

Redmine - Feature #7037

CSV export encoding and excel. UTF-8 and BOM

2010-12-03 19:18 - John Yani

<b>Status:</b>	Closed	<b>Start date:</b>	2010-12-03
<b>Priority:</b>	Normal	<b>Due date:</b>	
<b>Assignee:</b>	Jean-Philippe Lang	<b>% Done:</b>	0%
<b>Category:</b>	Issues	<b>Estimated time:</b>	0.00 hour
<b>Target version:</b>	3.1.0		
<b>Resolution:</b>	Fixed		
<b>Description</b>			
Redmine doesn't add BOM to the UTF-8 exported csv file, so I have to manually enforce csv encoding to windows-1251. In other way Excel will not display it correctly.			
Maybe there is some setting to enforce BOM while doing csv export?			
References:			
<ul style="list-style-type: none"><li>• <a href="#">Stackoverflow Q&amp;A</a></li><li>• <a href="#">Fix for this</a></li></ul>			
<b>Related issues:</b>			
Related to Redmine - Defect #7461: Special chars / umlauts broken in CSV / PD...		<b>Closed</b>	
Related to Redmine - Defect #7039: CSV export. Line breaks. Excel		<b>Closed</b>	<b>2010-12-03</b>
Related to Redmine - Defect #19523: problem in exporting Arabic text to csv		<b>Closed</b>	
Related to Redmine - Defect #11116: CSV export encoding problem when issues a...		<b>Closed</b>	
Related to Redmine - Patch #17902: CSV encoding should be UTF-8 in French locale		<b>Closed</b>	
Has duplicate Redmine - Feature #11301: Add "byte order mark" to the export C...		<b>Closed</b>	

Associated revisions

Revision 14302 - 2015-06-13 09:33 - Jean-Philippe Lang

Adds a class for handling CSV generation (#7037).

Revision 14303 - 2015-06-13 09:55 - Jean-Philippe Lang

Add BOM to UTF-8 encoded CSV (#7037).

History

#1 - 2011-01-25 11:40 - Terence Mill

Is that fixed?

#2 - 2011-01-25 12:39 - John Yani

Well, the fix exists. But AFAIK it is not merged yet.  
The correct way would be to add some setting to enable/disable BOM instead of just hard-coding it.

#3 - 2011-06-07 12:18 - Alexander Kriventsov

John Yani wrote:

Redmine doesn't add BOM to the UTF-8 exported csv file, so I have to manually enforce csv encoding to windows-1251. In other way Excel will not display it correctly.

How did you do this?

#4 - 2011-06-08 21:11 - John Yani

Alexander Kriventsov wrote:

John Yani wrote:

Redmine doesn't add BOM to the UTF-8 exported csv file, so I have to manually enforce csv encoding to windows-1251. In other way Excel will not display it correctly.

How did you do this?

I've edited the "general\_csv\_encoding" value in corresponding "/config/locales/\*.yml" file.

**#5 - 2011-06-10 10:21 - Etienne Massip**

- *Category set to Issues*

**#6 - 2013-03-27 10:20 - Nick Tan**

hi,

i have the exactly same issue, with the plugin [https://github.com/sprosin/redmine\\_csv\\_with\\_bom](https://github.com/sprosin/redmine_csv_with_bom) now the csv export works as expected.

but i think this patch should be merge into the redmine core, not in form of a plugin with monkey patching style

**#7 - 2015-01-29 10:36 - Yuri Rumega**

The problem is still actual in 2.6. Had to adapt "redmine\_csv\_with\_bom" plugin.

**#8 - 2015-04-02 09:50 - Toshi MARUYAMA**

- *Tracker changed from Defect to Feature*

**#9 - 2015-04-02 09:51 - Toshi MARUYAMA**

- *Related to Defect #19523: problem in exporting Arabic text to csv added*

**#10 - 2015-04-02 10:37 - Toshi MARUYAMA**

- *File 7037-csv-bom.diff added*

- *Target version set to 3.1.0*

This is patch for trunk [r14155](#).

**#11 - 2015-04-02 10:42 - Toshi MARUYAMA**

- *File 7037-csv-bom-1.diff added*

Removed extra white space from previous patch.

**#12 - 2015-04-02 11:02 - Toshi MARUYAMA**

- *File 7037-csv-bom-2.diff added*

Previous patch has mistake.

This patch fixed.

```
-- encoding = Setting.csv_utf8_bom ? "UTF-8" : l(:general_csv_encoding)
++ encoding = Setting.csv_utf8_bom? ? "UTF-8" : l(:general_csv_encoding)
```

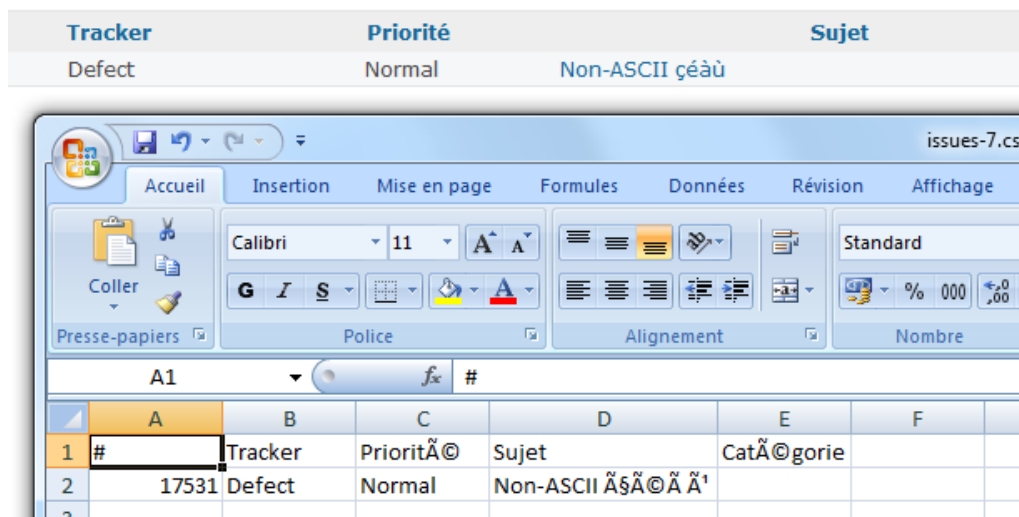
**#13 - 2015-04-18 05:35 - Toshi MARUYAMA**

- *Related to Defect #11116: CSV export encoding problem when issues and interface languages are not equal added*

**#14 - 2015-05-10 11:30 - Jean-Philippe Lang**

- *File excel2007.png added*

Doesn't work for me when opening the CSV with Excel 2007 (which is supposed to support UTF-8 with BOM), tested with 7037-csv-bom-2.diff applied :



#### #15 - 2015-05-10 11:55 - Toshi MARUYAMA

Could you check "csv\_utf8\_bom" is on in setting?

#### #16 - 2015-05-10 12:06 - Jean-Philippe Lang

Yes, it was checked. The csv file is recognized as UTF-8 with BOM by my text editor but not by Excel 2007.

#### #17 - 2015-05-10 12:48 - Toshi MARUYAMA

- File libreoffice.png added

```
$ hexdump -C issues.csv
00000000 ef bb bf 23 3b 54 72 61 63 6b 65 72 3b 53 74 61 |...#;Tracker;Sta|
00000010 74 75 74 3b 50 72 69 6f 72 69 74 c3 a9 3b 53 75 |tut;Priorit..;Su|
00000020 6a 65 74 3b 41 73 73 69 67 6e c3 a9 20 c3 a0 3b |jet;Assign...;|
00000030 4d 69 73 2d c3 a0 2d 6a 6f 75 72 0a 38 37 3b 42 |Mis-.-jour.87;B|
00000040 75 67 3b 4e 65 77 3b 4e 6f 72 6d 61 6c 3b 4e 6f |ug;New;Normal;No|
00000050 6e 2d 41 53 43 49 49 20 c3 a7 c3 a9 c3 a0 c3 b9 |n-ASCII .....|
00000060 3b 22 22 3b 31 30 2f 30 35 2f 32 30 31 35 20 31 |;"",10/05/2015 1|
00000070 39 3a 32 35 0a |9:25.|
00000075
```

```
$ hexdump -c issues.csv
00000000 357 273 277 # ; T r a c k e r ; S t a
00000010 t u t ; P r i o r i t 303 251 ; S u
00000020 j e t ; A s s i g n 303 251 303 240 ;
00000030 M i s - 303 240 - j o u r \n 8 7 ; B
00000040 u g ; N e w ; N o r m a l ; N o
00000050 n - A S C I I 303 247 303 251 303 240 303 271
00000060 ; " " ; 1 0 / 0 5 / 2 0 1 5
00000070 9 : 2 5 \n
00000075
```

```
$ libreoffice --version
LibreOffice 4.0.4.2
```

```
$ libreoffice issues.csv
```

**Text Import - [issues.csv]**

Import

Character set: Unicode (UTF-8)

Language: Default - Japanese

From row: 1

Separator options

☐ Fixed width

☒ Separated by

☒ Tab ☒ Comma ☐ Other

☒ Semicolon ☐ Space

☐ Merge delimiters

Text delimiter: "

Other options

☐ Quoted field as text

☐ Detect special numbers

Fields

Column type:

	Standard	Standard	Standard	Standard	Standard	Standard	Standard
1	#	Tracker	Statut	Priorité	Sujet	Assigné à	Mis-à-jour
2	87	Bug	New	Normal	Non-ASCII çéàù		10/05/2015 1

**#18 - 2015-05-10 13:01 - Toshi MARUYAMA**

- File Office2013Pro.png added

I can open on my Japanese MS Office 2013 Professional with general\_csv\_separator change.

```
diff --git a/config/locales/fr.yml b/config/locales/fr.yml
--- a/config/locales/fr.yml
+++ b/config/locales/fr.yml
@@ -158,7 +158,8 @@ fr:
   general_text_no: 'non'
   general_text_yes: 'oui'
   general_lang_name: 'French (Français)'
-  general_csv_separator: ';'
+  # general_csv_separator: ';'
+  general_csv_separator: ','
   general_csv_decimal_separator: ','
   general_csv_encoding: ISO-8859-1
   general_pdf_fontname: freesans
```

Office2013Pro.png

**#19 - 2015-05-17 08:32 - Toshi MARUYAMA**

[#7039#note-15:](#)

[German wikipedia](#)

The delimiter semicolon is expected when file is in ANSI.  
The delimiter TAB is expected when file is in Unicode.

I don't know about French.

**#20 - 2015-05-30 10:50 - Jean-Philippe Lang**

Is there any situation where we want to use UTF-8 **without** BOM?  
Because I'm not sure that a new setting is required, we could add the BOM when CSV is encoded in UTF-8 just like the redmine\_csv\_with\_bom plugin does.

**#21 - 2015-06-13 09:55 - Jean-Philippe Lang**

- Status changed from New to Closed
- Assignee set to Jean-Philippe Lang
- Resolution set to Fixed

**#22 - 2015-08-16 13:57 - Go MAEDA**

- Has duplicate Feature #11301: Add "byte order mark" to the export CSV file added

**#23 - 2016-01-21 10:47 - Toshi MARUYAMA**

- Related to Patch #17902: CSV encoding should be UTF-8 in French locale added

**Files**

7037-csv-bom.diff	4.53 KB	2015-04-02	Toshi MARUYAMA
7037-csv-bom-1.diff	4.53 KB	2015-04-02	Toshi MARUYAMA
7037-csv-bom-2.diff	4.54 KB	2015-04-02	Toshi MARUYAMA
excel2007.png	38.4 KB	2015-05-10	Jean-Philippe Lang
libreoffice.png	46.7 KB	2015-05-10	Toshi MARUYAMA
Office2013Pro.png	41.6 KB	2015-05-10	Toshi MARUYAMA