

Redmine - Defect #10590

WikiContent::Version#text return string with #<Encoding:ASCII-8BIT> when uncompressed

2012-04-03 10:16 - Alexander Oryol

Status:	Closed	Start date:	
Priority:	Normal	Due date:	
Assignee:		% Done:	0%
Category:	Wiki	Estimated time:	0.00 hour
Target version:	1.4.0	Affected version:	
Resolution:	Fixed		
Description			
<ul style="list-style-type: none">• ruby 1.9.3p125• postgresql 8.4.11• rails 2.3.14• redmine trunk			
Wiki history compression: none			
<pre>\$ script/console Loading development environment (Rails 2.3.14) /home/asor/.rvm/gems/ruby-1.9.3-p125-fast/gems/activesupport-2.3.14/lib/active_support/inflector.r b:3:in `<top (required)>': iconv will be deprecated in the future, use String#encode instead. NOTE: Gem.source_index is deprecated, use Specification. It will be removed on or after 2011-11-01 . Gem.source_index called from /home/asor/.rvm/gems/ruby-1.9.3-p125-fast/gems/rails-2.3.14/lib/rails /gem_dependency.rb:21. >> WikiPage.find(:first).content.text.encoding => #<Encoding:UTF-8> >> WikiPage.find(:first).content.versions.first.text.encoding => #<Encoding:ASCII-8BIT></pre>			
We should always force_encoding to UTF-8, if data stored as binary type. The patch attached.			
When applied:			
<pre>>> WikiPage.find(:first).content.text.encoding => #<Encoding:UTF-8> >> WikiPage.find(:first).content.versions.first.text.encoding => #<Encoding:UTF-8></pre>			

Associated revisions

Revision 9329 - 2012-04-06 09:21 - Jean-Philippe Lang

WikiContent::Version#text may return ascii-8bit encoded string (#10590).

History

#1 - 2012-04-03 10:27 - Toshi MARUYAMA

- Assignee deleted (Toshi MARUYAMA)

What database do you use?

mysql1 has problems on Ruby1.9.

You need to use mysql2 instead of mysql1.

[source:trunk/Gemfile@9244#L44](#)

#2 - 2012-04-03 10:31 - Toshi MARUYAMA

Sorry, in description "postgresql 8.4.11".

#3 - 2012-04-03 10:38 - Alexander Oryol

Patch is (in attachment):

```
commit b0ec4111730809bc2f4bafd071ba2271e290e9e1
Author: Alex Eagle <eagle.alex@gmail.com>
Date: Tue Apr 3 11:58:05 2012 +0400
```

```
Force encoding string from data binary always (not only compressed)
```

```
diff --git a/app/models/wiki_content.rb b/app/models/wiki_content.rb
index 430e9a5..d5a3870 100644
--- a/app/models/wiki_content.rb
+++ b/app/models/wiki_content.rb
@@ -91,14 +91,16 @@ class WikiContent < ActiveRecord::Base
   end

   def text
-     @text ||= case compression
-       when 'gzip'
-         str = Zlib::Inflate.inflate(data)
+     @text ||= begin
+       str = case compression
+         when 'gzip'
+           Zlib::Inflate.inflate(data)
+         else
+           # uncompressed data
+           data
+         end
+       str.force_encoding("UTF-8") if str.respond_to?(:force_encoding)
+       str
-     else
-       # uncompressed data
-       data
-     end
-   end
end
```

#4 - 2012-04-03 10:58 - Toshi MARUYAMA

It seems acts_as_versioned or PostgreSQL problem.

#5 - 2012-04-03 11:08 - Alexander Oryol

Not PostgreSQL exactly.

DB return byte-array. Rails interpretate column (t.column :data, :binary) in string context as ASCII-8BIT unless otherwise.

Code for transformation field 'data' into text defined in WikiContent::Version#text.

IMHO, this is model Version problem.

#6 - 2012-04-03 11:29 - Etienne Massip

Might be tied to [#6941](#); I just saw I had this old issue back and that upgrading pg gem to latest version in Gemfile (+bundle update) solved it.

#7 - 2012-04-03 11:49 - Alexander Oryol

Updated pg gem to pg (0.13.2) do not solve it.

```
$ cat Gemfile|grep pg
  gem "pg", "~> 0.9.0"
$ cat Gemfile.lock|grep pg
  pg (0.13.2)
  pg
```

Result:

```
>> WikiPage.first.content.versions.last.data.encoding
=> #<Encoding:ASCII-8BIT>
```

#8 - 2012-04-03 15:21 - Etienne Massip

- Tracker changed from Patch to Defect

- Target version set to 1.4.0

- Affected version (unused) set to devel

Potential release blocker.

Aside note: as explained above, required pg gem version should be raised to prevent [#6941](#).

#9 - 2012-04-03 19:50 - Etienne Massip

- Status changed from New to Confirmed

#10 - 2012-04-06 09:22 - Jean-Philippe Lang

- Status changed from Confirmed to Closed

- Resolution set to Fixed

Patch applied in [r9329](#).

Files

force_encoding_always.patch	1.06 KB	2012-04-03	Alexander Oryol
-----------------------------	---------	------------	-----------------