

## Bugs & Wish list

Disallow tiki-tell\_a\_friend.php in robots.txt | Tiki Wiki CMS Groupware :: Development

Disallow tiki-tell\_a\_friend.php in robots.txt

### Status

Closed

### Subject

Disallow tiki-tell\_a\_friend.php in robots.txt

### Version

3.x

4.x

5.x

### Category

- Usability
- Less than 30-minutes fix

### Feature

Calendar

All / Undefined

Search engine optimization (SEO)

Tell a Friend / Social Bookmarking

### Resolution status

Fixed or Solved

### Submitted by

Jan Krohn

### Lastmod by

Daniel Cedilotte

### Rating

(0)

### Description

Calendar URLs are currently excluded for spiders and bots, for example:

/calendar?focus=1268089200&viewmode=week

excluded by entries:

Disallow: tiki-calendar.php

Disallow: /tiki-calendar.php

However, the URL /tiki-

tell\_a\_friend.php?url=%2Fcalendar%3Ffocus%3D1268089200%26viewmode%3Dweek would be spidered instead.

### Solution

The following line added to robots.txt:

Disallow: /tiki-tell\_a\_friend.php

Seems fixed in Trunk.

Importance

6

Easy to solve?

8

Priority

48

Demonstrate Bug (Tiki 19+)

Please demonstrate your bug on [show2.tiki.org](http://show2.tiki.org)

Version: trunk ▼

Demonstrate Bug (older Tiki versions)

Please demonstrate your bug on [show.tikiwiki.org](http://show.tikiwiki.org)

Version: 18.x ▼

Ticket ID

3077

Created

Thursday 11 March, 2010 10:59:11 GMT-0000  
by Jan Krohn

LastModif

Saturday 05 February, 2011 02:41:55 GMT-0000

Comments



Jan Krohn 11 Mar 10 13:12 GMT-0000

The URLs like </calendar?todate=1268089200> also need to be disallowed...

New line in robots.txt:

Disallow: calendar

Disallow: /calendar



luciash d' being ☐ 13 Mar 10 04:14 GMT-0000

but for your comment about the SEFUrl version of the calendar link:

What if you have a wiki page named "calendarium" which you want to have indexed by search bots ?

This rule would exclude it too (or any other starting with calendar string).



Jan Krohn 16 Mar 10 07:48 GMT-0000

Very true... It's a shame there's no "allow" command for robots.txt

Alternatively, a noindex meta tag could be possible. But this would work only partially, as the pages would still be visited by the spiders.

Addin a nofollow tag to the hyperlinks pointing to calendar pages might be helpful as well:  
<http://en.wikipedia.org/wiki/Nofollow> - this would also be only a partial solution, as some bots don't respect this tag...

Maybe some other idea comes to my mind later...



Jan Krohn 16 Mar 10 10:34 GMT-0000

Ok, I think I have an idea. What about these two lines?

Disallow: calendar?todate

Disallow: /calendar?todate

Jan



luciash d' being 16 Mar 10 17:29 GMT-0000

but i will add this instead:

/calendar?todate

/calendar&todate

1) short URLs support both ? and & as first param delimiter

2) i haven't found any evidence search engines recognize the version without slash in the beginning ( robots.txt checkers complain about it, see <http://tool.motoricerca.info/robots-checker.phtml> )



Jan Krohn 18 Mar 10 13:45 GMT-0000

It's clearing up now, but browsing my log files, I found two more dangerous set of calendar related URLs being spidered. They could be excluded by:

/calendar?viewmode

/calendar&viewmode

/calendar?focus

/calendar&focus

(I have no idea though why MSN bot puts the viewmode parameter before the todate parameter...)

## Attachments

filename	created	hits	comment	version	filetype
----------	---------	------	---------	---------	----------

No attachments for this item

The original document is available at

[https://dev.tiki.org/item3077-Disallow-tiki-tell\\_a\\_friend-php-in-robots-txt](https://dev.tiki.org/item3077-Disallow-tiki-tell_a_friend-php-in-robots-txt)