GSOC 2009: phpBB and MediaWiki-TikiWiki importer

This page is intended to document the development of the GSOC 2009 project "phpBB and MediaWiki-TikiWiki importer". Ideas and suggestions are welcome!

If you are looking for instructions on how to use the importer please see Mediawiki importer

You can download here the application sent to GSOC 2009 and you can view here my project page on GSOC site.

> **ATTENTION:** the original project was to create a Mediawiki and phpBB importer to Tikiwiki. But in the middle of the project (6th week) we decided to support only Mediawiki. phpBB was kept in the name of the project for historical reasons as this is the name of the project sent to Google.

Sumary

Student: Rodrigo Sampaio Primo (rodrigosprimo (at) gmail (dot) com)
Mentor: Nelson Ko
Project: phpBB and MediaWiki-TikiWiki importer
Abstract: I'm a big fan of TikiWiki and I think that a phpBB and MediaWiki importer to Tiki might be a good feature to incentive more people to join the community. My plan is to create an importer interface with native support to phpBB and MediaWiki. The importer should be as generic as possible, so it will be easier to develop support for other softwares in the future. If I'm able to finish the project before the deadline I'm interested in developing the importers from other wiki based softwares to Tiki.

Supported versions

The script might work with other versions but it is tested with the following versions of each software:

- Tikiwiki 4.0 (trunk)
- MediaWiki 1.14

Expected results

The expected results are:

- Create a generic import interface for TikiWiki with support for MediaWiki
- The MediaWiki importer must consider wiki page content, wiki page history, wiki page attachments and users
- Documentation on how to use the script
- Tech documentation on how to code importers for other softwares using the same structure

What information is considered by the importer?

A list of MediaWiki information that we want to be able to import to TikiWiki.

- wiki page content (including wiki syntax, wiki images and attachments). See comparative table between MediaWiki and TikiWiki syntax for a detailed list
- wiki page history
- users (the Mediawiki XML file does not have information about the users so we use a external extension to export users)

Project evolution

Period	Sumary
1st week	1) Decision to start with the MediaWiki importer and leave phpBB to a second phase 2) comparative table between MediaWiki and TikiWiki syntax 3) Initial study of the older import scripts
2nd week	1) Research of the MediaWiki export formats to choose which one is the best to use in the importer 2) Evaluation of older import scripts to define if we are going to reuse some parts of then 3) Text_Wiki will be used as a base for the MediaWiki importer 4) Ideas for a XML Interchange Format 5) Importer planning game
3rd week	1) Study of Text_Wiki and first simple contributions (1, 2, 3) 2) Read relevant parts of PEAR manual
4th week	1) Writing tests for the packages Text_Wiki, Text_Wiki_Tiki and Text_Wiki_Mediawiki 2) While writing the tests a few bugs have been fixed or reported: 1, 2, 3
5th week	1) Started coding the importer interface for TikiWiki based on the planning game
6th week	1) GSOC TikiFest in London 2) Decision to remove the phpBB importer from the scope of this GSOC project
7th week	1) Still coding the importer interface for TikiWiki based on the planning game. I wrote tests for the code I have written in the 5th week and included a feedback for the user showing the stats after the importation
8th week	1) Added to the user interface of the Tiki Importer feedback during the importation and information about relevant PHP settings 2) Added Mediawiki parser and Tiki render for redirect rule to Text_Wiki 3) Fixing and adding code to Text_Wiki_Mediawiki and Text_Wiki_Tiki and updating the comparative table between MediaWiki and TikiWiki syntax
9th week	1) Tiki Importer handle Text_Wiki errors when parsing the wiki page syntax 2) Continuing the work from last week based on the comparative table between MediaWiki and TikiWiki syntax 3) Fixed Text_Wiki_Mediawiki bug 16455 related with parsing italic and bold
10th week	1) Support for wiki attachments 2) Start of the TikiFest Barcelona

11th week	1) Mediawiki extension to export users 2) Organizing all Tiki unit tests in a single directory (lib/test on trunk) 3) Improving the performance of the importer script by only parsing the revisions that are going to be imported 4) Testing the importer with a real Mediawiki XML file (from one old Jonny's site) 5) Bug reported on the Tikiwiki parser 6) Add an option to Tikiwiki import users to force user change password on first login
12th week	1) New release of PEAR packages Text_Wiki_Tiki and Text_Wiki_Mediawiki 2) User documentation for the importer: Mediawiki Importer

Development

- Comparative table between MediaWiki and TikiWiki syntax
- MediaWiki export formats
- Evaluation of older import scripts
- Ideas for a XML Interchange Format
- Importer planning game

Know issues

- The first letter of file attachments on Mediawiki is case insensitive and on Tikiwiki it is case sensitive.
- Wiki page name on Mediawiki is case sensitive and on Tikiwiki is case insensitive, so two different pages for Mediawiki can be the same page for Tikiwiki
- Mediawiki syntax does not differentiate between images and files so the importer consider them all to be images (this means that both images and files in wiki pages will be imported but only images will be properly rendered on Tikiwiki pages)

Alias

gsoc2009rodrigo