

Generation 2 search engines – backlinks rule

Google was launched in 1998 as a project in Stamford University using a new breed of ‘spider’. This kind of spider was not the hairy scuttling thing that you have to remove from the bathroom, but the automated program that analyses links on webpages, then stores the addresses of the linked pages to analyse at a future date for inclusion into its index. Spiders appeared on the internet in the early 1990’s but Google’s spider was different from previous ones in that it used the information about links to web pages to influence the position of the web page in the search engine results pages (SERPs). The more ‘backlinks’ a site had, the more it was assumed that that site was an authority on the particular subject and so the better its relevancy score. Other search engines quickly adopted this method.

This then changed the tactics of SEO so that their focus was no longer just on what was on their own web pages, but also on getting back links. Lots of backlinks. Link exchange schemes appeared and link directories started sprouting up. The more back links you could get to your site, the better your rankings were. Automated software was developed to create dynamic databases of links (we even made one or two such programs ourselves) and webmasters were emailing other site owners to ask for links to be added. SEO software even had tools to automatically generate polite sounding emails to other site owners to request a link on their site.

Foreign companies then got into the action. In particular, Indian SEO companies were being set up whose only product was to create fake blog sites to give your site more backlinks. They would typically promise 100 backlinks per dollar with a minimum commitment of a few dollars per month. Experiences with these companies were not generally very good. As soon as you stopped paying them, those backlinks would disappear or be redirected to other sites.

Because of the increase in spam websites floating to the top of the search engines, Google added a new feature to the backlinks called Page Rank (PR). PR was named after Larry Page, one of the founders of Google. I guess Brin Rank just didn’t sound right. Originally the concept was used in the still popular Baidu search engine in China.

Page rank worked by giving backlinks different ‘weightings’ according to how popular the page containing the backlinks was. This then meant that creating thousands of spam backlinks on ‘ghost pages’ and ‘doorway pages’ no longer had such an effect.

The way PR worked was like this: if a page has lots of links to it, it gets more ‘link juice’ (Google’s word for it, not ours). The juicier a page is, the more powerful its links became. However, at the same time, the more outbound links a page had, the more ‘leaky’ the page became, so the amount of link juice transferred to the pages that it linked to was reduced. To get a really juicy page, what you need is a few juicy pages linking to yours which don’t have dozens of leaks out of them.

How Link Juice is Transferred



Fig 4: Link Juice Flow diagram - source : <http://blog.woorank.com/2013/05/the-flow-of-link-juice/>

In the diagram above, we have 2 pages, A and B. A has twice as many inbound links than B so should normally have twice the amount of page rank. However, the external webpages that link to A are much more leaky than those that link to B so they pass on less juice. This effect passes on down the linking chain, so you will find that sites with good links from high PR pages have much higher PR themselves. That's why backlinks from high PR pages were much more valuable.

Since A is much more juicy than B, its links have much greater value. Note, there are ways to plug the link juice leaks in your page by using 'nofollow' links.

Another aspect of link juice is that the links have 'flavours' – each flavour is associated with a keyword, so for example if you have lots of high PR pages linking to yours with the keyword 'socks' then your link juice for that keyword will be high, however your ranking for other words like 'hats' will not be increased as much by these links.

Google used to publish the page rank data for websites and made it available in their free browser toolbar. However, they realised that having this data made it easier for spammers to manipulate the rankings, so they stopped updating the database and PR is now a secret part of the search engine's ranking systems.