Bugs & Wish list

TikiWiki without AdoDB lacks the capability of writing correct UTF-8 data to MySQL | Tiki Wiki CMS Groupware :: Development

TikiWiki without AdoDB lacks the capability of writing correct UTF-8 data to MySQL

Status
❌ Closed

Subject
TikiWiki without AdoDB lacks the capability of writing correct UTF-8 data to MySQL

Version
3.x
4.x
5.x

Category
- Error
- Regression
- Consistency

Feature
Database MySQL (MyISAM)

Submitted by
amette

Volunteered to solve
Philippe Cloutier

Lastmod by
Philippe Cloutier

Rating
★★✦★★★★★✦★★★★★✦★★★★★✦★★★★★✦★★★★★✦★ (0) ❓

Description
I have just migrated a Tiki from one server to another and ran into
UTF-8 Problems. The browser showed Umlauts as the dreaded diamonds with
question marks in them.

Both servers are fully utf-8 compliant as described in
http://tikiwiki.org/UTF-8

The big problem is that on the new server PDO is available while it
isn't on the old one. According to the mailing list PDO is not default,
but as far as I can see in the code PDO will be default, if available
and not set otherwise. Setting

$api_tiki='adodb';

in db/local.php fixes the special character problem.

The issue seems to be that PDO gives a silent fuc# about encodings and
just pumps everything to the database and back as you hand it to it.
Additionally PHP/PDO doesn't read your my.cnf and therefore doesn't know
which encoding to use.
A quick fix is telling PDO which encoding to use - line 39 in
db/tiki-db-pdo.php :

$dbTiki = new PDO("$db_tiki:$db_hoststring;dbname=$dbs_tiki",
$user_tiki, $pass_tiki,array(PDO::MYSQL_ATTR_INIT_COMMAND => "SET NAMES
utf8"));

But I think that this is not safe to commit...
e.g. someone started with Tiki on MySQL 4.0 which used latin1 as
default, so all her tables are latin1. She upgrades MySQL to 5.1, but
doesn't change the Tiki database to default to utf-8 and hasn't
converted the database. Then the PDO-fix above will most probably
produce the same problem, just the other way around...

No idea, what to do actually, but wanting to raise awareness here!

This is a dangerous issue as people with utf-8 databases upgrading their
system or Tiki and switching automagically over to PDO will end up with
mixed stuff in their database: clean UTF-8 and double-byte
ansi-control-characters which MySQL converted to their right unicode
codepoints (freaky, hm? ;)).

I hope, I was capable of describing the problem in a way that's
understandable....

Solution
SET NAMES utf-8 is used since r30000. There was a lot of work on UTF-8 since March. The installer
should propose to fix upgrade issues.

If you see any problem left in 6.1, feel free to reopen.

Importance
7
Priority
35
Demonstrate Bug
Please demonstrate your bug on show2.tikiwiki.org
Version: trunk ▼ Create show2.tikiwiki.org instance
Ticket ID
3068

Created

Monday 08 March, 2010 14:42:55 GMT-0000

LastModif

Thursday 16 December, 2010 15:09:39 GMT-0000

Comments

majorarcanum 16 Jul 10 16:44 GMT-0000

is local.php documented???

Attachments

filename	created	hits	comment	version	filetype

No attachments for this item

The original document is available at

https://dev.tiki.org/item3068-TikiWiki-without-AdoDB-lacks-the-capability-of-writing-correct-UTF-8-data-to-MySQL