

Redmine - Defect #35036

Markdown text sections broken by thematic breaks (horizontal rules)

2021-04-05 15:57 - Martin Cizek

Status:	Closed	Start date:	
Priority:	Normal	Due date:	
Assignee:	Go MAEDA	% Done:	0%
Category:	Text formatting	Estimated time:	0.00 hour
Target version:	4.2.2	Affected version:	4.2.0
Resolution:	Fixed		

Description

A [thematic break](#) composed of hyphens (e.g. "---") breaks the division of Markdown text into individually editable sections.

Steps to reproduce

Configure markdown text formatting, create a Wiki page in a web browser and enter the following content:

```
# Title
## Heading 2
Preceding CRLF is the default for web-submitted data.

---

End of thematic breaks.

## Heading 2
Nulla nunc nisi, egestas in ornare vel, posuere ac libero.

More in the unit tests in the enclosed patch.
```

Cause

The reason is that it is confused with a [setext heading](#). Although the current regexp in `extract_sections` actually tries to restrict setext headings in a way that it must follow a non-empty line, it does not account for a whitespace-only line or even plain CRLF. And as long as the text originates from a web browser, there is always a CRLF. So the problem is pretty common even for carefully formatted text.

Fix

Attaching a patch with a fix and corresponding unit tests.

Broader context

The current approach to section extraction is inherently fragile - as shown in the other (skipped) unit test enclosed in the patch. I'd suggest to keep the skipped test there to mark it as a known issue. Will create a dedicated issue for this.

Related issues:

Related to Redmine - Feature #35037: Make wiki text section extraction less f...

New

Associated revisions

Revision 20998 - 2021-05-14 11:39 - Go MAEDA

Markdown text sections broken by thematic breaks (horizontal rules) (#35036).

Patch by Martin Cizek.

Revision 21007 - 2021-05-25 04:13 - Go MAEDA

Merged r20998 from trunk to 4.2-stable (#35036).

History

#1 - 2021-04-05 16:05 - Martin Cizek

- File *0001-markdown_formatter-extract_sections-fix.patch* added

#2 - 2021-04-05 16:14 - Martin Cizek

Just to make it clear - the skipped test address an already existing error, which is probably not worth fixing with the current approach in `extract_sections` implementations. See [#35037](#) for more details.

But the particular and common error in thematic break mistreatment is solved by the patch "as is", it can be just applied. :)

#3 - 2021-05-12 06:46 - Go MAEDA

- Status changed from *New* to *Confirmed*

- Target version set to *4.1.4*

Setting the target version to *4.1.4*.

#4 - 2021-05-12 06:48 - Go MAEDA

- Related to Feature [#35037](#): *Make wiki text section extraction less fragile* added

#5 - 2021-05-14 11:39 - Go MAEDA

- Status changed from *Confirmed* to *Resolved*

- Assignee set to *Go MAEDA*

- Resolution set to *Fixed*

Committed the fix. Thank you for your contribution.

#6 - 2021-05-25 04:15 - Go MAEDA

- Status changed from *Resolved* to *Closed*

- Target version changed from *4.1.4* to *4.2.2*

Files

0001-markdown_formatter-extract_sections-fix.patch

2.76 KB

2021-04-05

Martin Cizek