Curriculum@UA - online xml based personal curriculum

C. Teixeira, J. S. Pinto, J. Santos

IEETA – Instituto de Engenharia Electrónica e Telemática de Aveiro
Universidade de Aveiro, Portugal
{claudio, jsp, johnny }@ieeta.pt

Abstract. This paper describes a project in development at Universidade de Aveiro concerning online filling and central storage of researchers' curriculum vitae (CV). This system enables anyone to have an up to date CV always available for use anywhere. The major features in this site are its use of XML as the storage format for each individual CV and PDF CV generation.

1. Introduction

Giving its flexible structure, XML [1] has been used not only as a standard for data interchange, but also for data storage. XML databases are being studied [2] in order to understand its advantages and limitations when matched against relational database systems. The choice to use XML technology instead of relational database is related with the need to present the same information in several ways. XML seems to be a more flexible approach to deal with very wide - but small - information sets. Each CV may have several sets of information, but each person will not have many elements in each set, and some of them may even be empty.

Online CVs are a good idea, but if the CV is only for use inside the institution where it was filled, with no possibility of being used in other organizations, then those users tend to insert their CV only when they need it, and after that, there is no updating of the their information (why would they?). We propose an online CV system whose goal is the standardization of CVs at Universidade de Aveiro, but that users can use to present their CV anywhere. With this system, users are invited to maintain an updated CV, since they may also benefit from this.

This paper describes an online tool that uses XML to store CV's according to international standards described later on this paper. The remainder of this paper is organized as follows: Section 2 presents related work, Section 3 presents the data model used in the project and Section 4 explains the overall architecture. Finally, experimental results, conclusions and future work are presented in Section 5.

2. Related Work

The majority of online services that allow users to fill their CVs are paid services. [3] lists several examples of such services available in the UK.

There are some research organizations and public institutions that also use online CV's: two of these are the Brazilian *Sistema CV-Lattes* [4] and the Portuguese *Fundação para a Ciência e Tecnologia* (FCT) [6].).

CV-Lattes was developed by CNPq [5] in 1999, and is used by several Brazilian institutions and organizations, including the Brazilian Ministry of Science and Technology. In order to scholars and researchers to apply to appropriate funding, they are obliged to have an updated CV in this system. With this policy since 2002, the CV-Lattes has grown to over 460.000 CVs.

Portuguese FCT [6] promotes national scientific research and technological development by sponsoring scholarships, projects and scientific research institutions. Again, in order to apply to these funding, online CVs must be filled into the system by researchers and their coordinators. There is no public information about the amount of CVs that are in the system at this moment, but as in CNPq, they are very important in the evaluation phase of funding applications.

3. Architecture

Fig. 1 represents the functional layers present in Curriculum@UA. The top level is where the user interaction takes place.

At the bottom of the system there are three store systems: XML, XSL-FO [8] and PDF [7]. Each one is responsible for storing a specific set of files: "XML Store" keeps the XML CVs, "XSL-FO Store" keeps the transformation rules for presenting a PDF of the CV and finally, "PDF Store" keeps all the generated PDFs.

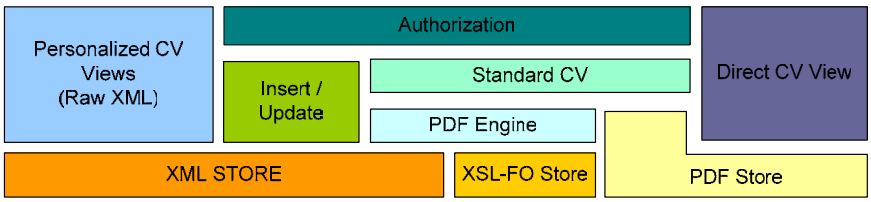

Fig. 1. Application Functional Layers

The "Authorization" module is responsible for validating the user and for granting access to the "Insert / Update" and "Standard CV (generator)" modules. "Insert / Update" module interacts directly with the "XML Store" to perform the desired operations. The "Standard CV" module creates the PDF CV at the "PDF Engine" and stores it in the "PDF Store" module. "PDF Engine" selects the XML CV and the XSL-FO to use to transform the XML.

As shown in Fig. 1, there are three main user interaction modules with the system: "Authorization" (already described), "Personalized CV Views" and "Direct CV

View". "Direct CV View" grants direct access to the "PDF store" (by using a specific URL), allowing universal access to the stored CVs. This feature can save paper and unnecessary printouts of CVs, if the users can give an URL (for a given application, job interview …) instead of a paper copy. The "Personalized CV View" allows everyone to access to the XML CV file. Since this application is built using public XSDs [12], anyone can use the updated XML to produce their own visual presentation for a given CV. This feature can be used by expert users to link a customized presentation of their always updated CV in their home pages. Organizations using HR-XML Resume schema can also use this to look at these CVs as if they were looking at "*house produced*" CVs.

Fig. 2, Fig. 3 and Fig. 4 represent different operations that can be performed in the system and the workflow involved.

Fig. 2 represents a user that needs his CV to file for application. The process starts whit the user's request of a PDF version of his CV to the system's Web Server. The system fetches it from the "PDF Store" and sends back both file and URL to the user. At this point, the user can choose between two scenarios depending on the application bureaucracy: either he prints out a copy and delivers it, or he simply delivers the URL for the services to see.

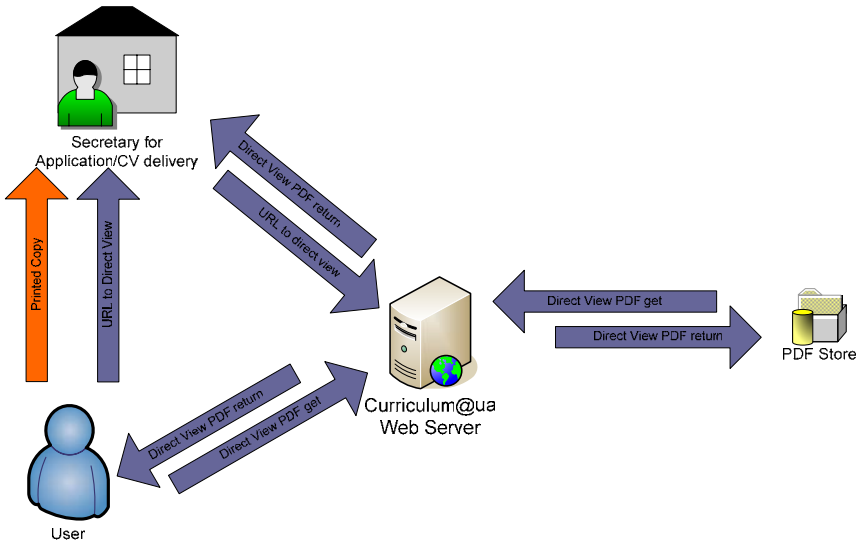

Fig. 2. Message flow for a PDF version of a CV to be shown

Fig. 3 represents the internal processing needed to complete an order of creating a PDF version of the user's CV. The messages flow as follows: the user requests the PDF generation; then the Web Server verifies the user's authorization on the "Authentication Server". If the user is allowed to generate the CV, then the Web Server requires from the "XSL-FO PDF Transformer" (XSL-FOT) the generation of the PDF file. The XSL-FOT gathers the required information - the CV in XML from the "XML Store" and the XSL-FO rules' file from the "XSL-FO Store" - and generates

the PDF CV and stores it in the PDF Store. Finally, the PDF Store sends both URL and PDF to the Web Server as referred before. For the sake of completeness, there must be made two remarks about this process: first, the PDF generation is a very expensive operation for the CPU. Therefore, it is only carried out at the user's request and not by a Direct View request. Second remark: as mentioned before, the goal of this system is standardization in information storage and presentation. This may seem inconsistent with the "XSL-FO Store" available, however, at the moment this Store has one single presentation format for CVs.

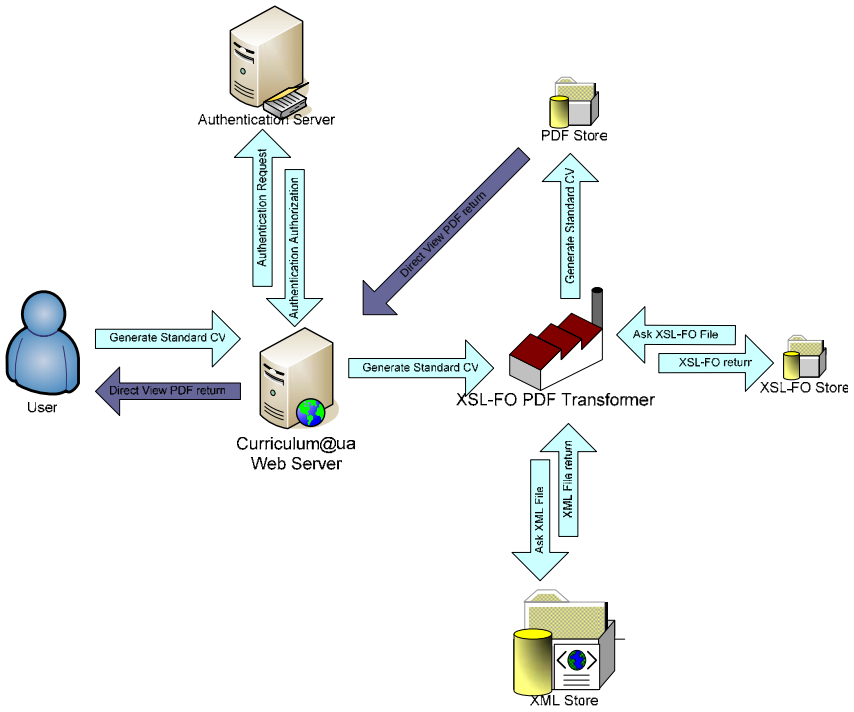

Fig. 3. Message flow to generate a PDF version of a CV

Fig. 4 represents an "Insert/Update" operation. The operation starts with the user requesting the operation to the "Web Server". Again, the "Web Server" verifies the user's authorization on the "Authentication Server". If the user is allowed to perform such operation, then the "Web Server" requests the XML CV from the "XML Store", makes the necessary changes and finally returns the updated XML CV to the Store.

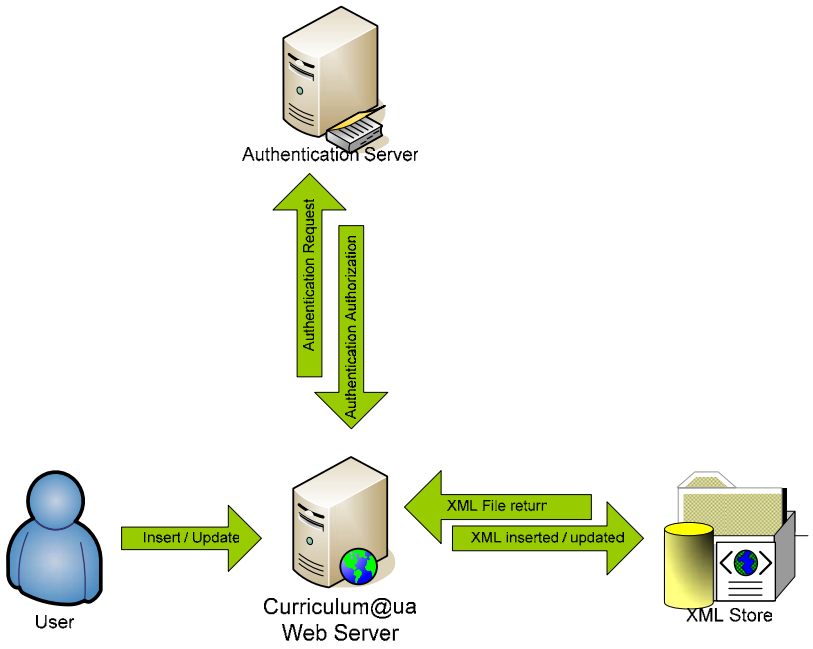

Fig. 4. Message flow for an insert/update to a given CV

4. Data Model

One of the major tasks on this project was the definition of which CV model to use. At the moment, there is no standard that states which information must or must not be placed in a CV or even in which place in the CV. After the analysis of three CV models (European [11], CNPq and HR-XML), the choice felt on the HR-XML Resume. It's a well documented model (that modelled more than enough information for our purposes) and is supported by a non governmental, international organization. Fig. 5 shows the XML-HR Resume schema adopted.

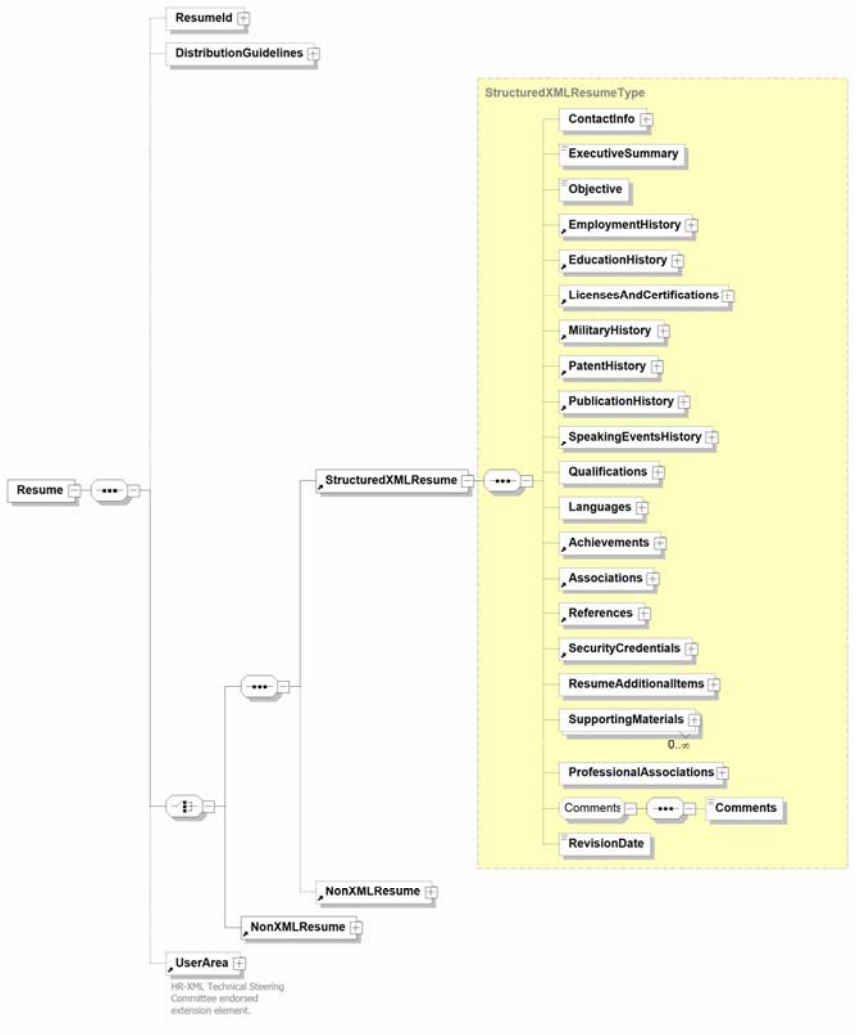

Fig. 5. Full XSD for Resume, as specified by HR-XML

As shown, the Resume is a very comprehensive list of structured information. For logistic reasons, not all this schema has been implemented. Only the subset that seemed to be fundamental for researchers and scholarship applications was implemented. This includes Contact Info, Executive Summary, Objective, Employment History, Licenses and Certifications, Patent History and Publication History.

5. Conclusions

As mentioned, this is still an on going project, but has already been used for PhD scholarship application held at Universidade de Aveiro during last summer. As not all of the possible candidates could understand Portuguese, the site developed is multilingual. XML technology was also used to implement such feature. This application also required two more features: an online work plan and a way to finalize each user's process. These features were developed according to the system's architecture and are also XML based.

There were about fifty candidates applying for scholarships and they had to register, fill in their CV and work plans in the system, print them, finalize the process and delivery the documents in the academic services in charge of the scholarship application.

An online management interface that enabled administrators to monitor the documents in the system was also developed. After the application period expired, the administrators could email a selected collection of PDFs to designated juries to evaluate the candidates.

According to the people in charge of the scholarship applications, this project has saved time in the evaluation phase, since the document layout was always the same, and every subject was exactly in the same place for all the candidates, enabling a quicker cross-checking analysis and evaluation. Therefore, the standardization has proven to be profitable.

From the user's point of view, he can continue to keep an updated record of his CV on the site, in order to use it whenever and wherever he wants to.

The results achieved by the scholarship application have encouraged the development team to increase the supported subset of information available.

With the increase of interoperability opportunities among applications developed at Universidade de Aveiro, this project will focus some attention in the development of interfaces to enable the automatic update of bibliographic references by existing specialized services, in a seamlessly way for the user.

As the HR-XML has a certification site and logo [10], tests are being made in order to make this site eligible for the certification logo.

References

[1] Bray, T., et al (editors), "Extensible Markup Language (XML) 1.0 (Second Edition)", W3C Recommendation 6 October 2000, W3C, online, last accessed in 10th November 2004, http://www.w3.org/TR/2000/REC-xml-20001006

[2] K. M. Win, W. K. Ng, and E. P. Lim. An Architectural Framework for Native XML Data Management. 2nd International Conference on Cyberworlds (CW 2003), 3-5 December 2003, Singapore. IEEE Computer Society 2003, 302-209 pp, ISBN 0-7695-1922-9

[3] "All About CVs", Regard-It, online, last accessed 16th November 2004, http://www.regard-it.com/cv.htm

[4] "CNPq – Plataforma Lattes", Conselho Nacional de Desenvolvimento Científico e Tecnológico - Brazil, online, last accessed in 16th November 2004, http://lattes.cnpq.br/curriculo/

[5] "CNPq - Conselho Nacional de Desenvolvimento Científico e Tecnológico", Conselho Nacional de Desenvolvimento Científico e Tecnológico, online, last accessed in 16th November 2004, http://www.cnpq.br/english/aboutcnpq/index.htm

[6] "Home Page FCT", Fundação para a Ciência e Tecnologia - Ministério da Ciência e Ensino Superior, online, last accessed in 16th November 2004, http://www.fct.mces.pt/

[7] "Adobe PDF: Secure, reliable electronic document distribution and exchange", Adobe Systems Incorporated, online, last accessed in 17th November 2004, http://www.adobe.com/products/acrobat/adobepdf.html

[8] "Formatting Objects", In "Extensible Stylesheet Language (XSL) – Version 1.0 – W3C Recommendation 15 October 2001", W3C, online, last accessed in 17th November 2004, http://www.w3.org/TR/xsl/slice6.html#fo-section

[9] "hr-xml.org -- Home of the HR-XML Consortium, Inc.", HR-XML Consortium Inc., online, last accessed in 28th October 2004, http://www.hr-xml.org/channels/home.htm

[10] "HRcertify.org: A Service of the HR-XML Consortium, Inc.", HR-XML Consortium Inc, online, last accessed in 17th November 2004, http://www.hrcertify.org/

[11] "CEDEFOP: The European Centre for the Development of Vocational Training - Cedefop in Brief", CEDEFOP, online, last accessed in 17th November 2004, http://www.cedefop.eu.int/transparency/cv.asp

[12] Priscilla Walmsley, "Definitive XML Schema", Prentice Hall PTR, 07 December, 2001. 560 pages. ISBN: 0130655678