your content has changed since it last crawled your site. **Check that your server outputs a 404 header even if you use custom error pages.** Get a reliable hosting provider; if your server goes down during a crawl, you could flush away months of hard work.

PROMOTION

- **Blog**

Start blogging even before you launch your service or product. There are thousands of great resources, tutorials and cheat sheets for you to learn how to do it properly. Some of the most important aspects for you to remember: **be transparent, be informal, be remarkable.**

- **Online Reputation**

The Web is such a depersonalized environment that people crave for a human presence behind the screen. **Become the spokesperson for your business.** Put your paranoia aside, drop your nicknames and create a professional virtual identity. Sign every article or comment with your real name. Attach a smiling photo of yourself to your accounts. People will be googling for your name. You should manage what they find.

- **Networking**

Comment on the blogs you read and leave your URL all over the place. Put your blog & website address on your email signatures; on your letters, business cards and everything else you print. **Join a Social Networking Community**, like LinkedIn, ask for connections and give recommendations.

- **Buzz Generation**

If you're not worthy of getting digged, **write your own Press Releases and spread them with free PR websites.** Try your luck by asking for a review with well known bloggers in your field.

D O N ' T S

- **Black Hat**

Cloaking, KeyWord stuffing and *invisible text* are probably the easiest way to get delisted, sandboxed or be put in *Supplemental Results* hell. Most amateurs start with the pick-up lines instead of the basic body language control. Study the fundamentals, don't let yourself fooled by some *secret techniques* your neighbour's kid told you.

- **Duplicate Content**

Useless to say you must not steal other people's content, for it will hurt your karma and ranking. But make sure you **don't show the same content with different URLs and titles**; like a product configurator, where a URL parameter only changes the shoe color; or Session IDs.

- **Page Rank Mania**

Focusing on increasing your PageRank is a narrow-minded approach to SEO and web marketing. It's not only inaccurate, as Google keeps is outdated for some months after it has really changed, but **it is irrelevant**. Bringing visitors to a poorly crafted page or dull content is suicidal. **You want conversions, not traffic.**

- **Link Exchange**

Reciprocal linkage is a thing of the past. Google already ignores reciprocal link accumulation on real estate websites. And never liked spammy/link farms anyway.

R E M E M B E R

- **“You** will never really know exactly how Google works (unless you work there)”
- *“Search Engine Optimization: An Hour a Day”* By J. Grappone, Grady Couzin
- **“Most** of the right choices in SEO come from asking, What’s the best thing for the user?”
- *Matt Cutts*
- **“Good** HTML titles, good body copy, great content, ensuring that your site doesn’t have roadblocks to crawling — these have worked for nearly a decade.”
- *Danny Sullivan*

Many thanks to [LifeClever](#) for the design tips and [SEOMoz](#) for their great material.

For suggestions, corrections and thoughts, please email me at mihaigheza@gmail.com

Downloaded from http://mihaigheza.wordpress.com/files/2007/08/seo_cheat_sheet.pdf

This document is licensed under a Creative Commons Attribution-Noncommercial-Share Alike 3.0 License

