

Automatic language detection in syntax highlighting

2022/05/09 18:41 - Admin Redmine

ステータス:	New	開始日:	2022/05/09
優先度:	通常	期日:	
担当者:		進捗率:	0%
カテゴリ:	Text formatting_26	予定工数:	0.00時間
対象バージョン:		作業時間:	0.00時間
Redmineorg_URL:	https://www.redmine.org/issues/30429	status_id:	1
category_id:	26	tracker_id:	2
version_id:	0	plus1:	1
issue_org_id:	30429	affected_version:	
author_id:	332	closed_on:	
assigned_to_id:	0	affected_version_id:	
comments:	3		

説明

As of Redmine 4.0, you can use >100 languages in syntax highlighter (#24681). However, as you can use so many languages, it is difficult to memorize which languages you can highlight and what language name you should describe highlighting the snippet.

To resolve the problem, I propose a language name "auto". When you use the language name "auto" like the following, Redmine automatically detects the language and highlight the code snippet.

!{width: 205px;}auto-keyword-textile.png!

Probably we can use "Rouge::Lexer.guess_by_source":https://www.rubydoc.info/gems/rouge/Rouge%2FLexer%2Eguess_by_source to implement this feature.

journals

Go MAEDA wrote:

Probably we can use "Rouge::Lexer.guess_by_source":
https://www.rubydoc.info/gems/rouge/Rouge%2FLexer%2Eguess_by_source to implement this feature.

Unfortunately, I found that guess_by_source is not so useful because it checks only a shebang line for most languages. It means that the following snippet that obviously looks like C language is not detected as C but plain text.

```
#include
```

```
int main()
```

```
{
    printf("Hello, world! /n");
    return 0;
}
```

```
2.6.0 :012 > text
```

```
=> "#include /n /nint main() /n{/n    printf("/Hello, world! /n /");/n    return 0; /n} /n"
```

```
2.6.0 :013 > Rouge::Lexer.guess_by_source(text)
```

```
=> Rouge::Lexers::PlainText
```

+1 I am thinking it would probably be better to have the syntax highlighter guess the language only in case a user forgot to define what language it is, else the manual definition to always override the automatic detection. :)

It's for sure a nice improvement for Redmine.

+1

Max Johansson wrote:

I am thinking it would probably be better to have the syntax highlighter guess the language only in case a user forgot to define what language it is, else the manual definition to always override the automatic detection. :)

I would also prefer to have auto detection enabled by default as it would be more convenient and less error prone.

履歴

#1 - 2022/05/10 17:03 - Admin Redmine

- カテゴリ を Text formatting_26 にセット