

ChiliProject - Bug # 972: Non-latin anchors doesn't work in wiki

Status:	Open	Priority:	Normal
Author:	Vladislav Poluhin	Category:	
Created:	2012-04-10	Assignee:	Felix Schäfer
Updated:	2012-05-23	Due date:	
Remote issue URL:			
Affected version:	3.1.0		
Description:	I create wiki-page with content: <pre><pre> h2. ðŸŒ€ð, ð½Œ, ðµŒŒ h2. Printers </pre></pre> This is rendered as: <pre><pre><code class="html"> <h2>ðŸŒ€ð, ð½Œ, ðµŒŒÂ¶</h2> <h2>PrintersÂ¶</h2> </code></pre></pre> Anchor for russian header is empty.		

Associated revisions

2008-04-01 09:40 pm - Jean-Philippe Lang

Add predefined date ranges to the time report in the same way as the details view (closes #972). It nows defaults to 'All time'.
This patch also fixes time report periods (columns) computation.

git-svn-id: http://redmine.rubyforge.org/svn/trunk@1318_e93f8b46-1217-0410-a6f0-8f06a7374b81

History

2012-05-15 08:21 am - Ilya Turkin

The problem is the following: ApplicationHelper#parse_headings parse the strings with \w metacharacter: @anchor = item.gsub(%r{[^\w\s\.-]}, "").gsub(%r{\s+(\-+\s*)?}, '-')@.
AFAIK after ruby 1.9.1p378 \w is ASCII, so to correct the work with non-latin characters you need to substitute \w with \p{Word}.

2012-05-17 08:10 pm - Felix Schäfer

Can anyone confirm this is a problem on 1.9 but not on 1.8?

2012-05-22 09:39 am - Ilya Turkin

- Assignee set to Felix Schäfer

```
<pre>
<code class="Python">ilya@ubuntu:~$ cd chiliproject/
ilya@ubuntu:~/chiliproject$ rvm use 1.8.7
Using /home/ilya/.rvm/gems/ruby-1.8.7-p358
ilya@ubuntu:~/chiliproject$ bundle exec script/console
Loading development environment (Rails 2.3.14)
```

NOTE: Gem.source_index is deprecated, use Specification. It will be removed on or after 2011-11-01.

Gem.source_index called from /home/ilya/.rvm/gems/ruby-1.8.7-p358/gems/rails-2.3.14/lib/rails/gem_dependency.rb:21.

```
>> a = "Ñ±Đ,Đ»Đ,Đ¿Ñ€Đ¼Đ´Đ¶ĐµĐ°Ñ,"
=> "Ñ±Đ,Đ»Đ,Đ¿Ñ€Đ¼Đ´Đ¶ĐµĐ°Ñ,"
>> a.gsub(/^w/, "")
=> ""
>> exit
```

ilya@ubuntu:~/chiliproject\$ rvm use 1.9.2

Using /home/ilya/.rvm/gems/ruby-1.9.2-p320

ilya@ubuntu:~/chiliproject\$ bundle exec script/console

Loading development environment (Rails 2.3.14)

NOTE: Gem.source_index is deprecated, use Specification. It will be removed on or after 2011-11-01.

Gem.source_index called from /home/ilya/.rvm/gems/ruby-1.9.2-p320/gems/rails-2.3.14/lib/rails/gem_dependency.rb:21.

```
>> a = "Ñ±Đ,Đ»Đ,Đ¿Ñ€Đ¼Đ´Đ¶ĐµĐ°Ñ,"
=> "Ñ±Đ,Đ»Đ,Đ¿Ñ€Đ¼Đ´Đ¶ĐµĐ°Ñ,"
>> a.gsub(/^w/, "")
=> "Ñ±Đ,Đ»Đ,Đ¿Ñ€Đ¼Đ´Đ¶ĐµĐ°Ñ,"
>> a.gsub(/\p{Word}/, "")
=> ""</code>
</pre>
```

2012-05-22 10:44 am - Holger Just

The following patch solves the issue:

```
<pre><code class="diff">
diff --git a/config/environment.rb b/config/environment.rb
index 84d6d1f..600597c 100644
--- a/config/environment.rb
+++ b/config/environment.rb
@@ -24,9 +24,12 @@ ENV['RAILS_ENV'] ||= ENV['RACK_ENV'] if ENV['RACK_ENV']
# Specifies gem version of Rails to use when vendor/rails is not present
RAILS_GEM_VERSION = '2.3.14' unless defined? RAILS_GEM_VERSION

+# this is replaced by config.encoding = "utf-8" in rails3
if RUBY_VERSION >= '1.9'
  Encoding.default_external = 'UTF-8'
  Encoding.default_internal = 'UTF-8'
+else
+ $KCODE='UTF-8'
end

# Bootstrap the Rails environment, frameworks, and default configuration
</code></pre>
```

Two minor potential issues:

- * unicode stuff might be invalid in anchors
- * It might break other regexes relying on @\w@ being only @[a-zA-Z0-9_]@

2012-05-23 06:01 am - Ilya Turkin

Holger Just wrote:

> The following patch solves the issue:

Sorry, but this is not the solution because the problem not with 1.8 but 1.9.