

Tuesday, August 24. 2004

Cloaking like a search engine would want you to

A recent entry in our patSpiderizer forum initiated me to renew the discussion about the ethics of cloaking. patSpiderizer enables you to create spider pages with contents tailored for a spider's needs, with the motivation to help websites with few indexable content (like flash-sites) to get listed in search engines. At the time of the initial release of patSpiderizer in 2001, about everybody was involved in cloaking, and everyone had his own more or less successful cloaking recipe. Can you remember those pages with a giant keyword block at the bottom with black text on black background? At the time, this worked, as the spiders did not filter those pages out yet.

What has changed since then? Not much - the cloaker/spider war goes on, each side inventing new ways to screw each other. Did you know that some spiders actually check the grammatical structure of your website's content? Forget keyword lists. All this nonsense awakens my urge to break a few necks. Imagine all that fake content automatically generated from keyword lists, just to get to the top of search engine rankings until the spiders filter those out too... I would not want to have to write the search algorithms needed to get intelligible search results from that heap of garbage.

This is not only about ethics - it's also about better ways to success. I have several patSpiderizer installations that have been running for about three years now, and apart from updating the content from time to time, I did not have to change anything to get very good results. As I see it, there is only one way to cloak - the way a search engine would want you to. So let's see how that can be done:

- Cloaking: this implies that you hide something from the spiders, but why would you have to? Don't interfere with the way a spider indexes your site - on the contrary, help it by providing links to relevant pages, and especially those which may be hidden quite far in the navigation.
- Automatic content generation: admit it, you are just lazy. You will only get good and persistent rankings if you don't invent any new content, but mirror that of your website, in a more organized fashion. Group your content by theme, and create spider pages for each theme so that it is clear the user will find that particular information on your website (example: if you sell vacuum cleaners and soap, make one page for vacuum cleaners and one for soap). Also, always add links to the corresponding pages in your site so they get indexed too - after all, those are the pages your visitors will want to see!
- Link rating: no secrets here - the more external links you have to your site, the better your ranking. If you want a good link rating, you will have to make sure you have the best information so that everybody links to you.
- Structuring your content: if you are not happy with your search engine rankings, don't blame the search engine, blame yourself. There is a lot of time to be invested in finding the right wording to get good results, and also in finding out what people search for to find you. I have one really search-engine obsessed customer that has a patSpiderizer installation, and she spends a lot of time changing the wording of her spider pages, with a hell of success. Cloaking simply gives you the possibility to structure your content in a way you cannot in your website.

My conclusion is that there is no miracle solution, no hack, no trick, no joker that will shazam! get you a pole position in search engines without giving it some thought, work, and respecting the way the web works. And with the risk of getting your neck broken by a reactionary spider-friendly cloaker like myself, I would not dare try the opposite

P.S. Comments welcome in the forum

Posted by argh in PHP at 13:40