

BDTD – The Brazilian National ETD Project

Ana M B Pavani

DEE, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, Brazil

Sueli A Maffia

BDTD, Instituto Brasileiro de Informação em Ciência e Tecnologia, Brasília, Brazil

Keywords: National projects, national metadata standard.

ABSTRACT

This paper addresses BDTD – Biblioteca Digital de Teses e Dissertações, the Brazilian national ETD project. It shortly presents the beginning of the activities and then the current challenges. The challenges are related to increasing the number of institutions with ETD projects, the quality of metadata to conform to the national standard, the support of funding agencies and continued growth of ETD activities in the institutions. The roles of IBICT – Instituto Brasileiro de Informação em Ciência e Tecnologia and of the institutions are stated.

1. INTRODUCTION

Institutional ETD projects started in Brazil between 1995 and 2000 and, by the end of 2000, 3 universities had consolidated projects. They were: PUC-Rio – Pontifícia Universidade Católica do Rio de Janeiro (“PUC-Rio”), UFSC – Universidade Federal de Santa Catarina (“UFSC”) and USP – Universidade de São Paulo (“USP”).

The institutions are very different in terms of organization – the first is private and belongs to the Roman Catholic Church, the second is public and belongs to the federal government, and the third is public and belongs to the government of the State of São Paulo. At the same time, they share the geography – they are located in the Southern and Southeastern regions of Brazil, which are the most developed and are responsible for the highest percentage of the GNP of the country.

IBICT – Instituto Brasileiro de Informação em Ciência e Tecnologia (“IBICT”) is an institution of the Ministry of Science and Technology. It is the legal deposit of all theses and dissertations presented by Brazilians who go abroad for graduate programs and receive financial support from a federal agency. IBICT is also responsible for the union catalog of referential records of Brazilian theses and dissertations.

In January 2001, IBICT took the leadership to start a national project. The 3 universities were invited to join the efforts, along with representatives from federal institutions and an international library, BIREME – Biblioteca Virtual em Saúde (“BIREME”), located in Brazil.

The national project was named BDTD – Biblioteca Digital de Teses e Dissertações and received a grant from the Ministry of Science and Technology. It had two main objectives. The first was to increase the number of institutions with ETDs and the second was to create a union catalog of metadata records. BDTD is managed by IBICT with the support of a technical committee, whose members are from the 3 universities, the international library and federal agencies responsible for graduate education and research.

ETD2005

The project has been going on since that date and some of its results and methodology have already been presented by Marcondes, Sayão, Triska and Pavani (2002) and by Pavani and Southwick (2004).

The following sections address results, problems, solutions and challenges.

2. ACTIVITIES IN 2001 – 2003

One of the first tasks of BDTD was to define a national metadata standard for ETDs that would: (1) allow a national union catalog to be created; (2) comply with international standards/practices; (3) yield information on published theses & dissertations x financing, geography, areas of knowledge, etc. The third objective was a result of the needs of federal funding agencies.

In 2001, a draft version of the metadata standard was established and in December the first union catalog was created with data from the 3 universities; XML files were sent to IBICT after adjusting local metadata sets to the draft version.

Other institutions were needed to make up a national project. To help them enter the ETD world, IBICT developed an open and free software solution; training sessions were held to start activities. The solution is called TEDE – Sistema de Publicação Eletrônica de Teses e Dissertações (Electronic Publishing System for Theses and Dissertations) and the first institutions received it in June 2003; currently, there are 8 TEDE users in Brazil.

TEDE was selected by ULA – Universidad de los Andes (“ULA”), in Venezuela, to be used locally. ULA translated it into Spanish and customized it to implement a digital library of senior projects too. The two Venezuelan customizations are available to other universities.

In 2002, v1 of MTD-Br – Padrão Brasileiro de Metadados de Teses e Dissertações (Brazilian Metadata Standard for Theses and Dissertations) was published as a result of the work of the technical committee lead by IBICT. At the same time, it was decided that the union catalog would be updated by metadata harvesting. Thus, OAI-PMH – Open Archives Initiative Protocol for Metadata Harvesting (“OAI”) was implemented. The first harvesting happened in Dec 2002.

In 2003, efforts to stimulate more universities to join the project lead to many training sessions. More institutions joined.

3. ACTIVITIES IN 2004 – 2005

TEDE was developed based on the Virginia Tech (“VT”) model – students participate in the various steps of publishing ETDs. The use of TEDE showed that an alternative model was needed to fulfill the needs of institutions which were not comfortable with student participation. This yielded a second free and open software solution – simplified TEDE. It keeps part of the functions but the cataloging and uploading of the works is done by the library. Simplified TEDE is already available.

After 2 years of operation, it was concluded that MTD-Br should be reviewed. In 2004, MTD-Br was reviewed and v2 was published in 2005.

The efforts to bring new institutions continued in 2004 – 2005, and currently there are 14 contributors to the BDTD union catalog. The total count is a little over 12 thousand metadata records.

ETD2005

4. PROBLEMS, SOLUTIONS AND CHALLENGES

A national project requires a lot of work to put together institutions with different local cultures, specially in a country that has very sharp differences in terms of development. Brazil has areas where human development, as defined by the United Nations Development Agency (“UNDP”), is similar to that of developed nations and other where it is very low.

One of the first problems to be faced was the change in culture of the universities – some seemed to believe that ETDs were not necessary, while others thought they were for “developed institutions”. The solution to this problem came with the intensive training program that started in 2001. This program was sponsored by UNESCO – Regional Office for Science in Latin America and the Caribbean (“UNESCO”), by IBICT and by some universities and librarians associations.

A second problem was the simultaneous existence of records on OPACs, for traditional T&D, and metadata for full-text ETDs. IBICT had both types of records and a first solution was to use one database to host them – the search could be in the whole set or the user could choose between ETDs or traditional works. This proved not to be satisfactory because of the huge difference between the 2 sets of records. At the moment, the solution under consideration is to have 2 separate sets and to use a federated search engine for the common fields, in case the user wishes to find works in either type of media. The development of this solution is planned for the second semester of 2005.

A third (and serious problem) is the quality of metadata for ETDs when records are generated from OPACs. In general, they lack translations of titles, abstracts and key words into foreign languages. Other data that are not standard in the MARC format are the discipline, the level, the title, the contributors, the information on funding and links to other repositories. So, the institutions that only want to use (only) OPACs have a hard time to adjust and their ETDs cannot be added to BDTD.

There is another problem related to OPACs, it is the lack of the data provider facility to allow harvesting by an OAI-PMH service provider. International OPACs have it, but local solutions do not.

In order to solve both problems, the BDTD team is trying to show how important it is to have a digital library associated with ETDs. It becomes, specially, easy because there are 2 TEDEs to choose from and both are complete in terms of national and international requirements. The resistance to the use of a digital library is one of the cultural aspects in traditional libraries in Brazil.

A problem that has also occurred is lack of continuity in the projects. Institutions start them but stop after some months. The solution to this problem lies on making the ETD project an institutional commitment.

A last problem is the lack of enthusiasm of faculty. Most advisors are not motivated to electronic publishing of works. At PUC-Rio, this problem has been addressed by developing access statistics. When the system offered access statistics, the number of both national and international visits to ETDs became public. At that moment, faculty started discussing ETDs.

At the national level, it is necessary to have the support of funding agencies. They can add ETD publishing clauses to the contracts with students. IBICT has held meetings with the agencies and ETDs are under consideration to become mandatory for students who receive grants.

The most obvious conclusion of these 4 and half years of work is that an ETD national project is an effort in many levels – from the universities to the national administrative group. It requires coordination, collaboration and integration. The goal is to achieve a union catalog of

ETD2005

metadata according to the standards and, if possible, with records from as many graduate programs as possible.

5. REFERENCES

BIREME – Biblioteca Virtual em Saúde: <http://www.bireme.org/>

IBICT – Instituto Brasileiro de Informação em Ciência e Tecnologia: <http://www.ibict.br/>

Marcondes, Carlos E., Sayão, Luiz F., Triska, Ricardo & Pavani, Ana M. B. (2002), Brazilian Electronic Theses and Dissertations Consortium, *ETD 2002 – Fifth International Symposium on Electronic Theses and Dissertations*, USA: <http://www.wvu.edu/~thesis/proceedings.html>

OAI – Open Archives Initiative: <http://www.openarchives.org/>

PUC-Rio – Pontifícia Universidade Católica do Rio de Janeiro: <http://www.puc-rio.br/>

Pavani, Ana M. B. & Southwick, Sílvia B. (2004), Some Practical Points on Organizing ETD Consortia, *ETD 2004 – Seventh International Symposium on Electronic Theses and Dissertations*, USA: <http://www.wvu.edu/~thesis/proceedings.html>

UFSC – Universidade Federal de Santa Catarina: <http://www.ufsc.br/>

ULA – Universidad de Los Andes: <http://www.ula.ve/>

UNDP – United Nations Development Program: <http://hdr.undp.org/hd/>

UNESCO – Oficina Regional de Ciencia para América Latina y el Caribe: <http://www.unesco.org.uy/>

USP – Universidade de São Paulo: <http://www.usp.br/>

VT – Virginia Institute of Technology and State University: <http://www.vt.edu/>