# 8 User Experience Design and Search Engine Optimization

User Experience Design's Fundamental Role in Successful SEO

Search engines are the cornerstone of the interactive economy. Everything that we do as "interactivists" is ultimately connected to the world at large through Google, Yahoo, MSN, Ask, and the myriad minor engines that make up the infrastructure for finding things online. Information architecture is a critical component of how Web sites are interpreted by search engines. This chapter is designed to give you some basic understanding of why UX design is critical to search engine optimization and what you must take into account so that the environments you create will have a fighting chance on Google.

Jonathan Ashton
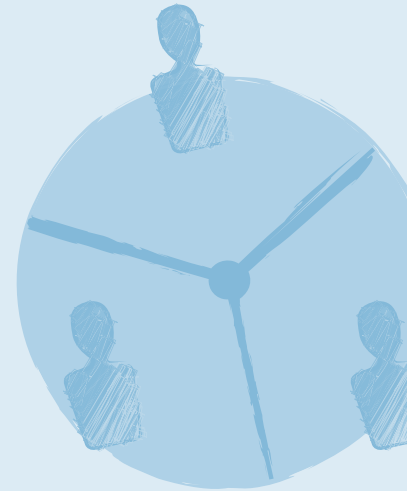
Introduction to SEO

Simply put, *search engine optimization* is the process of developing and maintaining a Web asset with the intention of gaining and keeping top place-ment on public search engines for specifically targeted keyword phrases. Search engine optimization (SEO) is like a martial art, a process of learning and doing that is never complete. Even a master can progress further using observed behavior or learned method. As long as there are search engines and Web sites interested in selling something to the people searching, there will be a role for search engine optimization.

SEO relies on three fundamental areas for improvement and influence:

▶ The critical group of things that the professional user experience designer can influence—site infrastructure, technology, and organizational principles

▶ Content and all the keyword issues that relate to optimized words which the search engines can see

▶ Links, or link popularity—the quantity and quality of links that point at your site from other sites, as well as the organizational structure of the links inside the site

We will take apart each of these three areas and examine them from the UX designer's perspective, to better equip you for the optimization chal-lenges that lie ahead.

Why Is SEO Important?

It is interesting that even today we need to explain the relevance of search engine optimization. Clients tend to understand on some level that it is important for their Web sites to attract targeted visitors from the natural search results of the main search engines, but beyond that it is difficult for most interactive marketers to understand the impact SEO can have.

Data on global search volume is available from a variety of sources, but what is most important to understand is that, whatever the source, the numbers are simply huge, and the year-over-year increases are always in double dig-its. For the most part, every quarter the global volume of searches increases. When Google first launched in 1998, 10,000 searches a day was a huge vol-ume and placed an incredible burden on the beta version of the system.

Hitwise (www.hitwise.com) reports that Google and its affiliates (including AOL and YouTube) own the lion's share of searches globally, with nearly 72 percent of U.S. searches performed in November 2008. Yahoo is a distant second, with nearly 18 percent, and MSN and Ask.com trail in at 4 percent and 3 percent, respectively. Internationally, Google is even more dominant: Its market share reaches more than 80 percent in many markets.

> **Note** For more background on Google's early days, see The Google Story, *by David A. Vise and Mark Malseed (Delta, 2008).*
>
> *According to comScore (www.comscore.com), in 2008 there were easily more than 60 billion searches per month performed globally by 750 million people, with over 18 billion searches performed in the United States alone. To put it another way, 95 percent of Internet users use a search engine at least once a month, with a global average of more than 80 searches per month.*

Aside from these remarkable volume numbers, what does this really mean on a practical level to interactive marketers? Simply put, if you are not reaching your target customers when they are searching for your products or services, your competition is getting the opportunity to sell to them.

Look at your site analytics and think of the issue this way: How much more revenue would the site generate if there was a 10 percent increase in strategically targeted traffic? What about a 100 percent increase? Or 1000 percent? If your site is not generating meaningful traffic through natural search, then SEO is a requirement.

A little investment in SEO can go a very long way, particularly if the interactive marketing effort to date has focused on purchasing clicks through sponsor listings. We have seen sites achieve a return on investment of 35 to 1 on monthly SEO expenditures. If you are paying the search engine companies for traffic from sponsored listings but you are not investing in natural traffic, you are really limiting yourself to about 10 percent of the opportunity. Think of your own search behavior: When was the last time you clicked through more than one or two of the paid sponsor listings in a search result?

Any discussion of why SEO is important and why it is here to stay could go on for chapters. Suffice it to say that Google is not going anywhere but up, and that effective interactive marketing must include search engine optimization as a core component of competent execution.

Important Basic Resources

Expertise emerges from a well-rounded education. The professional who simply focuses on his or her specialty loses perspective on everything else around. That is why it is imperative that every interactivist spend at least a few minutes learning about SEO. Although there is no official set of guidelines, Google has been kind enough to provide some very salient resources. If you're interested at all in getting better search engine performance from your efforts, check out these links:

▶ Webmasters/Site Owners Help: Search Engine Optimization: www.google.com/support/webmasters/bin/ answer.py?hl=en&answer=35291

▶ Webmasters/Site Owners Help: Webmaster Guidelines: Quality Guidelines: www.google.com/support/webmasters/bin/ answer.py?hl=en&answer=35769

▶ Search Engine Optimization Starter Guide: www.google.com/webmasters/docs/ search-engine-optimization-starter-guide.pdf

If that is not enough, drown yourself in newsletters and blogs. Start at SEOmoz.org and dig down. Just remember, as in all other things in life, if it sounds too good to be true, then it probably is.

Site Technology, Design, and Infrastructure

Search engines are essentially Web 1.0.5 technology that is firmly implanted in the Web 2.0+ world. The basic premise of the search engine has changed little since the World Wide Web Wanderer was launched in 1993 to crawl the Web and build the first Web search engine. Essentially every search engine has an application alternately called a *crawler*, *spider*, or *bot* that finds and follows links, sending back to the database a copy of the assets that it can see. The database is then analyzed according to the search engine's proprietary algorithm. Using these rules, a Web asset is indexed and then ranked according to how well it scores on that search engine's particular score card. In this rather straightforward process are a myriad of pitfalls for the UX designer.

Understanding these core relationships will enable you to see your site through the eyes of the search engines. An optimized Web site relies on a structure and technology that facilitates the movement of the search engine spiders. Likewise, many decisions about handling content determine how well the search engines ranks the resulting efforts. As a result, much of this is predetermined by the decisions that are made in wireframes and in the discussions that take place around how to style and manage content.

Flash, Ajax, JavaScript, and Other Scripted Content

Today's dynamic and interactive Web design relies on technologies that are not at all friendly to the needs of the search engines. There is a widening gap between what search engines can see and what designers can do. It is up to the UX designer to be sure that the strategic plans for dynamic, design-intensive sites are deployed so that both the search engines and the users get the best possible experiences.

Having a fundamental understanding of how search engines interact with this kind of content will help you to decide when to deploy it and where to compensate for its weaknesses. It is entirely possible to build an optimized site that relies heavily on scripted content if the appropriate compensations are in place at the beginning of the process. It is substantially more difficult to build static or indexable content once the site is built and live. So make a forceful argument for static content, on the grounds of usability and for the sake of the search engines' crawlers. It may seem like extra work up front, but the return on investment is exponential.

Flash

Flash content is technically "indexable." There have been some recent advances in the ability of the search engines to see into Flash files to find the text and links that are built into these assets. Although this content is indexed, have you ever seen an all-Flash asset win top placement in the search results? You probably haven't because it's risky for search engines to open themselves up to full compatibility with Flash. Let's assume that the search engines could completely see all of the links and text content that is embedded inside the SWFobject. What prevents an unethical (or "black hat") optimizer from putting apples in the text layers of the object while showing

oranges to the human user viewing the fully compiled assets through a browser? How can you deep link into a Flash asset without it being fully compiled? These fundamental vulnerabilities will remain until the search engines can reach some level of artificial intelligence that can tell that an image is a picture of a horse without some associated text that says "this is a picture of a horse."

To architect a Flash Web site that is compatible with search engines, you must add a static layer of content that duplicates the Flash content. Leaving aside the needs of search engines for the moment, a static layer of content is a key for compliance with usability requirements. Think of the search engine as the person who is viewing Web content over a dial-up connection or is using a screen reader browser. These people may be the lowest common denominator, and it is possible that the strategy behind your Web development discounts this very small percentage of human users. But when you discount this handful of people, you also discount GoogleBot and Yahoo Slurp—the two most important visitors to your site, since they are the crawlers that will enable the major search engines to index your site. If no text words or spiderable links are visible to the search engines, your content will inevitably not be findable through meaningful search results.

A static layer can be accomplished in a number of ways. To comply with search engine requirements, the static layer of content needs to mirror the Flash content. *This is not an opportunity to show the search engines something different than what is deployed in the Flash*; if you do that you are violating the spirit of the game and standing squarely on the dark side.

The ideal way to embed Flash content into a static layer is to use SWFobject so that both the Flash and static content can live on the same URL. This will allow the search engine to find the static content and the Flash-enabled browser to display the animation instead of the static content. If at all possible, do not use redirecting so that you can conserve the popularity of the link that is pointing to the Flash content. Google Code provides a simple set of instructions for implementing this straightforward piece of JavaScript at http://code.google.com/p/swfobject.

There is another option that runs on the gray side of SEO. *Cloaking* can be a dirty word to SEO purists, but if you approach the following challenges from the right side you can have some cake and eat it too.

Cloaking takes advantage of user agent detection, detecting search engine crawlers as they visit a Web site and routing them to static pages to index. But when a human visitor sees the same page in search results and clicks the link, the Web site detects that the user agent is a human with a Flash-enabled browser and shows that person the dynamic experience on a completely separate URL. The crux of the issue remains the same as with the SWFobject method: You have to show the search engines the exact same things in your cloaked content as you do in your Flash content.

Ajax, JavaScript, and Other Scripted Content

A powerful driver of Web 2.0 content, Ajax provides Web developers with the ability to build pageless content. However, the problems that search engines have with Ajax are multifold and require good planning to avoid big mistakes.

Ajax stands for Asynchronous JavaScript And XML, which hints at the difficulties search engines have with this technology. Search engines essentially can't deal with JavaScript; the efficiencies that JavaScript brings to developers are the problems that search engines have with dynamic content. An additional problem search engines have with Ajax is the asynchronous nature of the technology. A search engine can only see the contents of the initial page load, and any content that is loaded through a script that takes place after the initial shell loads will not be visible for indexing. Because Google can't extend a session beyond the initial page load and doesn't have a mouse or external agent to activate a script, any pageless content that is activated by the user will be invisible unless the text content is included in the pre-loaded shell. It is up to the UX designer to be certain that the three-dimensional modeling necessary to structure pageless design also includes the requirement that text and links all preload in the page shell. Anything else, and your cool design is invisible.

Scripted Navigation

One of the most common problems that will hamper optimization is the use of JavaScript in the core of site navigation. This is a very common condition and is the result of the way many site development and content management tools work. The scripted navigation looks cooler, so people tend to be interested in using it. But if JavaScript is the technology that drives the site navigation, the result is that search engines can't properly build a model of the link

relationships within the site: They simply can't see the link structure of the site. And if the search engines can't model the link relationships in the site, deep content will be invisible or will not be assigned the right link popularity.

Content Management Systems

Content management systems have been built for the convenience of humans—but many of these systems make it difficult for search engines to deal with their output. Following are some typical problems that need to be avoided, either by using work-arounds or choosing a content management system that is more search friendly:

▶ **Dynamic URLs.** Search engines don't understand a "page" of content; it understands the path to that content. A change in the path, or URL, leading to that content causes the search engines to accidentally clone the content multiple times. This condition substantially impairs the ability of a site to do well. If the content management system has a system that creates session IDs in URLs, you could be in real trouble. Track with mature analytics, not session IDs.

▶ **Multiple URL paths.** A typical problem with e-commerce content management is that as a product progresses through its lifecycle, it accrues multiple URLs. Again, since the search engine can only understand a page of content based on the URL where it finds the content, when a product appears in a category *and* is a part of a gift basket *and* is a weekly special (and on and on), pretty soon the search crawlers have followed a bunch of different links to find the same piece of content. Do whatever you can to ensure that each piece of content exists only on one URL and that multiple paths actually rely on one URL, regardless of where the links are deployed. Rely on mature analytics systems to parse channels.

▶ **Unintentional cloning.** When you come to the realization that a piece of content should only be accessible through a single URL path, it is easy to see other conditions in content management systems that cause content to be unintentionally cloned. Suffice it to say that the architecture must only have a single URL path to a single piece of content.

▶ **Infinite loops.** A corollary to the unintentional cloning issue is the infinite loop. Make sure that you do not put the search engine spiders into

a potentially endless task of following "next" links in a calendar or some similar situation. If the search engine spider can traverse a next link onto the next day of a calendar where it can find another next link, it will follow that link to the next page, and on and on. Prevent this kind of situation by using a scripted link that the search engines can't follow so that the crawlers can spend their time on the content that you want to have indexed.

▶ **Old URL structures.** The first thing that many site redevelopment projects do is to replace the old URL structure. The trouble is that the search engines have probably already indexed the content at these old URLs, and as soon as you change all of them you are essentially sending your indexing back to square one. In addition, any deep links that the site has accrued over time are pointing at the old URL structure. At all costs, preserve as many of the old URLs as you can. It is probable that when you replace the content management system you will have to change all the URLs, so if this is inevitable, be sure to recommend that the old URLs are given a status code of "301 Moved Permanently" and redirected on a one-to-one basis from the old URLs to the new URLs. The 301 redirect is the only acceptable redirect for search engine purposes.

Domains, Directory, and URL Structure All Matter

If you are starting from scratch, and if the restrictions of branding issues allow, try to select a domain that contains a keyword or two. It is difficult these days to get a .com domain that has quality keywords, but if you do, separate those keywords with hyphens.

An important part of how UX affects SEO is in a site's directory structure. It has a critical influence on how link popularity is distributed throughout the site. Simple is better. Avoid having extraneous files in the directory structure at all costs. Some content management systems will automatically insert a subdirectory; prevent this if at all possible. This condition dilutes the relevance of the entire site. The search engines understand the hierarchy of the site based on the way the site directories are structured, so be sure that the most important directories are at the top of the architecture.

If your environment allows it, use keywords in the URL structure that are relevant to the section of the site. Separate keywords with a hyphen, and

don't use too many keywords in one filename. Go for something like this: sitename.com/widget-catalog/aquatic-widgets/underwater-submersible.html.

In addition, be sure that you have redirects set up for http://site-in-question. com to 301 Moved Permanently redirect to http://www.site-in-question.com. If a site will resolve with and without the www, search engines (particularly Yahoo) will index content at both URLs, opening the entire site up for accidental duplication. This condition tends to propagate when a third party links to the site without the www and the site contains a dynamic link structure.

Content: The Once (and Current) and Future King

Although generating content is someone else's problem, the groundwork that is laid in site architecture has a lot to do with making the right content available to search engines.

As with all forms of keyword-driven search, you need to understand the actual search behavior of the people you want to view an asset. Search engines are still very "primitive," in that they rely on users typing in keywords to connect them with assets that are more or less relevant for these words. Picking the right phrases has everything to do with whether your site is relevant in the right context.

In a perfect world, your SEO partner will provide you with a set of keyword phrase targets before you begin and will collaborate with you throughout the wireframing process. If there is no such competent partner involved with your process, read up on the Google AdWords Keyword Research Tool (https://adwords.google.com/select/KeywordToolExternal) and do a bit of investigation into the actual search behavior of people exploring your category. Then spend some time with this input to figure out the phrases that potential customers are searching, and use those phrases as appropriate throughout the site. Search engines look for keywords in a number of places throughout their analysis of a site. Optimization relies on making sure that the right words are in the right places. By understanding the role of keywords in the UX design process, you will establish the framework necessary to enable future success.

So why is content king? It is the very core of what a Web site is designed to deliver. Search engines need text content that they can see and index. Site visitors need engaging content that is worthy of their attention. Bloggers and Webmasters need content that is linkworthy. Without the right content in the right places, search engines cannot connect the right visitors with your site.

Naming Conventions and the Battle Against Jargon

It is essential that keyword targets are reflected in the taxonomy developed for a site. Using keyword phrases in the main site structure makes the whole site more relevant for the things that you are selling. If you're selling widgets, don't call the online product list the Catalog, call it the Widget Catalog. Likewise, use your keyword research to make decisions against jargon. For example, use the words *laptops* as opposed to *notebooks* in your structure because people search for *laptops* 10,000+ percent more frequently than they search for *notebooks*.

Metadata, Headers, and Keywords

It is pretty remarkable that we have gotten this far into the chapter before digging back into basic issues of metadata. A myriad of meta tags are available, but only a handful really have much influence because all the others are susceptible to spamming. Relevant tags are these:

▶ **Page title.** Please note that this is not the <meta title> tag, but is the actual <title> tag in the page header. This tag contains the page's actual title, and it is the most important 65 characters on the page. Think of the title as the little tab sticking up in the old-fashioned library card catalog, which says "Clements, Samuel" and indicates that all the cards behind that tab are books by Mark Twain. Each page of the site must have a unique page title. Do not stuff keywords in the title, and be sure to front-load the title with the words that matter most.

▶ **Meta keywords.** This tag has virtually zero influence on the search engines because it is so vulnerable to spamming. The exceptions appear to be that Google AdSense syndication looks at the meta keywords tag and that Yahoo is influenced in a very tertiary way by it. Meta keywords

need to match the content of the page, and this tag is actually a good place to insert potential misspellings. It should be different for each page.

▶ **Meta description.** As with the page title and meta keywords, be certain that the meta description is unique to each page. This description is just that: a summary of what is contained on the page in question. Tell it, don't sell it, in about 150 to 160 characters. This content is critical because it is probably what search engines will display under the link to your page. If the page does not contain a meta description, the search engine will look for a snippet of text or other content that contains the keywords searched and display that in its results. The meta description is more about usability than SEO, so be certain that each page is properly tagged.

▶ **"Noindex" meta tag.** If you have any pages you do not want to include in search engine results, use the noindex meta tag. Just be certain that pages you do want to be indexed do not inadvertently contain this tag.

▶ **Headers.** Search engines recognize the headers <h1>, <h2>, and so on as influencing factors so long as you are not spamming with them. Take care to allow for section headers that are both descriptive and contain the relevant keywords for that page.

▶ **Link anchor text.** Link anchor text is an important influencer of what search engines think about the page on the other side of the link. This is the factor that creates the "GoogleBomb." If enough links point at a page with the same link anchor text, Google interprets the destination as relevant for the phrase in the anchor text. For instance, if you search on Google for "click here," the Adobe site will show up in the top results. There are hundreds of thousands of links that point at Adobe and read "click here to download Adobe Reader" or something similar. Use this to your advantage; anchor text should not be "More" or "Click Here." Instead, it should contain keywords that are relevant to the destination page.

Split the Hairs

It is to your advantage to have separately indexed pages for both your left-handed corrugated widgets and your right-handed corrugated widgets. This level of granularity gives your pages a better chance to be an exact match for the legendary long-tail searches. A page that is all about one thing has a

better chance of winning for that one thing than a page that is about multiple things (all other factors being equal of course). And who is interested in reading a page that is hundreds of words long anyway?

Use Site Maps

In recent years it has become popular to omit the classic site map page. This is a mistake for usability and a mistake for SEO. Find your way through to the fact that any site needs a site map. It may not be cool but it is necessary. Also, include site map files at /sitemap.xml and /sitemap.txt. Although this structure does not help the site rank better, it does help the search engines understand the directory structure and find new and updated content.

Keep Content Fresh

A key component to gaining and keeping top placement in search results is constantly refreshing the site content. This doesn't mean editing all the content in the site all the time; it means that the site must constantly grow. Build the directory structure so that adding content will be easy and intuitive, and anticipate that the site will grow over time.

Other Content Issues

A basic challenge in dealing with the UX of a content-rich site is to prevent cloning or duplicate content. Look out for creating duplicate pages with seemingly innocuous conveniences such as "printer friendly" content that is an exact duplicate of a page in a PDF or similar document type. Shield these kinds of pages with scripted links or use the rel="nofollow" link attribute.

Don't discount optimization for the wide array of digital assets that search engines can index. This topic would make almost another chapter in itself, but suffice it to say that PDFs, videos, images, and other non–Web page assets are clearly a part of natural search results. Structuring the wrappers around these assets is critical, because search engines need pointers to this kind of content. For example, search engines can't tell that an asset is a horse-race video unless there is a link pointing to the video with anchor text that reads "video of a horse race" placed near text about horse-racing videos in the page code.

Using alt attributes is another way to help identify nontext assets to the search engines and is always a good idea for the sake of usability.

Don't create dead-end content pages. Make sure that even the very bottom of the structure has links back into the main site, so the search engine spiders don't get stuck in a dead end. Breadcrumb links are a straightforward way to accomplish this if a page type does not contain the main site navigation.

Link Popularity Explained

If there is a Holy Grail of SEO, it is link popularity. It is the cornerstone of why Google worked so much better than the other search engines when it emerged on the scene. *Link popularity* is a determination of the quality and the quantity of links pointing at a Web asset from other Web pages. Google uses the term *PageRank,* and it is the über factor that can overcome many other deficiencies. Links are essentially votes for a Web asset, and it is generally assumed that something that is interesting or valuable to others will have links pointing at it from other trusted Web assets. Over time this concept has proven invaluable to overcome spamming efforts and is at its core a fundamental principle of quality search results. This principle is critical for the UX designer to comprehend because of the way that link popularity will distribute into the structure of a Web site.

Typical Link Popularity Distribution

Similar to the Richter scale used to measure the strength of seismic activity, Google's PageRank system (named by Larry Page for himself) is a logarithmic scale that ranges from 0 to 10. This means (in wildly general terms) that if one page has a PR of 4 and another has a PR of 5, the PR 5 page has 10 times the link popularity of the page with the 4.

It is important to understand this because PageRank distributes through a site based on the hierarchy of links and the structure of the directories. Generally speaking, if your home page has a page rank of 5, your primary section pages should have a PR of 4, the secondary pages PR 3, the tertiary pages PR 2, and so forth.

Pages with the most internal links pointing at them tend to have the greatest link popularity. The pages that are the most important to win need to have the most internal links pointing at them. This includes links in the main site navigation, site map, footer, and inline links embedded in text. Because link popularity is critical to ranking well in search results, you need to be as deliberate as possible in getting as much of it as you can into the pages that contain the "buying proposition."

Each page has a finite amount of link popularity that it can contribute to the other pages in the site. A page that has one link on it pointing at one other page is sending 100 percent of its available value to the recipient. A page that has a hundred links to a hundred other pages is sending 1 percent of its value to each of those hundred pages.

The home page tends to have the most link popularity, because the home page of a site tends to have the most links pointing at it from third-party Web sites. This means that the home page has the most value to contribute to other pages of the site. If there is a critical page that is a part of the "selling proposition," put in a direct link to it from the home page so that the search engines can understand how important this particular page is in comparison to the rest of the site. If possible, build a feature that can rotate links to deep content from the home page.

Footer Links

As we look for ways to marshal and control the distribution of link popularity throughout the site, remember that text links in the footer of each page are both a blessing and a curse, and they will have some bearing on the distribution of link popularity throughout the site. Typical footer links point at the privacy policy and other nontransactional pages, so if it is required that these links be in the footer, hide them behind some sort of scripting, or better yet, "nofollow" these links using the rel="nofollow" link attribute. This will prevent the link popularity of each page from leaking out to pages that do not really need to rank in search results. It is also better to prevent the passing of link popularity than to fully exclude the pages using robots.txt.

In-Content Cross-Linking

Search engines eat up links that are embedded in text. Just don't overdo it. Some schools of thought maintain that after the first few links in a block of text the search engines do not provide advantageous weighting. Put your most important links in the beginning of the text and use link anchor text that contains keywords that are related to the destination page.

Gaming the System

Who says that search engine optimization is all work and no fun? Search engines can contribute real dollars to Web site owners, and in certain environments, there is a real no-holds-barred approach to gaining top rankings at any cost. More than a few search engine optimizers have taken advantage of their clients, charging big bucks for spurious techniques that may have positive results in the short term but a devastating impact over time.

Over the years, a variety of optimization techniques have been employed by Webmasters looking for top results. One of the core evolutions in search engine technology has been work on engineering out the clever ways that have been found to game the system. From the search engines' perspective, their users' best interests are served by clean, highly relevant results at the top of any query. From the perspective of many search engine optimizers, all's fair in love and SEO.

White Hat Versus Black Hat

It is easy and fun to characterize SEO methods as being "white hat" or "black hat," but it's far more difficult to discern which is which. Many white hat optimizers are purists, saying in strong, declarative terms that certain technical management, content and link manipulation, and other approaches are simply off limits. The black hats look at the issue as a contest that has nothing to do with cheating: How can something be cheating if there is no specific written rule book or court of adjudication? Their approach is more along the lines of a game of cat and mouse where the cat holds all the cards and the mouse can stand to make some serious cash: Take a risk, get a win, and the payoff is big. But once the search engines catch up to you (and they nearly

always do) be prepared for your site to be banned or at the very least unable to perform when the methods are revoked.

Spamming with Meta Keywords

Many of the "cheating" techniques have been based in the principles of UX. An early method to game the system was *meta keywords stuffing*—essentially filling the meta keywords tag with hundreds of occurrences of *apples* when the site content is all about oranges. At its root, the meta keywords approach was created to help with the taxonomy of the early Web. Today, because of all the keyword spamming, this piece of a Web page has virtually no influence on search placement. The search engines easily detected this technique and were quickly able to engineer around it. The next generation of spam was a bit more difficult to unravel, and also had its roots in UX issues.

Cloning and Doorway Pages

Both *cloning* and *doorway pages* are methods used to make a Web site look bigger or different to the search engines. By cloning a page over and over, Webmasters could essentially manufacture minutely targeted content that could quickly dominate for a specific search phrase. Because of this practice, it is important to be sure your architecture prevents inadvertent cloning, since search engines are sensitive to duplications, intentional or otherwise.

Doorway pages are another method for influencing search results that straddles the gray space between white hat and black hat methods. On the one hand, Google's quality guidelines for Webmasters say "doorway pages … are in violation of our Webmaster guidelines" (www.google.com/support/webmasters/bin/answer.py?answer=66355). The guidelines identify doorway pages as poor-quality pages that have been specifically optimized for a set of keywords that may not be relevant to the actual site or that are spammy.

On the other hand, how do you help search engines gain access to content that is trapped in a non-spiderable database or is obscured because of a technology that search engines don't like? In its positive definition, a doorway page is high-quality static content that search engines can see and understand while providing the visitor a door into dynamic content. Today's content management systems are getting better at the core issues that

have necessitated this approach, but it can still be very useful to create extra pages for the express purpose of showing the search engines a static representation of content that they would otherwise be unable to deal with.

Link Spamming

Recent methods for gaming the system have centered on manipulation of link popularity, a concept that is core to the way that the modern Web search engines work. As discussed above, search engines determine the relevance of a particular Web asset by analyzing the quantity and quality of links pointing at that page from other places. Search engine optimizers have worked to manipulate this part of the puzzle through sneaky redirecting, overuse of subdomains, making every page of a site "/index.html," and a variety of other subtle machinations.

Some Final Thoughts

It is doubtful that this is your first exploration of search engine optimization issues. By now it is clear how much a site's architecture and related issues influence search engine performance. The current search environment is a quantum leap ahead of simple taxonomy or structure.

Thoughtful search engine optimization starts with quality UX. The architecting of a Web site is the critical point in its life cycle where it can either be destined for search engine success or set up for imminent failure. Search engine optimization is a strategy that never really ends, but quality SEO will never really get started without the careful attention of the UX designer.

Jonathan Ashton is vice president of SEO and web analytics for Agency.com and runs its Center of Excellence for SEO. His team provides SEO services for the entire company, ensuring that the process of designing and building rich interactive experiences results in sites that can be found on search engines. His monthly column, "Industrial Strength SEO," can be found at http://searchengineland.com/lands/columns/industrial-strength.php.