

Redmine - Feature #33658

robots.txt: disallow crawling login, register, and lost password form

2020-06-23 18:05 - Go MAEDA

Status: Closed	Start date:
Priority: Normal	Due date:
Assignee: Go MAEDA	% Done: 0%
Category: SEO	Estimated time: 0.00 hour
Target version: 4.2.0	
Resolution: Fixed	

Description

The attached patch adds robots exclusion tags (<meta name="robots" content="noindex,follow,noarchive" />) to login, register, and lost password form.

Currently, Redmine allows web crawlers such as Googlebot and Baiduspider to index those forms. I think it is undesirable for the following reasons:

- There is no meaningful information there
- If the Redmine instance requires login, probably the admin doesn't want to publish any information to the internet. It is so bad to be indexed by Google

That can be resolved by the attached patch by adding robots exclusion tags to the pages.

Associated revisions

Revision 20532 - 2020-12-01 03:03 - Go MAEDA

Robots meta tag to prevent web crawlers from indexing login, register, and lost password form (#33658).

Patch by Go MAEDA.

Revision 20577 - 2020-12-05 08:10 - Go MAEDA

Use robots.txt instead of robots meta tag to prevent web crawlers from indexing login, register, and lost password form (#33658).

Patch by Go MAEDA.

History

#1 - 2020-11-28 09:07 - Go MAEDA

- File 33658.patch added
- Target version set to Candidate for next major release

Added tests to the patch.

#2 - 2020-11-30 05:39 - Go MAEDA

- Target version changed from Candidate for next major release to 4.2.0

Setting the target version to 4.2.0.

#3 - 2020-12-01 02:05 - Go MAEDA

- Subject changed from *Prevent web crawlers from indexing login, register, and lost password form* to *Robots meta tag to prevent web crawlers from indexing login, register, and lost password form*

#4 - 2020-12-01 03:03 - Go MAEDA

- Status changed from *New* to *Closed*
- Assignee set to *Go MAEDA*
- Resolution set to *Fixed*

Committed the patch.

#5 - 2020-12-01 15:41 - Holger Just

- Status changed from *Closed* to *Reopened*

Redmine already generates a robots.txt in `welcome#robots`. Wouldn't this be better suited there, specifically in `source:trunk/app/views/welcome/robots.text.erb` so that we have a single location where we define behavior for robots?

#6 - 2020-12-02 09:34 - Go MAEDA

- File `33658-use-robots_txt-instead-of-robots-meta-tag.patch` added

Holger Just wrote:

Redmine already generates a robots.txt in welcome#robots. Wouldn't this be better suited there, specifically in source:trunk/app/views/welcome/robots.text.erb so that we have a single location where we define behavior for robots?

You are absolutely right. I forgot about the existence of robots.txt for some reason.

The attached new patch uses robots.txt instead of robots meta tag.

#7 - 2020-12-02 15:00 - Holger Just

Looks good to me, thanks!

#8 - 2020-12-05 08:11 - Go MAEDA

- Subject changed from *Robots meta tag to prevent web crawlers from indexing login, register, and lost password form* to *robots.txt: disallow crawling login, register, and lost password form*
- Status changed from *Reopened* to *Closed*

Committed patch #33658#note-6.

Files

robots-exclusion-tags.patch	1.43 KB	2020-06-23	Go MAEDA
33658.patch	2.92 KB	2020-11-28	Go MAEDA
33658-use-robots_txt-instead-of-robots-meta-tag.patch	4.17 KB	2020-12-02	Go MAEDA