

"The Story Behind Google's Bigdaddy Update..."

In the beginning of 2006, Google began publicly testing something that Matt Cutts named the **Bigdaddy Update**. Unlike previous algorithm updates, this update was much, much bigger, although the results weren't immediately visible.

Google uses a network of data centers with different IP addresses to answer search queries. These decentralized servers share the workload of indexing web sites. In effect, Bigdaddy was what you would call an upgrade to how Google managed its search indexes – an update to its data center infrastructure. It contained new code for sorting and examining web pages, and was billed to handle technical indexing issues much better.

Bigdaddy in effect was what you would call a massive **software infrastructure update** to Google's datacenters.

The official word from Google (through their virtual spokesman, [Matt Cutts](#)), was that the Bigdaddy update was a major step towards delivering **more relevant results** to search queries. In English, that just means that if you're trying to rank for your keywords on a website by following Google's webmaster guidelines, it's going to get easier for you. On the other hand, if you're spamming / using black hat techniques to rank in the top 10, your websites may take a massive hit.

The update went into full implementation at all of Google's data servers towards the end of March 2006/ beginning of April 2006.

What Bigdaddy Is All About

To quote Matt on this:

"...this data center improves in several of the ways that you'd measure a search engine."

and:

"the new infrastructure... will let us tackle canonicalization, dupes, and redirects in a much better way going forward compared to the current Google infrastructure."

Redirects

Most specifically, the Bigdaddy update deals with redirects as they pertain to **result relevance** and page hijacking. There are two types of redirects – permanent and temporary. Permanent redirects are known as **301 redirects** whereas temporary redirects are **302 redirects**. A redirect can be to another page on the same site, or it can be to a different website.

While permanent redirects tell search engines that a page has permanently moved to a new location, temporary redirects are a bit trickier to handle.

The resultant discussion on this topic (how search engines treat redirects) gets pretty technical, but the point here is that if you are using temporary (302) redirects, the new upgrade will help Google treat those redirects more effectively.

URL Canonicalization

Despite the difficult sounding name, this is actually quite a straight-forward concept. (but one that has caused search engines – and by extension, webmasters – lots of problems).

To start off with, let's take the example of Keyword Elite. Now there are several ways to write the web address:

- <http://keywordelite.com>
- <http://www.keywordelite.com>
- <http://www.keywordelite.com/index.htm>
- <http://keywordelite.com/index.htm>

For most of us, all four urls mean the **same thing** – the web address of Keyword Elite.

However, technically all four urls are **different**. You can judge that by considering a situation where your web host returns four different pages for these four urls. Canonicalization is the process of picking the url that is the **best option** for being the website's home page. It's an area search engines have had serious problems with (and is the reason why you might see a different PageRank value for your www and non-www versions of the same domain).

It helps to consolidate your inbound links and by extension, your search engine rankings.

To make sure that Google picks the url that you want to use (it really doesn't matter if you pick 'www' or non-www, at least for now), and make sure that you use that url consistently throughout your website. That is, if you are using www.example.com, then make sure you use that format and not any of the three other options as I showed above.

A second, and more fool-proof method, is to use 301 redirects on your web server to ensure that instead of having two websites (with and without the 'www'), you can keep just one version.

Here's a [301 redirection tutorial](#) if you need the code with examples.

New Google Spider

The Bigdaddy update was based on creating a brand new search index. This index has been built using a new spider, now known as Mozilla Bot. The reasons for using a new spider? Faster crawling, **smarter indexing**, and most likely, the ability to index different content in more depth.

Improved Results

The new infrastructure will allow Google to develop more advanced algorithms and larger databases.

Another reason for the new data center infrastructure is that Google wants to be able to index different content types. As I mentioned earlier, Google has a new spider (Mozilla Bot), which is based on the Mozilla browser. The new spider should be able to index **more** than traditional search engine spiders, possibly links within images, JavaScripts or Flash files (but this won't happen for quite some time after the update).

How Will Bigdaddy Affect Your Rankings?

This is the question most people have asked about Bigdaddy – how will this latest update to Google change rankings (yours and mine)?

Supplemental results

Supplemental results in Google are from an alternate index. These are only used when Google cannot find relevant results in its main index.

With the Bigdaddy update, Google was re-crawling the web using their new spider, Mozilla Bot. The problem with this crawl/index cycle was that as results from datacenters went live on Google (these were going on and off during the testing phase), the index did not contain all the pages from some websites.

Thus, when webmasters were searching for how well their websites were indexed by Bigdaddy (using the 'site:' operator e.g. `site:seoelite.com`), they sometimes found that their previously well-indexed and highly ranked websites had almost **disappeared** from the index! Usually only the main page would be listed, and the rest were shown as 'supplemental' results – results Google was pulling from an older index.

This situation has more or less resolved itself by the time Bigdaddy went live. Some people speculated that this **supplemental results** issue was due to a bug in Google's new spider – that's partially true. To understand why this happened, you have to understand how Google crawls and indexes pages. One of the main criteria for indexing and crawling is PageRank – Google's measure for how popular a page is. This, along with other factors, creates an **indexing threshold** for pages – if a page is above the threshold, it will get indexed, otherwise it won't.

A current search on Google (`site:seoelite.com`) gives me a figure of 536 pages – pretty good considering that this is a 1-2 page website, with several dozen lessons and a forum. Of course, the forum could stand to be indexed further, but that will change as Mozilla Bot crawls deeper.

Rankings going up

If your rankings went up as a result of the Bigdaddy update, congratulations – you must be doing something right. Still, if you know what you did right, it will help you rank better in the future as well.

I'll be doing a lesson soon on search result relevancy (how search engines determine the **most relevant results**) – needless to say, understanding this is mega-important on predicting what your next moves should be regarding search engine optimization.

Rankings taking a nosedive

Did your website rankings take a drop with the Bigdaddy update? Most of the problems webmasters have had in their rankings have been due to the rebuilding of the index. Many websites had their pages 'disappear' from Google's index overnight and panicked – in fact, this was just the effects of the Bigdaddy update in progress. As the index is fully rebuilt, any ranking issues that took place because of lost pages should be dealt with.

However, in certain cases your websites may still suffer a drop in rankings. This is mainly due to certain tweaks in the algo as well as part of Google's efforts to make search results more relevant.

Recap

Bigdaddy is a software upgrade to Google's infrastructure that provides the framework for a lot of improvements to core search quality post-Bigdaddy (starting April 2006) such as smarter redirect handling, improved canonicalization, etc.

This upgrade features a new Google spider, **a brand new index**, improved spam filtering (or at least what Google considers spam), and a massive restructuring of how Google's datacenters process and index websites. This should allow Google to index different types of content, as well as scale better.

Most importantly though, this update will help Google improve the **quality of search results**.

That's all for today – I hope that this lesson will help explain what the Bigdaddy update was all about (especially after the fuss it caused in SEO forums).

All the best,



Brad Callen
Professional SEO
<http://www.seoelite.com>

**Get Your Free Copy of 7 Days To Massive Website Traffic!
by Optimization Expert, Brad Callen**

If you liked the lesson above and haven't yet signed up for my SEO course, simply submit your first name and email address in the form below to receive many more powerful lessons from me... **RIGHT NOW!** (Check your email in a few seconds) If you're already signed up, you'll automatically receive more lessons from me without signing up again.

First Name:

E-mail Address:

Your privacy is SAFE with us.