

Time for a change: the Italian CRUI-Open Access Working Group's action for a national e-theses provision service

Paola Galimberti, Marialaura Vignocchi

Università degli Studi di Milano, CIB – Alma Mater Studiorum – Università di Bologna

Abstract

The paper intends to present the status of e-theses collections in Italy, focusing on the major drivers for change that can contribute to create the basis of a national e-theses provision service able to collaborate with other international services. Finally, it illustrates the e-theses project of Bologna University as a case study of a viable integrated system offering added-value services.

Until very recently electronic doctoral theses provision has been an initiative of a few universities or of single departments in Italy. So coverage has been poor and the vast majority of current theses and older ones exist only in paper. Discovery and access services have not been much satisfactory and doctoral theses have been largely underused.

The situation is now changing rapidly. The major initiative has been taken by the Open Access Working Group sponsored by CRUI (Conference of the Rectors of the Italian Universities). The OA CRUI Working Group intends to provide Italian universities with guidelines and a tool-kit offering practical and legal support for managing and making theses available via Open Access IRs. Indeed, technology to set up IRs is relatively simple and most Italian universities have, or are planning to implement, one. The new legislation on legal deposit of publications is another driving force towards the development of e-theses collections. The lobbying action of the OAI CRUI Working Group is accelerating the adoption of new procedures for the legal deposit of electronic theses at the National Libraries based on OAI-PMH. These procedures will improve central discovery and access services and will guarantee accurate long-term digital preservation programs for e-theses. Finally, they will bring about a simplification of administrative work, which is greatly contributing to a reconsideration and a repositioning of IRs within the academical workflow and administrative information systems. Another driver for change is represented by the new spreading culture of research evaluation. Research assessment has begun to determine the funding system of departments and faculties generally appreciate initiatives that increase the visibility of their outputs, provided that IP protection is guaranteed. A national e-theses provision service will also enable Italy to participate in the larger context of European research.

In September 2006, Bologna University has launched the “AMS Tesi di Dottorato” project that has achieved a functional integration between the administrative information system of the Doctoral Management Office and the E-prints facilities developed to host the collections of the University Digital Library (Alma-DL). Doctoral students do the self-archiving into the repository reducing data-entry operations of administrative staff, while bibliographic metadata are enhanced by librarians. Administrative staff retain control tasks in order to guarantee that patents and IP rights are respected. Legal deposit is performed via OAI-PMH, while print-on-demand is provided as an added-value service to authors.

Introduction

In Italy doctoral theses are documents with a double nature: they have both an administrative and a scientific value. As administrative documents, doctoral theses complete a study course and must be properly archived by university administration offices together with documents

certifying a student's career. At the same time, as outputs or milestones in research projects, doctoral theses lay between pre-prints and scientific publications. They are not subjected to a formal peer-review as it is normally performed in the publication process, but they are subjected to Faculty tutoring and to the final approval of an examination commission, appointed on a national level, that certifies the originality and the quality of the contribution. Moreover, even though they are not normally intended for publication, doctoral theses have been deposited according to legal disposition at the two National Central Libraries in Florence and in Rome since 1987, as it normally happens to publications in Italy. [1]

Doctoral theses collections in Italy

The results of a recent survey carried out by the Italian OA-CRUI WG on behalf of the European ETD group have shown that, in Italy until very recently, electronic doctoral theses provision has been an initiative of a few universities or of single departments.[2] Treatment of doctoral theses varies greatly. In a few very advanced cases, all theses are collected in electronic format in an institutional repository. In some isolated cases nothing is being done at all and theses are collected and stored in paper copy by administrative offices that do not normally grant discovery and access services. Most universities however appeared to be in a start up phase at the time of the survey, planning to set up an institutional repository or adapting existing ones to include doctoral theses. Indeed, technology to set up IRs is relatively simple and most Italian universities have, or are planning to implement, one by themselves or with the support of inter-university consortia offering technological infrastructure and know-how. [3] Finally, in a few universities, digital format for doctoral theses is mandatory and the students self-archive their work, in others, librarians do the archiving.

The results of the survey have shown that coverage of e-theses has been poor and the vast majority of current and older theses exist only in paper or microfiches, which is beginning to cause problem of storage and preservation. The increasing cost of paper copies storage has become a great problem for Italian universities and National Central Libraries that have the preservation of doctoral theses among their institutional and administrative duties. On the other end, microfiches used to reduce storage costs have rapidly deteriorated and have often become unreadable.

Discovery and access services have not been much satisfactory as well. At a national level, the National Central Libraries in Florence and Rome should guarantee bibliographic and access services for all deposited theses. Unfortunately cataloguing has never been provided promptly as a consequence of the huge annual production. In addition, bibliographic records can be searched in the OPAC of the National Central Library of Florence [4] only since doctoral theses, regarded as grey literature, are not admitted in the National Library Service on-line catalogue. The National Central Library in Rome maintains a print catalogue only. Finally, access to the full content of theses is possible only in the National Libraries rooms since theses are excluded from interlibrary loan, lending and reproduction services.

At a local level, discovery and access services greatly differ from university to university. In some cases the bibliographic data are available through the OPAC of the institution, in others through internal hardly visible catalogues or lists. In most cases cataloguing and access services are an isolated initiative of single departments. As a consequence doctoral theses have been largely underused and their public, full-text availability has begun to be perceived as a goal to be achieved with urgency.

Drivers for a change: the action of the OA-CRUI WG

The situation seems to be mature for a change. The international discussion about doctoral theses and their increasing importance within the Open Access movement, together with the emerging role of e-science and data sharing in research practice, have brought forth new perspectives. In Italy, the major initiative for a change has been taken by the Open Access Working Group that was created within the Library Commission of the Conference of the Rectors of the Italian Universities (CRUI) in 2006. [5]

The interest of Italian universities in European innovative projects, and the need for a central coordination of single initiatives taken in favour of open access to research publications have led to the creation of the working group formed by delegates from the most important Italian universities. The task of the OA-CRUI WG is to develop guidelines and provide recommendations to encourage Italian universities to guarantee an open access to the knowledge they produce and to induce institutional stakeholders to go beyond mere commitment to Open Access principles and to formulate concrete official policies in support of their IRs and open access publications - as it is recommended in the roadmap redesigned by the Berlin 3 Open Access Meeting in Southampton. [6] The agenda for 2006-2007 included issues such as management and sustainability of open access journals, open access research impact and evaluation, e-theses archiving in institutional repositories. The WG split into smaller groups to work on all different strands simultaneously and carry out its task in a relatively shorter time.

The e-theses group has intended to provide Italian universities with the guidelines and a legal tool-kit offering practical support, for managing, collecting, preserving and making e-theses available via institutional repositories. Through an active collaboration with the two National Central Libraries in Florence and Rome, the group has tried to establish a new procedure for legal deposit of theses in digital format. This initiative is going to create the basis for a national e-theses provision service able to interoperate with other international services.

Legal issues

Legal issues have been judged of primary importance and have required a qualified legal advice. Italian legal provisions define doctoral theses as “public documents” subjected to legal deposit, as a consequence institutions do not need to be authorized to publish doctoral theses in an OA IR. [7]

Nonetheless there remain a few critical aspects related to the public availability of doctoral theses:

1. ownership of copyright;
2. theses reporting information relating to patents;
3. third parties involved (i.e. sponsoring companies);
4. use of third party owned materials;

The adoption of electronic theses requires some measures regarding the rights that authors intend to give away (i.e. non-exclusive license). As a consequence, it is important to establish who is the copyright owner - whether the author, the institution or eventually a third party.

The archiving, preservation and dissemination of doctoral theses have always been a librarian’s task. Libraries have always asked authors for permission to publish the abstract and the bibliographic data of theses independently from their format (paper, microfiches, CDs). Doctoral students have also been asked to authorize access to the full content of their

theses, even though public availability of all paper doctoral theses at the two National Central Libraries is mandatory by legal provision. The theses have however remained largely inaccessible since, as we have seen, the discovery and access services maintained by the National Central Libraries have not been very effective.

The electronic environment needs also to take into account some other aspects that are not covered in the agreements regarding paper theses treatment. For example the universities will have to maintain the right to copy and transform files for preservation and dissemination on the Internet.

The legal advisor of CRUI has prepared a document containing guidelines for the Italian universities. This document can be summarized in the following points:

- as most universities have signed the Berlin Declaration [8] and as doctoral theses can be regarded as research outputs, it is recommended an institutional mandatory policy for the archiving of e-theses in IRs in order to make them accessible and increase their impact;
- university regulations for doctoral courses should be changed to inform applying students that their works will be stored in IRs;
- it is reasonable to admit exceptions to the immediate OA availability. For example, when theses contain patentable discoveries, students should be given the chance to apply for a patent. In addition, if research is financed by a third party, copyright issues should be clarified in the contracts. In these cases, students will be asked to archive both metadata and full-text, but metadata only will be immediately available while the full text after a 12 month embargo;
- for ongoing courses, students will have to declare to have been informed about the new procedures adopted by universities and no permission will have to be required from them;
- it is possible that theses contain third party material under copyright. In these cases students will have to either ask for third party permission or give up using the material.

The new legislation on legal deposit of publications is another driving force towards the development of e-theses collections in Italy. [9] The law dates back to 2004 and it includes digital documents of scientific and cultural interest produced by research institutions and distributed via Internet in the list of documents subjected to legal deposit.

The lobbying action of the OA-CRUI Working Group is accelerating the adoption of new procedures for the legal deposit of electronic theses at the National Central Libraries. The OA-CRUI group and the National Central Libraries have established a submission protocol for doctoral theses in digital format. Metadata and full-content of doctoral theses can be directly uploaded into the National Central Libraries repository or alternatively can be harvested from existing IRs by the central service via OAI-PMH. To guarantee files integrity and authenticity a MD5 hash is to be produced and attached to each deposited document, while notifications of legal deposit procedures will be sent via certified e-mail.

These procedures will improve central discovery and access services and will guarantee accurate long-term digital preservation programs for e-theses. Conservation and preservation are mandatory duties for National Central Libraries. Their active

participation into international projects for digital long-term preservation guarantees that the problem, even though complex, is being dealt properly. The creation of a central trusted repository managed by the National Central Libraries and updated automatically or manually by single institutions will provide more efficient and effective bibliographic services granting metadata coverage and access at least to current production.

In order to create a central, national gateway to Italian e-theses production, metadata standardization is of vital importance to ensure interoperability and integration of the different systems both on a national and international level. The CRUI group has asked the National Central Libraries to produce a proposal for a common national metadata set for e-theses which should include policy and rights management statements. On this matter the CRUI group recognises the importance of the experiences made at a European and international level in order to adopt shared solutions and contribute to the standardization process. In particular, simple DC does not seem to be enough and a richer and more specific profile - as it is for example contained in the Knowledge Exchange metadata set - will have to be adopted. [10]

The E-theses Group recommendations

At the moment, the e-theses group has agreed on the following recommendations:

- Italian Universities have the faculty and should include self-archiving of doctoral theses in IRs within the set of rules establishing institutional policies on doctorate courses;
- all universities should deliberate similar solutions;
- self-archiving is preferable to archiving by librarians, so authors remain responsible for their work integrity;
- all doctoral theses will be submitted to, or harvested by, the National Central Libraries which are responsible for legal deposit, long-term preservation and national discovery and access gateway;
- technical solutions for integration and interoperability should be guaranteed by the adoption of standard protocols and metadata.

The group has also expressed a generalized need for European intervention and common guidelines. The group believes that syntactic and semantic interoperability is very important as well as search/browsing systems by country, language, discipline, subject, keywords. The “next step” in the roadmap of the Italian OA-CRUI e-theses group is to join the European e-theses group to contribute to a broader European e-theses project.

“AMS Tesi di Dottorato”: Bologna University project for open access doctoral e-theses

The project for the creation of a collection of current e-born doctoral theses developed at Bologna University within the services of the University Digital Library (Alma-DL) [11] with the collaboration of the Research Management Office and the Information System Services is an example of how times are mature for a change. The convergence of positive drivers has indeed contributed to start a reform process from which main institutional stakeholders - researchers, administration managers and librarians - can benefit along with the international scientific community.

Started in September 2006, the “AMS Tesi di Dottorato” project [12] has achieved a functional integration between the administrative information system of the Doctoral Management Office and the GNU E-prints infrastructures developed to host the collections of the University Digital Library (Alma-DL). The metadata relevant to the description of doctoral e-theses are imported via web services from the administrative information management system into the GNU E-prints 3.0 repository dedicated to doctoral theses. Doctoral students do the self-archiving into the repository, further reducing data-entry operations of administrative staff, while metadata, especially bibliographic ones, are enhanced by librarians. Administrative staff retain control tasks in order to guarantee that patents and IP rights are respected.

Legal deposit may be performed by automated harvesting via OAI-PMH as soon as the National Central Libraries have completed the technical implementation of the service. At the moment legal deposit is performed “at request” on account of an agreement signed with the National Central Library of Florence. The National Library launches the harvesting procedure of the metadata and full-text of documents deposited in the IRs after receiving an e-mail request containing details and the URLs of the documents intended for legal deposit.

“Print on demand” is offered as an added-value service. At the moment it is provided by local publishers who have signed an agreement with the University. According to the agreement publishers can associate their trademark with the University logo on publications and include them in their catalogues but are obliged to renounce to copyright claims and are asked to offer a reduction on print service price. Editorial support is excluded from the agreement and files should be provided ready for print by authors.

The project has received a good acceptance from university managers and administrative staff since it brings about a simplification and a reduction of administrative work. The project has demonstrated how IRs can integrate with, support and add value to the activities of the institution, which is greatly contributing to a reconsideration and a repositioning of IRs within the academic workflow and administrative information systems. As a consequence we hope that it will receive long-term financial and institutional support beyond test phases.

The open access repository, still in a test phase, has been also appreciated by doctoral schools and research groups who consider it an opportunity to increase visibility for their activities, provided that IP protection is guaranteed. Visibility and research impact and evaluation are indeed important issues for faculty who have been lately called to account for outputs of their research projects more and more frequently. In Italy, as in many other European countries, research assessment has begun to determine the funding system of institutions and departments. In 2006 the assessment body of the University of Bologna has established the open access availability of doctoral theses as one of the evaluation criteria for quality of doctoral schools. Hopefully, “AMS Tesi di Dottorato” will provide doctoral schools with the technological solution and the organizational support to meet the requirement.

References

[1] Legge n. 28, 1980

[2] <http://www.ukoln.ac.uk/repositories/european-e-theses/index/Italy>

[3] There are 24 institutional repositories in Italy. Theses and dissertations cover the 66 % of the total amount of content types. <http://www.opendoar.org/index.html>. CILEA is the consortium offering infrastructural support and services. <http://www.cilea.it/>

[4] <http://www.bncf.firenze.sbn.it/cgi-opac/opac.cgi?Lingua=ITA&unicode=F>

[5] <http://www.cruis.it/link/?ID=1812>; <http://bfp.sp.unipi.it/moinwiki/CategoryCruis>

[6] <http://www.eprints.org/events/berlin3/outcomes.html>

[7] Legge n. 28, 1980; DPR n. 382/1980; D.M. n. 224/1999

[8] <http://oa.mpg.de/openaccess-berlin/berlindeclaration.html>

[9] Legge n. 106/2004; DPR n. 252/2006

[10] <http://www.knowledge-exchange.info/>

[11] <http://almadl.cib.unibo.it/>

[12] <http://almadl.cib.unibo.it/aree-di-sviluppo/open-access/dottorato/>