15 Years of ETDs – The Experience of Two Brazilian Institutions

Ana Maria Beltran Pavani
Sueli Mara Ferreira
15 Years of ETDs – The Experience of Two Brazilian Institutions

Ana Maria Beltran Pavani
Sueli Mara Ferreira

CREDITS

Publisher:
MAXWELL / LAMBDA/CCPA/VRAc
Sistema Maxwell / Laboratório de Automação de Museus, Bibliotecas Digitais e Arquivos
http://www.maxwell.vrac.puc-rio.br/

Organizers:
Alexandre Street de Aguiar
Delberis Araújo Lima

Cover:
Ana Cristina Costa Ribeiro
This article was presented at ETD 2015 India - 18th International Symposium on Electronic Theses and Dissertations, 04-6 November 2015. It was published in the proceedings of the symposium pp. 29-39, ISBN 978-93-84869-84-7. Jawarharlal Nehru University, Excel India Publishers; New Delhi India.
15 Years of ETDs—The Experience of Two Brazilian Institutions

Ana M.B. Pavani and Sueli Mara Ferreira

Abstract—Brazil has had a National Consortium of ETDs for almost 15 years. Currently, the Consortium Union catalogue hosts over 370K metadata records. When the Consortium was established, two institutions had had ETD programs for a year and supported the deployment of the Consortium. Pontifícia Universidade Católica do Rio de Janeiro, in Rio de Janeiro, and Universidade de São Paulo, in São Paulo, are these institutions. They had different backgrounds in digital publishing and also followed different tracks after working together to create the Consortium. This paper addresses the experience of the two institutions. It begins by giving a brief description of each and then tells the histories of the ETD programs. It continues by presenting the current status of the programs and their follow-ons. Some data related to the publishing of ETDs and also the accesses to the works are presented. The two institutions have solid ETD programs and also strong and varied production of other digital contents.

INTRODUCTION

The history of ETD—Electronic Theses and Dissertations in Brazil goes back to the second half of the 1990s when UFSC: Universidade Federal de Santa Catarina, http://www.ufsc.br/) started an ETD program. In 2000, two other universities implemented their ETD programs: PUC-Rio (Pontifícia Universidade Católica do Rio de Janeiro, http://www.puc-rio.br/) and USP (Universidade de São Paulo, http://www.usp.br/). The three universities represent different actors in the Brazilian higher education scenario—USF is a large public institution that belongs to the Federal Government, USP is also large and public, and belongs to the Government of the State of São Paulo, and PUC-Rio is a small confessional university that belongs to the Roman Catholic Church. The universities have something in common though, all of them are located in the most developed areas of Brazil—the Southern and Southeastern regions.

In January 2001, IBICT (Instituto Brasileiro de Informação em Ciência e Tecnologia, http://www.ibict.br/), an agency of the Federal Government, was assigned the mission of implementing a national consortium of ETD programs and of running a union catalog of metadata records. In January 2001, the three universities were invited to join other agencies to create BDTD (Biblioteca Digital de Teses e Dissertações, http://bdtd.ibict.org/). A fourth institution, BIREME/ENSP/FIOCRUZ, joined some months later. At the end of the year, the institutions sent XML files with their metadata recordsto deploy the first union catalog—the number of records was 1,047 (Marcondes, 2002).1

In 2002, BDTD adopted the OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting, http://www.openarchives.org/) to automate the transference and the uploading of records into the database. A free and open source software solution,
Sistema de Publicação Eletrônica de Teses e Dissertações (TEDE) was developed and distributed among Brazilian higher education institutions to implement their digital libraries; an intense training program came along with this solution. The result of all efforts was the dissemination of ETD programs all over country.

Currently, there are 101 cooperating institutions and the number of metadata records is 371,216 (132,805 DSc and 234,778 MSc). Some remarkable numbers, indicating the evolution of the ETD programs, are shown in Table 1. It is important to observe that from a little over 1,000 in December 2001, the number of ETD metadata records has grown to almost 400,000 in July 2015.

<table>
<thead>
<tr>
<th>Date</th>
<th>Number of ETD Programs</th>
<th>Number of Metadata Records</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dec 2001</td>
<td>4</td>
<td>1,047</td>
</tr>
<tr>
<td>Dec 2004</td>
<td>10</td>
<td>5,394</td>
</tr>
<tr>
<td>Oct 2010</td>
<td>95</td>
<td>145,200*</td>
</tr>
<tr>
<td>Jul 2015</td>
<td>101</td>
<td>371,200*</td>
</tr>
</tbody>
</table>

An interesting characteristic of the programs is that the 5 with the largest collections account for 240,400* ETDs (64.8%) of the total number. USP is the second largest cooperating institution with 54,235 records (14.61%) and PUC-Rio is in the ninth position with 7,414 records (2%).

BDTD is a service provider to Brazilian institutions and a data provider to other national and international union catalogues of ETDs.

ABOUT THE UNIVERSITIES

In the year 2000, two Brazilian universities–Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio) and Universidade de São Paulo (USP)–implemented ETD programs. Both are located in the Southeastern region of Brazil. This is the only physical similarity they have–they are very different in size, in the numbers of courses and graduate programs they offer, and in the funding of their activities. They started their ETD programs with different experiences in digital publishing too.

At the same time, they are very similar in the commitment they have to their ETD programs and the variety of digital contents they host on their institutional repositories. Both USP and PUC-Rio are members of NDLTD (Networked Digital Library of Theses and Dissertations, http://www.ndltd.org/). The following sub-sections introduce the two institutions.

Since the names of the titles and the designations of the works vary from country-to-country, in order to be unambiguous, the authors use the names used in Brazil. The titles are: MSc (Master in Science) and DSc (Doctor in Science). In Brazil, a dissertation refers to the MSc level work and a thesis to the DSc level.

PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO DE JANEIRO

PUC-Rio is a small confessional private university located in Rio de Janeiro; it is 75 years old. It is divided in four centres: Science & Technology, Social Sciences, Humanities &
Theology and Biological Sciences & Medicine. The total number of undergraduate and graduate students is 15000+ (approximately 83% are undergraduates); extension and specialization students are not included in this number.

The first graduate program, Electrical Engineering, started in 1963. It was followed by the Graduate Program in Mechanical Engineering in 1964. The two oldest graduate programs are a little older than 50 years.

CAPES (Coordenação de Aperfeiçoamento do Pessoal do Ensino Superior, http://www.capes.gov.br/) is the Federal Government agency responsible for the accreditation of graduate programs. Programs are rated in a scale of seven levels.

Currently, PUC-Rio offers 30 Master level programs and 24 Doctoral level programs. Among them, 20 (37%) are rated by CAPES in the top two levels.

**Universidade de São Paulo**

The University of São Paulo (USP), created in 1934, is a public university, maintained by the State of São Paulo and affiliated with the State Secretary of Economic, Scientific and Technological Development.

USP, as the major institution of higher learning and research in Brazil, is responsible for educating a large part of Brazilian MSc and DSc graduates. This achievement, the result of more than 75 years of an intense search for excellence, allows USP (founded in 1934) to be part of a select group of institutions of high world standing. This year, for instance, the QS World University Ranking classified USP between the 100 best universities in the world in 21 topics and also the best one in Latin America. The Times Higher Education classified USP in the World Reputation Ranking as one of the 60 best universities in the world. In another classification from 2013, that of the Webometrics Ranking of World Universities, considered important by the world scientific community, the University was placed at 31 (USP, 2015).

The graduate program is composed of 239 programs (with 332 Masters degree courses and 309 Doctoral courses), more than 30,000 students. Its undergraduate program consists of 249 courses, dedicated also to all fields of knowledge, distributed in 42 units of learning and research and offered to more than 58,000 students.

In order to carry out its activities, USP counts on various campi, distributed in the cities of São Paulo, Ribeirão Preto, Piracicaba, São Carlos, Santos, Pirassununga, Bauru and Lorena, besides units of learning, museums and research centres located outside these spaces and in different municipalities.

**HISTORIES OF ETOS AT THE UNIVERSITIES**

The two universities started their ETD programs in the same year-2000, though their backgrounds were different.

**Pontifícia Universidade Católica do Rio de Janeiro**

The digital library activity at PUC-Rio was implemented to host and make available courseware in Electrical Engineering. In 1995, the first version of the Maxwell System,
http://www.maxwell.vrac.puc-rio.br/) was deployed with this mission. This type of digital content is deeply related with ICT (Information and Communication Technology) supported learning. For this reason, the system started offering LMS (Learning Management System) functions so that students could use contents in their learning activities. In 1999, PUC-Rio registered the Maxwell System in the Brazilian Patent Office (n° 99003015) with this integrated characteristic.

After 1995, new types of digital contents started being published and not only in Electrical Engineering; new functions started being added to the LMS side of the system too. A very important step was the publishing the first ETD in May 2000. At that time, ETDs were not mandatory and graduate programs and/or authors were free to submit their works (or not). The first graduate programs to join the ETD program in 2000 were Business Administration, Civil Engineering and Electrical Engineering. In 2001, Industrial Engineering and Mechanical Engineering joined too. The numbers of ETDs made available in the first three years are shown in Table 2.

In August 2002, ETDs became mandatory to all graduate programs in both the Master and the Doctoral levels. Since 2000, the number of published ETDs has grown to 7,400+. The average number of new ETDs is between 500 and 600 per year. But the numbers also show results of other related activities:

- The Graduate Program in Electrical Engineering has retrospectively digitized all printed ETDs;
- The Graduate Program in Mechanical Engineering is retrospectively digitizing all printed ETDs;
- Some authors of printed ETDs have voluntarily submitted the digital versions of their theses and dissertations.

An important aspect of the ETD collection is that 76% are in Open Access. The other 24% have some type of restriction; restrictions can be temporary and in most cases they are a one-year embargo. Other types or periods of restriction must be submitted to a committee for approval.

Digital preservation had been a concern for many years. An important step was joining the Meta Archive Cooperative (http://www.metaarchive.org/) to address the problem of digital preservation of ETDs; this happened in 2009. The ETD program is consolidated and stable at PUC-Rio.

From June 2004 on, new products have been added to the Maxwell System and the ETD Program was impacted. Two of them are worth mentioning:
Statistics

As the collection started to grow (the number of ETDs was 1,100+), ETDs became part of the graduate programs and numbers associated to them began being examined. Statistics related to the number of published ETDs were made available so that graduate program staffs could keep track of them. At the same time, statistics related to accesses by country, graduate program and by each individual work were also implemented. The numbers are used for reports.

Figure 1 shows an example of the time-series of presented ETDs in the Centre of Science and Technology; this is one type of information made available concerning published ETDs. There are 12 options of statistics available to the public that are about ETDs, some concern the publishing and other the usage.

![Time-series of Presented ETDs](image)

**Fig. 1: Time-series of Presented ETDs in the Centre of Science and Technology**

The statistics related to accesses show some interesting numbers. Table 3 shows the numbers of countries that accessed ETDs, the numbers of accesses, the numbers of ETDs on the system, the average number of partitions (files) of the ETDs and, finally, the average number of accesses the ETDs had. Table 4 shows the numbers related to the 10 most accessed ETDs. Both tables have data of the last four years. Data were gathered from the system set of applications that generate statistics.

| Table 3: Numbers of Accesses to ETDs in 2011–2014: Robots were Filtered Out |
|-----------------|-------|-------|-------|-------|
|                 | 2011  | 2012  | 2013  | 2014  |
| Countries       | 148   | 187   | 186   | 186   |
| Accesses        | 696,294| 4,023,830| 3,440,763| 4,862,684|
| ETDs            | 5,972 | 6,644 | 6,926 | 7,138 |
| Average Partitions| 7.2   | 7.1   | 7.0   | 6.9   |
| Average Accesses per ETD | 16.19 | 86.60 | 70.97 | 98.75 |

The numbers of countries accessing ETDs are quite large with respect to the size of the collection and with the fact that 98.9% of the works are in Portuguese and 0.95% are in English (http://www.maxwell.vrac.puc-rio.br/Esta_por_lingua.php). The percentages of accesses from Brazil are 68.4% (2011), 81.0% (2012), 74.5% (2013) and 61.5% (2014).
Table 4: Numbers of Accesses of the 10 Most Accessed ETDs in 2011–2014, Robots were Filtered Out

<table>
<thead>
<tr>
<th></th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
<th>2014</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accesses</td>
<td>29,339</td>
<td>165,828</td>
<td>132,071</td>
<td>243,761</td>
</tr>
<tr>
<td>Average Partitions</td>
<td>7.2</td>
<td>7.1</td>
<td>7.0</td>
<td>6.9</td>
</tr>
<tr>
<td>Average accesses per ETD</td>
<td>407.49</td>
<td>2,335.61</td>
<td>1,886.73</td>
<td>3,535.77</td>
</tr>
</tbody>
</table>

It is important to mention that the 10 most accessed ETDs had the following numbers of accesses: 55,804 (2014), 14,849 (2013), 24,679 (2012) and 4,859 (2011) accesses. The first has 7 partitions, the second and the third have 10, and the fourth has 9, yielding the average numbers of accesses per ETD of 7,972 (2014), 1,485 (2013), 2,468 (2012) and 540 (2011).

Accessibility

In March 2009, the fourth version of the Maxwell System was deployed. It is accessible to the blind and the visually impaired. The development of the new version was motivated by a request of blind and visually impaired students who wanted to access ETDs.

After March 2009, all course wares have also been developed to be accessible whenever possible.

Universidade de São Paulo

In June 2000, the project of the Digital Library of Theses and Dissertations of the University of São Paulo (http://www.teses.usp.br/) was started by direct request of the Rector of USP at that time. In June 2001, it was officially inaugurated. It aimed at making available on the Internet the knowledge produced at the University at the graduate level. It has allowed the Brazilian and the international community to freely access the full digital versions of theses and dissertations.

The average numbers of Doctoral theses and Masters dissertations at USP each year are, respectively, 1,500 and 2,600. These two numbers in addition to the collection of theses and dissertations presented before ETDs started being published required a robust system; open access meant many readers from Brazil and worldwide. Just for the sake of curiosity, the oldest digitized work dates back to 1942.

Starting in April 2007, ETDs became mandatory. Embargoes are possible for parts or the complete work under some conditions and no more than four years, if properly justified (USP, 2007).

Currently, though the University has already granted the titles of Master and Doctor to more than 100,000 students, its digital collection accounts for just over 54,000 items. This situation is justified given the difficulty in getting authorization from past authors.

The development of the ETD program at USP involved a multidisciplinary team, and the main steps of the initial implementation process were: technological prospecting, definitions of graduate activities, software engineering, a proposed pilot project, training, dissemination plan, operation and evolution (Masiero, 2001). After choosing to use the system developed by Virginia Tech and made available by NDLTD (Networked Digital Library of Theses and Dissertations, http://www.ndltd.org/), many customizations were required. The changes were needed in the source code in order to adapt it to the
specificities of the complex and diverse graduate system of USP as well as its integration
with other enterprise systems already in use: the postgraduate system and the online
public access catalogue of the Integrated Library System of USP” (Kondo, 2009).

A new version of BDTD-USP was under development in 2009 when new updates
were implemented. The Ajax technology was introduced to enable increased
interactivity of users and to facilitate system maintenance. This also required the
rewriting of PHP code and the addition of new module (Kondo, Sueli).

Table 5 shows the Digital Library of the USP growth from 2011 to 2014. In June
2015 the total number of documents was 54,8832. Table 6 shows the numbers of
documents with respect to the availability to the public. Table 7 shows access to
statistics and yields, and gives the clear idea of the growing importance and the social
impact of the ETD collection.

<table>
<thead>
<tr>
<th>Table 5: Growth of the Digital Library of ETDs in 2011–2014</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>MSc</td>
</tr>
<tr>
<td>DSc</td>
</tr>
<tr>
<td>Other documents</td>
</tr>
<tr>
<td>Total documents</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 6: Numbers of Documents According to their Access Criteria in 2011–2014</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Open Access</td>
</tr>
<tr>
<td>Restricted Access</td>
</tr>
<tr>
<td>Embargoed</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 7: Access Statistics in 2011–2014</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Visits</td>
</tr>
<tr>
<td>Robots</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

The last three tables show the steady growth of the ETD collection as well as of the
importance it has as a reference to research.

CURRENT STATUS OF THE ETD PROGRAMS AND AFTER ETDs

As mentioned before, the ETD programs are consolidated in both institutions. So, the
publishing of ETDs is no more under discussion. At the same time, new features have
been incorporated to the programs.

ETDs were important starting points in digital publishing at many institutions
worldwide. This happened in Brazil and at PUC-Rio and USP too. Some of the spin-offs of
the ETD programs at the two institutions are presented in the following sub-sections.

2All statistics were extracted from
Pontifícia Universidade Católica do Rio de Janeiro

The ETD program is under permanent enhancement in order to keep in pace with international best practices. The newest features added to the ETD program were:

Digital Object Identifiers

PUC-Rio decided to assign DOIs to all ETDs stored on the Maxwell System. All ETDs on the system have retrospectively received DOIs and all new ETDs have DOIs assigned to them when they are deposited. This means that all ETDs have DOIs.

Other digital contents—articles, journals, courseware, books and datasets—will have DOIs too.

Research Data

Making research data available along with the ETDs that used them is a worldwide trend (UBOGU, 2008). A survey was submitted to theses and dissertations supervisors concerning their interest in publishing data used in ETDs. Quite a few seemed interested. For this reason, the decision was made to include research data on the system. A new module was developed for the system so that additional metadata elements could be added to describe the datasets.

In June 2015, the first two datasets were stored and made available from the Maxwell System. They were collected during the development of a MSc work and are available in Open Access. It is expected that others will follow.

The three most important spin-offs of the ETD program at PUC-Rio were: (a) the publication of journals (2003), books (2005) and series (2009); (b) the publication of senior projects (2003); and (c) the publication of monographs of specialization courses (2011). All contents of journal articles, books and series items are authored by at least one person that belongs to the university.

Courseware has been a focus since the Maxwell System started in 1995. New types have been developed and the quality of the courseware has been enhanced due to the evolution of ICT in the last 20 years.

After the digital preservation of ETDs, other digital contents are under consideration to be preserved.

Universidade de São Paulo

The development of digital libraries at USP after ETDs has been staggering. It covers diverse document typology—electronic publication of journals, digital copies of rare and special documents, educational learning objects, monographs of undergraduate courses and other scholarly publications (scientific, artistic, technical and academic).

USP is a university of large dimensions in all senses and it is spread all over the State of São Paulo. For this reason, the decision was made to create independent digital
libraries and institutional repositories. They are integrated to Integrated Library Systems at USP (SIBiUSP); though they are local, they are nationally and internationally interoperable.

Ferreira (2013) shows the various initiatives that have been successful at the university and that are ‘based on open source software, OAI/PMH, Dublin Core metadata standard, Sword protocol for cooperative services and other features.’

The second platform for digital content was created at USP in 2008; it is called the Journals Portal (www.revistas.usp.br). Currently, this portal is the largest in the genre in Latin America. A reason for its rapid growth was the adoption of the Open Journal System (OJS) which started in 2012.

As in August 2015, this portal hosts 137 journals published at USP by several research groups, teaching units or graduate programs in various fields of knowledge; the collection has 4,477 booklets and 66,496 articles. Only in the year 2015 (January-June) this collection has recorded 10,083,005 hits and 5,918,113 downloads.

Other initiatives that need to be mentioned are the Digital Library of USP Intellectual Production (http://www.producao.usp.br) launched in October 2011 in partnership with IIBICT Institutional Repositories Program. It uses DSpace and was launched with more than 28,000 items. As in August 2015, it hosts 44,622 documents including articles, books, chapters, events proceedings and other documents.

The Rare, Historical and Special Books Library (http://www.obrasraras.usp.br/) was launched in 2013. It followed the same methodology used in BPDI but its contents are diverse library materials such as books, brochures, magazines, newspapers and other types published between the fifteenth century and the twentieth century. These documents were selected according to parameters to classify them as rare or precious; definitions were made by SIBiUSP with advice of experts in the field. Items may be ancient works, unique unpublished works, or part of special editions; luxury bindings, special illustrations or even autographs of famous personalities were considered too (Ferreira, 2014). Such Library resulted from a digitization project developed in a partnership among the three São Paulo State universities UNESP, USP and Unicamp with support from FAPESP (Fundação de Amparo a Pesquisa do Estado de São Paulo).

Other initiatives are still under development: the newly created Digital Library of Academic Publications aimed at undergraduate monographs of the Campus of São Carlos, the IPTV Portal for videos, Portal eAulas for educational learning objects among others.

All these initiatives and proposals were ‘crowned’ with the implementation, in March 2012, of the Integrated Search Portal. It is an aggregator and thus a single interface and aggregator of all content referenced in the Integrated Library System.
OPAC. It gives access to items stored in libraries and digital repositories of USP, and also to the electronic contents directly acquired from publishers or through the CAPES Portal of Journals.

This integrated portal uses the PRIMO system. It is a product of ExLibris, the Israeli Company, and belongs to a new category of library systems known as Web-Scale Discovery Systems. According to Ferreira (2014), this new type yields access to disperse information contents by incorporating multiple sources and data exchange protocols. It also has other highly desirable features such as semantic web principles; web 2.0 tools; FRBR function; the filtering of searches results by facets; intuitive interface providing a better experience to the end users; reduplication of duplicate records; and others.

In two years of its existence, from March 2012 to December 2013, this portal has aggregated 20 million documents and has received more than 380,000 visits, with more than 2.2 million page views.

CONCLUSION AND COMMENTS

The conclusions and comments start by addressing the importance of the ETD programs in Brazil. In most universities there were no activities related to digital publishing or to the creation of digital contents prior to ETDs. ETDs were the first contact with digital contents in most of them. The fact that ETDs are theses and dissertations was of paramount importance to motivate institutions because they wanted to show the products of their graduate programs. In most of them, many other digital publications followed.

In 2002, DSpace (http://www.dspace.org/) was made available. The following year, Lynch (2003) introduced the concept of Institutional Repository (IR). This concept went far beyond ETDs, but did not exclude them. In 2003, many universities in Brazil were running digital libraries of ETDs. Table 1 shows that in December 2004 there were 10 ETD programs in the country. From 2003 on, IRs started to be implemented.

There are two models of scholarly publications management in the country. The first is similar to the Maxwell System—one single repository hosts all types of digital scholarly publications. This is quite natural in the Maxwell System case since the system was deployed as a digital library of courseware that added ETDs five years later. An example of a solid integrated IR is Lume (http://www.lume.ufrgs.br/). The second model follows the one adopted by USP—there are separate digital libraries/IRs for different types of scholarly publications. An example of university that adopted this model is Universidade de Brasília (http://www.unb.br/). It has the Universidade de Brasília Repository (http://repositorio.unb.br/) and the Biblioteca Digital de Monografias de Graduação e Especialização (http://bdm.unb.br/).

A visit to Webometrics—Ranking Web of Repositories (http://repositories.webometrics.info/en) yields the information on the digital publishing of scholarly
communication in Brazil. There are more than 40 ranked repositories; 9 belong to institutions that are not universities.

At PUC-Rio and USP, ETDs have easily been accepted—in two years and five years respectively the deposit became mandatory. This does not mean that many problems did not arise. The most important of them are related to the formatting of the works and the quality of the information contained in the introductory parts where metadata information is contained.

At the same time, the acceptance is shown by the use of access data in reports that are submitted to government agencies.

Since a digital preservation activity is under way, both universities are considering quit requiring the printed copies still in use, especially in some fields.

Whereas both universities studied here are among the best in the country and it is fair to say that the access statistics prove the great social impact that ETDs programs have been playing. Moreover, although empirically, it can also be mentioned, that these developments greatly influence the appearance and implementation of similar initiatives in various other private and public Brazilian universities.

REFERENCES


