

---

## Avoid duplicate content -nofollow on useless links

Posted by silexian - 2007/12/28 06:41

---

Hello

I would like to know if something can be done with Fireboard, in order to avoid having duplicate content in search engines.

I explain : I would like to set links as nofollow the following URLs :

- reply links (why crawl the reply or new thread pages ?)
- edition links (idem)
- display mode links (example : "page1.html" is the same as "page1.html?view=flat" AND "page1.html?display=threaded")
- useless links that are not contents and shouldn't be crawled.

Do you think something can be done ?

The philosophy behind this is "Less pages crawled but only content pages (indexes,threads, posts,...) have to be crawled".

What do you think that ?

regards  
lolo

=====

## Re:Avoid duplicate content -nofollow on useless li

Posted by AlexandreFrance - 2008/03/12 06:23

---

Hello,

I am afraid of a so bad SEO forum and no answer to your question since 2 months:

I am a professional webmaster and migrated from phpBB+SEO mod rewrite on october to Fireboard on one of my websites:

Results is 3500 to 5000 unique visitor per day before migration.

And now less than 3000 visitors per day: since november.

I lost almost all my Google incoming traffic on my big forum because Google do not like my new fireboard duplicate content and bad SEO, and at the same time my others pages outside of the forum have better results.

So I guess after this test, professional webmasters should not use firebard because of SEO.

I am curenly looking to migrate from fireboard to phpBB3, but what a work!

Really SAD the Fireboard team does not care about the duplicate content HUDGE problem, this sabotage their very nice work.

However, if a FB team member read this topic before I migrate, I would be please to give money to sponsor a real fix for this problem.

Please note also that duplicate content generate THOUSANDS of rewrited urls with SH404.

This is really insane for my database on a dedicated server, and I don't imagine how this would be on a shared server.

All the best,

Alexandre

=====

## Re:Avoid duplicate content -nofollow on useless li

Posted by fxstein - 2008/04/18 10:42

---

You are correct that there are lots of duplicate links. Had some time tonight to fix most of these issues inside SVN version 673.

---

It adds nofollow to the following types of links:

user created links inside posts and signatures  
threaded vs flat  
latest posts (4h, 8h, .... 1 week)  
user and admin profiles  
user avatars  
"report to moderator" link

If you are familiar with SVN you can simply grab the files in SVN 673.

Let me know if I missed some.

Hope this helps!

Oliver

---

## Re:Avoid duplicate content -nofollow on useless li

Posted by fxstein - 2008/04/18 11:53

---

Forgot view.php - Is now also fixed in SVN 676

---

## Re:Avoid duplicate content -nofollow on useless li

Posted by rickert - 2008/07/08 15:03

---

Hi fxstein,

Thanks for your work. About possible ommisions: shouldn't the links to the latest messages on the fireboard homepage be nofollowed as well? The identical message will be found by the bot under the forum category but with a different url (and will result in duplicate content).

Rickert

---

## Re:Avoid duplicate content -nofollow on useless li

Posted by fxstein - 2008/07/08 19:00

---

Possibly. Please check the latest nightly build (only for testing), as lots of changes have been made since the early SEO work:

<http://joomlancode.org/gf/project/fireboard/frs/>

In general there is a much larger issue that FireBoard creates too many unique URLs pointing to the same content. I have focused on deduping that a lot but there is still work left. For example: FB allows you to address each post by its post ID, rather than by the thread and page it sits on. Very nice and simple for anybody who wants to reach a certain post, but very bad for SEO.

Even though I have rewritten most of those links to be properly paginated, every external tool or plugin uses the 'easy' way of pointing to a post within a thread. You only need one of those installed and search engines get thrown for a loop.

I am contemplating on providing an external API of sorts for other tools to get the URL to a particular post, that is properly paginated and formatted and to create 303 (permanent redirect) codes for all other URL formats, so search engines know that these are redirect rather than unique pages.

Have not started the work on this as I am helping to get 1.0.5 out the door.

Comments?

---

## Re:Avoid duplicate content -nofollow on useless li

Posted by rickert - 2008/07/08 23:48

I think this is a better way than using nofollow to exclude duplicate urls. As I understand nofollow has never been never intended to do this.

<http://en.wikipedia.org/wiki/Nofollow> says that while Google will not index nofollowed pages it will not exclude duplicate content that is already indexed. Other searchengines like Ask ingnore the nofollow tag altogether (and will follow AND index).

So if you say you've already rewritten most links to be properly paginated does that mean that you're advising not to use the code in SVN673 and wait till 1.0.5?

Rickert

---

## Re:Avoid duplicate content -nofollow on useless links

Posted by pimpag - 2008/07/09 00:10

Yeh nofollow will only help for google, i think, most/if not all other engines ignore the tag and index what they like lol.

---

## Re:Avoid duplicate content -nofollow on useless li

Posted by rickert - 2008/07/11 19:46

@fxstein,

So if you say you've already rewritten most links to be properly paginated does that mean that you're advising not to use the code in SVN673 and wait till 1.0.5?

Could you please give me a comment on this?

Thanks upfront

---

## Re:Avoid duplicate content -nofollow on useless li

Posted by fxstein - 2008/07/11 20:09

rickert wrote:

@fxstein,

So if you say you've already rewritten most links to be properly paginated does that mean that you're advising not to use the code in SVN673 and wait till 1.0.5?

Could you please give me a comment on this?

Thanks upfront

Yes SVN 673 is way outdated by now - I believe the SVN is over 800 by now. Please test the most current nightly builds instead.

Thx!

---

=====

## Re:Avoid duplicate content -nofollow on useless li

Posted by rickert - 2008/07/12 23:59

---

@fxstein:

I don't know how to download the complete nightly build but I managed to download the latest versions of the files you mentioned earlier (like listcat, showcat etc).

I did a small test to look at the url's and it looks better indeed. However, could you please check on the links that are created for pages in flat-view with a pagenumber > 1. The links for 'begin', 'previous' (on page 2) and pagenumber '1' end with '/limit,x/limitstart,0/' which will result in duplicate links. The category number follows the id number for these links as well.

Ofcourse I could be wrong because I did not try the full nightly build but if you could check that would be nice.

I saw that the links from flat view to threaded view vv are still nofollowed. Are you sure this will stop Google to index the threaded pages if the default is flat view?

Thanks,

Rickert

---

## Re:Avoid duplicate content -nofollow on useless li

Posted by fxstein - 2008/07/13 00:32

---

Here is the link to find the latest nightly builds:

<http://joomlancode.org/gf/project/fireboard/frs/>

The page links MUST have the limit-start format and so should all other links except for the very first post of a thread. It was not like that before and that cause for huge amounts of URLs pointing to the very same content - not something google & Co like.

---

## Re:Avoid duplicate content -nofollow on useless li

Posted by rickert - 2008/07/13 02:41

---

@fxstein:

Maybe I misunderstand but

[http://www.bestofjoomla.com/component/option,com\\_fireboard/Itemid,38/func,view/id,36033/catid,3/](http://www.bestofjoomla.com/component/option,com_fireboard/Itemid,38/func,view/id,36033/catid,3/) (this url of this post is found f.e. in the recent discussions tab and most other places)

and

[http://www.bestofjoomla.com/component/option,com\\_fireboard/Itemid,38/func,view/id,53648/catid,3/limit,10/limitstart,0/](http://www.bestofjoomla.com/component/option,com_fireboard/Itemid,38/func,view/id,53648/catid,3/limit,10/limitstart,0/) (this url is found on page 2 of this post)

represent the same content in my opinion. Why does the pagelink to the first post of a thread have a limit-start format (with limit-start equal to 0)?

Rickert

---

---

## Re:Avoid duplicate content -nofollow on useless li

Posted by fxstein - 2008/07/13 02:47

---

rickert wrote:

@fxstein:

Maybe I misunderstand but

[http://www.bestofjoomla.com/component/option,com\\_fireboard/Itemid,38/func,view/catid,3/id,53648](http://www.bestofjoomla.com/component/option,com_fireboard/Itemid,38/func,view/catid,3/id,53648) (this url is found f.e. in the recent discussions tab and most other places)

and

[http://www.bestofjoomla.com/component/option,com\\_fireboard/Itemid,38/func,view/id,53648/catid,3/limit,10/limitstart,0/](http://www.bestofjoomla.com/component/option,com_fireboard/Itemid,38/func,view/id,53648/catid,3/limit,10/limitstart,0/) (this url is found on page 2 of this post)

represent the same content in my opinion. Why does the pagelink to the first post of a thread have a limit-start format (with limit-start equal to 0)?

Rickert

I am not in charge of BOJ. I only contribute to the code line of FB. The latest builds of FireBoard are ahead of BOJ. BOJ does not have the mentioned SEO optimization in place. The correct format for content on multi page threads is the second one with limit and limitstart. Tools like sh404SEF will convert this to URLs with the ending -PageX

Here is an example of a long multi page thread on my site that has pretty much the latest SEO optimization in place:  
<http://www.starvmax.com/Forum/General-VMAX-Discussion/4821-Claim-number-sequence-where-are-you-in-line/Page-36.html#7210>

---

## Re:Avoid duplicate content -nofollow on useless li

Posted by rickert - 2008/07/13 12:16

---

Yes but now sh404SEF solves the duplication problem. Indeed urls that end with limit-start equal to 0 are rewritten by sh404SEF to urls WITHOUT a '/Page-1.html' extension, which would otherwise be urls that would introduce duplicate content.

If you deactivate sh404SEF for a moment you will see my point.

Example:

[http://www.domainname.com/component/option,com\\_fireboard/Itemid,38/func,view/catid,9/id,155/](http://www.domainname.com/component/option,com_fireboard/Itemid,38/func,view/catid,9/id,155/)

and

[http://www.domainname.com/component/option,com\\_fireboard/Itemid,38/func,view/id,155/catid,9/limit,6/limitstart,0/](http://www.domainname.com/component/option,com_fireboard/Itemid,38/func,view/id,155/catid,9/limit,6/limitstart,0/) (in links on subsequent pages > 1)

Both urls are rewritten by sh404SEF to

<http://www.domainname.com/descriptionforumcategory/155-descriptionsubject.html>

For FB without sh404SEF the problem still exists and my point is that urls that end with limit-start equal to 0 should not exist (or the first url from the example should have a limit-start equal to 0 as well).

By the way: I know that FB is ahead of BOJ, I tried to make an example .. which was not to clear I admit.

Rickert

---

## Re:Avoid duplicate content -nofollow on useless li

Posted by fxstein - 2008/07/13 12:22

---

---

I am not sure I understand.

The latest nightly build do not create limitstart =0

Here is the code from the class that generates the links:

```
function GetThreadPageLink($func, $catid, $threadid, $page, $limit, $name, $anker="", $rel='follow', $class="")
{
    if ($page == 1 || !is_numeric($page) || !is_numeric($limit))
    {
        // page 1 is identical to a link to the top of the thread
        $pagelink = fb_link::GetSefHrefLink(JB_LIVEURLREL.'&func='.$func.'&catid='.$catid.'&id='.$threadid
            .($anker?('#'.$anker):"), $name, $rel, $class);
    }
    else
    {
        $pagelink = fb_link::GetSefHrefLink(JB_LIVEURLREL.'&func='.$func.'&catid='.$catid.'&id='.$threadid
            .&limit='.$limit.'&limitstart='.(($page-1)*$limit).($anker?('#'.$anker):"), $name, $rel, $class);
    }
    return $pagelink;
}
```

As you can see for page=1 it generates a link without any limit and limitstart parameter. Only for pages other than 1 that info gets created.

---

## Re:Avoid duplicate content -nofollow on useless li

Posted by rickert - 2008/07/13 16:26

The code you showed me is the same I use. Nevertheless, I made a fresh install of the latest nightly build to test the urls but too many bugs/problems. I think I give up for now and wait for 1.05 production to show up.

Thanks for your help.

Rickert

---

## Re:Avoid duplicate content -nofollow on useless li

Posted by rickert - 2008/07/14 14:47

I figured out that (in the latest version of view.php) GetThreadPageLink is NOT called to create the page navigation links on the top and bottom of the page. In stead Joomla's writepageslinks is called. This causes the limitstart,0 to be included (f.e. on page 2) in the url's and causes url's with duplicate content.

Rickert

By the way: I saw that the links from flat view to threaded view vv are still nofollowed. Are you sure this will stop Google to index the threaded pages if the default is flat view?

---

## Re:Avoid duplicate content -nofollow on useless li

Posted by fxstein - 2008/07/14 19:35

---

rickert wrote:

I figured out that (in the latest version of view.php) GetThreadPageLink is NOT called to create the page navigation links on the top and bottom of the page. In stead Joomla's writepageslinks is called. This causes the limitstart,0 to be included (f.e. on page 2) in the url's and causes url's with duplicate content.

Rickert

By the way: I saw that the links from flat view to threaded view vv are still nofollowed. Are you sure this will stop Google to index the threaded pages if the default is flat view?

Oh - I see. I must have missed them. I was searching for hrefs in the code, skipped these two. Need to look into this as soon as I have the upgrade process taken care of.

=====