Using TEX in a wiki

Tomasz Łuczak
Katowice, Poland
tlu (at) technodat dot com dot pl
http://team-tl.livenet.pl

Abstract

This article describes the use of wikis as sources and presentations of texts, with TEX as a hidden engine for typesetting the wiki content.

1 The preliminaries

When one wants both a printed version of a document (usually in PDF format) and in the form of a web page, one usually starts from a TEX document, which is converted to HTML. However, this requires that the author has at least basic TEXnical skills. If several people who do not know anything about TEX work on a document and the printed version is the last step in the chain, then starting from a web page might be a more appropriate direction.

The process of creating technical documentation thus can involve the editing of web pages — we have an instant presentation of the current state of the document. The next step is conversion to the TEX format followed by compilation. It is important to note that the converted documents should not need to be manually adjusted or corrected.

2 The need

One of the practical uses for such a "from web to TEX" process is creation of a system's quality management documentation. The documentation is being created by employees who usually have no TEXnical skills. This leads to the idea of providing such tools with which they are familiar, but of course we will not accept any loss of quality of the typeset result.

3 The execution

If one insists on ease of editing combined with keeping a change history, as well as the ease of converting into TEX, it turns out that a wiki is the simplest solution. We only have to choose the proper program.

Dokuwiki was used for this project (`http://wiki.splitbrain.org/wiki:dokuwiki`), thanks to its automatic table of contents generation and page fragments editing. *Dokuwiki* keeps pages UTF-8 encoded without any other control codes. This enormously simplifies the realization of an automatic converter.

To restrict authors' invention, a set of "wiki complementing rules" was prepared which listed the allowed markup tags along with examples of use. During the preparation of these "rules" it turned out that the following features suffice: text emphasis; tables; drawings; itemization and enumeration without nesting; and headings for three levels (chapters, sections and subsections).

Unfortunately, the wiki does not automatically number chapters or sections, so to achieve a uniform look we decided to include chapter and section numbers directly into the titles. This solution is not a most elegant, but proved to be effective. Incidentally, this also gave section numbers in the automatically generated table of contents for a given wiki page.

Limiting the allowed tags made writing the converter easy. The code written in TCL fits into about hundred lines, mostly thanks to the following limitations: the PDF document is only for printing so the wiki's hyperreferences were not converted; all tables had the same structure; and similarly, all drawings had the text column width.

Due to the deficiency of the converter authors were required to put an empty line to mark the end of a list. Unfortunately this was not always adhered to, leading to the biggest issue during conversion and compilation.

4 The summary

The use of the wiki and a web browser proved to be a hit: documents were quickly and willingly edited by the authors. The documentation not being up-to-date proved not to be an issue. Last but not least: the readers liked the easy to read and nicely typeset documents.