CENTRE DE SUPERCOMPUTACIÓ
DE CATALUNYA

CESCA

Catalan Policies and Experiences on Cooperative Repositories

Miquel Huguet, Lluís Anglada and Ricard de la Vega • 12-03-07

Summary

The Centre de Supercomputació de Catalunya (CESCA) together with the Consorci de Biblioteques Universitàries de Catalunya (CBUC) started in 1999 a cooperative repository, named TDR, to file in digital format the full-text of the read thesis at the universities of our country to spread them worldwide in open access preserving the intellectual copyright of the authors. This became operational in 2001 and today it is a service fully consolidated not only among the Catalan universities, but also used by other Spanish universities.

Since then, there are four additional cooperative repositories which have been created: RECERCAT, for research papers; RACO, for scientific, cultural and erudite Catalan magazines; PADICAT, for archiving Catalan web sites; and MDC, for Catalan digital collections of pictures, maps, posters, old magazines...

These five repositories have some common characteristics: they are open access, that is, they are accessible on the internet for free; they mostly comply with the Open Archive Initiative interoperability protocol for facilitating the efficient dissemination of content; and they have been built in a cooperative manner so that it is easy to adopt common procedures and to share the repository developing and managing costs, it permits more visibility of the indexed documents throughout the search engines, and a better provision for long-term preservation can be made.

In this paper we present the common policy established for the Catalan cooperative repositories, we describe the five of them briefly, and we comment on the results obtained of our 6-year experience since the first one became operational.

Keywords: e-science, institutional repository, research, open access, web archiving.

1. Introduction and Objectives

Back in 1998 the Generalitat de Catalunya, the local Government of Catalonia, and Localret, a consortium of Catalan municipalities, started a thinking process for promoting the development and the use of the information and communication technologies in several fields (administration, education, culture, health…). The result of this process was submitted [1] to the Catalan Parlament on April 14th, 1999. As a fruit of this project, on September 8th, 1999, the Catalan Universities, together with the Centre de Supercomputació de Catalunya (CESCA) [2] and the Consorci de Biblioteques Universitàries de Catalunya (CBUC) [3], signed an agreement [4] to create a digital cooperative repository for doctoral theses, to allow remote consultation over the internet. This repository, named TDR [5], became operative on February 2001 ant it was the first repository created in Spain, as it will be described afterwards.

By the end of 2003 the German Max Planck Society promoted the Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities [6] which supports the wide and transparent diffusion of the scientific knowledge and the human thinking by means of new possibilities provided by internet. To reach this goal, the Declaration proposes to spread the knowledge, not only by the traditional system but also by the open access; that is, by permitting the free and open access through the net to provide an alternative to the paradigm of having to pay for accessing information obtained with public funds.

In addition to its promoters, the Declaration has already been signed by 226 organizations from all over the word. Among then, there is the Catalan Government. Although it did not sign it formally until March 24th, 2006 the Government was, as we said, an early promoter for open access, so that today there are five cooperative repositories available to the scientific community (TDR, RECERCAT, RACO, PADICAT and MDC).

The objective of this paper is to present these repositories and our experiences developing and managing them. First, we describe the policies established for these repositories. Second, for each one, we explain what is their goal and the software used to implement them. Finally, we comment our experiences on its current usage.

2. Policies on the Cooperative Repositories

The Scholarly Publishing and Academic Resources Coalition (SPARC) [7] defines the institutional digital repositories as digital collections that capture and preserve the intellectual results from several institutions. The following are some remarkable characteristics:

- Contain documents created by the institutions.

- Academic content.
- Accumulative and persistent.
- Open access and interoperability.

Clifford A. Lynch [8] has another definition of the repositories, an “institutional repository is a set of services that a university offers to the members of it’s community for the management and dissemination of digital materials created by the institution and its community members. It is most essentially and organizational commitment to the stewardship of this digital materials, including long-term preservation where appropriate, as well as organization and access or distribution”.

Thus, all our repositories are open access and we have the commitment to preserve both the documents and its URL reference. In addition, they should comply with the interoperability protocol created by the Open Archives Initiative (OAI) [9].

2.1. Open Access

Authors, libraries, universities and all the institutions which finance the research have promoted the open access (OA) movement to the scientific information. OA’s objective is to create an alternative to the paradigm of having to pay in order to have access to the information that has been elaborated in the own institution.

As we said, at the end of 2003 the Berlin Declaration was developed. Its objective is to defend and to keep the new possibilities of diffusion of the knowledge, not only through classical forms, but also through the paradigm of the open access by internet, and this open access is defined like the universal core of knowledge and the cultural patrimony that has been approved by the scientific community.

Berlin Declaration says that the open access contributions should carry out two conditions. First, the author (or authors) and those who retain the rights about the collaborations should guarantee to all the users the right to the open access […], with permission to copy, to use, to spread, to transmit and to put forward the works publicly […] with some responsible purpose, and the engagement of naming correctly the responsibility. On the other hand, a complete version of this work […] will have to be placed online in an appropriate electronic format. This file will be administered and keep by an academic institution, a research institution, a public administration or an institution that can ensure the open access.

2.2. Open Archive Initiative

The repositories use an interoperability protocol created by Open Archives Initiative (OAI) which increases the visibility of the e-information by sharing the metadata of the repository with other international repositories, like OAIster [10].

The OAI develops and promotes interoperability standards that aim to facilitate the efficient dissemination of content. OAI has its roots in the open access and institutional repository movements. Over time, however, the work of OAI has expanded to promote broad access to digital resources for e-Scholarship, e-Learning, and e-Science.

One of these standards is the OAI Protocol for Metadata Harvesting (OAI-PMH), a low-barrier mechanism for repository interoperability. There are two possible actors with the protocol, on one hand, the *data providers* are repositories that expose structured metadata via OAI-PMH. On the other hand, the *service providers* then make OAI-PMH service requests to harvest that metadata.

2.3. Cooperation Benefits

A third characteristic of our repositories is that they are built in a cooperative manner, that is as a collective initiative of most of the Catalan universities which are part of the Government of both consortia, CESCA and CBUC (universities of Barcelona, Autonomous of Barcelona, Politechnic of Catalonia, Pompeu Fabra, of Girona, of Lleida, Rovira i Virgili, Open, and Ramon Llull).

The first benefit for the Catalan University System is that the participation of different institutions and universities eases the adoption of common procedures and allows to share the repository developing and managing costs. In this case, our universities are able to concentrate in doing its own task (research) rather than designing the e-infraestructures required to sustain it.

The second benefit is that the cooperative repositories permit more visibility of the indexed documents throughout the search engines. This facilitates a better diffusion of the research activity produced and help the development of e-Science and the Information Society of our country.

Finally, a third benefit is that a better provision for long-term preservation can be made. The integration of thesis, research documents and articles produced in Catalonia in a few set of catalogues is an important and innovator difference with other similar initiatives that helps to ensure its preservation.

3. Catalan Repositories

CESCA and CBUC started in 1999 a cooperative repository to file in digital format the full-text of the read thesis at the universities of our country to spread them worldwide in open access preserving the intellectual copyright of the authors. This repository, called Tesis Doctorales en Red (TDR), started 6 years ago and today it’s fully consolidated.

This success has made possible that the Generalitat de Catalunya has sponsored two new open access repositories: the Dipòsit de la Recerca de Catalunya (RECERCAT) [11] and the Revistes Catalanes amb Accés Obert (RACO) [12]. The first one is a cooperative repository of digital documents that includes literature of the universities and research centres of Catalonia, like preprints, proceedings, research reports, working papers, thesis, etc. RACO is a repository where the full-text from articles of scientific, cultural and erudite Catalan journals can be open access acceded.

In September of 2006, the Biblioteca Nacional de Catalunya [13] has started another ambitious repository in cooperation with CESCA: the Patrimoni Digital de Catalunya (PADICAT) [14]. This repository was set up to ensure permanent access to the Catalan web production in digital format. This guarantees for access includes the assurance that the traditional document cycle (compiling, handling, preservation and dissemination) is maintained for bibliographic material published on the internet.

The last one to become operative has been the Memòria Digital de Catalunya (MDC) [15], coordinated by CBUC and the BC. Since last November it allows the consultation of digital collections of old Catalan magazines, pictures, maps, posters, ex-libris, etc.

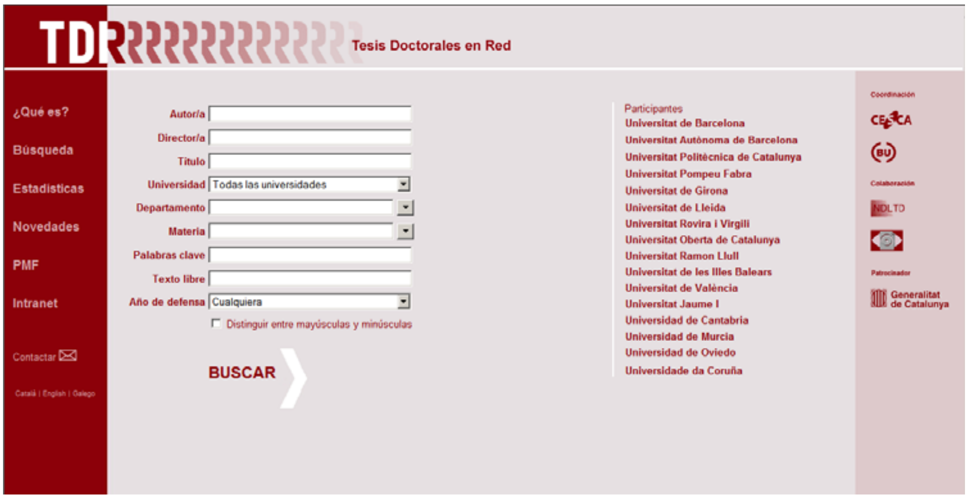
3.1. TDR

Tesis Doctorales en Red (TDR) is a digital cooperative repository of doctoral theses presented at 16 Spanish universities. It allows for remote consultation of the complete text of theses over the internet, additionally allowing the user to construct searches by author, advisor, title, knowledge area, university and department of publication, year of defense, etc. This repository has the following goals:

- To publicise, around the world and on the internet, the results of university research.
- To offer the authors of theses a tool to increase public access to their work, enhancing its visibility.
- To improve the cataloguing and bibliographic processes.
- To encourage electronic publishing and the use of digital libraries.
- To stimulate scientific productivity.

The universities that take part in TDR are responsible for publicising this repository to their doctorate students, providing them with recommendations and a list of accepted electronic formats for theses and (once these are presented and approved), editing and uploading them to the TDR. An additional goal for the next few years is to digitise earlier theses presented prior to the existence of the TDR in other formats (microfiche or paper).

Author's copyright privileges are protected by contract. The integrity of the text is also guaranteed by the security options encoded in the format used for data storage: PDF.



Autor/a
Director/a
Título
Universidad Todas las universidades
Departamento
Materia
Palabras clave
Texto libre
Año de defensa Cualquiera
□ Distinguir entre mayúsculas y minúsculas

Participantes
Universitat de Barcelona
Universitat Autònoma de Barcelona
Universitat Politècnica de Catalunya
Universitat Pompeu Fabra
Universitat de Girona
Universitat de Lleida
Universitat Rovira i Virgili
Universitat Oberta de Catalunya
Universitat Ramon Llull
Universitat de les Illes Balears
Universitat de València
Universitat Jaume I
Universidad de Cantabria
Universidad de Murcia
Universidad de Oviedo
Universidade da Coruña

Figure 1. The TDR homepage

The TDR repository is a member of the Networked Digital Library of Theses and Dissertations (NDLTD) [16].

The Electronic Theses and Dissertations (ETD)[17] of Virginia Tech University has been adapted for the management of the theses. Also Glimse and WebGlimse [18] are used for indexing and searching.

3.2. RECERCAT

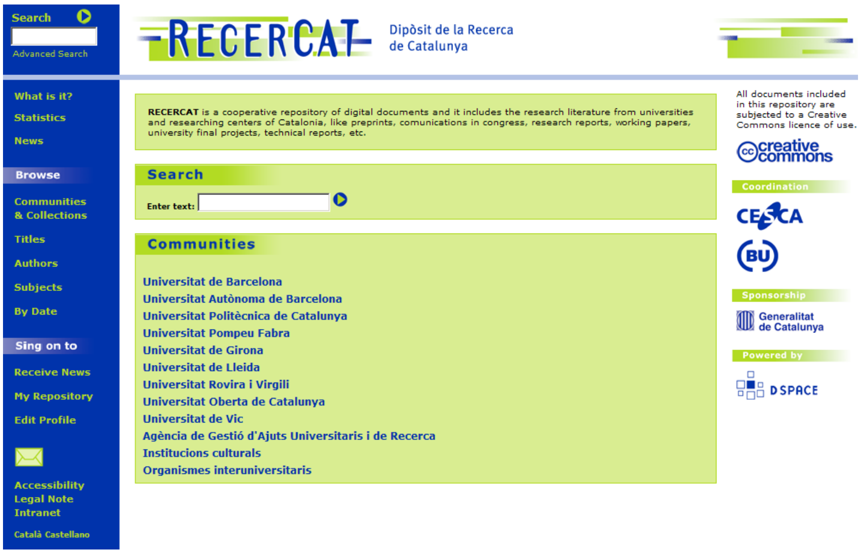
The Dipòsit de la Recerca de Catalunya (RECERCAT) is a cooperative archive of digital documents that includes the research literature of the universities and research institutions of Catalonia, such as heretofore unpublished articles (preprints), conference papers, research reports, working papers, final theses and technical reports.

The main objective of RECERCAT [19] is to increase the visibility of research carried out in Catalonia and to contribute to the world movement for a free archive of academic output and research on the Internet, organized by the institutions that finance research, in order to create alternatives to the model of paying for access to information that has been produced by the institution itself.

The objectives of RECERCAT are:

- To increase the visibility of documents, authors and their institutions, and research produced in Catalonia in general.
- To facilitate publication via self-loading of documents.
- To add value to documents by means of standardised citations, consultation statistics, permanent addresses and preservation mechanisms.

All documents included in RECERCAT are open-access and are subject to the Creative Commons [20] Attribution-NonCommercial-ShareAlike license. This license establishes that the work may be copied, distributed and publicly displayed, provided that the original authors and their institution are cited and that no commercial use is made of the work or any further work built upon it.



RECERCAT is a cooperative repository of digital documents and it includes the research literature from universities and researching centers of Catalonia, like preprints, comunications in congress, research reports, working papers, university final projects, technical reports, etc.

All documents included in this repository are subjected to a Creative Commons licence of use.

Figure 2. The RECERCAT homepage

RECERCAT works with the open-source program DSpace [21]. This software was created by the Massachusetts Institute of Technology (MIT) and Hewlett Packard (HP). DSpace is a groundbreaking digital repository system that captures, stores, indexes, preserves, and redistributes an organization's research data.

This application allows searches by author, title, university, research centre, etc. It also allows you to subscribe to the alert service for a collection, whereby subscribers are informed by e-mail each time a new document is entered.

3.3. RACO

The Revistes Catalanes amb Accés Obert (RACO) is a cooperative repository where can be consulted, in open access, the full-text from articles of scientific, cultural and erudite Catalan journals. While the previous two repositories were built and are coordinated by CESCA and CBUC, for this third one the Biblioteca de Catalunya (BC) joined us.

The main purpose of RACO is to increase the visibility and consults of the journals included and to spread the scientific and academic production published in Catalan journals. This purpose makes specific in three aims:

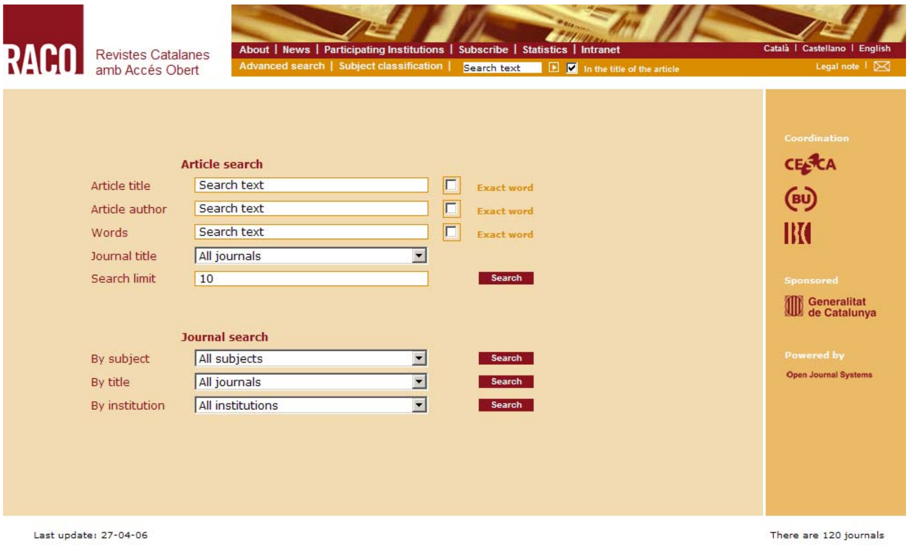
- To encourage the electronic edition of Catalan journals.
- To be the interface that allows the whole search of all the journals.

- To provide the instruments for its preservation.

Journals included in RACO are those that contain a table of contents, that have articles of over 3/4 pages and are signed and those that are published inside the field of a scientific, cultural and/or erudite Catalan institution.

The table of contents and the full-text of articles are introduced in RACO by the own publisher institutions. Most of the journals offer the full-text of all the published issues is offered. Nevertheless, in some journal it can have a delay between the introduction of the table of contents and the full-text.

The full-text articles included in this repository are free access and property of their authors and/or publisher institutions, and therefore, any act of copy, commercialisation, public communication or total or partial transformation needs the express and written consent.


RACO
Revistes Catalanes amb Accés Obert
About | News | Participating Institutions | Subscribe | Statistics | Intranet
Català | Castellano | English
Advanced search | Subject classification |
Legal note
Coordination
CESCA
BU
Sponsored
Generalitat de Catalunya
Powered by
Open Journal Systems
Search
Search
Search
Search

Search text ☑ In the title of the article

Article search

Article title: Search text □ Exact word

Article author: Search text □ Exact word

Words: Search text □ Exact word

Journal title: All journals

Search limit: 10

Journal search

By subject: All subjects

By title: All journals

By institution: All institutions

Last update: 27-04-06

There are 120 journals

Figure 3. The RACO homepage

RACO works with the open source program Open Journal Systems (OJS) [22], a software developed by Public Knowledge Project (PKP) with the objective to promote the access to the investigation, making easy the management and publication of scientific journals. This application allows to do, among others, searches by author and/or title of article and by title, subject or the publishing institution of the journal. Of each article it is possible to consult the recommended bibliographic citation, statistics, metadata that describe it, recommend readings to colleagues, etc. It also offers the possibility of subscribing to the alert service of anyone of the journals included to receive by e-mail the notice of the new published issues.

Two new modules have been implemented at CESCA: a module to classify journals by subject, and second one to simplify the editor's publishing process.

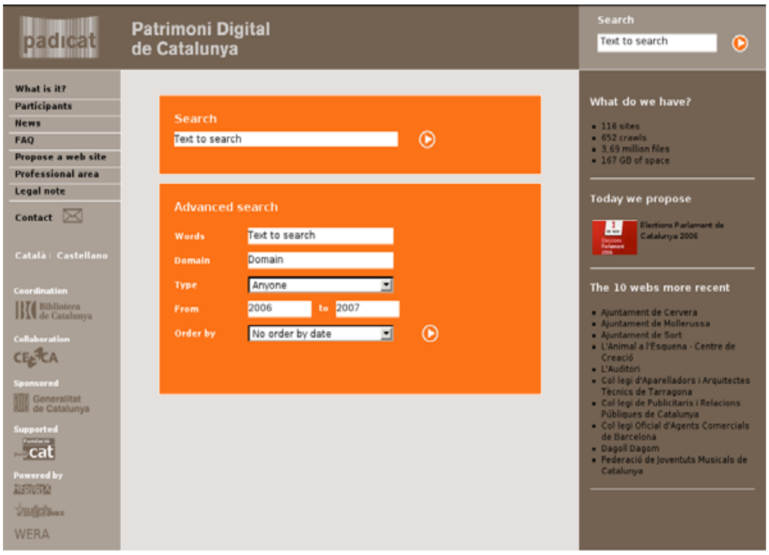
3.4. PADICAT

The main objective of our fourth repository, the Patrimoni Digital de Catalunya (PADICAT) [23], is to archive Catalan web sites. That is, PADICAT collects, processes and provides permanent access to the entire cultural, scientific and general output of Catalonia in digital format. In this case, the repository manager is the Biblioteca de Catalunya (BC), as the institution responsible for compiling, processing and distributing the bibliographic heritage of Catalonia, while CESCA is its technology partner.

In some countries, similar projects are known as "national digital archives" or "web archives". The best known of these are the giant Internet Archive [24], Australian's Pandora [25] and Sweden's Kulturaw3 [26]. PADICAT, as well as these others, is a member of the International Internet Preservation Consortium (IIPC) [27].

In accordance with the general trend among national libraries, the archive model used by BC is a hybrid system consisting of the following:

- Mass compilation of open-access digital resources published on the internet.
- Systematic archiving of the web site output of Catalan organizations.
- Fostering of lines of research through themed integration of the digital resources pertaining to specific events in Catalan public life.



Search: Text to search

Search: Text to search

Advanced search

Words: Text to search
Domain: Domain
Type: Anyone
From: 2006 to 2007
Order by: No order by date

What do we have?

- 116 sites
- 652 crawls
- 3,69 million files
- 167 GB of space

Today we propose

Elections Parlament de Catalunya 2006

The 10 webs more recent

- Ajuntament de Cervera
- Ajuntament de Mollerussa
- Ajuntament de Sort
- L'Animal a l'Esquena - Centre de Creació
- L'Auditori
- Col·legi d'Aparelladors i Arquitectes Tècnics de Tarragona
- Col·legi de Publicitaris i Relacions Públiques de Catalunya
- Col·legi Oficial d'Agents Comercials de Barcelona
- Dagoll Dagom
- Federació de Joventuts Musicals de Catalunya

Figure 4. The PADICAT homepage

On 21 July 2006, work began on the automated collection of web sites that were candidates for becoming part of PADICAT. The first of these were those of the town councils of Berga and Palafrugell, and the professional associations of Quantity Surveyors and Technical Architects of Tarragona and of Social Workers of Catalonia. On September 11th, 2006 PADICAT went into operation for the general public, with some thirty web pages archived.

PADICAT is based on the application of a number of computer programs that allow web pages published on the Internet to be collected, stored, organized, preserved and permanently accessed. Heritrix [28] is the Internet Archives's archival-quality web crawler that harvests and stores, in compressed files, the crawled web pages. Then, NutchWAX [29] generates the indexes that will then be used to search the data. Finally, the Web Archive Access (WERA) [30] is a freely available solution for searching and navigating the archived web document collections.

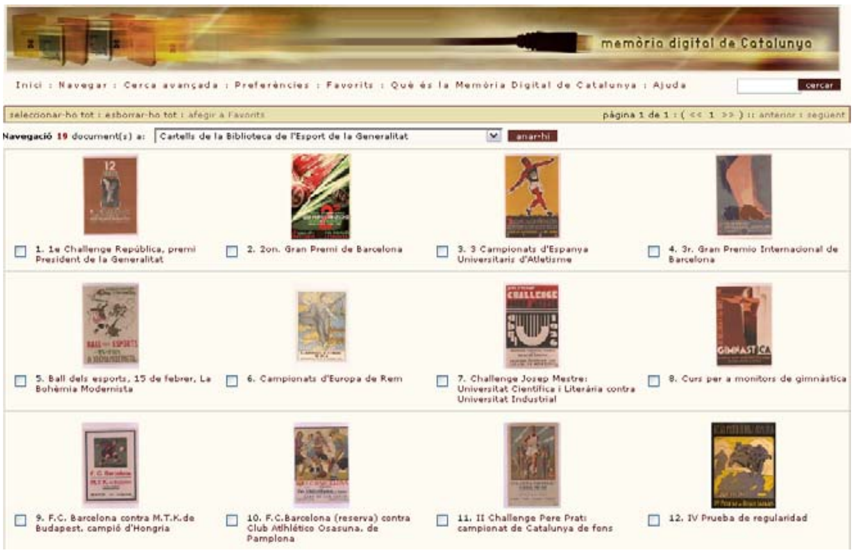
By 2009, PADICAT should be in an optimum position, whereby this system --a pioneer in Spain and a benchmark in Europe- operates at full capacity, with quantitative indicators of 100,000 web pages captured in different editions. This may include some 50 million files and 30 terabytes of data. Furthermore, cooperation agreements are scheduled to be signed with 300 institutions of all kinds and online open access to a considerable part of the collection will be available.

3.5. MDC

The Memòria Digital de Catalunya (MDC), an open-access cooperative repository to preserve digital collections such as old Catalan magazines, pictures, maps, posters, ex-libris… became operational in November 2006.

The main objective of MDC is to increase the visibility and remote consultation of the Catalan cultural heritage. This goal includes:

- To promote the digitalization of the Catalan heritage.
- To became the common interface for this type of consultations.
- To facilitate the instruments for it's preservation.


memòria digital de Catalunya
Inici : Navegar : Cerca avançada : Preferències : Favorits : Què és la Memòria Digital de Catalunya : Ajuda
cercar
seleccionar-ho tot : esborrar-ho tot : afegir a Favorits
pàgina 1 de 1 : (<< 1 >>) :: anterior : següent
anar-hi

Navegació 19 document(s) a: Cartells de la Biblioteca de l'Esport de la Generalitat

□ 1. 1e Challenge República, premi President de la Generalitat
□ 2. 2on. Gran Premi de Barcelona
□ 3. 3 Campionats d'Espanya Universitaris d'Atletisme
□ 4. 3r. Gran Premio Internacional de Barcelona
□ 5. Ball dels esports, 15 de febrer, La Bohèmia Modernista
□ 6. Campionats d'Europa de Rem
□ 7. Challenge Josep Mestre: Universitat Cientifica i Literària contra Universitat Industrial
□ 8. Curs per a monitors de gimnàstica
□ 9. F.C. Barcelona contra M.T.K.de Budapest. campió d'Hongria
□ 10. F.C.Barcelona (reserva) contra Club Athlético Osasuna. de Pamplona
□ 11. II Challenge Pere Prat: campionat de Catalunya de fons
□ 12. IV Prueba de regularidad

Figure 5. The MDC homepage

MDC works with CONTENTdm [31], a digital collection management software developed by DiMeMa of Online Computer Library Center (OCLC).

4. Experiences

After six years of starting our first repository, our experience has been positive and our objectives have been accomplished. In summary,

a) TDR, with 16 participant universities (7 from outside Catalonia), is fully consolidated: It has more than 4,400 theses and it has received more than 3.5 million searches in 2006, 30% of South America, with Mexico in the fist position.

b) RECERCAT has more than 3,100 research documents from 13 institutions.

c) RACO has more than 29,000 articles of 117 journals which belong to 24 editorial institutions.

d) PADICAT has already captured 646 instances of 110 webs.

e) MDC has more than 65,000 images distributed in 11 collections of 7 differents institutions.

If we have a retrospective look at the past, we should confess that the beginning for our first repository (TDR) was not easy: it took more than 16 months to reach the first 500 thesis, while nowadays it takes only about 6 months (see Figure 6).

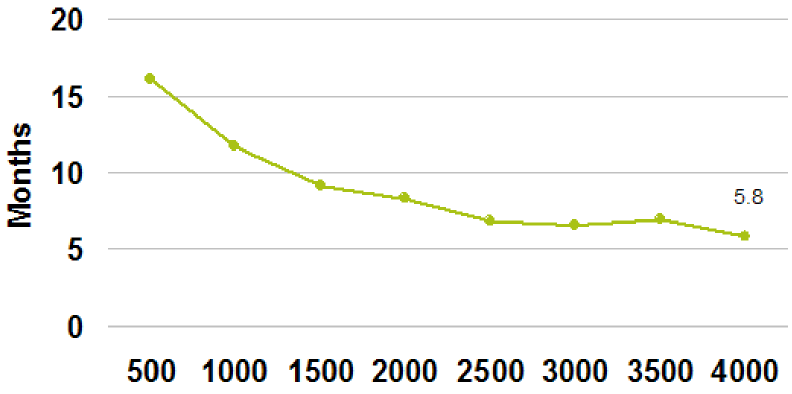
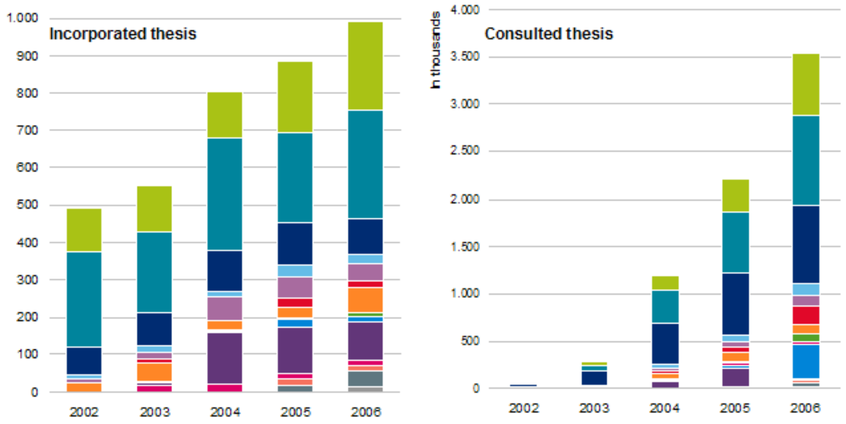


Figure 6. The rhythm of introduced thesis at TDR

As you can see in Figure 7, the ratio of thesis introduced year after year has been improved, not only because of the new participants, but also because the universities succeed in collecting and incorporating them into TDR. Right now, on the average, one of each two thesis being read in the Catalan universities is being introduced into the repository.



%	UB	UAB	UPC	UPF	UdG	UdL	URV	UOC	URL	UIB	UV	UJI	UC	UM	UDC
Incorporated	20	36	14	3	5	3	5	0	0	1	9	2	1	1	0
Consulted	17	28	29	3	3	3	5	0	0	1	9	2	1	1	0

Figure 7. The TDR evolution (2002-06)

The figure also shows the increasing usage rate of thesis looked at. While when we started, TDR was getting about 5,000 requests a month, last November it got the record of 378,930.

In addition, we can also verify that this usage increased has benefited all areas of knowledge. If we look at the TOP30 thesis requested (see Figure 8), during the first three years of operation (2001-03) the most requested theses were from technical areas (27 of the 30), whereas in 2006 this ratio has been near two thirds (19 of 30). Thus, TDR shows the progress and the positive penetration of the Information Society across knowledge areas.

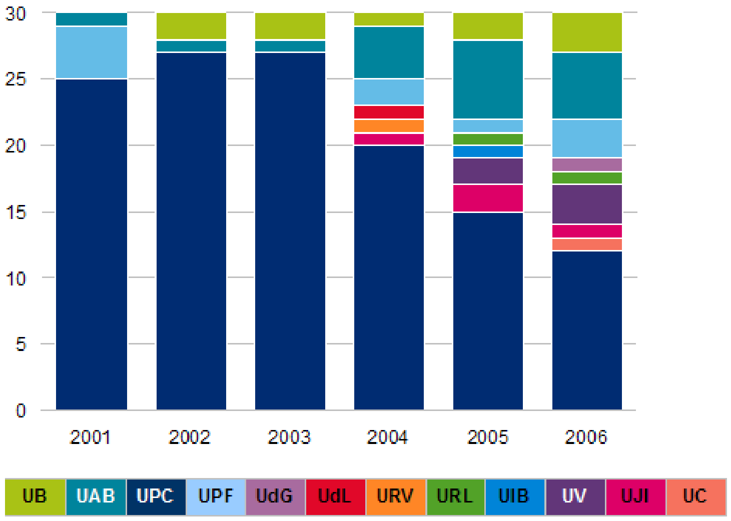


Figure 8. The Information Society penetration factor at TDR

5. Conclusions

By having these five repositories available, it seems that most of our work has been accomplished. This is not true.

Only the ‘easiest’ part of the work has been done. Legal mechanisms have been established to protect the intellectual property of knowledge as well as the necessary e-infrastructures have been set up to allow to spread it through the internet.

From now on, there’s the ‘most difficult’ part to accomplish: that researchers became the best users of this cultural heritage and that contribute to it with their work, that universities and research centers stimulate researchers to do it by establishing the procedures to make it possible; and that the Government and the financial research organizations promote the open access to scientific and humanistic information as the basis of the Knowledge Society and its presence in the internet.

If the information cannot be found on internet, it is as if it had never existed.

Acknowledgments

The developing of these repositories would not have been made possible without the collaboration of many people, both internal and external to our consortia. Internally, the work has been performed by an enthusiastic team who has been successful in bringing each repository alive when was needed in spite of its technical difficulties. Thanks to Marc Boix, Joan Cambras, Nuria Comellas, Maite Novo, Jordi Prats, Sandra Reoyo, Ramon Ros, Marc Terrones, Natalia Torres, and Ricard de la Vega. Also to the BC staff for PADICAT: Eugènia Serra and Ciro Llueca.

Externally, there are too many people to mention them individually. Without the generosity of all the universities that participate in this project, we are certain that we could not have made it possible.

References

[1] *Catalunya en Xarxa: Pla Estratègic per a la Societat de la Informació*, Generalitat de Catalunya, 1999.

[2] Centre de Supercomputació de Catalunya, *http://www.cesca.es*.

[3] Consorci de Biblioteques Universitàries de Catalunya, *http://www.cbuc.es*.

[4] "La universitat, un laboratori de la Societat de la Informació", *Teraflop*, núm. 44, October 1999.

[5] Tesis Doctorales en Red (TDR), *http://www.tesisenred.net.*

[6] Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities, *http://www.zim.mpg.de/openaccess-berlin/berlindeclaration.html.*

[7] Scholarly Publishing and Academic Resources Coalition (SPARC), *http://www.sparceurope.org/Repositories*, *http://www.arl.org/sparc/ir/ir.html.*

[8] Clifford A. Lynch. "Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age", ARL Bimonthly Report 226, 2003.

[9] Open Archives Initiative (OAI), *http://www.openarchives.org.*

[10] OAIster, *http://oaister.umdl.umich.edu/o/oaister.*

[11] Dipòsit de la Recerca de Catalunya (RECERCAT), *http://www.recercat.net.*

[12] Revistes Catalanes amb Accés Obert (RACO), *http://www.raco.cat.*

[13] Biblioteca Nacional de Catalunya, *http://www.bnc.cat.*

[14] Patrimoni Digital de Catalunya (PADICAT), *http://www.padi.cat.*

[15] Memoria Digital de Catalunya (MDC), *http://www.cbuc.cat/mdc.*

[16] Networked Digital Library of theses and dissertations (NDLTD), *http://www.ndltd.org.*

[17] Electronic Theses and Dissertations (ETD), *http://etd.vt.edu.*

[18] WebGlimpse, *http://webglimpse.net.*

[19] Sandra Reoyo et altri. "RECERCAT: El Dipòsit de la Recerca de Catalunya", 10es Jornades Catalanes d'Informació i Documentació. Col·legi Oficial de Bibliotecaris-Documentalistes de Catalunya, 2006. *http://hdl.handle.net/2072/2195.*

[20] Creative Commons (CC), *http://creativecommons.org.*

[21] DSpace, *http://www.dspace.org.*

[22] Open Journal Systems (OJS), *http://pkp.sfu.ca/ojs.*

[23] "Memòria del plantejament del projecte PADICAT", Biblioteca de Catalunya, 2006. *http://hdl.handle.net/2072/1757.*

[24] Internet Archive, *http://www.archive.org.*

[25] Australia's web archive (PANDORA), *http://pandora.nla.gov.au.*

[26] Kulturalw3, *http://www.kb.se/kw3/ENG.*

[27] International Internet Preservation Consortium (IIPC), *http://netpreserve.org.*

[28] Heritrix, *http://crawler.archive.org.*

[29] NutchWAX, *http://archive-access.sourceforge.net/projects/nutch.*

[30] WERA, *http://archive-access.sourceforge.net/projects/wera.*

[31] CONTENTdm, *http://www.dimema.com/products.*