Maturing towards the digital library: Implementation of the electronic thesis and dissertations

Victor M. Ferracutti
Jefe de Sistemas, Biblioteca Central, Universidad Nacional del Sur, Bahía Blanca, Argentina

Luis A. Herrera
Director, Biblioteca Central, Universidad Nacional del Sur, Bahía Blanca, Argentina

Natalia I. Piccotto
Directora, Servicio de Información y Documentación, Planta Piloto de Ingeniería Química, CONICET, Argentina

ABSTRACT

Thesis and dissertations constitute the foundation and evidence of the ingenuity and scientific production in universities. In the latinamerican region various actions have taken place attempting to bring into a proper perspective, aided by current technologies, all of the efforts put into years of research, teaching, and human resources development at universities and research institutions.

The implementation of Open Source software for the development of digital libraries of thesis and dissertations is recent in the region, and particularly in Argentina. Current technologies favor integration and cooperation but national and regional policies are in the process of discussion and definition while there is a current of Governmental regulations which tend to support the Open Source paradigm as a valid alternative for continued improvement and sustainable growth.

In this context, the administration of the "software process" has become critical. As a result, the Capability Maturity Model (CMM) was adopted.

At the beggining of 2001 the Biblioteca Central of the Universidad Nacional del Sur started to use web technologies to provide improved access to the various information resources in higher education. The first electronic thesis were published using the TEDE Open Source software furnished by IBICT (Instituto Brasileiro de Información en Ciencia y Tecnología), receiving training and support from the ETD-Net project by UNESCO.

During 2003, other institutions of the public sector in the research and teaching area, joined the Universidad in the project of strengthening and giving renewed impulse to the electronic thesis and dissertations project.

We present here the preliminary results of the project in this set of critical success factors

- interinstitutional cooperation
- Standars adoption
- Attitudes of users co mmunity
- Software tools and processes.

1. INTRODUCTION

The Universidad Nacional del Sur (UNS) has a Biblioteca Central (BC) (Main Library) and eight special libraries in the subjects of Management, Agronomics, Computer Science, Law, Economics, Geography, Humanities and Mathematics. One of the functions of the Biblioteca Central is to preserve and give access to the thesis and dissertations and the memory of the academic production of the institution.

In this context, the Systems area of BC participated of the III Course of Training of Directors of ETD-Net projects, which was held at the Catholic University of Pernambuco, Recife, Brazil, and was sponsored by UNESCO, on April 9 -11, 2001.

The preliminary Project for the development for the Digital Library of Thesis and Dissertations came about as a result of the training activity. This was presented to the University`s authorities and other library people in June of 2001. Since then, a number of actions related to the digital collection of Thesis began at the University.

During 2002, the BC purchased hardware destined to the ETD project and began the process of software selection. By then, contacts had already been established with BIREME, due to the implementation of a Scielo Project with one of the Journals published by the University. All of these actions were giving form of our Digital Library Project.

In 2003, the BC purchased a dedicated web server and participated in the workshop sponsored by UNESCO through its ETD-Net program and organized by Instituto Brasilero de Información en Ciencia y Tecnología (IBICT) and was held between September 22 and 24, 2003 in Brasilia, Brazil. That particular activity was aimed to the transfer of technology and methodology for the implementation of the thesis and dissertations library.

Since end of 2003 the BC has advanced in the implementation and adaptations of the software. This was worked out jointly with technical people from CRIBABB, (Center for Regional Basic and Applied Research), which is funded by CONICET. This was given the status of test bed, and had the permanent assistance of technical people from IBICT.

In March of 2004 the BC team started to run a pilot test in association with the Servicio de Información y Documentación de PLAPIQUI, one of the oldest and well respected chemical engineering basic and applied research facility in Argentina. Most of the master´s and PhD thesis and dissertations from the UNS come from the people in PLAPIQUI.

The objectives of this pilot project were to inform and interest thesis candidates and directors about the overall concept of a digital library and institutional repository of thesis and other documents, to test the software TEDE with a greater volume of information, to develop a guide for authors in Spanish and survey new functional requirements for TEDE.

Various legal aspects of the copyright in the digital environment were compiled. The creation of interdisciplinary teams is included among the alluded goals. A special paragraph should be devoted to this aspect, since this is probably the best indication of continuing development for this project.

Human resources involved in the Project include people from the Systems Area, which has a good reputation and experience in advanced programming languages for searching and setting up data bases for internet access. The collaborative work realized together with the Systems people of CRIBABB, which covered the configuration of the new UNIX/LINUX servers, resulted in a sustainable framework for future development. The people at PLAPIQUI provided the data entry and managed the interaction with authors, running surveys about generalized use of the product and answering all the pertinent questions users may have.

Other institutions in the country have initiated actions related to the adoption of a ETD system. The lack of participation of their technical staff in the workshops sponsored by UNESCO has produced much confusion and diverse attitudes which tended to postpone the adoption of any system.

In the case of the UNS, participation in various other projects related to digital libraries such as Scielo Argentina and Alfa-Revistas have helped in the decision making processes involved.

Through these experiences, it is pertinent to observe that the role of UNESCO in the region is absolutely necessary to produce any significant result. The technical expertise is gained in the context of the workshops, while the sluggish and cumbersome -at times- process of adoption of the Open Source Software (OSS) concepts, may prevent a smoother and safe systems development process. UNESCO may perhaps call on the higher levels of university government by providing sensitizing workshops regarding recommended information policies in the academic settings, as well as giving support to the generation of policies which have to do with institutional repositories and digital libraries management infrastructure of equipment and human resources. Secondly, UNESCO should play an active role in favoring the creation of strong technical teams which may receive help in specific training only by showing evidence of a strong commitment to do collaborative work for the development and adoption of OSS.

2. METHODS

This paper contains applied research information, based on document studies and field research.

In the face of the definition of the Digital Library as an institutional objective, the BC identified a standard of software development to obtain guidance in the definition of all the characteristics of the process which allowed for evolution from an immature stage (unrepeatable, unreliable) towards well-managed software and to a mature stage. (Paulk, et al. (**a**),1993)

Institutional capabilities and the cumulative experience of the Capability Maturity Model (CMM) version 1.1 provided the appropriate software development and adaptation framework for this first stage. The CMM model has recently evolved to CMMI (Capability Maturity Model Integration) and it will be applied for the future stages.

In summary, CMM allows for evaluation of the stages of implementing any system. This, together with the "Agile Modeling" (http://www.agilemodeling.com) as a support methodology in software development constituted the initial tools.

Another valuable tool in this pilot test has been the semi-structured interview with authors of masters and doctoral thesis and dissertations in PLAPIQUI were surveyed using this technique.

Table 1 Summary of used techniques and methods

Table 1. Methods

Method/Technique	Procedure
Literature review	
Semi-structured interview	Universe: researches at the Planta Piloto de Ingeniería Química, authors of graduate thesis Population: researches and research associates at the Planta Piloto de Ingeniería Química, authors which completed the MS or PhD requirements of graduate thesis between 1997-2005.
Selection of the software	Basic Criteria: Open Source Software Work flow analysis and complience Technical support (staff and external)
Inter-institutional Cooperation	Distribution of information through technical meetings Attendance to scientific meetings and seminars Consulting Participation in cooperative projects.

3. RESULTS

Inter-institutional cooperation and participation y various events promoted by UNESCO in the ETD-Net Project have given the conceptual framework to initiate the development of the Digital Library starting out with the digitization of thesis and dissertations.

Interaction with NDLTD, IBICT (Brazil) and professor Edward Fox, through a visit he paid to Argentina sponsored by UNESCO, have helped to:

- Define the work flow and schedule;
- Requirements analysis for the digitization of thesis and dissertations
- Modelling above requirements;
- Select the Open Source Software (http://tede.ibict.br/) which provided the required functionality, complying with the quality variables which are present in commercial softwares.

Interviewing authors allowed us to sensitize the academic community about the importance of the project and also to compile the necessary information to test the correct functioning of the software within the institutional workflow. This activity was led by people at PLAPIQUI (a chemical engineering research institute at UNS with a long tradition for graduate studies and research and international recognition). This is why the first digital thesis and dissertation documents originated an this institution

The Capability Maturity Model for Software provides software organizations with guidance on how to gain control of their processes for developing and maintaining software and how to evolve toward a culture of software engineering and management excellence. The CMM was designed to guide software organizations in selecting process improvement strategies by determining current process maturity and identifying the few issues most critical to software quality and process improvement (Paulk et al. **(b),**1993).

CMM 1.1 was adopted early in 2004 using ETD as a pilot Project. At this time, UNS was at level 1 in reference to the CM Model for quality. In measuring the acquired capabilities through the actual implementation of this pilot project we observe that the majority of the key practices are complied with, defining a step forward to level 2 in the model.

The set of key practices for Level 2 are (Paulk et al.**(a)** 1993):
- Requirements Management
- Software Project Planning
- Software Project Tracking and Oversight
- Software Subcontract Management
- Software Quality Assurance
- Software Configuration Management

Documents on Structured Analysis and Sommerville`s Software Engineering 1992, provide support for first three key practices.

Adoption of a software developed by IBICT and selected on the basis of quality heuristics for software, guarantee the product quality. It remains as a challenge to complete and receive a final approval of the regulations which sustain and support organizational control and audits.

There is a proposal which contemplates the possibility of a Grupo de Estudio e Investigación en Sistemas de Información (Study and Research Group on Information Systems) from the Computer Science and Engineering Department of UNS, be the auditing body and may monitor the QC routines of the software developments of the Digital Library. During the second half of 2005, this initiative may become a reality.

The last part of key practices Software Configuration Management remains as an objective.

The concept of Digital Library is strongly attached to the idea of collaboration and intra- and inter-institutional cooperation. Complying with CMM facilitates cooperation and interaction with other institutions, providing a common way for administering systems and a common language throughout all the pertinent developments.

4. DISCUSSION

In our context we perceive that we lack the experience to help in the decisions related on how to best make the rest of our systems work together. MARC21 was recently adopted basically on the grounds that institutionally, the UNS should make a strong case of adoption of an international standard for bibliographic description and records deployment and sharing with other institutions.

We then understand that the various considerations related to the creation of catalogs of international, national and regional reach is directly related to the adoption of standards.

We studied and accepted the 5S Model (Gonçalvez et al.) for the development of digital libraries as an essential tool to help step up into level 3 of CMM. This has been the theoretical model for digital libraries which integrate the various digital products in the following concepts:

- Societies   (Sociedades)
- Scenarios   (Escenarios)
- Spaces   (Espacios)
- Structures   (Estructuras)
- Streams   (Cadenas)

With the 5S model of a digital library we may elaborate abstractions of the digital elements, such as digital objects, metadata, collections, services, and they can be treated rigorously and usefully described by compositions of basic and higher level mathematical objects.

This model provides a vision such that a Digital Library is not only defined as a collection of electronic products but it is in essence a service. In this sense, the user is perceived as the element which generates the demand. This model takes into consideration the fact that data bases can be described, modeled, designed, implemented, used and evaluated.

5. CONCLUSION

Projects such as ETD-Net are one of the many links in the chain for building a university Digital Library. It is natural to insert electronic thesis and dissertations in Universities, but it is not so natural to develop consciousness and policies about the adoption of Open Source Software.

It has become urgent for countries and universities in the Latin America region to manage their own production in science and technology. This has been for sometime an strategic position of the developed countries. Following this, there is also a need to have the different systems to interact and exchange data and information, to cooperate more actively in the development of common platform base, and in the management of data-warehousing which may help in the analysis and calculations of indicators of Science and Technologies.

6. ACKNOWLEDGMENTS

7. REFERENCES

Paulk, M. C.**(a)**, Weber, Ch. V., Garcia, S. M., Beth, M. & Bush, C. M. (1993). Key Practices of the Capability Maturity ModelSM, Version 1.1. Software Engineering Institute Technical Report, CMU/SEI-93-TR-025, ESC-TR-93-178. Retrieved March 10, 2005, from
http://www.sei.cmu.edu/pub/documents/93.reports/pdf/tr25.93.pdf

Paulk, M. C.,**(b)** Curtis, B., Chrissis, M. B. & Weber Ch. V. (1993). Capability Maturity Model for Software, Version 1.1. Software Engineering Institute Technical Report No. CMU/SEI-93-TR-024. Retrieved March 30, 2005, from
http://www.sei.cmu.edu/pub/documents/93.reports/pdf/tr24.93.pdf

Sommerville, I. (1992). Software Engineering. dissertation (4th ed.). Wokingham, England: Addison-Wesley

Gonçalvez, M. A., Fox, E. A., Watson, L.T. & Kipp, N. A. Streams, Structures, Spaces, Scenarios, Societies (5S): A Formal Model for Digital Libraries. Retrieved January 10, 2004, from
http://www.dlib.vt.edu/projects/5S-Model/5s6.pdf