

UPCommons: an institutional repository and the public domain

Ruth Iñigo (Universitat Politècnica de Catalunya, ruth.inigo@upc.edu)

Anna Rovira (Universitat Politècnica de Catalunya, anna.rovira@upc.edu)

This article is licensed under a Creative Commons Attribution 3.0 unported license available at <http://creativecommons.org/licenses/by/3.0/>

ABSTRACT

Institutional Repositories give the opportunity to faculties and researchers from universities and research institutes to freely publish and facilitate open access to their publicly funded research activities results. There is also a good chance for scholars and research communities to highly increase their visibility in the world and their impact. For University libraries this represents the opportunity to document, organize and preserve the intellectual heritage of the institution at the same time as it increases its prestige.

Furthermore, publishing in *UPCommons* is one of the indicators used to evaluate the performance of strategic plans of the Research and Academic Units.

UPC libraries have developed different repositories to offer a tool to the university community a tool to publish their academic and scientific works in open access.

- E-prints UPC (<https://upcommons.upc.edu/e-prints/>) collects documents generated by academics in their research activities. Content is organized around communities which can correspond to departments, research groups or institutes.
- Revistes i Congressos UPC (<https://upcommons.upc.edu/revistes/>) accommodates full text of e-journals articles and proceedings published by any unit of the UPC (institutes, departments, etc.).
- Theses and dissertations Online is a digital cooperative repository of doctoral theses presented at some Spanish universities managed by the Consortium of University Libraries of Catalonia (CBUC). Universities taking part are responsible for editing and uploading theses and dissertations to the repository.
- Academic works collects, in digital format, the final academic works (final degree projects/works, minor theses, recognition of foreign diplomas tests, etc.) presented by university studies at UPC.
- Opencourseware is a repository inspired by the MIT Opencourseware. It is a web-based electronic publishing initiative with the goal to provide free, searchable access to UPC's course materials for educators, students and self-learners and extend the impact of UPC opencourseware and all the opencoursewares around. The repository grants access and preservation to the course material from now on.
- The Digital Video Library contains a selection of the available video recordings of the University (academic lessons, conferences, etc.)

- The Graphic Archive of the School of Architecture of Barcelona preserves part of the documentation generated through the academic activity of the institution throughout its history. The collection includes both architectural projects and drawings dated from the 19th century and current academic works.

Several services developed by the UPC Library through recent years have become strategic within UPCommons Project. One of them is the Intellectual Property Service (SEPI): UPC libraries offer information and guidance about rights and copyright policies to the authors (academics, students, etc.) of the documents published at UPCommons. SEPI website provides authors with answers to FAQs regarding common aspects of creation, dissemination and publication of academic and research works.

TABLE OF CONTENTS

1. Introduction
2. Extending “open access” at UPC
3. Creating the institutional repository and the digital library of orphan works at UPC:
 - Theses
 - Videos
 - Student’s works
 - Journals and congresses
 - E-prints
 - Course materials
 - Old collections
4. The Intellectual Property Service
5. Problems, questions
6. Conclusions
7. References

1. Introduction

The libraries' mission is to gather, preserve and disseminate knowledge to ensure life-long information to citizens, not only in printed but also in electronic form. In the last two decades and because of the explosion of information technology, lots of digital libraries and institutional repositories projects have arisen. But copyright aspects and other topics related to intellectual property have reduced the capacity of putting all the documentation collected by libraries on the web and disseminating it.

This paper describes the experience of the UPC libraries in the creation of an institutional repository hosting the output of academic activities and the development of a digital library of orphan works after their digitalization.

After the presentation of both projects, the institutional repository and the digital library of orphan works, an explanation of the problems and solutions to topics related to copyright and the re-use of information will be realised.

2. Extending “open access” at UPC

Since 2002, the University Library has led various initiatives aimed at enlarging the movement for open access to scientific information into the University community.

In 2002, UPC became a member of SPARC (Scholarly Publishing and Academic Resources Coalition), and the University Library, together with the University of Barcelona (UB), took on the job of translating into Catalan and publishing the *Create Change* leaflet. This brochure was based around the idea that the way the faculty shares and uses academic research results would change rapidly and irreversibly thanks to the open access movement. It also offered practical ways faculty could look out for their own interests as researchers.

In 2003, the Third REBIUN (Red de Bibliotecas UNiversitarias) Workshop was held at the University. It also focused on changes in scholarly communication: it was attended mainly by librarians, but there were also some lecturers and managers from Spanish universities. Noteworthy speakers were included like Jean-Claude Guédon (University of Montreal), David Prosser, (director of SPARC-Europe) and Christopher Gutteridge (an IT specialist from the University of Southampton and an expert in e-prints software), among others. This workshop also allowed UPC librarians to be engaged with the OA movement; from this point, they started to act promoting the open access to scientific information.

The following year, the University Library planned a new action concerning the dissemination of the Open Access movement in the University: a touring exhibition at the UPC campus called “Towards a new scholarly communication” was organized. Using text and images, the idea was to show the situation of scholarly communication especially to the academics, the philosophy of open access and also worldwide initiatives to bring about changes in the current publishing system (World Digital Mathematics). Therefore, the aim of the exhibition was to raise the awareness of the most important people in the communication scientific system: the writers, who are very often university lecturers. Once the exhibition had been set in all campuses, its contents

could also be visited online at http://bibliotecna.upc.edu/e-portals/comunicacio_cientifica/.

In 2005, the University Library decided to create an institutional repository. Since 2002 some electronic publications (such as theses, videos, academic works) written by authors of the university have been collected and hosted in the libraries' webpages. At this point, it was necessary to create a project which could host all types of academic publications, as will be described in the following part of the paper. It was also important to get political support for this project. Because of that, we presented it to all the vice-rectors to involve them with the projects and to get more support to convince faculties. Furthermore, in 2007, the Rector of UPC signed the Berlin Declaration (http://oa.mpg.de/openaccess-berlin/berlin_declaration.pdf) to institutionally give support to the Open Access movement.

3. Creating the institutional repository and the digital library of orphan works at UPC:

Institutional Repositories give the opportunity to faculties and researchers from universities and research institutes to freely publish and facilitate open access to their publicly funded research activities' results. With the same objectives, the main goals of *UPCommons* (<http://upcommons.upc.edu/>) are to organize and preserve the intellectual heritage of the Universitat Politècnica de Catalunya and to offer a tool to publish their academic and scientific works in open access to the university community.

UPCommons facilitates a unique access point to all different open access repositories (theses, e-prints, journals and conferences published/organized by UPC, course materials, academic works and videos) but the different parts of the repository were created independently, launching a new repository every year: theses in 2002, videos in 2003, academic works in 2004, revistes i congressos in 2005, e-prints in 2006 and learning materials in 2007.

2006 was also the year of the launch of *Upcommons*, the portal which gives access to all open repositories and which allows jointly and separately searches in all of them.

Referring to the technology used, and after having analysed E-prints and DSpace (an open-source digital archiving system designed by MIT libraries and Hewlett Packard to capture, manage and share research in digital formats) the second software was selected for most of the UPC digital repositories. Managing interface was complex but this software offered different profiles or roles with different tasks: editor, reviewer, etc. which was very useful in this context.

In the following lines, some technological and organizational characteristics are described:

- The DSpace system's information model is built around the idea of organizational communities –natural sub-units of an institution that have

distinctive information management needs. In the case of UPC, communities are defined to be departments, schools, labs and research groups....

- The process of submission includes filling out information about the item on a metadata form and uploading the file(s) comprising digital item. Each community sets its own submission policy.
- Reports and other academic documents are filed by teaching staff via an automated system. The repository has permanent accumulative open access and allows different levels of visualisation.
- The Dublin Core and additional administrative metadata are used to label documents, following the specifications set out by the Consorci de Biblioteques Universitàries de Catalunya (CBUC).

Each community can include a description of their activity (research, learning, etc.) their logo and the structure of their collections.

Until now it has not been mandatory for authors to archive their works in any repository. However, UPC commons is being integrated with other UPC information systems or services (like the virtual campus, the information system for the research output, academic managing tools, electronic ID systems etc.). This fact makes UPC repositories very useful and we are encouraging the university community (professors, researchers, students) to archive their works. By now, there are about 9.800 documents in the repositories hosted in *UPCommons*. It is recommended that authors use a Creative Commons License.

In the following part of this paper, there is a description of each part of the institutional repository. The last repository described, which is the newest one, does not actually collect institutional output; its aim is to host all the books digitalised by UPC libraries, which normally are orphan works.

- Theses

TDX (Tesis Doctorals en Xarxa) was the first document storage facility that UPC participated in: theses and dissertations Online is a digital cooperative repository of doctoral theses presented at some Spanish universities (20 in total) managed by the Consortium of University Libraries of Catalonia (CBUC).

The aims of the TDX are to make the results of university research available around the world via the Internet, to offer the authors of the theses a tool that increases the profile of and access to their work and to encourage the creation and use of the University's scientific output and to improve bibliographic control of the theses.

Universities taking part are responsible for editing and uploading theses and dissertations to the repository. The works are copyrighted by contract: every author who

publishes his theses in the repository signs an agreement and after that, the theses is incorporated into the server by the staff of his university.

Last year a digitalization project started within TDX: among 400 theses presented in recent years which previously only existed on paper, were converted to digital format and archived in this repository. Up to now, there are more than 6.000 theses in the repository.

In this case, -which is an exception-, the software has been designed by the Networked Digital Library of Theses and Dissertations (NDLTD) of which TDX is a member. CBUC is also participating in the DART-Europe project (<http://www.dart-europe.eu/>) the objective of which is to improve global access to European research theses, through the creation of a web portal.

- Videos

The digital Video Library contains a selection of the available video recordings of the University. Some lecturers ask the library to record a course lesson. Or when a conference takes place there is also a videorecording service. After the signature of an agreement, videos are archived in the repository and made available through the web. Up to now, there are 226 videos in this repository.

- Student's works

This institutional repository collects the final academic works (final degree projects/works, minor theses, recognition of foreign diplomas tests, ...) presented by university students at UPC. All University schools and faculties are participating in this project. Up to now, there are about 2,194 academic works in the repository.

- Journals and congresses

This part of the institutional repository (<https://upcommons.upc.edu/revistes/>) accommodates full text of e-journals articles and proceedings published or written by any unit scholar or research department of the university. Contents are organized in e-journals or items and can be searched or browsed by author, title or date. By now, there have been located about 3,808 articles or congress papers in this repository.

- E-prints

The E-prints repository (<http://upcommons.upc.edu/e-prints/>) aims to provide access to the work of UPC research groups. The material is collected, organised and stored in the repository with no restrictions on access rights, so it can be shared with other researchers working within the same specific field (for example, broadband networks, aerial vehicles, etc.)

E-prints collects documents generated by academics in their research activities: reports, articles, conference reports or lectures, etc. Content is organized around communities which can correspond to departments, institutes or research groups.

Each research community receives support from a subject specialist librarian who informs the faculties about the internal organisation of the collections in the repository, provides advice on publishing rules and copyright and updates the metadata.

Up to now there have been about 1,300 research publications in this repository.

- Course materials

UPC learning objects repository (<http://e-md.upc.edu/home/>) is part of the OpenCourseWare consortia, promoted by the Massachusetts Institute of Technology (MIT) and Universia (hispano-portuguese universities network <http://www.universia.es>). It is a web-based electronic publishing initiative with the goal to provide free, searchable access to UPC's course materials for educators, students and self-learners. It is also trying to extend the impact of UPC opencourseware and all the opencoursewares around. The repository grants the access and preservation to the course material from this point on.

There are about 5,882 educational resources in the repository, mainly exams from previous courses and lecture materials.

In addition, there is another initiative which is being held, unrelated to the intellectual heritage of the institution, but to the old books collected in the libraries. Since last year a project of creating a digital memory repository has generated more doubts and questions to add to the ones related to the institutional repository as we will see in the very last part of the paper.

- Old printed collections

As it was already been said, UPC libraries also have important collections of old documents, not only books but also journals, maps and other minor documents. The majority of them are orphan works. For different reasons (document delivery, for example) sometimes those books are digitalised and the library archives the electronic copy in the digital library.

But digitalization projects are very expensive for libraries. Because of this, it is important to get extra funding. Last year, the UPC library received a grant from the National Government. This gave us the opportunity to start the creation of a digital library for old books with the following objectives:

- To start a project that enables the digitalization of all orphan works
- To host all the old scientific collection of documents in a digital library
- To promote access and use of old scientific collections

It is planned that, at the end of the year this digital library will contain about 250 digitalized books, for about 70.000 pages.

4. The Intellectual Property Service

Several services set up by the UPC Library over recent years have become strategic lines of action within the *UPCommons* Project. One of them is the Intellectual Property Service (SEPI), through which UPC libraries offer authors (academics, students, etc.) of documents published by *UPCommons* information and guidance about rights and copyright policies. The SEPI website provides authors with answers to common FAQs on the creation, dissemination and publication of academic and research works.

UPC's Intellectual Property Service (SEPI) was set up in January 2005 and is run by library staff. The main objective of the Service is to provide information and guidance to the members of the university community on the basic principles of copyright rules and regulations, especially with regard to the information that is made available to them via UPC's library services.

There is clearly a need for guidance on the subject of copyright, as well as answers to some of the doubts that currently exist about the use of scientific and technical information: basic guidelines on copyright legislation and the uses members of the university community are allowed to make of the information available on *Biblioteca*, UPC's digital library.

The service is complemented by other UPC services, such as the Legal Office (legal advice), the Technology Transfer Centre (industrial property) and Edicions UPC (which deals with the copyright of the material it publishes).

It is important to highlight that the Service provides general guidance and makes the necessary documents available (regulations, jurisprudence, protocol, etc.) with regard to copyright, but it does not provide legal advice.

The SEPI is staffed by one librarian from each of the UPC libraries, as well as a librarian from the Library Service.

Users of the SEPI generally send their requests for information via e-mail by filling in a specific form. Users who ask library staff—either directly or over the phone—for information on topics related to this area are redirected to the SEPI.

The SEPI always sends a report in response to a request for information. This contains recommendations and guidelines for the case in hand. The reply also includes references to current legislation and to the regulations on exploitation rights and confidentiality that apply to UPC research and teaching works.

5. Problems, questions

Digital publishing has led to organisations taking on new roles, for example universities now act as editors, etc. Consequently, authors, the institutions in which they work, and consumers of information must find new solutions to existing problems.

The establishment of institutional repositories has meant that libraries, which traditionally managed collections and provided services, have had to take on editing

tasks. The process of learning about electronic publications and the job of disseminating information about the new channels for publishing have raised library staff's awareness of many of the most common issues that affect open access publication.

Below is a summary of the main questions received by the SEPI in its two years of operation. The responses are also provided. The questions are arranged into different areas.

- Use of excerpts and parts of other works when you are writing teaching materials and creating derivative works
 - The Spanish Intellectual Property Law (IPL) establishes that under certain circumstances¹ citations or reviews may include excerpts from original works. These circumstances include teaching and research activities. The purpose of the citation or review must coincide with those defined by the law.
 - A derivative work is created when contents are translated. Therefore, works may only be translated with the author's permission.
 - The SEPI recommends using materials that have licences that allow them to be used. In addition, always ask authors for permission to use their information and always cite sources.

- Authorship and ownership of exploitation rights for works created by UPC members
 - The Spanish IPL clearly states that students have sole exploitation rights for their dissertations and other assignments undertaken throughout their studies. The SEPI has informed both authors and the different university departments of this fact. In particular, it has stressed that authors must give their authorisation or transfer their rights to be able to publish any kind of academic work in open access repositories.

- Protection against plagiarism
 - Many authors are concerned about plagiarism. In fact, the chance of being plagiarised is one of the main reasons why authors are unwilling to publish their work in open access repositories.
 - On occasions, UPC teaching materials have been plagiarised or used commercially by private academies. However, these practices were also common before the advent of digital formats, when materials were disseminated in printed form only. Although the misuse of teaching materials has been detected more frequently than the inappropriate use of other UPC documents, the repository of teaching materials contains open

¹ Article 32. It is legal to include excerpts from other written, sound or audiovisual work, as well as isolated examples of plastic art, figurative photography, or similar, when these works have already been disseminated. In addition, such works must be cited and may only be included for the purposes of analysis, comment or criticism. This law only applies to works created for teaching or research purposes and the number of excerpts included must fall within the bounds of reason. The source and name of the author of a quoted work must also be included.

access documents as well as documents that are only available within the UPC domain.

- Some new authors (mainly doctors who have just defended their doctoral theses) are particularly concerned about the illegal use of their materials, and are not as motivated by the advantages of open access, such as greater visibility of their research results.
- Authors often ask about the scope of intellectual protection. In addition to providing information about the possibilities of the current legal framework and the use of licences, the SEPI advises authors on how to proceed in areas that are not covered by legislation. In addition, authors frequently ask about protecting “ideas”, including ideas that are the starting point of works that have not yet been undertaken; ideas developed within published works (such as the hypothesis of a thesis, an experiment carried out as part of research, etc.); or patentable inventions (this issue is related to industrial protection). One of the SEPI’s recommendations to authors who wish to demonstrate the originality of ideas that have not yet been developed is to publish them in work-in-progress papers.
- Licences specify and restrict the permitted uses of materials. Authors are increasingly aware of the impact of free access on their work. Nevertheless, there is a real risk of plagiarism and the misuse of open access documents. Once inappropriate use has been detected, authors can take legal action. However, this is a long and expensive process that often involves several countries’ legislation and administration of justice. The results are always uncertain. Thus, the SEPI is promoting a procedure through which UPC will seek legal advice and intervene, if so requested by an author.

- Requesting and using identifying numbers

- In the Spanish university system, calls for applications (for tenders, funding, grants, etc.) usually require researchers to send their curricula in a standard format, with a detailed list of their published works. Each publication must be listed with an identification number, such as the ISBN, the ISSN or the DOI. The different characteristics of each of these systems and the lack of specific information on the criteria used by the evaluating agencies to assess curricula leads to these identifiers frequently being erroneously considered as “quality indicators”.
- In addition, many people believe that identifiers serve to “register” the authorship of an article and to “protect it” against plagiarism and misuse.
- When the SEPI receives a question about identifiers, whether this is a request for information or for initiating the process of attaining an identifier, the Service’s role is initially to provide information about the different types of identifiers, the reasons why they were created, and the cases in which each one should be used. In addition, the SEPI provides information on new identifiers such as those related to publishing in web repositories, e.g. handle identifiers.

- Use of Creative Commons licenses
 - Most UPC repositories use Creative Commons licences, even though it is still not compulsory to do so. The SEPI provides information about the different uses of these licences and on how to apply them to different document types.

- Publication of works in *UPCommons*
 - The SEPI informs all authors about journals' self-archiving policies. The Service either checks the Sherpa/Romeo website, or puts authors in touch with the editors of the journals in question.
 - In addition, the Library Services is creating a wiki to store information on the self-archiving policies of journals, conferences and publishing companies that are not included on the Sherpa/Romeo website (for example, any non-English journal). This wiki will also include the authorisations received by UPC to deposit published articles written by UPC members.
 - In addition, the SEPI has drawn up forms that can be used to authorise the dissemination of contents via *UPCommons*. These forms must be signed by authors and the holders of rights.

- Publication in the public domain and orphan works
 - As mentioned above, UPC owns historical collections of documents that are gradually being digitalised and made available on *UPCommons* or through other websites managed by the UPC libraries (<http://bibliotecnica.upc.edu/bib160/colleccions/poesia/home/home.asp>). This is the case of Science and Technology historical collection of the School of Industrial Engineering of Barcelona (ETSEIB) <<http://bibliotecnica.upc.edu/bib240/serveis/fhct/inici1.asp>>, the Arxiu Gràfic of the School of Architecture (ETSAB) <<http://eprints.upc.es/adminbustia/bib210/>> and the historical collection of the Faculty of Nautical Studies of Barcelona (FNB) <<http://bibliotecnica.upc.edu/e-portals/fons/>>. These collections cannot be digitalised until we have checked whether copyrights are still in force.
 - Obtaining biographic information about the authors of some of the works in these collections is a laborious task. It can be difficult to determine whether a work is in the public domain, who currently owns the copyright and how to get in touch with this person.
 - One interesting case involves the Rector Gabriel Ferraté Library's collection of Catalan poetry.
<<http://bibliotecnica.upc.edu/bib160/colleccions/poesia/home/home.asp>>
The historical core of this collection comprises 5,000 works that were mainly published between the late 19th century and the early 20th century. In addition to books by major authors, the collection contains works by little-known authors. These authors cannot be found in

reference works and no experts have studied them. In addition, many of the archives of the publishing companies or presses that published these works have disappeared. In response to this problem, the SEPI aims to set up a procedure for declaring a work as “orphaned”.

- Company archives are a type of collection that have been donated to our libraries by certain companies based in Catalonia. However, there are no plans to digitalise these collections as yet. Before disseminating these works, we must determine who owns the exploitation rights.
- Video recordings
- These materials include recordings of classes, conferences and institutional events that are available in open access repositories in *UPCommons*. In all cases, the consent of the speakers must be sought, and, if necessary, that of the other people who appear in the images.

6. Conclusions

As they provide a public service, the mission of university libraries is to make their collections as accessible as possible. Institutional repositories and online libraries can help to bring the academic world closer to society, by making scientific and teaching output, as well as heritage collections, available to the general public.

Repositories also help to attain another main goal of libraries: to preserve information for the future by including it in the public domain.

As public institutions, universities must be committed to both the general public and to members of the university community.

The digital environment is bringing about a redefinition of the public space, and is focusing attention on the boundaries between the common and individual interests that converge on the Internet. These are open and permeable boundaries; individuals may find themselves on one side or the other, depending on the role they are playing at any particular time (author, consumer, etc.). Through the management of repositories and online libraries and the questions asked of the SEPI, the libraries are in a position to observe situations in which conflicts of interest arise.

- Conflicts of interest between the public and private sector
 - Universities cannot always disseminate the knowledge generated by their members. When projects are carried out by universities in collaboration with (or cofinanced by) private companies, the companies usually insist that the project remains confidential for a certain period. In addition, they normally demand the transfer of exploitation rights from the author.
 - In their regulations, universities can set maximum periods of confidentiality, during which works cannot be disseminated.

- Conflicts of interest between authors
 - Members of the same institution can reuse teaching materials created in their department, thus optimising the work of the teaching staff. However, all of the authors involved must reach a consensus on who will be the official authors and how the materials will be reviewed.

- Conflicts of interest between institutions and authors
 - Two situations may arise in relation to the publication of academic works and exploitation rights:
 - Normally, only some dissertations are published (the most interesting, those with the highest marks, etc.). Therefore, many students do not have the opportunity to publish their work.
 - Some public universities have decided to make it compulsory to publish all academic papers, doctoral theses, etc. produced by university members in open access repositories. According to Spanish legislation, this cannot be done unless rights are transferred to the universities by the authors. Nevertheless, departmental regulations, which must be signed by students on registration, may contain a point on open access publication. This has led to the practice of obliging students to give their consent to make their work freely available.

- Conflicts of interest between creators and consumers of information
 - Finally, the attitude of users of open information may vary when they change roles. Internet is the main source of information in the academic world, which is often done through the exchange of information. It is also a way of accessing material about entertainment and leisure.
 - However, what happens when consumers of information on the Internet become creators?
 - According to recent studies, 35% of Internet users in the United States also post materials that they have created on the Web, such as photos, videos, music and blogs. However, how is the public domain constructed? Do users contribute as much information as they take?

7. References

Jones, R., Andrew T., MacColl J., 2006. The institutional repository, Chandos Publishing, 2006.

Spain . *Law 1/1996, of April 12 Consolidated Law on Intellectual Property, regularizing, clarifying and harmonizing the applicable statutory provisions (and amended by Law 5/1998 of March 6, 1998, incorporating Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the Legal Protection of Databases).* <http://www.wipo.int/clea/en/details.jsp?id=1373> [June 2008]

Create Change <http://www.createchange.org/index.shtml> [june 2008]

European Union *Directive 2001/29/EC of 22 May 2001 of the European Parliament and of the Council on the harmonisation of certain aspects of copyright and related rights in the information society*

<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32001L0029:EN:NOT>
[june 2008]

Lenhart, A. Content Creation Online. Pew Internet & American Life Project Home
Broadband Adoption: 2006 report from Pew Research
http://www.pewinternet.org/pdfs/PIP_Content_Creation_Report.pdf [june 2008]