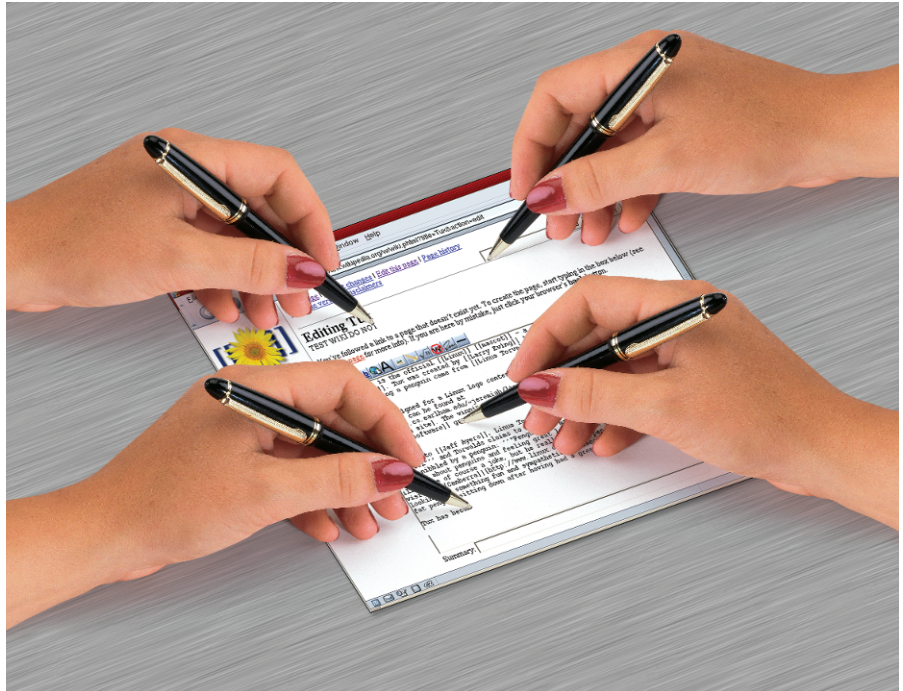


## Installing and Maintaining Mediawiki

# Collective Authoring

The free Wikipedia encyclopedia uses the Mediawiki engine to collect an enormous knowledge database. If you want to launch your own wiki, this PHP-based software is a good choice that provides rich functionality and plenty of power.

BY ERIK MÖLLER



**W**iki means quick in Hawaiian – and the Web-based wiki technology has established itself extremely quickly. Ward Cunningham had the idea back in 1995, when he was looking for a method to collect examples of common programming tasks amongst a group of developers. The simple idea was to allow modifications of

each wiki page in the browser. Square brackets are all you need to make a text string into a link. If the page already exists, the link is created; if not, the user is provided access to an edit box where he can create a new page with the provided heading (see Figure 1).

An easily understandable syntax makes light work of formatting. Typing

an asterisk at the start of a line creates a list; text surrounded by equals signs turns into a heading.

There are more than 100 wiki engines [1] at present. Mediawiki [2] is a popular example. It provides the underlying technology for the Copyleft encyclopedia, Wikipedia, the role model for most other wikis (see *A Journey to the World of Wikis*). In its first year, Wikipedia used the usemod wiki [3], which is written in Perl. But the task proved too challenging for usemod, which threatened to collapse under the load of thousands of articles and visitors.

This led the wikipedians to develop the Mediawiki engine in PHP, using a MySQL back-end. PHP skills are useful, if you need to set up and modify a wiki. But even if you don't speak PHP, Mediawiki should still allow you to set up a wiki that performs well, and is easy to use.

## Stable or Experimental

Mediawiki is available in two flavors. The stable version is available as a download from the homepage at [2]. The

## A Journey to the World of Wikis

Although it was specially developed for Wikipedia, Mediawiki is not restricted to encyclopedias. Besides Wikipedia, the Wikipedia Community launched the Wiktionary <http://www.wiktionary.org> in December 2002. This dictionary supplies an English definition and a translation into a number of languages for each word. The wiki model is ideal for this kind of task as it allows users from all over the world to supply translations.

### Quotes and Schoolbooks

Wikiquote <http://www.wikiquote.org> is a categorized collection of quotes; Wikisource <http://www.wikisource.org> adds original source material to supplement Wikipedia. The material is either public domain or

released under a free license. The Wikibooks project at <http://www.wikibooks.org> is unrelated to the Wikipedia as regards content, and provides free teaching material for arbitrary subjects. The project could develop into a central archive for lecture scripts for university teaching.

### Propaganda Encyclopedia

A list of Mediawikis is available at Wikipedia [10]. Some noteworthy examples are the propaganda encyclopedia, Disinfopedia <http://www.disinfopedia.org>. Mediawiki co-author Evan Prodromou runs a travel guide wiki called Wikitravel at <http://www.wikitravel.org>, and the author's open source documentation knowledge base, Openfacts, is available at: <http://openfacts.berlios.de/>

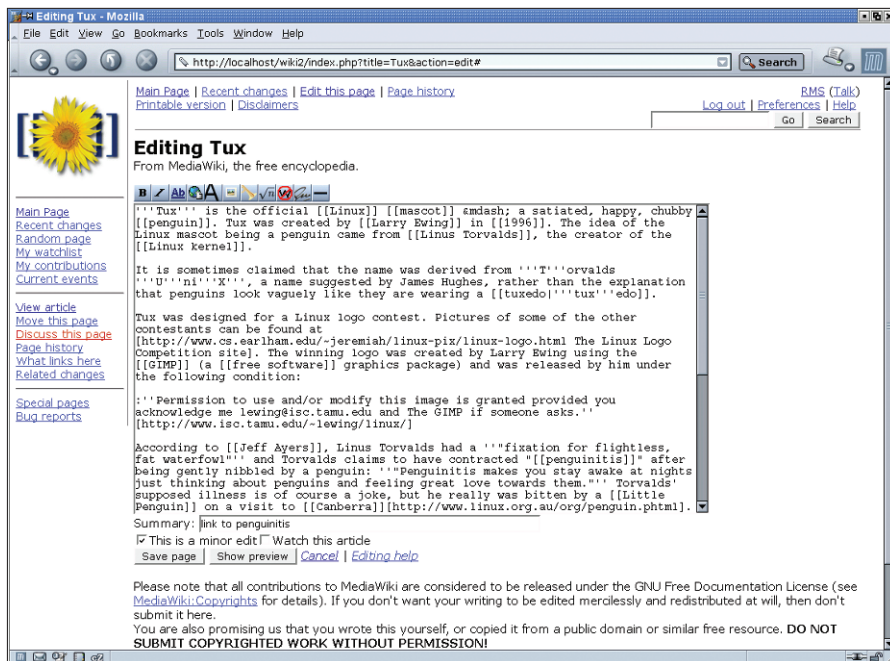


Figure 1: The Mediawiki edit window has a Javascript toolbar to help wiki newbies with formatting tasks. Double square brackets indicate a link to another page.

developer version should run, but it does contain experimental functions. You can download it with CVS:

```
cvs -d:pserver:anonymous@cvs.
sourceforge.net:/cvsroot/
wikipedia login
cvs -z3 -d:pserver:anonymous@
cvs.sourceforge.net:/cvsroot/
wikipedia co phase3
```

The “phase3” directory contains the software, which you can update at any time by entering *cvs update*. Mediawiki requires Apache 1.3.27, MySQL 4.0.13, and PHP 4.3.2 including the command interpreter.

It is compatible with PHP 5. If you want the software to reduce images automatically, you can opt for the *Image-magick* toolkit, or add the *gd* library from the current PHP versions. *Image-magick* supports far more file formats than *gd*.

The *install.php* script in the program root directory starts the install. Before installing, you will need to perform a few configuration steps. The admin should copy the sample configurations, *LocalSettings.sample* and *AdminSettings.sample*, to *LocalSettings.php* and *AdminSettings.php*. Table 1 shows the *LocalSettings.php* variables that need to be configured.

## Choice of Languages

Wikipedia supports over 50 languages. It is one of the most widely translated wikis. This allows you to create a wiki in Arabic, Chinese, Japanese, Hebrew, Hindi, Russian or Vietnamese, including Unicode support. To use these languages, the wiki operator simply needs to set the *\$wgInputEncoding* and *\$wgOutputEncoding* variables to *UTF-8*.

The installation script creates three MySQL users and assigns them database privileges. Mediawiki uses the *\$wgDB user* account for normal database opera-

tions, whereas the SQL user *\$wgDBsql user* has read access only. This account is used for a special page that enables sysops to perform SQL queries, for example, to locate and welcome recently registered users. The wiki operator should also add the *\$wgDBadminuser* MySQL admin user's account and password to the *AdminSettings.php* file. This account has write access and is used by scripts for installation and maintenance tasks.

To install the wiki, make sure you are root, and enter *php install.php*. Some distributions call the command interpreter *php4* rather than *php*. Among other things, the script prompts you for the MySQL root password, and creates the database, tables, and the MySQL users. Finally, it offers to create two wiki user accounts, a sysop and a developer. This is recommended. Without this option, you will need to assign the access privileges manually at a later stage.

To configure your Web server, add *phpml* as the PHP extension to *httpd.conf*:

```
AddType application/x-httpd-php
.php .phtml
```

In the *upload* directory created by the installation (where Mediawiki stores uploaded files) PHP execution or HTML rendering should be prohibited:

```
<Directory "/path/to/
uploaddirectory">
```

Working outside the atmosphere has advantages because the atmosphere obscures images and filters out electromagnetic radiation at certain wavelengths, mainly in the infrared.

**Table of contents** [hide]

- 1 Technical description
- 2 Discoveries
- 3 Launch and initial disappointment
- 4 Servicing Missions
- 5 Hubble's future and beyond
- 6 External links

**Technical description**

The unit weighs about 11,000 kilograms, is 13.2 meters long, has a maximum diameter of 4.2 meters and cost US\$ 2 billion (2 × 10<sup>9</sup> dollars). The telescope is a reflector with two mirrors; the main mirror has a diameter of about 2.4 meters.

The light collected and focused by the telescope ends up in one of several instruments. The instruments have been changed with different (and better) ones during maintenance missions.

The current (as of 2004) complement of instruments is:

- Near Infrared Camera and Multi-Object Spectrometer (NICMOS)
- Advanced Camera for Surveys (ACS)
- Wide Field and Planetary Camera 2 (WFPC2)
- Space Telescope Imaging Spectrograph (STIS)

Each of which has some capability as a spectrometer. In addition, the telescope's Fine Guidance Sensors (FGS) can and have been used for science.

It uses two solar panels to generate electricity, which is mainly needed to power the cameras and the four large flywheels used to orient and stabilize the telescope. The telescope's infrared camera and multi object spectrometer also need to be cooled down to minus 180 degrees Celsius for operation.

Hubble Space Telescope as seen from the Space Shuttle Discovery on mission STS-82.

Figure 2: A typical Wikipedia article shows that wikis don't need to be ugly. The table of contents is automatically generated from the headings.



```
AllowOverride None
AddType text/plain .html
.htm .shtml
php_admin_flag engine off
</Directory>
```

You need to enable a few global variables for the script directory. *php\_value register\_globals 1* in the directory statement of your Apache configuration will take care of this.

As an alternative, you could modify the *php.ini* file, which will typically reside in */etc* or */etc/php4*.

Hardening

The install script copies all the PHP files to the Web server directory. The more scripts with external access you have there, the greater the risk that an attacker might gain access to internal functions. This is particularly true of *register\_globals*, as URL parameters allow you to set internal script variables.

Only really essential files should be stored in the Web server directory, that is, anything ending with *.phtml*, images and stylesheets. Best practice is to move any files ending in *.php* to a separate directory outside the *document root* after installing. After moving the files, now replace the links to *./LocalSettings.php* in the *phtml* files with links to *LocalSettings.php*, and add the new path to the *include\_path* in the *php.ini* file.

Note that in the current unstable version, which may be stable by the time you read this, the *.phtml* files are deprecated – in this case, copy *index.php* and *redirect.php* to the web server directory, and copy all other *.php* files to a safe directory from which they can be included.

Manual Install

If you do not have root access to the server, or if the install script fails, you can install Mediawiki manually. The first step is to create a database. You should be aware that name conflicts with existing tables can occur if you use an existing database – Mediawiki does not use a prefix.

Import the tables from *tables.sql* and *indexes.sql* in the maintenance directory:

```
mysql -u root -p databasename
< filename.sql
```

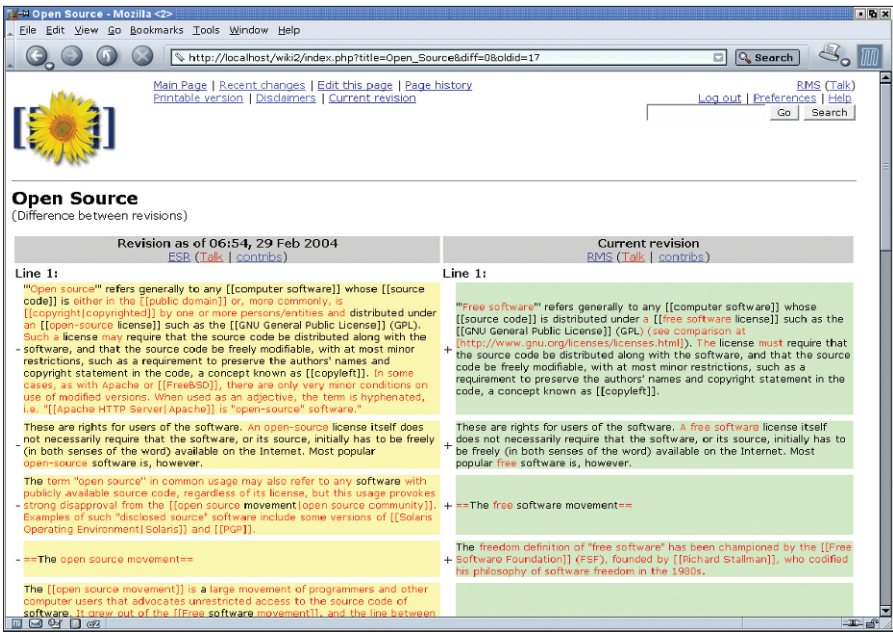


Figure 3: The diff view highlights the differences between two revisions, showing modifications and deletions. Sysops can point and click to reinstate previous versions.

You will need a minimal collection of settings in *LocalSettings.php*; leave out the SQL user, and ignore *AdminSettings.php*. Of course, you must have a main database user. Copy any *.php* and *.phtml* files from *includes* and *languages*, and the *stylesheets* and *images* directories, to the Web server's main directory.

You can then run the *wiki.phtml* file with the path chosen in the previous step to access the main page of the newly installed Mediawiki in your browser. Take a look around and try out the software's features (see the Mediawiki 101 box). If the Web server does not recognize *phtml*, rename the *index.php* file. In this case you will also need to edit the *\$wgScript* and *\$wgRedirectScript* variables. *DefaultSettings.php* provides a sample.

User Privileges

Most wikis allow anonymous users to edit their pages. This poses the question as to how the wiki operator will handle undesirable content or even defacement. Mediawiki has a few defense mechanisms. You can restrict reading or editing content to specific user groups. The *\$wgWhitelistEdit* and *\$wgWhitelistRead* flags in *LocalSettings.php* help you do so. The *\$wgWhitelistAccount* array defines the groups allowed read or write access:

```
$wgWhitelistAccount=array(
    "user" => 0, "sysop" => 1,
    "developer" => 1)
```

This setting restricts modifications to sysops and developers. The following will add a new user type called *editor*:

Table 1: Basic Configuration	
Variable	Local environment
\$IP	Local path on the server where you are copying the wiki, for example <code>/var/www/wiki</code>
\$wgServer	Server address element before the first slash, for example <code>http://www.mywiki.com</code>
\$wgScriptPath	Subdirectory with PHP files, for example <code>wiki</code> ; leave this blank ("" ) if the files are in the main directory
\$wgEmergencyContact	Email address of the administrator to be displayed in case of problems
\$wgDBserver	Name of the MySQL database
\$wgDBuser and \$wgDBpassword	MySQL user for normal database access
\$wgDBsqluser and \$wgDBsqlpassword	MySQL user for SQL queries via Web interface; only has read access
\$wgLanguageCode	Two letter code for language to use, for example <code>en</code> for English

```
$wgWhitelistAccount=array(
"editor" => 1, "user" => 0,
"sysop" => 1, "developer" => 1)
```

Having created the user type, the operator can go on to assign read or write privileges by setting the value of a user's *user\_rights* database field to *editor*.

Mediawiki does not have a generic privilege management tool at time of writing. This means formulating an SQL query to make a user a *sysop*, *developer* or *editor*, for example:

```
USE Databasename;
UPDATE SET user_rights='sysop'
WHERE user_name='username';
```

The developer version now has a *bureaucrat* user type. Bureaucrats use a special page to designate other users as sysops. To do so, the user needs sysop and bureaucrat privileges.

A change history allows you to monitor recent developments in your wiki. Mediawiki provides a *diff* feature that highlights the difference between two revisions (see Figure 3). The change history records the modifications to a page right back to the original version. The version history allows you to reload previous versions, and save them again, if required. Finally, users have a *User contributions* function that allows them to view all the content created by a specific user. Sysops can undo entries from this list, allowing them to repair defacements within a few seconds.

If a dispute over the content of a page arises, the sysop can temporarily protect the page. Admins can ban persistent trouble-makers by blocking their IP addresses. Unfortunately, unregistered users with dialup connections constantly change their IPs.

For this reason, an IP block will only last for 24 hours by default (*\$wgIPBlockExpiration* option). On the whole, defacement is an over-

rated issue. Social issues are more difficult. What kind of pages and what kind of behavior will the wiki allow? Operators definitely need to put some thought into policies.

Wiki operators should define the default options, the *\$wgDefaultUserOptionsEn* array in the *Language.php* file or its localized variant, carefully. Options such as *quickbar* (0 = no navigation bar, 1 = navigation bar left, 2 = right) are important. *editondblclick* will open the editing window when a page is double-clicked. *showtoc* generates a table of contents for pages with more than three headings, and *showtoolbar* enables the Javascript toolbox in the edit window (see Figure 1).

The *editsection* option displays an *Edit* link next to each heading. This allows direct editing of specific sections, and is particularly convenient for longer pages, as it removes the need to navigate the Edit window. On the downside, links of this type can spoil your layout. If you enable the *editsectiononrightclick* option, right-clicking a heading has the same effect. This caused Konqueror to crash occasionally in our labs.

## Namespaces

Wikipedia quickly faced the issue of separating information about the wiki, policies, or discussions from encyclopedia articles. For example, there is an

article about FAQs, and an official Wikipedia FAQ. To separate the two sensibly, Mediawiki introduced so-called namespaces, which are defined in the *Language.php* or the local translation. By default, these are the main namespace, multiple discussion namespaces, a namespace for images, one for user pages, a special namespace, and the very special Mediawiki namespace. Articles outside the main namespace will always have a prefix. *[[Discussion:Mainpage]]* points to the discussion page that belongs to the main page, *[[User:Troll]]* is the personal page belonging to the user Troll.

If you do not intend to run a gigantic wiki, you can save yourself the trouble of the meta namespace. By default, this namespace assumes the value of the variable *\$wgSitename* in *DefaultSettings.php*. The Wikipedia FAQ is stored below *[[Wikipedia:FAQ]]*. This resolves the conflict referred to previously, but makes things difficult for smaller sites. If namespaces 4 and 5 are deleted from the *\$wgNamespaceNamesDe* array in the *LanguageDe.php* file, for example, they are also removed from the user interface.

One feature that Mediawiki retained after migrating from usemod is the sub-page. For example, an article called *Linux* could have a sub-page called *Linux/Kernel Tips*. The Kernel Tips page would automatically have a link back to *Linux*. Sub-pages can be enabled individually for a namespace by setting the *\$wgNamespacesWithSubpages* variable.

The *MediaWiki* namespace, which is enabled via *\$wgUseDatabaseMessages*, allows users to edit all the text elements of the software within the wiki, thus aiding translations. The names of the text variables are stored in *Language.php*.

## Shortcuts for Text Blocks

The content of any pages created in this namespace can be embedded at any location within the wiki. For example, to provide a default wel-

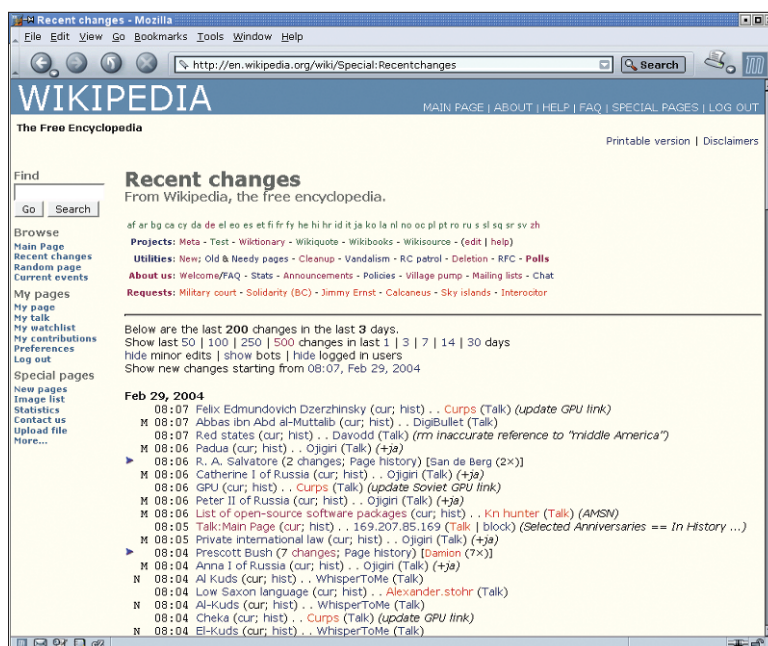


Figure 4: Viewing recent changes to the wiki in the CologneBlue skin. The *Recent Changes* link leads directly to the diff view for the current change.

come text, create a page called *Mediawiki:greeting*. `{{msg:greeting}}` will display the content of the page on other pages. Any changes to the Mediawiki page are immediately applied to any other pages that use `{{msg}}` to access it. The `{{subst:greeting}}` command adds the text at the current location but without the automatic update feature.

Note that pages in the Mediawiki namespace are not allowed to have non-standard characters or spaces in their headings. As each program text is retrieved from the database, you should use this feature only in combination with *memcached* (described later).

## Operating a Wiki

Running Mediawiki over a longer period of time is actually far less stressful than getting the installation right. In a perfect world, the wiki operator would only need to assign sysop and developer privileges. *mysqldump* or *phpMyAdmin* are recommended for backing up the database. When importing large datasets, such as the Wikipedia articles at [5], run the *maintenance/rebuildall.php* script in

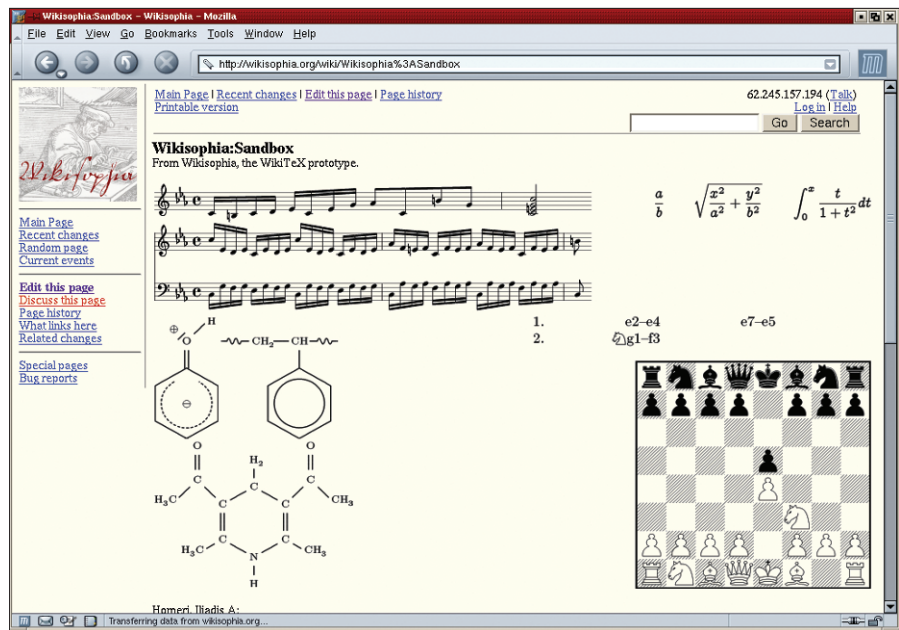


Figure 5: The Wikitex interface allows Mediawiki to display sheet music, chemical and mathematical formulae or even chess games. Clicking on the notes will automatically generate a midi file.

the install directory as your next step, to avoid queries such as *What links here*, and other requests failing.

Mediawiki has a number of tuning options – there are examples in *Default-*

*Settings.php*. Ensure you copy modified options to *LocalSettings.php* to avoid overwriting them during an update.

Mediawiki has the *\$wgMiserMode* option for larger wikis. Enabling the

## Mediawiki 101

The main navigational element in Mediawiki is the sidebar on the left of the screen (see Figure 1). *Recent changes* is one of the most important links for regular wiki visitors. It shows the pages the wiki contributors are currently working on (see Figure 4).

Registered users can click on the *Watch this page* link to add a page to their watchlist. The list shows the changes made to the pages in your watchlist over the past week. This allows contributors to monitor pages where they have provided content. *Move page* allows you to rename a page. The previous title still remains valid; users following links to the page will be referred to the new link.

The *Discussion* link points to the discussion

for the current page. A page called *Linux* would have a discussion called *Discussion:Linux* for example. A link called *Post a comment* is available in the side bar for this page type. You can follow the link to open an empty editing window and add your comment. After clicking on *Save Page* your comment is added to the end of the page.

### A Page and its History

To add a new page, simply create a link on an existing page, for example `[[My new page]]`, and then follow the red link. The *Page history* link takes you to a page with the previous versions of the current page (see Figure 4). This view can also be used to reinstate previous versions. To do so, select a version by

clicking on its timestamp, edit the page you are viewing, and store the page. The *Current revision* link is also quite useful, and shows the differences between an arbitrary revision and the current one.

*Upload* allows you to upload arbitrary files. Uploads are recorded in the logbook, and can be embedded using the `[[Image:file.jpg]]` or `[[Media:file.zip]]` syntax. Mediawiki displays image links on the fly; by contrast, media links point to the file.

### Recognizing Relations

One reason for the move to a database centric solution was the Wikipedians requirement for advanced searching. For example, Mediawiki uses special pages to display new or orphaned pages, that is pages that are not referred to by any links. Additional functions list articles alphabetically or by length.

The special page for *Most wanted articles* shows an overview of the most commonly used keywords that did not return a page. The *What links here* link shows a list of the links to the page. *Related changes* allows you to view changes to the articles that a page links to, and thus monitor specific topics. The official Mediawiki Manual at [4] provides more detail.

### Poker

Revision history

View (previous 200) (next 200) (20 | 50 | 100 | 250 | 500).

Legend: (cur) = difference with current version, (last) = difference with preceding version, M = mi

- (cur) (last) 23:31, Feb 3, 2004 . . Vikreykja (Added Esperanto link)
- (cur) (last) 06:09, Jan 28, 2004 . . Snoyes (mv intlinks)
- (cur) (last) 20:07, Jan 21, 2004 . . Revolver (lk)
- (cur) (last) 12:27, Jan 21, 2004 . . 62.64.160.101
- (cur) (last) 17:53, Jan 20, 2004 . . Raul654 (Move the quote; don't delete it)
- (cur) (last) 13:44, Jan 20, 2004 . . 213.10.155.53

Figure 6: The history allows the sysop to monitor the development of a Wikipedia article.



option will disable time-consuming queries such as a list of the longest articles. It definitely makes sense to enable the page cache option, *\$wgUseFileCache*. This will serve up static HTML pages to anonymous users and thus provide a considerable performance boost.

If you are prepared to put some time and effort into tuning your wiki, you might like to try memcached [6] and zlib. Memcached reduces database access by caching user data and link information. Zlib compresses older page revisions. This assumes that you compiled PHP with the *--enable-sockets* and *--with-zlib* options. Type `<?phpinfo()?>` on a PHP test page to find out.

### Memcached

The daemon should run in the background with the options *-d -l 127.0.0.1 -p 11000 -m 64* enabled. This assigns 64 MBytes of RAM as cache memory for local applications. Memcached does not provide an authentication feature. Local users have unrestricted access. Avoid

running memcached on a system with multiple local users. Without a firewall, or the *-l* parameter, external users can access the server and capture passwords.

You can now enable the *\$wgCompressRevisions*, *\$wgUseMemCached*, *\$wgSessionsInMemcached*, and *\$wgLinkCacheMemcached* options. *\$wgCompressRevisions* uses *gzip* to compress older page revisions, and thus saves a lot of space on the server.

### In the Pipeline: Wikitex

It is not in the official version yet, but check out [7] for a first glance at the Wikitex interface written by Peter Danenberg. This will allow Mediawiki to make use of numerous back-ends in future. For example, GNU Lilypond [8], which translates a simple syntax into neatly formatted sheet music layouts, various Latex macros that generate mathematical and chemical formulae, or chessboard images to illustrate game plans (see Figure 5). New users may find the mass of features confusing at first. The Mediawiki community is

pleased to lend a helping hand. If you run a Mediawiki, you should sign up for the *mediawiki-l* [9] mailing list, or join the *#mediawiki* IRC channel on *irc.freenode.net*.

#### INFO

- [1] Master list of wiki engines:  
<http://cz.com/cgi/wiki?WikiEngines>
- [2] Mediawiki: <http://www.mediawiki.org>
- [3] Usemod Wiki: <http://www.usemod.com>
- [4] Mediawiki manual:  
[http://meta.wikipedia.org/wiki/MediaWiki\\_User%27s\\_Guide](http://meta.wikipedia.org/wiki/MediaWiki_User%27s_Guide)
- [5] SQL Dumps of Wikipedia:  
<http://download.wikimedia.org>
- [6] Memcached:  
<http://www.danga.com/memcached>
- [7] Wikitex trial: <http://wikisophia.org>
- [8] Lilypond: <http://lilypond.org/web>
- [9] Project mailing list: <http://mail.wikipedia.org/mailman/listinfo/mediawiki-l>
- [10] Projects that use Mediawiki:  
[http://en.wikipedia.org/wiki/Wikipedia:Sites\\_using\\_MediaWiki](http://en.wikipedia.org/wiki/Wikipedia:Sites_using_MediaWiki)

## High Performance 64-bit Linux Clusters

At DNUK we can build a complete Linux computational cluster for you. When you purchase a group of rackmount servers from us we will build, configure and test your cluster for you at no extra charge.

Choose from 32-bit Intel Xeon or 64-bit AMD Opteron nodes and a variety of open source libraries including MPICH, LAM/MPI and PVM.

A range of hardware interconnects including Gigabit, Dolphin SCI and Myricom Myrinet are available.

Visit **[www.dnuk.com](http://www.dnuk.com)** and find out why corporate customers, small and medium businesses and most UK universities choose us for their IT requirements.

For more details on the clustering solutions we can build, please visit:

<https://secure.dnuk.com/store/clustering.php>

