Redmine - Feature #35037

Make wiki text section extraction less fragile

2021-04-05 16:00 - Martin Cizek

Status:	New		**Start date:**	
Priority:	Normal		**Due date:**	
Assignee:			**% Done:**	0%
Category:	Text formatting		**Estimated time:**	0.00 hour
Target version:				
Resolution:				

Description

The current approach to per-section text editing is inherently fragile:

- Sections links are generated based on rendered HTML headings (before macro injections).
- Section extraction from the markup source is based on simplified regexp-based parsing of the markup, independent on the fully featured markup parser.

The issue is going to be more significant after relaxing restrictions on markup syntax (#32424, #35035).
But as #35036 shows, even the current restrictions do not guarantee correct section extraction. See also the exmple below.

I can imagine two approaches to the solution:

1. For CommonMark - using the sourcepos feature of the renderer. Also applicable to other eventual formatters with similar feature.
2. For any markup - introducing validation of the section extraction results, which would detect that the extraction fails and would disable per-section edit links.

I can offer creating a PoC of the sourcepos approach for CommonMark format after #32424 is incorporated.

A difficult-to-solve example of Markdown with broken section extraction follows (copied from the skipped unit test in #35036):

```
# Title

## Heading 2

- item
one
-
not a heading

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

Related issues:

Related to Redmine - Defect #35036: Markdown text sections broken by thematic...	**Closed**

History

#1 - 2021-05-12 06:48 - Go MAEDA

- Related to Defect #35036: Markdown text sections broken by thematic breaks (horizontal rules) added