

Redmine - Defect #34235

Losing data due to unsupported characters

2020-11-08 15:37 - Ewan Makepeace

Status: Closed	Start date:
Priority: Normal	Due date:
Assignee:	% Done: 0%
Category: Issues	Estimated time: 0.00 hour
Target version:	Affected version: 4.1.1
Resolution: Duplicate	
Description	
<p>Occasionally I will write a long issue and when I submit I get an error page and my work is lost (or at least I can use back to recover the text, but I am obliged to start a new issue and recreate all the other fields).</p> <p>This has usually happened when I paste in text from other sources, Today this line caused the issue:</p> <p>Curve Line thickness scaled to absorbance of each individual channel. AL# ≤ 0 = no line, AL# ≥ 4 = maximum line thickness.</p> <p>[Update: Just testing on redmine.org and not seeing the same behaviour, so it may depend on database character set]</p>	
Related issues:	
Related to Redmine - Feature # 31921: Changes to properly support 4 byte char...	Closed
Duplicates Redmine - Defect # 10772: 4-byte utf-8 characters	Closed

History

#1 - 2020-11-08 15:39 - Ewan Makepeace

On my server submitting the above string in the task Description returns the error page:

Internal error

An error occurred on the page you were trying to access.

If you continue to experience problems please contact your Redmine administrator for assistance.

If you are the Redmine administrator, check your log files for details about the error.

Back

I feel at a minimum the error should be more explanatory?

#2 - 2020-11-09 18:00 - Holger Just

- Status changed from New to Needs feedback

The internal error page is the result of an unhandled error, that is: an unexpected issues which was not anticipated and thus could not be handled in a more user-friendly way. The error message send to the user in that case is deliberately vague to ensure that no internal settings can be leaked to malicious clients. As a server administrator, you should find more information (including a full error message and a stack trace) in your log/production.log file.

As the the issue itself: it might be you are using MySQL with a column encoding such as utf8 (or well: anything besides utf8mb4). In that case, MySQL

can not store 4-Byte unicode characters (such as most emojis or special characters), which might result in an error from the database. If that is the case, see #10772.

In any case, please check your log/production.log file and provide more information. Please see [[SubmittingBugs]] for details.

#3 - 2021-01-25 00:29 - Marius BALTEANU

- Related to Defect #10772: 4-byte utf-8 characters added

#4 - 2021-01-25 00:30 - Marius BALTEANU

- Related to Feature #31921: Changes to properly support 4 byte characters (emoji) when database is MySQL added

#5 - 2021-01-25 00:30 - Marius BALTEANU

- Related to deleted (Defect #10772: 4-byte utf-8 characters)

#6 - 2021-01-25 00:32 - Marius BALTEANU

- Duplicates Defect #10772: 4-byte utf-8 characters added

#7 - 2021-01-25 00:38 - Marius BALTEANU

- Status changed from Needs feedback to Closed

- Resolution set to Duplicate

I agree with Holger, most probably is the 4-Byte unicode issue.

In Redmine 4.1 we fixed this problem for new fresh installations (#31921), but for existing installations, before making that change, you need to migrate your database. In that ticket or on all the related issues, you can find various links to articles that describe the process. Please be aware that it's a sensitive process and I highly recommend to backup your data before proceeding.

We can reopen this issue if the logs point to another problem.

Later edit:

I've confirmed that is the 4-Byte unicode problem:

```
mysql2::Error: Incorrect string value: '\xE2\x89xA4 0 ...' for column 'notes' at row 1: INSERT INTO `journals` (`journalized_id`, `journalized_type`, `user_id`, `notes`, `created_on`) VALUES (6, 'Issue', 1, 'Curve Line thickness scaled to absorbance of each individual channel. AL# ≤ 0 = no line, AL# ≥ 4 = maximum line thickness.\r\n\r\n', '2021-01-24 23:40:24')
```