

Proceedings of the  
**SPDECE-2012. Ninth multidisciplinary  
symposium on the design and evaluation of  
digital content for education**

13–15 June 2011  
Universidad de Alicante  
Alicante, Spain

*Edited by*  
Manuel Marco Such  
Pedro Pernías Peco



Copyright 2011 The authors

Printed in Alicante, Spain



Available online at [www.sciencedirect.com](http://www.sciencedirect.com)



Procedia - Social and Behavioral Sciences 00 (2011) 000–000

Procedia  
Social and  
Behavioral  
Sciences

[www.elsevier.com/locate/procedia](http://www.elsevier.com/locate/procedia)

Spdece 2012

## "Multiformat educational resources"

MP.Escobar<sup>a</sup>, G.Candela<sup>a</sup>, M.Marco<sup>a</sup>, P.Pernías<sup>a</sup>, T.Giner<sup>a,\*</sup>

*<sup>a</sup>University of Alicante, San vicente, 03690, Spain*

---

### Abstract

This communication will present the services built by the Biblioteca Virtual Miguel de Cervantes, from now on BVMC, which aims to provide the documentary resources, in rich text formats, with scientific quality, that teachers can use in their practice. Created in 1999, The BVMC is the largest provider of information about Latin American classic literature in digital format. A high percentage of its works are in a standard, interdisciplinary and internationally format: XML (TEI 2). The XML TEI2 allows generating different rich formats, such as PDF's, EPUB, etc., for scientific or teaching purposes. Based on these XML files, it is possible to generate educational material accessible from the web as open educational resources.

© 2011 Published by Elsevier Ltd. Selection and/or peer-review under responsibility of Spdece.

"Keywords: BPM;SOA;TEX;web;service;XML;"

---

---

\*  
*E-mail address:* [pescobar@dlsi.ua.es](mailto:pescobar@dlsi.ua.es), [gcandela@dlsi.ua.es](mailto:gcandela@dlsi.ua.es), [marco@dlsi.ua.es](mailto:marco@dlsi.ua.es), [ppernias@dlsi.ua.es](mailto:ppernias@dlsi.ua.es), [tginer@dlsi.ua.es](mailto:tginer@dlsi.ua.es)

## 1. Introduction

This communication aims to highlight the importance of digital material in a format accepted by the scientific and educational community. In addition, this format can be generated (on demand), in a simple, transparent and accessible way from anywhere.

A web services-based architecture is presented as a process that generates digital documents in different formats, reproducing the strictest standards required by the community.

As is shown in Figure 1, combining the richness of available resources provided by the BVMC and the benefits of the XMLTEI2 format, in which a high percentage of these resources are encoded, the process generates the different formats required by the user.

The reason why TEI XML 2 has been chosen as encoding format is because it has become an international and interdisciplinary standard that has enabled digital libraries and research centers to store valuable information about its contents.



Fig. 1. Global generation process

Currently, the existing marked formats, as XML or HTML, conversion projects are in a totally disconnected development phase. There are applications that are capable, within its limitations in complexity in the source file (XML or HTML), to generate rich formats such as PDF, but there is not a management as powerful as BPM to orchestrate the execution of connection-oriented services (web). The goal of this research is to generate valuable Web 2.0 oriented formats for teaching and scientific application, providing web services that generate intermediate formats (TeX).

### 1.1. Goals

The main objective of this research is to provide a method of generating digital, simple and easily accessible formats. Considering the rise of new technologies and the increasing use of Internet with Web

2.0, the process is based on a set of Web Services that can be managed independently or orchestrated by a process management tool BPM (Business Process Management) .

This article proposes a solution based on a service-oriented architecture (SOA), but without leaving the management of these services. To perform this task, a process using a tool BPM Open Source, BONITA, has been created which is capable of orchestrating the services, creating an organized and monitored workflow, whose final result is a document in a digital format valid to satisfy the accepted standards for scientific and educational community.

On the other hand, the generation of the formats is done using TEX as an intermediate generation format. TEX is a language developed to facilitate the writing of papers, articles and scientific and scholars books in the same way as a printer do. LaTeX, in turn, adds functionality and usability of TEX when writing technical papers of high quality typography.

Over the years, the Open Source community has evolved and improved standard technologies in all areas under the premise to be redistributed for free. Developing open source helps other developers to collaborate and facilitates cooperation with other virtual libraries. As expected, all this research has been done under Open Source license.

## 2. Methodology

Below is treated in a more detail the complete process to generate different digital formats from the BVMC catalog, based on three premises: TeX as an intermediate format in the generation, use of a service-oriented architecture (SOA) and management by BPM tools.

### 2.1. TeX as a intermediate format

LaTeX is a typesetting system specially oriented to the creation of books, scientific and technical documents that contain or may contain mathematical formulas, images, etc. This format has been used for its power and the simplicity of coding, as well as being a language widely used by scientific and academic community at the time of writing articles, thesis and technical books.

As discussed, SOA consists of a series of services that perform a series of operations. Thus, the process begins by requesting the information of a work in XMLTEI format through its unique identifier to the BVMC, where there are a large number of documents in that format.

### 2.2. Web services

- Getting a XMLTEI format representation (ObtenerXmlTei): the requested work is searched in the BVMC database getting its XML template in TEI2 format.
- Generating the HTML file from XMLTEI (Xml2html): a file in HTML format is generated meeting the academic and scientific standards with a stylesheet XSL format. It is important to ensure that the XML is properly trained to avoid possible inconsistencies and problems in further processing.
- Generating an intermediate format TeX , from the above representation (Html2tex): This is the most complex service of the system. Once the correct HTML created in the previous phase is available, a

service is executed to get the intermediate TEX format which can be interpreted by the LaTeX compiler later.

- Getting a device-independent representation (Tex2dvi): once the TEX file is created in the previous phase, a file with a richly formatted machine-independent (DVI) is produced by the LaTeX compiler, which creates the final format desired by the user (PDF, EPUB, RDF, ...) using Open Source tools.

The set of operations mentioned above are integrated into a service-oriented architecture (SOA) that defines the use of services to support business requirements. This architecture facilitates exposure and invocation of services by letting the interaction between systems of different types.

### 2.3. BPM formats generation process

BPM is presented as a simple and powerful way to be able to orchestrate web services in order to obtain the desired result. Any user is capable to use the BPM process directly or invoking the web services independently [6].

This process generates a proper and optimized workflow for processing enriched formats from the original XMLTEI as seen in Figure 2.

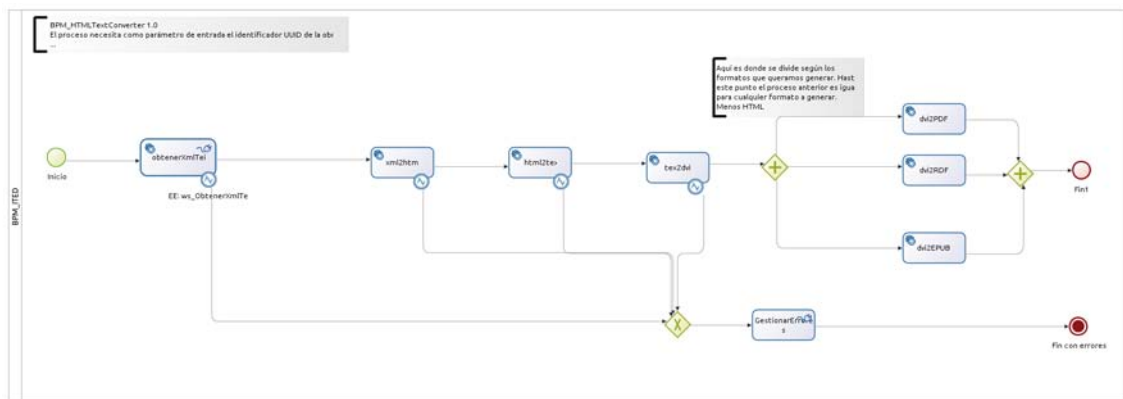


Fig. 2. BPM diagram

The process begins by requesting the information of the particular document to the BVMC repository. As seen in Figure 3, a web service is invoked, using the document identifier, to return the XMLTEI requested file, accessing the BVMC repository.

The way to access and view the repository through this web service is completely transparent to the user.

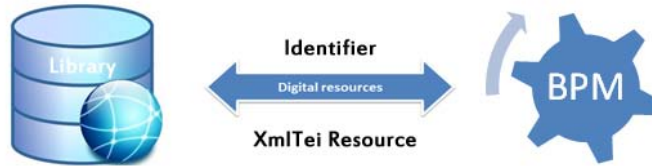


Fig. 3. Requesting the document

Once the XMLTEI file is provided, the next step in the workflow is executed. It must be remembered that the generation of enriched formats is thanks to the BVMC who has a lot of his works in XMLTEI format. This process is meaningless if the work does not have this marking, and would be subject to further investigation. The workflow checks the validity of the information obtained to ensure that the document is not malformed. To perform this task, the process uses a Document Type Definition or DTD that verifies and validates the XML file contents, as shown in Figure 4. In summary, there are a set of fields such as title, author or content that must be found.



Fig. 4. Verification of the document

Continuing with the process, the next step is to convert this file to a rich HTML format of final output which has been reviewed, agreed and accepted by the scientific community. This first transformation is carried out invoking the Web service xml2html, whose description was mentioned in the previous section.

Once the correct HTML is created, the intermediate format TEX is generated invoking the Web service html2tex. As mentioned in the previous section, the service is responsible for creating the .TEX file, receiving as a parameter an HTML file. Basically the web service explores the contents of the HTML file translating into its equivalent TeX. The process is described in Figure 5.

At this point the process, a document with logical markup is created and used in the following steps to generate more formats, without change.

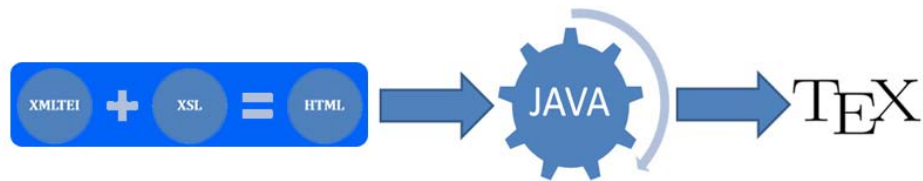


Fig. 5.Verification of the document

The TeX format result contains the information needed to view it on any device. Specifically, the next step is to obtain a device-independent format (DVI) using the LaTeX compiler, containing the text already prepared and the instructions for viewing the document in a device. The aim is to exploit the power and strength of LaTeX generating many rich formats requested by the user that can be used from any environment.

The last phase of this process is to generate precisely the final format requested by the user as shown in Figure 6. One of the formats to be generated, using TeX as an intermediate format, is PDF that is clearly the most common and compatible format for viewing documents. Each format will be generated by a Web service.

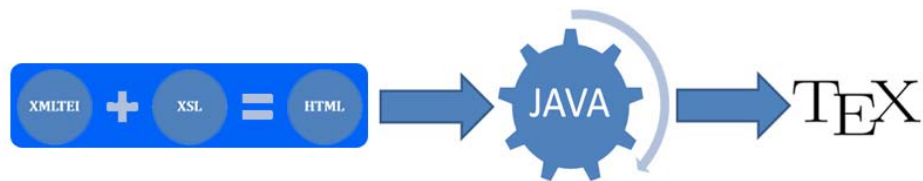


Fig. 6.Generated formats



### 3. Results

After this investigation, a series of documents have been obtained in various digital enriched formats. As an example, the following extracts from documents available in BVMC's catalogue are presented after being transformed by the process described above:

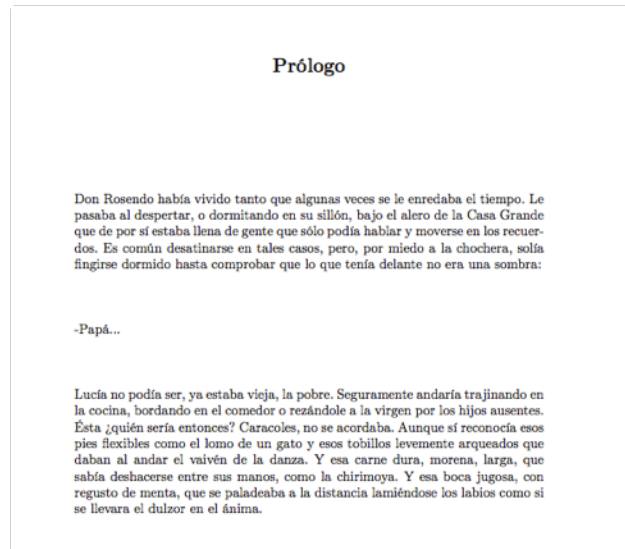


Fig. 7. Several paragraphs and section title

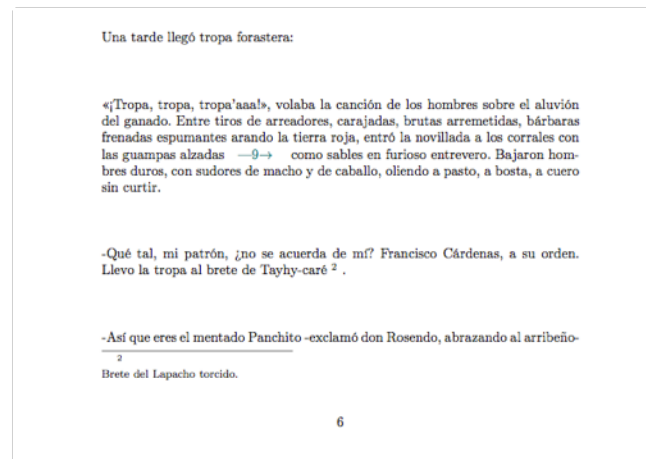


Fig. 8. Footnote

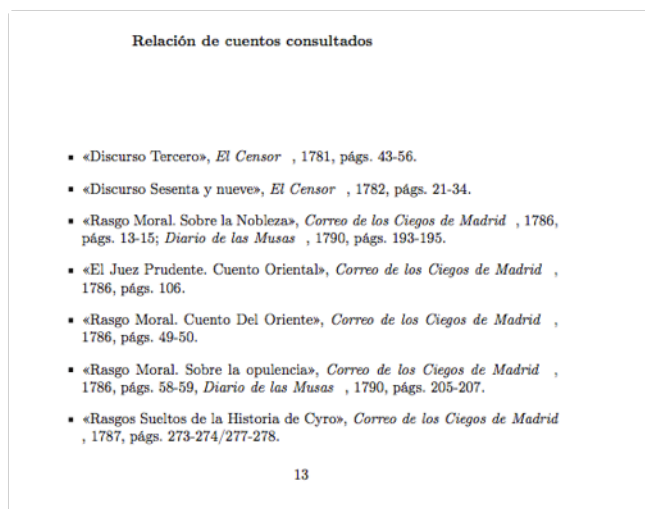


Fig. 9. Item lists

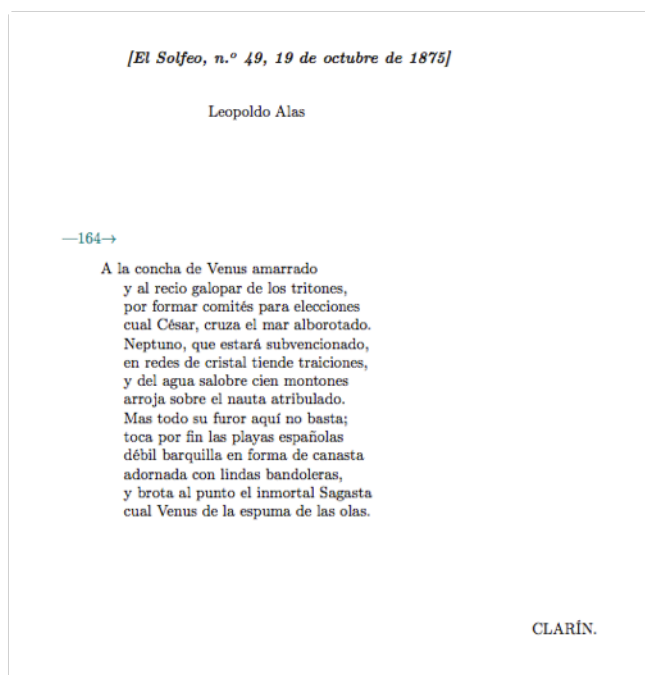


Fig. 10. Verse example

#### 4. Conclusions

This research has highlighted the need for a powerful intermediate format to generate others formats, while maintaining the standards required by the scientific and academic community. The Web Service flexibility and the TeX format chosen for this work, allows us to integrate the process in nay tool or environment, such as BPM tools.

Work continues on this line of research, improving the generation of the different formats and moving towards its use in semantic networks or Linked data.

#### References

A. Diller. LaTeX Line by Line: Tips and Techniques for Document Processing. John Wiley & Sons, 1999

S. Fear. Publication quality tables in LaTeX, 2005. <http://www.ctan.org/tex-archive/macros/latex/contrib/booktabs/booktabs.pdf>

Text Encoding Initiative (TEI). <http://www.tei-c.org/index.xml>

HyperText Markup Language (HTML). <http://www.w3.org/html>

BVMC, (1999) Biblioteca Virtual Miguel de Cervantes. <http://www.cervantesvirtual.com>

Marco Such, Manuel, Iriarte Navarro, Leonel, Pernías Peco, Pedro, Escobar Esteban, M<sup>a</sup> Pilar (2011). Servicios Web con utilidad educativa en Bibliotecas Digitales para PLE y LPM, SPDECE 2011.

Ayodeji Adesina, Derek Molloy (2010). Capturing and Monitoring of Learning Process Through a Business Process Management (BPM) Framework

Perkonigg, Fidelis (2006). Application of Business Process Management in Web-Based Education

Bonita Software. [www.bonitasoft.org](http://www.bonitasoft.org)

Marco M., Pernías Peco, P., Carrasco, Rafael.: Ámbito Tecnológico. Plan Estratégico Biblioteca Virtual Miguel de Cervantes 2006-2011. <http://hdl.handle.net/10045/16803>

Donald E. Knuth (2005). DVltype. <http://www.ctan.org/pkg/dvitype>