

TDX doing it together

Lluís Anglada

Consortium of Academic Libraries of Catalonia (CBUC, Consorci de Biblioteques Universitàries de Catalunya)

langlada@cbuc.es

C/Gran Capità, 2-4, Nexus building, E-08034 Barcelona, Spain

www.cbuc.es/

Keywords: electronic publishing, library consortia, academic libraries

Abstract

The Consortium of Academic Libraries of Catalonia (CBUC) and the Supercomputing Center of Catalonia (CESCA) have built and maintained the TDX server (TDX: Tesis Doctorals en Xarxa - Networked Doctoral Theses), which now hosts around 1,000 e-doctoral theses. The theses belong to 10 different universities from the Catalan context. The National Library of Catalonia also participates in the project, working on preservation topics.

The paper is not about technological developments, but is focused on:

- Advantages of putting e-theses on the Internet as a result of the cooperative work among organizations from the same territory
- Sinergies created thanks to cooperative work
- Relationship with other related projects, also devoted to put local contents on the Internet

This experience can be specially interesting for universities that would like to publish their theses on the Internet, but who have difficulties due to bureaucratic issues. The Catalan experience shows that cooperative work can save costs, it breaks inertia, and can also encourage creativity.

Background

In 1999 the then Information Society Commission of the Generalitat of Catalonia¹ and LOCALRET² put forward a strategic plan for encouraging and developing the information society in Catalonia.³ The plan was formulated with the participation of a wide array of experts and made a series of recommendations, including several which considered the university as an agent within the process. For the purpose of implementing the priorities established in the plan, the Department of Universities, Research and the Information Society of the Generalitat sponsored the signing of an agreement whereby the Department, universities, CESCA⁴ and CBUC⁵ committed

themselves to carrying out a series of tasks in support of the development of new technologies. One such task was the design of a server for doctoral theses in digital format (TDX⁶) for Catalan universities. The idea was to use remote consultation as a means to provide an incentive for using theses and to facilitate their conservation.

The first decision was that the CBUC should commission a report⁷ on the situation of doctoral thesis digitisation around the world. On the basis of this initial study, the decision was made to commence organisational work with the institutions involved, and conduct legal consultations related to the online publication and distribution of theses. In addition, at the beginning of December 2000, agreement was reached for the adhesion of TDX to the NDLTD project.

- The initial objectives of TDX in 2000 were:
- To make software available for accessing and viewing full text doctoral theses in electronic format.
- To generate a set of procedures and guidelines on how to write electronic theses for subsequent Internet publication.
- To harmonise public access to electronic doctoral theses with intellectual property and copyright law.
- To obtain a commitment from university managers and create a platform for collaboration between teachers, service personnel, students and libraries to guarantee access to and conservation of electronic doctoral theses.

The different functions of each organisation in TDX were defined as follows:

- The CBUC would be responsible for coordination and promotion.
- The CESCA would deal with the access software and storage.
- The universities would collect the electronic files of the theses, deal with the publishing contracts and transfer the theses in electronic format to the server.

1 Currently a secretariat within the Department of Universities, Research and the Information Society.

2 LOCALRET is a consortium composed of the majority of municipal areas of Catalonia. The objective of LOCALRET is to promote the development and use of new technologies.

3 Comissionat per a la Societat de la Informació, Catalunya en xarxa: Pla estratègic per a la societat de la informació, Barcelona, 1999.

4 <http://www.cesca.es>

5 <http://www.cbuc.es>

6 TDX: Tesis Doctorals en Xarxa (Networked Doctoral Theses)

7 Dijkstra, J.M.M., "Towards Storage and Access of Electronic Theses: A Proposal of Its Organisation for the Consortium of Catalan Libraries", unpublished manuscript, Barcelona, 1999.

- The Library of Catalonia would be responsible for the future conservation of the theses.

Preparing the software was a less arduous task than was initially thought, but the agreements within each university to allow them to start to upload full-text theses onto the TDX server required great efforts. The universities of Catalonia (the University of Barcelona, the Autonomous University of Barcelona, the Technical University of Catalonia, the Pompeu Fabra University, the University of Girona, the University of Lleida, Rovira and Virgili University and the Open University of Catalonia) approve about one thousand doctoral theses per year and they all had mechanisms for publishing them on microfiche. Furthermore, the project was launched from a cooperative platform—the CBUC—that was highly active in fields of librarianship such as the union catalogue and licensing electronic resources, so it seemed that the introduction of thesis would be a simple and immediate task.

However, though everything seemed to be well defined and ready, the project went through a phase of almost two years in which only slow progress was made. Below we will explain the main problems that were encountered and how they were solved.

Problems and solutions

In early 2000 everything seemed ready for the universities to automatically switch from collecting theses on paper and publishing them on microfiche to publishing them in electronic format through TDX. In this rather naive phase the CBUC informed the different universities of the steps they had to take and trusted that they would start to feed the server. This did not happen.

It was obvious that the principles of the project had to be revised, and after many visits and meetings the following problems were detected:

- The new framework for publication of theses (from microfiche to digital format) aroused many fears and uncertainties. The bodies that had to take the decisions had doubts about the best way to do so.
- The electronic publication of theses could involve the coordination of up to four bodies within a university: the academic service, the publications service, the computer service and the library. It became clear that the coordination between them was not easy, that the information did not circulate easily and that there could be confrontations between the different university departments involved in the process.
- In the academic area there were many examples of theses that could not be offered openly in digital format (theses involving confidential research, theses that would be published in book form, theses made from articles published in journals for which the reproduction rights were not available...) In general these cases were formal obstacles that concealed a certain sloth to change established processes.

The solutions that were found all involved joint work consisting of: complementariness, globalisation, promotion and emulation.

Though the CBUC had prepared the processes that it thought necessary, practice showed that far more than had initially been planned would be required. The CBUC had neither the means nor the experience to resolve them and it therefore decided to coordinate efforts in order to carry out all the processes complementarily. An intranet was created and fed with documentation from the different universities, showing how to fill in the ISBN form for electronic publications and the different types of publishing contracts.

The second mechanism used was to present the activity of publishing digital theses within a wider vision of digital universities and libraries. Information Technologies (IT) applied to university life currently enjoy great prestige, but publishing theses in digital format is seen as an administrative activity that is unrelated to IT. The fact that the universities of the Consortium were undertaking major projects involving the licensing electronic resources, including a table of contents database and digitisation of documents under the name of the Catalan Digital Library (Biblioteca Digital de Catalunya - BDC) represented an opportunity. It was therefore decided to include TDX within the overall project of the BDC and to link its results to a more general project involving the whole university.

The instruments of information had proven to be too weak to reach all the persons involved. Despite the web page of TDX and the official information, a large number of interested persons (directors of theses and doctorands) were unaware of the possibilities and advantages of publishing theses in digital format. It was attempted to solve this through the publication of bookmarks and leaflets which were distributed through doctorate programmes and in the libraries.

Finally, however, the mechanism that proved to be most effective was emulation. As in all regions, the universities of Catalonia are not equal in size, history or the courses offered. The "smaller" universities presented fewer theses per year, but could be more flexible and better able to control the mechanisms for changing from microfiche to digital publishing. The CBUC encouraged the more active universities and constantly showed their number of theses in TDX to the other universities (even with a thesis counter for each university on the main page of TDX). This finally led to a certain amount of competition from which contributed to the success of the project.

The present and the future

After this setback to the project in 2000 and 2001, theses began to be introduced at a normal rate, and between late 2001 and early 2002 the number of theses introduced in TDX by the eight universities that form part of the CBUC came to 500. From then on the figure rose more quickly and has reached the figure of 973 at the time of writing this paper. The cooperative project has now gone beyond the limits of the CBUC to include the Jaume I University of Castelló, the University of the Balearic Islands and the University of València, who are also participating in TDX.

The TDX server has the features of other NDLTD servers. Perhaps, however, three additional things should be pointed out. The CBUC manages a union catalogue with 2,300,000 bibliographic records, which also catalogues the electronic resources licensed by the CBUC. The digital theses have been entered in this union catalogue (CCUC)⁸ from the start and the catalogue links the bibliographic reference of each thesis to its description and full text in TDX. The experience of the CBUC with the use of the electronic resources told us that in order to increase their use it is very important for them to be in the catalogues of the libraries integrated with the other materials. After finding that the traditional mechanisms of access to theses could be insufficient for some purposes, TDX incorporated the possibility of browsing the theses by wide subject areas. After some discussion UDC was adopted as a basic classification SYSTEM and now, in addition to searching for theses by key words, one can see all those that fall under wide subject areas, such as "Catalan Literature" or "Algebra".

The third new feature of TDX is a statistics module that provides figures on major subject areas and on the number of theses introduced by each university. The statistical analysis must form the starting point for increasing the rate of introduction of theses in TDX. While it is still too early to assess growth in the dissemination of theses available in TDX, the data available at an international level demonstrate that theses available online are consulted far more than those on microfilm or in paper format. The benefits for the university and research communities are undeniable. Readers can locate relevant data in a highly specialised area, where in other circumstances the information might have remained out of their reach. The authors gain readers, making their research efforts worthwhile. Lastly, doctoral research acquires a new relevance which it previously did not have.

At present we are facing three future challenges:

- In the next few months the aim is that TDX should follow the OAI protocol so that it can supply meta-data to NDLTD's collective catalogues of theses.
- To achieve a higher level of incorporation of theses. Currently TDX incorporates approximately 50% of

the theses presented at the universities of Catalonia. These results may be good but we feel they are insufficient. The digital publication of theses has proven in this brief period to be easier, cheaper and more effective than the equivalent in other formats and we hope that the mechanisms of promotion and the figures of consultation of the database will lead to higher percentages.

- To start to incorporate retrospective theses in TDX. We are aware that this is not an equally interesting objective for all thematic areas. Some theses, however (especially in humanities and social sciences), are still valid instruments of study. In these cases we will be unable to use original electronic files and will have to digitise the originals on paper. We have already had meetings with some university departments to select the theses for incorporation, and here, as on other occasions, the technical and economic difficulties are less important than the bureaucratic ones and those arising from inertia.

Lessons that have been learnt

We have learnt two main things: firstly, that the main obstacle to universities playing an active role in electronic publishing is neither technical nor economic, but rather lies in the inertia associated with traditional ways of doing things; and secondly, that working together is one of the best ways of overcoming any problems that may arise.

In discussions of electronic publishing it is common to accentuate technical matters. In TDX the technical side caused only minor problems. Instead, the major issues have been related to changing the circuits of established processes and coordinating groups of people who do not normally work together (publishing services, academic and library services). The new form of dissemination proposed for theses has often involved changing views about publishing and even the formulation of contracts for electronic publishing. The preparation and functioning of the project has also involved modifying legal deposit submission circuits.

On the other hand, this apparent weakness, which has slowed down the implementation schedule, has turned out to be the best means for obtaining results at the collective level. The obstacles encountered at different points along the way have been overcome largely thanks to the fact that someone earlier on had found a practical solution to the problem in question. It should also be emphasised that working together, which was a constant ingredient in this project, facilitates a much greater visibility of the publishing work of universities and scientific societies. In the case of theses, the different sizes of the universities results in an unequal production of theses. Yet, in

8 <http://www.cbuc.es/ccuc>

spite of this, the ability to access them all together in one place has made consulting the server more cost-effective and has therefore increased its use.

On a more general note, in the next few years we think that the success of TDX should serve to strengthen the idea that universities must work to place digital contents at the disposal of the community of students and lecturers. Consortial licensing of databases and electronic journals and books must have a continuation in the

works generated within the universities. Digital theses are an example of how universities can set up institutional archives that conserve and give access to working papers, articles and teaching materials of different types. The CBUC wishes to use the experience of digital theses to promote electronic publishing in general and specifically that of journals published by the universities that form the CBUC.