Study on Perspectives Regarding Deposit on Open Access Repositories in the Context of Public Universities in the Central-Eastern Region of Argentina

Patricia Silvana San Martín, Paola Carolina Bongiovani, Ana Casali, & Claudia Deco
National University of Rosario

Abstract
This article presents the results of a survey and qualitative study of needs and practices of open access scholarly works dissemination. The survey and study focused on different issues related to usability, navigation, and accessibility of institutional repositories, which have guided the design of an experimental prototype in the context of a regional project that joins three universities in Argentina. In such academic practices, the teaching community offers various types of production and simultaneous application areas for digital objects. Thus, there is a technological requirement to enable users to carry out their self-archiving process of different objects, in a user-friendly way, allowing them to upload these objects in various collections. A prototype was implemented and tested in order to obtain an initial assessment of the proposed model.

Keywords
Higher education; Open access repositories; Educational digital objects; DSpace; Professors

Résumé
L'article présente les résultats d'un sondage et d'un diagnostic qualitative sur les besoins et les pratiques de publication en accès libre. Il fait le point sur différents problèmes d’utilisabilité, de navigation et d’accessibilité rencontrés dans les archives institutionnels qui ont motivé la conception d’un prototype expérimental dans le cadre d’un projet regroupant trois universités en Argentine. Dans ces pratiques académiques,
la communauté d’enseignants propose des types de production diverses et des champs d’application simultanés pour un même objet numérique. Il y a donc un besoin technologique pour l’enseignant-chercheur d’auto-archivage et de description de son objet numérique à travers une interface conviviale qui prenne en charge les champs d’application simultanés. La mise en place et le test du prototype permettront une première évaluation du modèle proposé.

Mots clés
Enseignement supérieur; Archives ouvertes; Objets numériques éducatifs; DSpace; Enseignants

Introduction
The development of open access (OA) institutional repositories (IRs) at public universities in Argentina is a priority in the framework of policies of the Ministry of Science, Technology and Productive Innovation and the Interuniversity National Council (Miguel, Bongiovani, Gómez, & Bueno, 2013). To this end, since the first quarter of 2012, six research and regional interuniversity development projects related to open access institutional repositories have been under way, involving professors from the majority of public universities in the country.

The project titled “Towards the Development and Use of Open Access Repositories for Educational Digital Objects in the Context of Public Universities of the Central-Eastern Region of Argentina” (PICTO-CIN) has been designed in the organizational framework of the National University of Rosario (UNR), Litoral National University (UNL), and National University of Technology (UTN), district of Santa Fe and Rosario. It focuses on designing and proposing an initial technological, methodological, and theoretical model for the IR of OA of the Educational Digital Object (EDO). This project contributes to the knowledge of:

Organizational, pedagogic, social, and technological aspects that involve the development of socio-cultural strategies that promote the appropriation of an IR starting from the study of difficulties such as usage, accessibility, information display, search, and other problems related to the use of those repositories in the referred community.

Innovative open source computing tools: processes and technical procedures dealing with descriptive, technical, and structural metadata representation, storage process, metadata automatic extraction, design of users interface, accessibility for the disabled, and interoperability.

The adopted methodology is based on concepts, methods, and epistemological grounds of interdisciplinary research in the framework of complex systems (García, 2007). Production processes and social construction of technology are inseparable and cannot be disassociated from the relevant procedures and local context, not only in terms of technological innovation, but also in terms of cognitive development (Thomas & Buch, 2008). The technological development premise holds that the use of a device or technological knowledge is present in the design of such a device and in the significant technological processes in which different relevant social groups participate. For this
study, the research team is made of 18 people, including professors, researchers, and advanced students of the participating institutions, who come from fields such as Education Science, Information Science, and Computer Science, among others. They constitute an interdisciplinary group with the appropriate profile to deal with the difficulties mentioned above, building a common theoretical framework.

The academic community’s teaching practices include the production of the scientific, technological, and artistic activities of the university, together with a growing variety of educational resources in different formats. Thus, our concept of EDO is inclusive of that variety. The reality of the Argentine university context is linked to the production and performance profile that the higher education system requires of academics, assigning the role of professor to those who also perform research in their fields. Given the different types of required production and the simultaneity of application fields of one object (scientific, technological, cultural, and educational), this project assigns the EDO to the meta-type level, representing the original meaning coming from the practices and educational objectives of the public university. The goal of the project contributes to the generation, preservation, and dissemination of open, free, and accessible knowledge for all Argentinian citizens.

Considering the three participating universities, only two have IRs, and their contents are mainly science and technology production. Thus, there is a need to properly integrate the general aspects of interactivity with the system, including the complex description of the object to be stored, and to support the needs and publication practices of OA in the academic community. It is important to mention that until the beginning of this project there had not been any other systematic studies of open access scholarly works dissemination in the Argentine context, which might have recommended specific ways to meet the expectations of the university community.

In this article, we first present our theoretical and methodological perspectives. This is followed by an analysis of the results of the survey performed on scholarly works dissemination needs and practices of OA. The study found various interactivity barriers within the IR system, which influenced the most important design decisions proposed for the development of a First Experimental Prototype (FEP). Finally, we present conclusions and opportunities for future research, taking into consideration that since the implementation of the FEP, a first evaluation of the theoretical, methodological, and technological perspective proposed may help to validate this model.

**Brief theoretical and methodological perspectives**

To systematize the needs and barriers for the development of OA IR in our region, considering their organizational, pedagogical, social, and technological aspects, studies were performed in a synchronic way, during the May to November term of 2012. The objective of the studies was to determine the level of information, needs, and practices regarding OA dissemination of works in the academic community and to detect barriers to the uptake of an IR. This research was undertaken via a qualitative sampling of professors at the participating universities, following the perspectives and recommendations suggested by Thomas, Fressoli, and Santos (2012), about the socio-technical dynamics of inclusion/exclusion. Referring to the specific IR software technologies, issues regarding usability, accessibility in relation to disability, deposit
process, and objects retrieval were addressed, analyzing and evaluating DSpace, since the only two existing repositories at UNR and UNL use such technology.

Researchers from Argentina have a positive opinion toward OA publishing (Bongiovani, Gómez, & Miguel, 2012), but there have been no existing local studies that could address other dissemination forms of different types of academic production. The configuration of multidisciplinary OA repositories, from the point of view of the practices of the professor, requires a conceptual framework, underdeveloped until this time in the context of Argentine universities. We observed the need to handle different types of heterogeneous resources along with the needs of different distribution, description, and treatment. These repositories, apart from helping in the successful management of EDO, could fulfill the needs of a strategy for the management, exchange, and reuse of OA resources.

With regard to the PICTO-CIN project being OA IR centred within academic production at the university, it was essential to undertake the design and implementation of a diagnostic study with the aim of outlining the promotion and training actions. These actions allow the development and appropriation of this kind of technological resource under the OA perspective by the university community, as well as carrying out the organizational and infrastructural development necessary to support the IR. The treatment of some of the problems related to the use of IRs, self-archiving, search and retrieval of EDOs, and the focus on accessibility standards are essential, since users’ basic interactive actions are set up with the system through a communicative interface space, which should enable all functions expected of an OA IR.

Following Bueno de la Fuente (2010), regarding OA IR that contains EDO, recommendations about the design of a quick, simple, and intuitive deposit process were taken into account, ensuring that the addition of metadata is as easy as possible, requesting authors provide the basic metadata that only they can offer. Other important aspects set in the international recommendations are as follows: the incorporation of automatic and semi-automatic metadata extraction assistants (Casali, Deco, Romano, & Tomé, 2013); the use of quality metadata standards to facilitate interoperability, preservation, and appropriate information from the EDO; and the maintenance of support services and help on specific aspects. The idea is to design intuitive and efficient interfaces, as well as provide a smart search of resources, and to offer tools and useful services to the individual authors and practice communities, strengthening the interactive cooperation and creating new benefits for the educational quality. Finally, a necessary condition for this cyber-infrastructure is the support of consistent and robust software and hardware.

In reference to the possibilities of integrating the physical-virtual context of the practices that constitute a free and public university, the theoretical-methodological concept of the Dynamic Hypermedia Device (DHD) addressed by San Martín (2008), which was implemented in the Virtual Campus of the National University of Rosario, offers a comprehensive view of the resources and a contributing working environment, based on the creation of socio-technical networks of diverse and responsible participation to produce, manage, and educate. This device combines a variety of appropriate technologies—systems with different uses—in direct relation to the
requirements, possibilities, and social-organizational aspects of the institution. In this way, the “DHD Campus Virtual UNR” model integrates different collaborative environments for education, socio-technological bonding, IR, and other systems that offer their services to the whole UNR community and other organizations.

**Diagnostic study and proposal**

This section discusses the most significant results of the survey of professors undertaken as part of our study. In relation to the diagnostic study performed, the proposals for the development of the FEP are addressed last. From the socio-technological framework, the model presented conceptualizes the professor as the actor performing the self-archiving process of works, which requires not only robust and accessible tools, but also the development of an interactive, open, collaborative, and plural academic culture.

**Survey of professors**

The first phase of the project was an online survey, which was designed, implemented, and aimed at the academic community in the central-eastern region. The questionnaire was made up of 28 questions, distributed in different sections:

- **General Information:** Institution, University School/Faculty, Education Level, Position, Age, Main Thematic Area (complying with the areas of the Evaluation System of Scientific and Technological Projects)
- **Online Publications:** Opinions about the online dissemination of academic works, sites where online availability is offered, types of disseminated works, person in charge of uploading contents to different sites, and reasons to offer free access to this content
- **IR:** Use of OA IR, motivations to submit works to IRs regularly, and valued services to be offered by IRs
- **Scientific Data:** Production, characteristics, types, formats and software used, access, motivations to share the data, and interest in support services
- **Final Questions:** Comments about the survey and the option to fill in a form with name and email address

Invitations to complete the survey were distributed via institutional mailing lists of the UNR, UNL, and UTN, including the National University of Entre Ríos (UNER). In total, 1,009 responses were received between July 5 and November 14, 2012.

The most significant results discussed here are related to the decisions made for the technological development of the FEP. Other complementary data were collected by the research group, analyzing various sites that offer OA publications related to the teaching and/or research activities in the framework of the above-mentioned universities.

Overall, out of the 1,009 responses obtained, 786 were from professors at UNR, 95 from professors at UNL, 106 from professors at UTN (Districts of Rosario and Santa Fe), and 25 responses from UNER. It is worth mentioning that UNR has organized several promotion activities of its “RepHip UNR” IR. Since 2009, weekly workshops have been open to the community, with seminars and other free academic activities in...
the framework of OA Week. Regarding academic units represented in the survey, professors may belong to more than one academic unit within universities, resulting in figures that exceeded 100%. Concerning education attainment of survey respondents, it was found that all the respondents hold a bachelor’s degree and nearly 50% of them had obtained postgraduate degrees as well. Most of the survey respondents were between 26 and 55 years old.

After the general information questions, the survey focused on queries referring to online availability of academic production. The meaning of “free online accessibility” was clarified within the questionnaire as referring to free access in a digital format to complete academic works on the Web (for instance, the whole text of a published article, and not just a summary). The respondents were asked whether they would like their works to be freely accessible online, and 68% answered positively. This number, added to the people whose works were already freely accessible (13%), represents an 81% positive attitude toward this form of the dissemination of academic work. However, 16.5% indicated that they were uncertain and 3% that they would not like their work to be freely accessible online (Figure 1). Professors who answered that their works were already freely accessible online were asked where these works were available: 35% stated that they were in open access journals; 29.5% offered access through the research centre, laboratory, or group website; and 28% used the university website to offer free access to their documents.

In addition, these same respondents were asked what kinds of work were freely accessible online. Among the most common types were articles from scientific-academic journals (63%), works presented at conferences (55%), class notes (42%), guidelines (37%), and course programs (36%), among others. In an exploratory study, it was verified that works disseminated via websites of universities, groups, or academic classes lacked, to a high degree, the basic metadata necessary to identify and/or retrieve such digital objects. However, a great number of links were not persistent, and some of the work or class material identified by the students or the professors of the group were based only on additional information and not provided clearly on the website. A vast majority of the surveyed objects did not take into account accessibility standards for people with disabilities.

As for the person who submits the works, 71% of the respondents said it is a personal responsibility, whereas 30% delegated that job to members of the work team; the administrative staff (15%) and other teams (journal editors, congress organizers, webmasters, and students) who also participate. This information is relevant because it indicates the reality regarding the low availability of library staff who support the academic work in this regional framework, and it also reflects the skills and issues of professors to perform such tasks. Empirical assessments conducted by our research team, over what has been deposited in different repositories by researchers, found a high percentage of incomplete, inaccurate, or erroneous metadata submission.

Representing 76% of respondents, the most important motivations driving authors toward offering OA to their works, as shown in Figure 2, was “I believe in the open

Figure 1: Consensus of free online availability of academic works (N = 776)
access principle (open and free of charge access to the academic material) for all the users.” This was followed by 36% who chose the option “Placing my works in open access is beneficial for my students.” This view is mainly focused on the benefit to the readers, inclusively of students, restating the adopted perspective about the importance of a free, open, and public university, one which supports the adopted IR profile in the context of participating institutions, as indicated in the first section of this article.

All the respondents were asked whether they had submitted academic works to any IR. To this question, 44% answered that they ignored the existence of an IR, 30% answered that they had not submitted any work to an IR, 15% answered that they have submitted works to an IR (or another person had on their behalf), and 9% answered that they did not have an IR available at their institution. These results show that even though promotional activities about the benefits of OA and IRs have been performed, there is still a great deal to do regarding the development and appropriation of these technologies in the given context.

The survey also asked about what would motivate people to regularly submit their works to an IR. The main driver (73% of the respondents) was that the repository should be more frequently used by professors, students, and the general public. Increased visibility of their works would motivate 52% of the respondents. Likewise, having a simple, easy-to-use system would motivate 34% of the respondents to regularly submit works to an IR (Figure 3).

Being able to choose several answer options, the respondents gave their opinion about services they considered were more valuable in an IR. Services were as follows: smart search (63%), reports including access and download statistics of the works (50%), automatic update of their publications lists (38.5%), links to the works in the repository from the academic website (35.9%), training in the use of an IR (35.6%), digital preservation services for the works (35.6%), among others (see Figure 4).

Usability and accessibility barriers
As mentioned before, it has been very important to establish the main barriers to authors’ self-archiving in the available IRs. Due to the context of this publication, we address here only the user’s experience as a
professor in the use of an IR. In this regard, we performed a case study in 2012 about the UNR institutional repository.

Our diagnostic methodology was based on the Contingent Model of the Analysis and Evaluation of the Interfaces Design (MC-AE is the Spanish acronym) (Rodríguez Barros, 2013). This model accepts as explicit aims the possibility of exploring, describing, comparing, taking into consideration the interface, and contributing to its design as part of the researchers' participation. The evaluations were carried out by members of the research group; by members of the Virtual Campus of the UNR; by members of the department of integration and inclusion for the disabled, associated with the National University of Rosario Extension Office; and by professors.

The tasks were focused on submission of different types of works, using a test repository (mirror) that had equal characteristics to the repository in production (DSpace version 1.6.2). Researchers identified the doubts and barriers that occurred during the process and accounted for the effectiveness, efficiency, and satisfaction of the use, including visual impairment, among others.

As regards the Web accessibility issue, the Reference Centre in Web Accessibility of the National Office of Information Technology issued an evaluation report requested by the research team. The evaluation was performed in the home page of RepHip UNR, using only the TAW and eXaminator automatic evaluation tools. It is worth mentioning that certain accessibility obstacles cannot be automatically detected and the results of these tools can be inaccurate in some cases. For this reason, these results should not be considered definite, but they should be present in the evaluation process.

Main barriers that users encountered while performing the submission tasks were identified as:
- Orientation and visualization impediments in the interface to carry out initial and essential steps to start the submission, for instance, login options and upload of files. These included orientation difficulties to complete file upload and locate the storage place, since an IR does not allow the visualization of the file on the screen, nor does it indicate the persistent file link; instead, this information is sent to the user via email.
- The interface introduces inconsistent terminology regarding the semantic comprehension for users in the regional context, for example, in the first step of submission, “multiple titles” options to add translations or the “Save/Cancel” button. There are also repetitions of terms in the top bar of the storing place, for instance, “license” and “description.”
- Disorientation regarding the question of what is the minimum meaningful required element in the process of describing the object. For instance, the first stage of the step “describe” presents essential metadata (Author, Title, and other key information) along with other less relevant metadata, such as Series/Number of Report. The metadata fields for teaching materials are scarce, for example, the proper indication of the age range of intended users is not allowed. The descriptive fields of the works are inaccurate, for instance, when editing or creating a collection, the legend “Copyright Text (html).” The aids are insufficient and at times ambiguous, for example, when editing a work the field “origin” is not clear and no explanation is provided.

Prior to the beginning of the development tasks in the installed version of the test server, the most recent demo version of DSpace was analyzed. This analysis included evaluating all the submission process pages. Apart from using two automatic tools (eXaminator and AChecker), manual revisions were made by software engineers, who enabled us to imagine what users would do with the interface, specifically. Problems were detected at level “A,” precluding a disabled person from uploading files on the repository.

Toward a proposal of technological development
In order to develop a first prototype, based on the theoretical, methodological, and technological frameworks analyzed in the preceding paragraphs, three areas of development have been outlined to work simultaneously and in interrelationship with one another.

Submission process: This refers to the redesign of the workflow of DSpace, incorporating an assistant for the automatic and semi-automatic download of metadata, able to contribute to the description process of the EDO.

Required (mandatory) and non-mandatory metadata: In this development area, we carried out an analysis of the metadata to describe and upload the EDO, taking into consideration the National System of Digital Repositories Guidelines (SNRD is the Spanish acronym) and the internationally accepted standards. In particular, Dublin Core and LOM (Learning Object Metadata) standards were used.

Interface design: It elaborates a proposal at the level of users for the submission process, meeting the surveyed requirements, possibly contextualized in the repositories of the
Next, we describe the principal changes introduced with regard to these development areas in some details.

Submission process: When initiating the submission, the pending and concluded submissions are visualized, and users are asked to choose from among the Communities and/or Sub-communities to which they belong, a site where they would like to upload their work. Afterwards, a collection will be selected to start the submission. For this OA IR, 12 pre-selected collections were settled on as the main purpose of EDO: Articles, Books, Book Chapters, Journals, Bachelor/Postgraduate Thesis, Communications, Reviews, Research Production (research reports), Technology Production, Extension Production, Art Production, and Teaching Material.

After the work is chosen, the process consists of five sequential steps to finish the submission successfully: 1) authorize the institutional licence and/or grant a Creative Commons licence type, if requested; 2) attach the main file and any related one(s), if applicable; 3) introduce mandatory metadata; 4) introduce non-mandatory metadata, optionally adding metadata to categorize the work as teaching material, if any of the other options were chosen (in this case, the remaining mandatory metadata and other options can be completed, if desired); otherwise, follow the last step: 5) check the submission (verify/change metadata).

The original submission workflow of DSpace has significantly changed. Users (professors) should be aware of legal frameworks and rights and give their consent before performing any action (standard procedures, both in physical and virtual spaces). Another change has been the introduction of a wizard that will automatically extract general metadata from the main file, such as title, authors, summary, keywords, and language. This wizard is an independent module that includes various tools for information extraction.

Based on previous results (Casali, Deco, Bender, Fontanarrosa, & Sabater, 2013), the research team decided to use ParsCit and Alchemy tools to implement the prototype. So far, only text files have been used with this wizard. Steps 3 and 4 enable the user’s validation of the extracted information, as a 35 to 40% error rate is still detected in some metadata fields.

Metadata process: for each type of work to be uploaded, following the SNRD guidelines, a list of mandatory and non-mandatory metadata has been analyzed, and other metadata, ensuring a successful search. This meticulous task has taken into consideration any work performed, pondering the unique and diverse profile that could arise; for example, in the production of art, technology, and teaching materials (Dublin Core and LOM-ES standards were used). As the type of work was already fixed in many cases in the selection of the default collection, where appropriate, subtypes have been added. Metadata collection tables have been developed to specify the requirements in the design of the system. For each metadata field, the following items
are included: metadata name, standard used, range of values, explanation about the fields and examples, among others.

**Accessible interface design:** Based on the survey results, the interface proposal for the submission process has been designed with a didactic orientation. The focus was on the simple understanding of the five loading steps, the mandatory requirements needed in each instance to continue, and the clear location of the sequence of steps in the process. The design interface consists of a top menu, which indicates the screen-step relation, the corresponding number on the left, and brief aids in the different metadata fields. Multiple option dropdown menus have been used to reduce uploading data errors, where applicable. Three levels have been analyzed in regard to to accessibility obstacles: those purely technological in nature, for example, complying with the specification of the HTML tags; those dealing with the interface use, for instance, to provide a tab order which makes browsing easier, colour contrast; and those obstacles that depend on the context of use. In the third level, real users work in specific contexts of use, while automatic and manual evaluations are performed as well. At this point, the team found it very valuable to receive input from a visually impaired member of the work group, who is an advanced user of assistive technology.

**Conclusions**
This article summarizes the real situation of the use and development of OA IR, in the context of public universities, located in the central-eastern region of Argentina. It also analyzed the barriers that the DSpace system use presented. Our research revealed that these barriers could be overcome, maintaining interoperability and adapting standards. These problems do not represent, in their majority, a technical difficulty, but a paradigm revision of technological development. In this sense, the common interdisciplinary theoretical framework, which constitutes the investigation and development group and the detailed socio-technical dynamics analysis, provides the foundations that enable the innovation processes in the context of practice. The diagnostic data of this project may also be worthy for governmental organizations, regarding future design and development of training policies, research, and social and technological linkages of the region.

The current online implementation of the First Experimental Prototype (FEP) is promoted through open and free workshop activities, mainly in the UNL and UNR, especially directed to the professors community in the region. These activities include in-depth observations and the participants’ interviews, which provide invaluable input for improvement of the FEP and actions that promote IR appropriation.

Regarding the development of institutional abilities, this regional project tends to consolidate the construction and effective operation of a new concept of physical-virtual campus. It integrates Communications and Information Technology (TIC is the Spanish acronym) services offered for the production, use, storage, and dissemination of academic works and forms of collaborative online work. In this sense, the progressive change in the organizational culture that promotes this type of integrated project, through a responsible and participating socio-technical network, constitutes a highly significant antecedent to consolidate supporting policies of the IR of OA and, in a wider sense, the appropriation of the TIC in all public universities of Argentina.
Acknowledgements

This work was partially supported by the PICTO-CIN II Project No. 0143 (ANPCyT-CIN), CONICET and UNR, Argentina. We would like to thank professors who participated in the studies and research of the project “Towards the Development and Use of Open Access Repositories for Educational Digital Objects in the Context of Public Universities of the Central-Eastern Region of Argentina” and the project “Towards a Theoretical, Methodological and Technological Model for the Open Access Institutional Repository of UNR Based on Its Academic Community Practices.”

Notes


2. The term EDO has been strengthened by its normative use by the Spanish Association for Standardization and Certification in the elaboration of Spanish profile application of LOM-ES educational metadata (Spanish Association for Standardization and Certification/AENOR, 2010).

3. This refers to the categories known by professors in the online forms of national management and evaluation systems (categorization, career accreditation, CV, etc.).

Websites

AlchemyAPI. http://www.alchemyapi.com
Campus Virtual UNR. http://www.campusvirtualunr.edu.ar
DSpace Demo Repository. http://demo.dspace.org/xmlui/
National University of Rosario (UNR). RepHip UNR. http://rephip.unr.edu.ar
National University of Technology (UTN). http://www.utn.edu.ar/default.utn
ParsCit. http://wing.comp.nus.edu.sg/parsCit

References


