

## Redmine - Defect #55

### Non-unicode repository content not displayed correctly

2007-05-03 09:07 - Nikolay Solakov

<b>Status:</b>	Closed	<b>Start date:</b>	
<b>Priority:</b>	Normal	<b>Due date:</b>	
<b>Assignee:</b>		<b>% Done:</b>	0%
<b>Category:</b>		<b>Estimated time:</b>	0.00 hour
<b>Target version:</b>		<b>Affected version:</b>	
<b>Resolution:</b>			
<b>Description</b>			
Hi, Look at the attached screenshot.  When I set the page encoding to Windows-1251 it shows the characters correctly.  Thanks, Nikolay			

#### History

##### #1 - 2007-05-05 16:47 - Jean-Philippe Lang

As files in the repository may be encoded to anything, I don't see any solution that would work in any cases.  
Do you have a proposition for this ?

##### #2 - 2007-05-07 05:41 - Nikolay Solakov

Unfortunately I don't :(  
Maybe there should have some sort of auto-detecting feature to say what is the encoding of the files, but I don't know how it is implemented for example in Firefox's View->Encoding.

Thanks,  
Nikolay

##### #3 - 2007-05-21 05:14 - Kyosuke Takayama

Hi.

It should use iconv.

I wrote patch for svn viewer.  
It is optimized for Japanese now. However, there is not influence towards the English zone.  
And It can cope with other character codes zone if you change CharSets.

Index: app/models/svn\_repos.rb =====

```
--- app/models/svn_repos.rb (revision 537)
+++ app/models/svn_repos.rb (working copy)
@ -17,7 +17,23 @
```

```
require 'rexml/document'
require 'cgi'
+require 'iconv'
```

```
class Iconv
CharSets = ['iso-2022-jp', 'euc-jp', 'utf-8', 'shift_jis',
'windows-31j']
```

```
def self.to_utf8(str)
+ return str if /^[r\n\t\x20-\x7e]*\Z/n.match(str) # for
us-ascii
```

```
+ CharSets.each do |char|
+   begin
+     return Iconv.conv("UTF-8", char, str)
+   rescue Iconv::IllegalSequence
+   end
+ end
+ str
+ end
end
```

module SvnRepos

```
class CommandFailed < StandardError #:nodoc:
@ -178,7 +194,7 @
  cat = io.read
  end
  return nil if $? && $? .exitstatus != 0
-   cat
+   Iconv.to_utf8(cat)
  rescue Errno::ENOENT => e
    raise CommandFailed
  end
```

- 
- for Cyrillic support.  
CharSets = ['windows-1251', 'utf-8']

Thank you.

#### #4 - 2007-05-24 02:49 - Nikolay Solakov

Hi,  
It's not working for me. I put the  
CharSets = ['windows-1251', 'utf-8']  
instead of  
CharSets = ['iso-2022-jp', 'euc-jp', 'utf-8', 'shift\_jis',  
'windows-31j']  
The result is the same. :(

Thanks,  
Nikolay

#### #5 - 2007-05-24 08:23 - Kyosuke Takayama

It does work only in FILE view.  
But does not work in DIFF view.

Probably, in the case of DIFF, it is so..

```
Index: app/models/svn_repos.rb =====
--- app/models/svn_repos.rb (revision 538)
+++ app/models/svn_repos.rb (working copy)
@ -310,7 +326,7 @
self << diff_table if diff_table.length

  1

diff_table = DiffTable.new type
end
-   a = diff_table.add_line line
+   a = diff_table.add_line(Iconv.to_utf8(line))
end
self << diff_table
end
```

#### #6 - 2007-05-24 09:20 - Nikolay Solakov

Thanks,  
I can now see the cyrillic chars.

Hope Jean-Philippe integrate it soon.

Nikolay.

**#7 - 2007-05-24 12:48 - Jean-Philippe Lang**

Thanks for your help.

I propose to add the list of possible charsets in lang files, so that they can be set for each language. The charsets defined for the default language of your redmine installation would be used for iconv conversion.

**#8 - 2007-05-24 13:49 - Jean-Philippe Lang**

One more thing. It should be ok for diffs, but trying to convert files to utf-8 (in cat method) in any case may cause some problems (eg. binary files). It should be done only for text files. I'll have to add a check.

**#9 - 2007-09-28 02:52 - Nikolay Solakov**

Hi,  
Is this bug-feature implemented?  
When I updated my redMine recently the incorrect characters appeared again :(

Thank,  
Nikolay

**#10 - 2007-09-28 03:00 - Nikolay Solakov**

Forgot to say that I was applied the Kyosuke Takayama's patch on my Redmine. But now it's gone due to the repository support rewrite from JP.

**#11 - 2007-09-28 12:10 - Jean-Philippe Lang**

It's not implemented in current trunk (otherwise this ticket would be closed). I'll have a look at it.

**#12 - 2007-10-07 11:23 - Jean-Philippe Lang**

Hi,  
This fix was finally committed in [r814](#).  
You can set the list of encodings in Admin -> Settings (Repositories encodings).

Thanks.

**#13 - 2007-10-07 17:01 - Nikolay Solakov**

Thanks, again and again.  
Nikolay

**Files**

---

cyrillic_repo_files.JPG	134 KB	2008-02-03	Nikolay Solakov
-------------------------	--------	------------	-----------------