Are SES URLs evil?

Posted At : 06 August 2009 11:08 | Posted By : Gareth
Related Categories: SES URLs

I was reading a recent article in .NET magazine about optimising Flash sites for Google. When I got to the box on *Friendly URLs*, I thought it ironic that they were perpetuating another web design myth. After all, their recent June issue had been dedicated to myth busting, and SES URLs are the biggest myth out there?

Checking back however, I realised that even .NET magazine are still taken in …

At this point, I should point out that I'm aware of the irony in this post. I'm well aware that I'm writing an article against SES URLs on a website that uses them. This blog is powered by Ray Camden's BlogCFC which has SES URLs by default. I've also implemented SES URLs on other projects.

The Myth

Every so often we get a customer complaining that their URLs aren't *search engine friendly*. Normally it's the result of some *expert* telling them that Google can't index dynamic URLs.

If you're not aware yet, dynamic URLs are often of the form http://www.mysite.com/product.cfm?productID=123. This is typically used when the page is produced dynamically. Often from a database. The *productID* and number bit are referred to as a *URL parameter*.

Now many people mistakenly believe that search engines such as Google can't index pages with these URLs. Often they will re-write this URL along the lines of http://www.mysite.com/product/productID/123.

The truth is that search engines have been able to index dynamic URLs for years without problem. Google even debunk this myth on their own **official blog**. If that's not enough to convince clients, we ask them to do a Google search such as

```
site:www.mysite.com
```

. That search will list every page indexed by Google for the site.

The Problems with SES URLs

Now maybe *evil* was a bit too string a word to use, but there are certainly disadvantages to using SES URLs.

- You need to use a rewrite module such as Apache's mod_rewrite. As well as the extra time and knowledge setting this up requires, it makes your application less portable if you want it to work on IIS as well.
- Relative URLs don't work - you need to link everything from the web root. There's a lot more work invloved if you wanted to change your directory structure.
- You need to take extra care not to duplicate content. While it's easy for Google to work out that www.mysite.com/product.cfm?productID=123&categoryID=456 is the same page as www.mysite.com/product.cfm?categoryID=456&productID=123, Google would consider www.mysite.com/productID/123/categoryID/456 and www.mysite.com/categoryID/456/productID/123 to be two separate pages with duplicated content. As soon as Google thinks you are duplicating content, your search ranking will plummet.
- You also can't put session or other client parameters in the URL, or it will confuse the crawlers. For example, Google is clever enough to work out that the CFID parameter here is tied to a client's session: www.mysite.com/product.cfm?productID=123&CFID=12345678. It can therefore ignore it. www.mysite.com/product/productID/123/CFID/12345678 however will confuse Google.

Are there any reasons left why I might use SES URLs

Now to be fair to SES URLs, there **may** be some benefits:

- Adding kewords in the URL **might** increase search engine ranking. Some SEO experts claim this and Google don't specifcally mention in it in their blog article. However, if Google are discouraging URL re-writing, it would seem strange.
- From a usability point of view, it's fairly obvious that mysite.com/contact is easier to remember than mysite.com/page.cfm?pageID=789. However, there's nothing to stop you setting up /contact as a permanent redirect to /page.cfm?pageID=789 then you have the best of both worlds. And Google likes this approach.

And Finally…

Use the correct tool for the job.

A few years back, Jeffrey Zeldman taught us to stop abusing table tags, and to use html for the purpose it was designed for. Everyone now sees the benefits.

URL parameters were designed for use by dynamic web pages, not complicated folder structures. Search engines understand this, and it gives them a better understanding of how your site works.

While the internet may have started off as a collection of static pages, it's now about online applications. As the likes of Google develop more intelligent crawlers, they will not only be able to read your static content, but interact with your application. Who knows what this may bring?