

Being a Bigdaddy Jagger Meister

By Jim Hedger, StepForth News Editor, [StepForth Placement Inc.](#)

It took a little while to start to figure it out. Such things almost always do. After months of observation, research, discussion and debate, Search Engine Optimization experts appear to be getting a better handle on the effects of Google's Bigdaddy infrastructure upgrades.

From mid-winter until this week, StepForth has strongly advised our clients to be conservative with any changes to their sites until enough time has passed for us, along with many others in the SEO community to observe, analyze and articulate our impressions of the upgrade. About ten days ago, the light at the end of the intellectual tunnel became eminently visible and SEO discussion forums are abuzz with productive and proactive conversations regarding how to deal with a post-Bigdaddy Google environment.

To make the long back-story short, in September 2005, Google began implementation of a three-part algorithm update that became known as the Jagger Update. Shortly after completing the algo update in late November, Google began an upgrading of their server and data storage network that was dubbed the Bigdaddy Infrastructure Upgrade. The Bigdaddy upgrade took several months to completely roll out across all of Google's data centers, which are rumoured to number in the hundreds.

In other words, the world's most popular search engine has, in one way or another, been in a constant state of flux since September. The only solid information SEOs had to pass on to curious clients amounted to time tested truisms about good content, links and site structure. Being the responsible sort we are, no good SEO wanted to say anything definite for fear of being downright wrong and misdirecting others.

Starting in the middle of May and increasing towards the end of the month, ideas and theories that had been thrown around SEO related forums and discussion groups started to solidify into the functional knowledge that makes up the intellectual inventory of good SEO firms.

Guided by the timely information leaks from Google's quality control czar [Matt Cutts](#), discussion in the SEO community surrounding Bigdaddy related issues has been led by members of forums [WebMasterWorld](#), SEW (1) (2) (3), [Threadwatch](#), SERoundtable (1) (2) (3), and Cre8asite (1) (2). SEO writers [Aaron Wall](#), [Jarrod Hunt](#), and [Mark Daoust](#) have also added their observations to the conversation in a number of separate articles.

By now, most good SEOs should be able to put their fingers on issues related to Bigdaddy fairly quickly and help work out a strategy for sites that were adversely affected by the upgrade. The first thing to note about the cumulative effects of Jagger and Bigdaddy is the intent of Google engineers to remove much of the poor quality or outright spammy commercial content that was clogging up their search results.

The intended targets of the Bigdaddy update go beyond sites that commit simple violations against Google's webmaster guidelines to include affiliate sites, results gleaned from other search tools, duplicate content, poor quality sites and sites with obviously gamey link networks. In some cases, Google was targeting sites designed primarily to attract users to click on paid-search advertisements.

After the implementation of the algo update and infrastructure upgrades, SEOs have seen changes in the following areas: Site/Document Quality Scoring, Duplicate Content Filtering and Link Intention Analysis.

The first area noted is site or document quality scoring. Did you know there are now more web documents online than there are people on the planet? Many if not most of those documents are highly

Search engine placement experts

StepForth Placement Inc.

2-1358 Gladstone Avenue
Victoria, BC V8R 1S1 Canada

T 250 385-1190
F 250 385-1198

1 877 385-5526
info@stepforth.com

professional and some are sort of scrappy. While Google is not looking for perfection, it is trying to assess which pages are more useful than others and attention to quality design and content is one of the criteria.

Quality design simply means giving Google full access to all areas of the site webmasters want spidered. Smart site and directory structures tend to place spiderable information as high in the directory tree as possible. While Google is capable of spidering deeply into database sites, it appears to prefer to visit higher level directories much more frequently. We have noted that Google agent visits do tend to correspond with update times set via Google Sitemaps.

Accessibility and usability issues are thought to make up elements of the Jagger algorithm update, marking the way visitors use a site or document and the amount of time they spend engaged in a user-session associated with a site important factors in ranking and placement outcomes. Internal and outbound links should be placed with care in order to make navigating through and away from a site as easy as possible for site visitors.

Another quality design issue involves letting Google know which "version" of your site is the correct one. For most, a website can be access with or without typing the "www" part of the URL. (ie: <http://www.stepforth.com>, <http://stepforth.com>, <http://www.stepforth.com/index.shtml>) This presents a rather funny problem for Google. Because links directed into a site might vary in the way they are written, sometimes it doesn't know which "version" of a site is the correct one to keep in its cache. To Google, each of the variations of the URL above could be perceived as unique websites, an issue known as "canonicalization", a subject Matt Cutts [addressed on his blog](#) in early January.

"Suppose you want your default URL to be *http://www.example.com*. You can make your webserver so that if someone requests *http://example.com/*, it does a 301 (permanent) redirect to *http://www.example.com/*. That helps Google know which url you prefer to be canonical. Adding a 301 redirect can be an especially good idea if your site changes often (e.g. dynamic content, a blog, etc.)."

Quality content is a bit harder to manage and a lot harder to define. Content is a word used to describe the stuff in or on a web document and could include everything from text and images, Flash files, audio or video and links.

There are two basic rules in regards to content. It should be there to inform and assist the site user's experience and it should be, (in as much as possible), original.

Making an easy to use site that provides visitors with the information they are looking for is the responsibility of webmasters but there are a few simple ways to show you are serious about its presentation.

Focus on your topic and stick to it. Many of the sites and documents that have found themselves demoted or de-listed during the Bigdaddy upgrade were sites that delved across several topics at the same time without presenting a clear theme. Given the option between documents with clear themes and documents without clear themes, Google's choice is obvious.

Google is working to weed out duplicate content. Google appears to be looking for incidents of duplicate content in order to try to promote the originator of that content over the replications. This has hit sites in the vertical search sector; affiliate marketing sector, real estate sites, and even retail sites that carry brand name products, especially hard. Several shopping or product focused database sites have seen hundreds or even thousands of pages falling out of Google's main index.

In many cases, there is little or nothing to do for this except to start writing unique content for products listed in the databases of such sites. Many real estate sites, for example, use the same local information sources as their competitors do and all tend to draw content from the same selection of MLS listings. It's

not that Google thinks this content is “useless”, it’s that Google already has several other examples of the same content and is not interested in displaying duplicate listings.

Many of the listings previously enjoyed by large database driven sites have fallen into a Google index known as the supplemental listings database. Supplemental listings are introduced to the general listings shown to Google users when there are no better examples to choose from to meet the users search query. This is the same index that is often referred to as the “Google Sandbox”.

The last major element noted in the discussions surrounding Bigdaddy is how much more robust Google’s link analysis has become. Aside from site quality and duplicate content issues, most webmasters will find answers to riddles posed by Bigdaddy in their network of links.

In order to ferret out the intent of webmasters, Google has increased the importance of links, both inbound and outbound. Before the updates, an overused tactic for strong placement at Google saw webmasters trying to bulk up on incoming links from where ever they could. This practice saw the rise of link farms, link exchanges and poorly planned reciprocal link networks.

One of the ways Google tries to judge the intent of webmasters is by mapping the network of incoming and outgoing links associated with a domain. Links, according to the gospel of Google, should exist only as navigation or information references that benefit the site visitor. Google examines the outbound links from a page or document and compares them against its list of inbound links, checking to see how many match up and how many are directed towards, and/or coming from pages featuring unrelated or irrelevant content. Links pointing to or from irrelevant content or reciprocal links between topically unrelated sites are easily spotted and their value to the overall site ranking downgraded or even eliminated.

The subject of links brings up an uncomfortable flaw in Google’s inbound link analysis that is being referred to as Google Bowling. As part of its scan of the network of links associated with a document or URL, Google keeps a detailed record of who links to who, how long the link has been established, if there is a recip link back, along with several other items.

One of those items appears to be an examination of how and why webmasters might purchase links from another site. While bought-links are not technically a rankings killer, a bulk of such links purchased from un-relevant sites in a short period of time, can effectively destroy a site or document’s current or potential rankings. In an article published at WebProNews last autumn, [Michael Perdone](#) from e-TrafficJams speculates on the issue.

Google has tried to deal with the predatory practice of “Google Bowling” by considering the behaviour of webmasters whose site have seen a number of inbound links from “bad neighbourhoods” suddenly appear. If a site that has incoming links from bad places also has or creates links directed out-bound to bad places, the incoming links are judged more harshly. If, on the other hand, a website has a sudden influx of bad-neighbourhood links but does not contain outbound-links directed to bad places, the inbound ones might not be judged as harshly.

The combination of the Bigdaddy upgrade and the Jagger algorithm update have made Google a better search engine and are precursors to the integration of video content and other information pulled from other Google services such as Google Base, Google Maps and Google Groups in the general search results.

Before the completion of both, Google’s search results were increasingly displaying a number of poor quality results stemming from a legion of scraped content “splog” sites and phoney directories that had sprung up in efforts to exploit the AdSense payment system. Bigdaddy and Jagger is a combined effort to

offer improved, more accurate rankings while at the same time, expanding Google's ability to draw and distribute content across its multiple arms and networks.

Moving forward, that is what users should expect Google to do. Google is no longer a set of static results updated on a timed schedule. It is constantly updating and rethinking its rankings, especially in light of the number of people trying to use those rankings for their own commercial gain.

The effect of the duo upgrades seems to be settling out. Credible, informative sites should have nothing to worry about in the post-Bigdaddy environment. As Google is trying to move into the most mainstream areas of modern marketing, credibility is its chief concern. The greatest threat to Google's dominance does not come from other tech firms. It comes from the users themselves who, if displeased with results being shown by Google, could migrate en masse to another search engine.