Edukalibre Project: Versatility in e-learning

Jaime Villate, Guilherme Dutra, Nuno Faria, Mário Lopes
Dep. de Física – Fac. de Eng. da Univ. do Porto
villate@fe.up.pt

Abstract

Edukalibre, a project funded by the Socrates/Minerva program of the European Union, aims at the promotion of information and communication technology in education. Its main goal is to transpose the advantages of the Free Software development model to the production of educational documentation.
This article focuses on the solutions found within the Edukalibre project to facilitate collaboration for the production of educational materials. The tools under development allow different authors to work on the same document, using different computer programs and therefore different document formats, without the need to rewrite existing content or to learn new formatting languages. The Edukalibre system also includes document version management and the possibility of writing documents on-line and/or off-line, via a user-friendly Web platform or through the Webdav protocol.
The results of a case study involving the elaboration of work reports by university students using the Edukalibre platform are also presented.

1. Introduction

The development of internet based communications has brought significant changes to the way software is developed. These are not merely technical changes. An entirely new concept of the information society has arisen. An example is the Free Software movement. More recently, these developments have led to increased interest in e-learning software tools such as Moodle (Moodle Project, 2004).

The Edukalibre project arises from these two concepts as it tries to introduce the free/libre software philosophy into the world of education document elaboration. Everyday, all around the world, different teachers teach the same course materials and write their own course manuals. All this effort could be better applied if these materials were created from cooperation between their authors. However, one of the main difficulties found is the wide range of typesetting tools available and their respective incompatible file formats. This article focuses on the solution found within the Edukalibre project to cope with these files.

2. The Edukalibre project

Edukalibre is a project financed by the Socrates / Minerva program (Edukalibre Project, 2005). Within this project, six European partners cooperate, trying to explore new ways to promote the elaboration of education materials in a way similar to the methodology generally used in communities devoted to the development of free/libre software.

One of the tasks of this project consists of developing software tools to facilitate the collaborative development of education-related documents like manuals and assignment reports. The idea is that a group of users should be able to work together, simultaneously on the same document, using the typesetting tool they prefer. The tool being developed is an interface to a collection of independent tools, namely:

- a document version control system;
- several document format conversion tools;
- several user interfaces among which are Moodle modules and OpenOffice macros.

Figure 1 illustrates the general architecture of the tool previously mentioned.

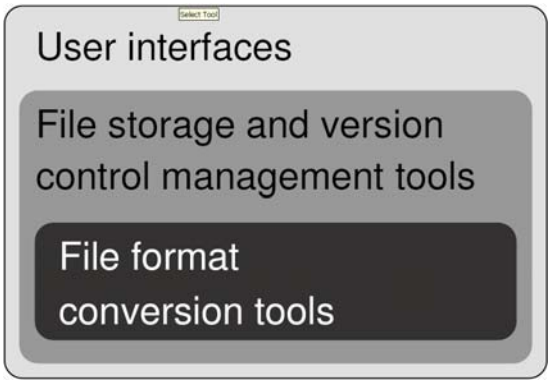

Figure 1. General architecture of the tools being developed under the Edukalibre project.

All software being used by the project is licensed as free/libre. Likewise, all tools developed by the project are made available to the community as free/libre software and all documents are published under a free/libre documentation license.

2.1. Document version management system

"Version control is the art of managing changes to information. It has long been a critical tool for programmers, who typically spend their time making small changes to software and then undoing those changes the next day. But the usefulness of version control software extends far beyond the bounds of the software development world", (Collins-Sussman et al., 2004).

In the Edukalibre system, version control is used to manage the documents. In fact, version control is the heart of the Edukalibre system. It deals with document storage, version management, creating and merging different version branches. The software chosen for this task was Subversion (Collins-Sussman et al., 2004). Some of the reasons that led to this choice were:

- Lock System: it forbids authors or collaborators to overwrite documents belonging other users;
- The documents may be cataloged using Name Spaces, thus it is possible to create collections of documents, even if they are in distinct places or have been erased.
- It is possible to add meta data, this information is stored in XML format transparently to the user.

2.1.1. Branches and hierarchy

The Edukalibre system assumes three levels of user hierarchies: Coordinator, Author, and Contributor/Reader. Figure 2 shows the hierarchical relation between people enrolled in a document editing.

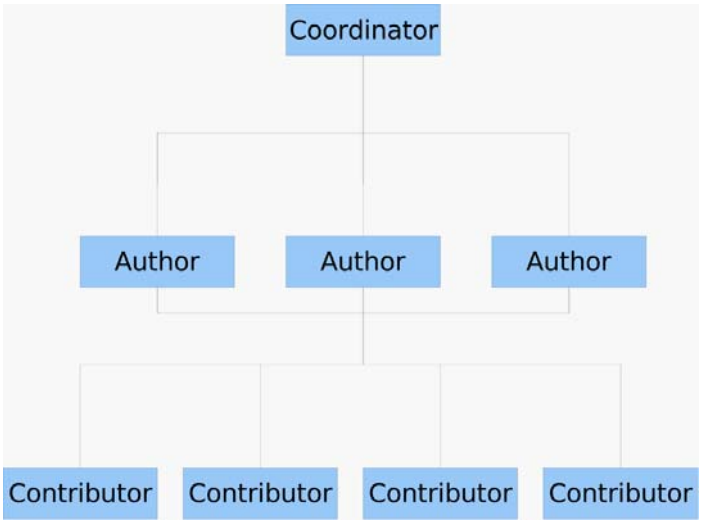

Figure 2. Hierarchy.

The Coordinator is responsible for managing the document edition. He decides who are the authors and when a new version of the document is released.

Authors can edit and create their own document versions. He decides whether or not to accept the contributions from Contributors. It can be a single author or a group of authors. Contributors/Readers are allowed to read, or edit and create their own document versions. These modifications start as new branches.

Each document has a main branch which is created and updated by a coordinator. When an author or contributor updates a document, a new branch is created. It is up to the coordinator to decide whether it will be merged or not into the main branch.

In figure 3 we can see an example that illustrates how the ramifications are treated. Some of them are merged in the main branch and others are not.

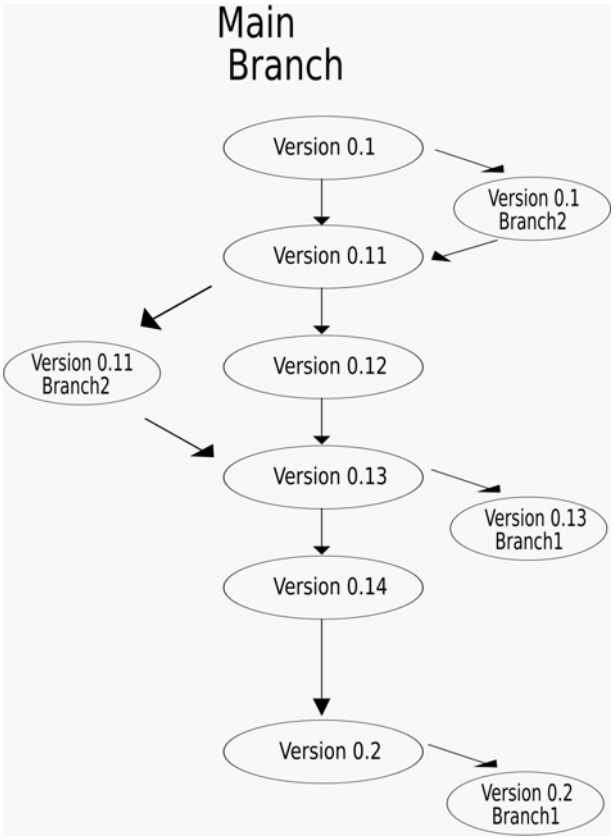

Figure 3. Branch schemas

2.2. Document typesetting tools and user interfaces

The users can access the repository through several alternative interfaces.:

- Command line, using an SVN client and SSL authentication;
- Direct access by Web-based Distributed Authoring and Versioning (WebDav), which is a set of platform independent HTTP extensions that allows users to edit and manage files in remote Web servers, (Webdav web site, n.d.);
- Transparent access through OpenOffice macros;
- Collab and Collab for Moodle - Web interfaces developed in PHP;
- Condor - A Moodle module which supplies a integrated set of tools useful for management of collaborative work.

Users can edit documents using OpenOffice, Wiki on-line editor, Lyx/Latex text editor, or even their preferred text editor, provided they change the XML source file itself. When a Wiki or OpenOffice document is uploaded to the repository, it is automatically converted to Docbook/XML format. Latex and other formats remain unchanged, see figure 4.

The Wiki on-line editor is available both on Collab and Condor Web interfaces. When documents are edited online, they are instantly posted into the repository. OpenOffice documents can be posted through OpenOffice macros, which are activated by menu, thus the post process is transparent for users.

Lyx/Latex and other formats must be posted by upload in Collab or Condor, or through an

SVN client.

Through the Collab interface it is possible to open and download all documents available in the repository. It is also possible to see the differences between the several versions of the documents. These capabilities are implemented using Webdav extensions.

The Condor module is integrated in the Moodle e-learning platform, (Moodle Project, 2005). This module includes, in the same interface, a message management tool, a real-time forum environment, and access to the Edukalibre repository.

3. Document format conversion tools

It was assumed from the beginning of the project that users should be allowed to choose the document typesetting tool they preferred. In order to do this, it was necessary to provide the system with a set of conversion tools for the formats allowed. This section describes the solutions found and the work still to be done. Figure 4 summarises the system architecture for document conversions.

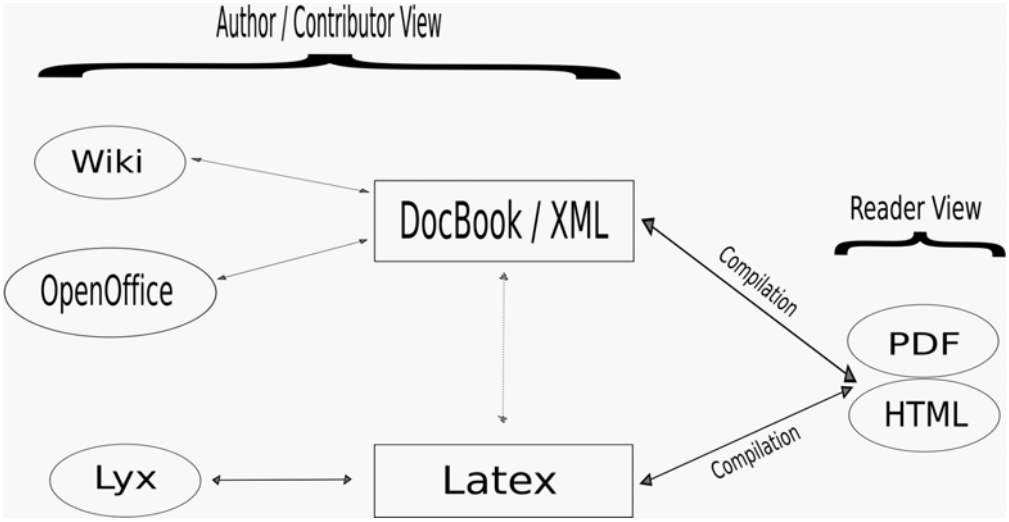

Figure 4. Conversions between formats.

3.1. Docbook XML - the base format

"Docbook is a very popular set of tags for describing books, articles, and other prose documents, particularly technical documentation. Docbook is defined using the native DTD syntax of SGML and XML", (Normal & Muellner, 2005).

Docbook is the Edukalibre system base format because it is open source, widely used for book and article writing, and it is possible, using OpenOffice, to write documents in Docbook format through a 'What You See Is What You Get' (WYSIWYG) interface. Moreover, there are conversions to several formats, and it is easy to write conversion scripts for XML (eXtensible Markup Language) documents.

From all the file formats allowed in the Edukalibre System, PDF, Postscript and HTML are only output formats. Everytime a new Docbook file is generated or uploaded, these formats are generated through the conversion tool FOP (The Apache XML Project, 2004).

3.2. Conversion to and from OpenOffice

OpenOffice.org (Sun Microsystems, 2002) is a free office productivity suite that differs from other tools in its way of handling and saving the documents data. Whereas some editors encode the content as a binary format, OO.org makes use of a XML-based format for storing it. This means that the information is always available in a human readable format.

One of the major advantages of this format is content, layout and meta information separation. The objects, such as images, are embedded into the final package. In Version 2 of the OO.org suite, the format has been adopted by OASIS (Organisation for the Advancement of Structured Information Standards) as a move to standardising it as a common format for saving office documents.

Although the file format is widely available, it is not yet being used widely. Fortunately, there are open and free software tools that convert to and from the OO.org document format.

In order to use them there are some tools available like OOoConv or BatchConv (Goddard, n.d.). Nevertheless, there are some other applications that manage to extend OO.org capabilities and convert to even more formats, like Writer2LaTeX (Just, n.d.) that is able to create a LaTeX file with BibTeX references.

As for the Edukalibre Conversion Tools framework, the default procedure is initially converting to Docbook by making use of a XSLT stylesheet and processor, and then making the other conversions from this (Figure 5).

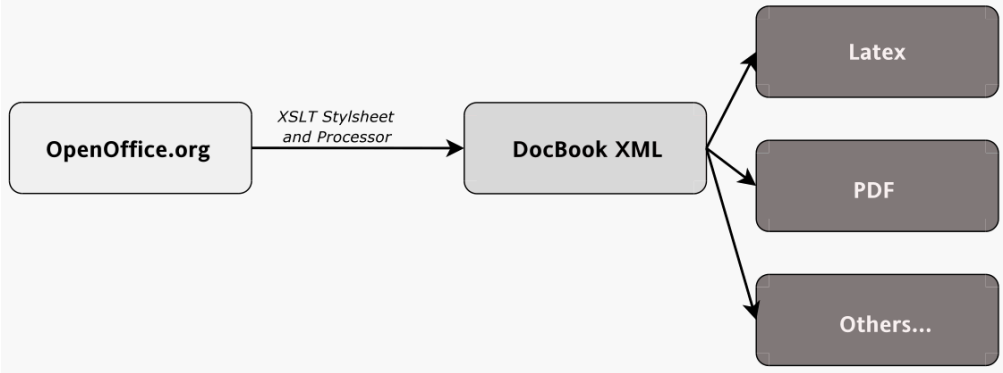

Figure 5. How Edukalibre handles OpenOffice.org

Unfortunately this cannot be done to OO.org files since Docbook is not as rich in formatting as OO.org. So, information and formatting styles would be lost in the conversion. Due to this, the conversion is made directly from OO.org using internal macros. This situation is represented in Figure 6.

The disadvantage of this method is that OpenOffice as well as a X Windows server (or a virtual frame buffer) must be running on the server. This will be addressed when OpenOffice 2.0 becomes a stable and more widespread platform.

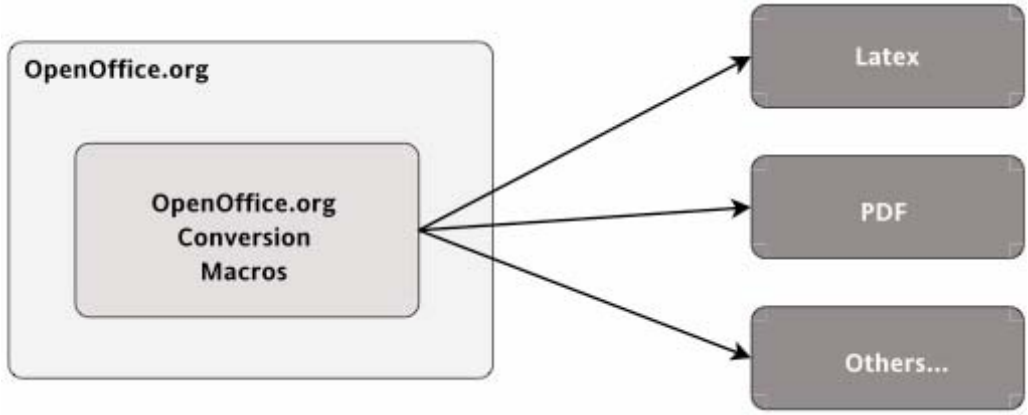

Figure 6. How Edukalibre handles OpenOffice.org conversions - detailed view.

3.3. Wiki - A format for online document typesetting

According to WikiPedia (Wikipedia, 2005), Wiki is an online collaboration model and tool that allows users to edit content through a simple browser. It is unnecessary to know tags, the syntax is comprised entirely of punctuation characters.

There is no single Wiki format standard. Instead, there are several similar formats but different enough to make it impossible to have one single fully compatible tool.

Within the Edukalibre project, the Wiki Markdown format (Gruber, 2004) was chosen. Its documents are easy to write and easy to read. The punctuation characters have been carefully chosen so that they convey, as much as possible, the formating information they represent - asterisks around a word actually look like **emphasis**, blockquotes look like quoted passages of text and so on.

The TextWiki (Jones, 2004) package, which is a part of the PEAR (PHP Extension and Application Repository) framework (The PHP Group, 2004), has been used to convert Wiki to Docbook. It is an object oriented framework organised as *Parsers* and *Renderers*.

The *Parsers* analyse the text and identify Wiki rules occurrences.

When a Wiki rule is found, the Parser replaces the matched text with a "delimited token" and creates an entry in the tokens array. This process generates an intermediate document. The Renderers transform these tokens into its destination format equivalents, in this case, the Docbook format. Figure 7 illustrates this process.

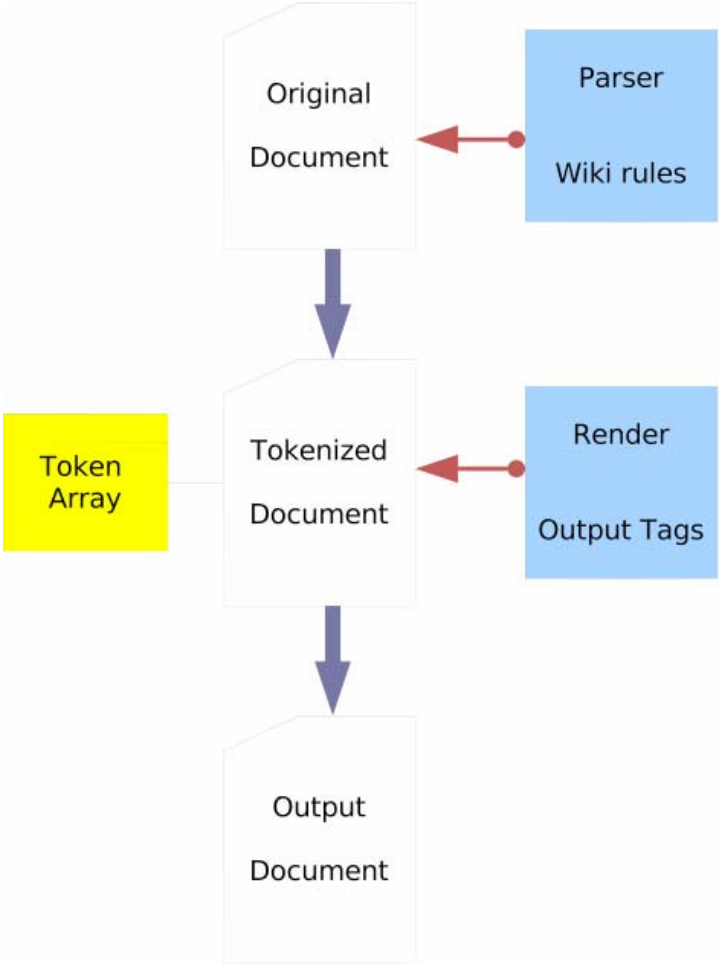

Figure 7. TextWiki schema.

In order to perform the Docbook to Wiki conversion, a new tool has been written from scratch. As the main language of the project is PHP, this tool was developed using the XML *Parser* and SimpleXML PHP toolkits.

As the Wiki format is simpler than the Docbook format, it is impossible to convert all Docbook text formating tags. In order not to lose information from files which might have been generated with more complete tools, such as OpenOffice, it was decided to include these tags in the Wiki document,exactly as they appear in the Docbook original file.

3.3. Latex - A special case

Latex is an extremely rich typesetting language. Therefore, it is extremely difficult to develop a tool capable of converting any Latex source file to Docbook XML.

This reason could justify leaving the Latex format out of the Edukalibre project. However, Latex is widely used to produce scientific documents, especially those which involve writing complex mathematical equations.

Within the Edukalibre project, it was decided to include Latex as one of the allowed document formats. Until now only the conversion between Docbook XML and Latex is performed, using the "db2latex-xsl" stylesheet (db2latex Web Site, n.d.). If the original file is submitted to the system in Latex format, then only the output formats HTML, PDF and Postscript are generated. The Latex processor is in charge of producing the postscript and PDF files. The HTML version is generated by the program "latex2html" (latex2html web site, n.d.).

In the future the system may contemplate the conversion from Latex to Docbook, but this will certainly require the usage of a strict subset of Latex instructions.

4. Case Study

This case study focused on students from the Engineering Faculty of the University of the Oporto, who have little previous experience in collaborative writing (CW).

A group of four students from Informatics Engineering were asked to write a report about Artificial Intelligence. The content was split into three subjects. This task occurred from 2005-06-16 to 2005-06-22. The participants' average age was about twenty years.

4.1. Objectives

The aim of this study was to show that the adoption of a tool developed specifically to support collaborative writing stimulates and, consequently, improves the performance in the writing process. Moreover, it shows that the Wiki format may be used to write scientific documents.

4.2. Design

The collaborative writing process was supported by a Learning Content Management System (LCMS) web tool - Moodle. The document was written on-line and asynchronously through a Moodle module named WikiEduka, which consists of a Wiki text editor and an interface to the Edukalibre system.

As the Edukalibre system implements an SVN platform, it was possible to investigate which changes had been done by each user. The contributions of the participants were monitored by analysing the document versions. This analysis was done through the SVN Diff command, which shows what has been changed in each version.

The suggested writing strategy was Parallel Partitioning (Sharples et al., 1991), where each collaborator is responsible for a section. However, in this case, the others could always edit any section. The writing occurred in a distributed way, members could write anywhere (home, office, college laboratory, etc.). Since it was intended to be a hybrid experience, face-to-face meetings were also allowed.

The students had also been encouraged to make annotations to the document. These annotations gave rise to discussion of ideas, resolution of doubts, and decision-making through mutual dialogue. That communication took place via the chat and forum modules provided by Moodle.

4.3. Methodology

A questionnaire based on Noël and Robert (2004), composed of twenty questions aiming to gain the participants' previous experiences, had been previously presented. It confirmed that the participants were people with little experience in collaborative writing and showed what they expected from a collaborative writing tool. The details enclosed in the questionnaire had been: technologies previously used, subject complexity, work plan, leadership, communication, relationship, satisfaction, writing strategies, problems during project, ideas about ideal collaborative tool.

Productivity was evaluated through qualitative analysis of synchronous and asynchronous communication (chat and forum), and post quantitative/qualitative analysis of member contributions. The subjects measured were:

- Chat and Forum content (qualitative);
- Number of interventions (quantitative);
- Type of interventions (qualitative - quantitative);
- Interventions by subject;
- Document extension; total number of words (quantitative).

The type of intervention was determined through its summary. It was possible to distinguish among insertion, correction and cosmetic interventions. For changes without summary, when the number of new words was smaller than the average word number in insertion interventions, it was assumed that it was a correction intervention. Interventions with less than 4 words were considered cosmetic changes.

4.4. Analysis

The answers to the questionnaire showed that all participants had just little prior experiences with collaborative writing. In average, they had written two documents. None of them had used specific tools for collaborative writing, two of the students had used Microsoft Word, one had used Open Office and one had used Lyx (Latex). All of them pointed out that the communication and text grouping had been done via e-mail.

Analyzing the forum and chat contents we realized that there was little redundancy of ideas and that immediately after an item had reached consensus the document was modified.

Figure 8 shows that the number of words grew abruptly by the beginning of the period. This occurred because the first contributions were mainly new text insertions. After that, the number of words grew continuously during the rest of the testing period. It can be concluded that the participants developed a constant effort during the rest of the time.

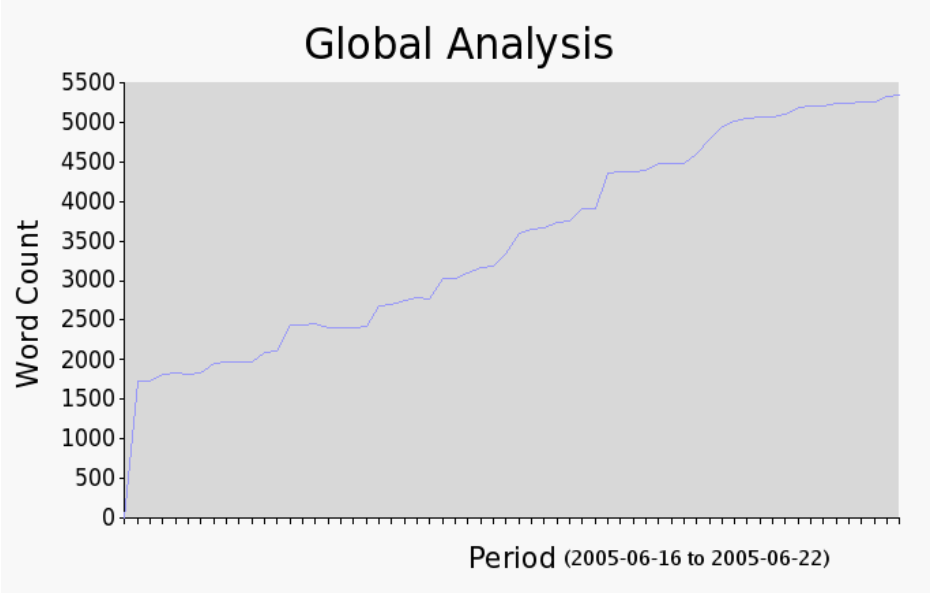

Figure 8. Global analysis

Figure 9 shows the evolution of the document by subject. Subject 1 had a beginning marked by new text insertions, but it was only improved after subject 2 had been sufficeiently developed. New text insertions in subject 3 occurred at same time as subject 2 development, but corrections occurred at a moment of less activity in subject 2. The students focused their work on one subject at a time. This observation is probably explained by the fact that this was a reduced group.

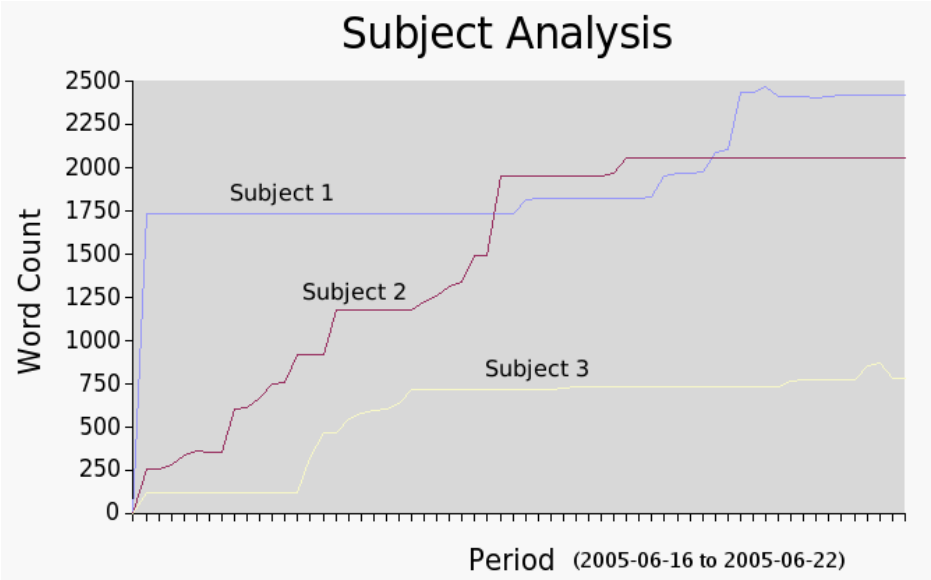

Figure 9. Subject analysis

Analyzing the nature of the interventions by each participant, it is possible to tell who preferred insertion of new text and who dedicated more time to corrections. Moreover it was possible to infer the amount of work dedicated by each participant (figure 10).

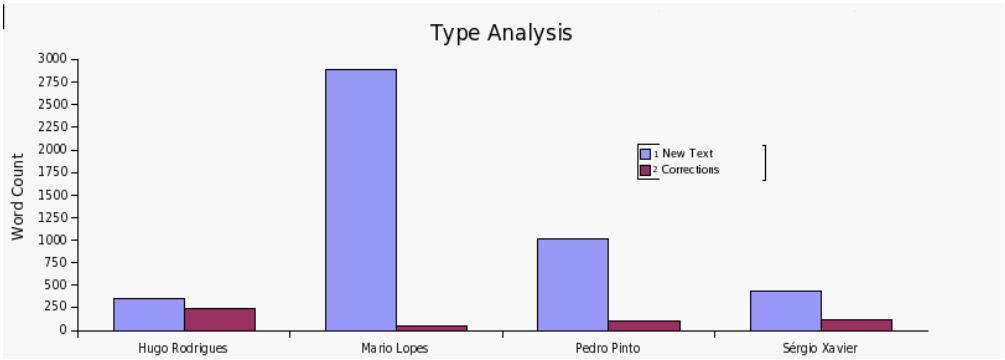

Figure 10. Type analysis

5. Conclusion

One of the most visible outputs coming out of the Edukalibre project is the Edukalibre System. This set of tools makes it possible for authors to collaborate over the Internet in the process of producing educational documentary materials.

Following the idea that led to the project in the first place, the choice of its underlying technologies took into consideration several criteria which came directly from the free/libre software philosophy. One of those guidelines consisted in using, as much as possible, tools previously developed and licensed as free software and also releasing all the tools developed within the project as free software. Moreover, in the spirit of the free software movement, it was decided from the beginning of the project that the authors should have the freedom to choose the typesetting tools they use. This led to the development of the file format conversion system described in this article.

The Docbook XML file format was chosen to be the main format from which all other formats are generated. This decision was taken, mainly due to the availability of several free software conversion tools, to and from this format. On the other hand, the Docbook format is oriented towards the document structure and lacks many of the formating concepts allowed by more advanced typesetting tools such as OpenOffice and Latex. In order to overcome this set back, authors are encouraged not to use formating information which is not included in the different templates developed for the Edukalibre system. Instead, the focus on the document structure should induce the production of standard formated material which is better suited for wide dissemination.

One of the possibilities offered to authors is the on-line editing of the documents in a Wiki format. The Wiki Markup format has the advantage of being simple and fast to learn. However, it does not implement all the formating options available in the Docbook XML format. In order not to lose formating information when converting documents from Docbook to Wiki, it was decided to pass into the Wiki document all those Docbook formatting tags that have no equivalent in the Wiki format.

Future work related to document format conversions should concentrate on three main tasks. On one hand, finding an alternative way to convert from the OpenOffice format to Docbook without the need to execute it at the server is essential to promote the reliability of the server which will run the system. On the other hand, the conversion between Latex and Docbook is still a task to be undertaken. Finally, a more complete set of case studies should be carried out. These will not only allow an evaluation of the system but also the evolution of its goals and strategies.

References

Collins-Sussman, B., Fitzpatrick W., B. & Pilato, C. M. (2004), "Version control with subversion", http://svnbook.red-bean.com/en/1.0/svn-book.pdf.

db2latex Web Site (n.d.), http://db2latex.sourceforge.net/.

Edukalibre Project (2004), "Edukalibre prject web page", http://www.edukalibre.org.

Godard, L. (n.d.), "OOoConv Web Page", http://oooconv.free.fr.

Gruber, J. (2004), "Markdown", http://daringfireball.net/projects/markdown/.

Jones, P. (2004), "Text wiki", http://wiki.ciaweb.net/yawiki/index.php?area=Text Wiki.

Just, H. (n.d.), "Writer 2 LaTex Web Page", http://www.hj-gym.dk/~hj/writer2latex/.

latex2html web site (n.d.), http://www.latex2html.org/.

Moodle Project (2004), "Moodle project web page", http://moodle.org.

Noël, S. & Robert, J. (2004), "Empirical Study on Collaborative Writing: What Do Co-authors Do, Use, and Like?", Comput. Supported Coop. Work 13, 1, 63-89.

Normal, W., Muellner, Leonard (2005), "Docbook 5.0: The definitive guide", http://www.docbook.org/tdg5/en/html/docbook.html

Sharples, M., J. Goodlet, E. Beck, C. Wood, S. Easterbrook, L. Plowman & W. Evans (1991), "A Framework for the Study of Computer Supported Collaborative Writing", School of Cognitive and Computing Sciences, University of Sussex. Research Paper CSRP 190.

Sun Microsystems (2002), "Open Office Org Web Page", http://www.openofice.org.

Tan, B. C. Y., Wei, K., & Lee-Partridge, J.E. (1999), “Effects of facilitation and leadership on meeting outcomes in a group support system environment”, Eur. J. Inform. Syst., vol. 8, no. 4, pp. 232–246.

The Apache XML Project (2004), "Fop (formating objects processor)", http://xml.apache.org/fop/.

The PHP Group (2004), "Pear - php extension and application repository", http://pear.php.net/.

Webdav web site (n.d.), http://www.webdav.org/.

Wikipedia (2005), "Wikipedia, the free encyclopedia", http://wikipedia.org.