

Bugs & Wish list

change the table definitions to support the new utf8mb4 character | Tiki Wiki CMS Groupware :: Development
change the table definitions to support the new utf8mb4 character

Status

✖ Closed

Subject

change the table definitions to support the new utf8mb4 character

Version

19.x

Category

- Feature request

Feature

Database MySQL (MyISAM)
Installer (profiles, upgrades and server-related issues)
i18n (Multilingual, l10n, Babelfish)

Resolution status

Fixed or Solved

Submitted by

waterglass

Volunteered to solve

Louis-Philippe Huberdeau

Lastmod by

Marc Laporte, luciash d' being ☐

Rating

★ ★ ★ ★ ★ (0) ?

Related-to

- ✖ 4-bytestring (most Emoji) Causes Wiki Cutoffs
- ✖ Changing (modernizing) Tiki smileys (we should support Emoji)
- ● Replace Emoticons with Emoji
- ✖ Emojis crash Tiki (error 500)
- ✖ UTF-8 warning when upgrading Tiki18 to Tiki19

Description

I was creating a wiki page entry and encounter a bug. When a special character is used in the text, text is deleted at the instance of the first special character and below. I just upgraded from 10.2 to 11.1 hoping to fix this problem. Both version delete text using special characters inserted using the character feature of the editor. mu, epsilon, pi, kappa, h bar were used. epsilon is the first special character and was the start of the deleted text.

--

This is now converted into a Feature Request for tiki 13, See comment thread for reasons why

Solution

Done for 19.x
Related commits:

- <http://sourceforge.net/p/tikiwiki/code/67620>
- <http://sourceforge.net/p/tikiwiki/code/67633>

Importance
7

Easy to solve?
2

Priority
14

Demonstrate Bug (Tiki 19+)
Please demonstrate your bug on show2.tiki.org

Version: trunk ▼

Demonstrate Bug (older Tiki versions)
This bug has been demonstrated on show.tikiwiki.org

Ticket ID
4885

Created
Wednesday 06 November, 2013 05:29:39 GMT-0000
by waterglass

LastModif
Monday 29 October, 2018 10:20:46 GMT-0000

Comments



Xavier de Pedro 06 Nov 13 12:32 GMT-0000

Could you please reproduce the error in [show.t.o?](http://show.tiki.org)

There is button in the tracker item that allows you to create an instance (tiki installation) in the branch of your choice, or several, where you are asked to reproduce the issue. This way it will be much easier that other devs can revise your bug and fix it.

Thanks in advanced!



Pascal St-Jean 06 Nov 13 14:37 GMT-0000

I did a show instance, if you look at the homepage history. You'll see I added 5 paragraphs and saved. Then added the Pi character after the 1st paragraph and everything below disappeared



Jonny Bradley 14 Nov 13 23:22 GMT-0000

Investigating...



Jonny Bradley 14 Nov 13 23:50 GMT-0000

Seems to occur inside the PDO execute function, so not sure what can be done about it, especially without messing up other encoding (add a call to `utf8_encode` on the data stops it breaking, but messes up the characters for instance).

Maybe there's a PDO option that can cope with this, but i can't find it...



Pascal St-Jean 15 Nov 13 23:53 GMT-0000

I'll have a chat with LPH once he gets back from vacation to see if he has an idea. This is a tough one to crack.



Louis-Philippe Huberdeau 27 Nov 13 14:29 GMT-0000

MySQL 'utf8' support is limited to 3 bytes. Somewhere along 5.5 they added a mode to support 4 bytes per character.

<http://dev.mysql.com/doc/refman/5.5/en/charset-unicode-utf8mb4.html>



Nelson Ko 27 Nov 13 16:01 GMT-0000

For the 4-byte characters to be supported, one would need to change the table definitions to support the new utf8mb4 character set and also add way to check that the user has MySQL advanced enough to support it.

This would need extensive testing to make sure it does not break any other content. (too late now before release).

I suggest to remove the special characters input tool from the editor until this can be tried for Tiki 13.



Pascal St-Jean 27 Nov 13 16:08 GMT-0000

Thanks guys for reviewing and coming up with a proposed solution. I think this makes sense, We are days away from releasing Tiki 12. This is too drastic of a change.

Lets push to Tiki 13 (early tiki 13) so that we can test



Pascal St-Jean 27 Nov 13 16:10 GMT-0000

I converted this into a Feature Request in order to support the new encoding.

Attachments

filename	created	hits	comment	version	filetype
----------	---------	------	---------	---------	----------

No attachments for this item

The original document is available at

<https://dev.tiki.org/item4885-change-the-table-definitions-to-support-the-new-utf8mb4-character>