

The management of an international open access repository: the case of E-LIS

Management eines internationalen Open Access-Archivs: das Beispiel E-LIS

Abstract

E-LIS is the largest open access repository in the field of library and information science and is maintained voluntarily by an international team of librarians and information professionals. As from April 2009, it contains at about 9000 full text documents in 37 languages from more than 5600 authors from 90 countries. Additionally to the provision of services to authors and associations in the field, the management of policy issues is crucial for the repository administration. Thus E-LIS has, inter alia, completed a policy audit and intends to formulate and communicate its policies in a standardized way.

Keywords: open access, subject repository, library and information science, organization, policies

Zusammenfassung

E-LIS ist das größte Open Access-Archiv für das Bibliotheks-, Informations- und Dokumentationswesen und wird von einem internationalen Team aus dem Bibliotheks- und Informationswesen ehrenamtlich betrieben. Mit April 2009 sind rund 9000 Dokumente im Volltext von über 5600 AutorInnen aus 90 Ländern in 37 Sprachen verzeichnet. Neben dem Anbieten von Services für AutorInnen und Fach- und Berufsverbänden ist das Management bezüglich der Grundlagen und Richtlinien von besonderer Bedeutung. E-LIS hat sich deshalb unter anderem einem entsprechenden Audit unterzogen und bemüht sich um eine standardisierte Kommunikation der Richtlinien und Grundsätze.

Schlüsselwörter: Open Access, thematisches Archiv, Bibliotheks- und Informationswissenschaft, Organisation, Richtlinien

Introduction

The open access (OA) movement is probably the most prominent and promising approach to achieve a sustainable, alternative system for scholarly communication. In the narrower sense OA refers to scholarly, peer reviewed literature that is made freely available on the Internet. In the broader sense it means providing free access to all scholarly and academic material, including presentations at conferences, datasets, theses, books, etc. Apart from the publication in OA journals, the depositing of scholarly materials in OA repositories by the authors (self archiving) or on behalf of them (mediated archiving) is a basic strategy to achieve OA. Whereas institutional repositories contain the intellectual output of one or more institutions or institutional units, subject or disciplinary repositories include documents from authors from around the world covering one or more subjects [1].

In the following article we exemplify the management of a centralized disciplinary repository by E-LIS (<http://eprints.rclis.org>). Because general information on E-LIS has already been given in the relevant literature [2], [3] we concentrate on recent efforts and developments. First we present an up-to-date snapshot of the content deposited in E-LIS. Following that we illustrate the organizational model of E-LIS. Next we show how authors, publishers and professional or academic associations in the field of library and information science (LIS) can use E-LIS for their purposes. Finally, we address some issues concerning policy management that are important for a successful repository administration.

Antonella De Robbio¹
Michael Katzmayr²

¹ Member of the E-LIS administrative board; Università degli Studi di Padova - Center for University Libraries, Padua, Italy

² E-LIS editor for Austria; WU Vienna University of Economics and Business - University Library, Vienna, Austria

E-LIS: disciplinary repository for library and information science

E-LIS is an abbreviation for E-prints in Library and Information Science. It is a centralized disciplinary repository, was established 2003, and is the world's largest open archive in the field of librarianship, information science and technology, and related disciplines. In its beginning it was partly funded by the *Spanish Ministry of Culture*, and it is hosted by the *AEPIC* team on servers of the *Italian Consorzio Interuniversitario Lombardo per Elaborazione Automatica (CILEA)*. E-LIS is part of the *Research in Computing, Library and Information Science (RCLIS)* project, an international effort to organize and disseminate scholarly papers in librarianship and related fields. The applied software is GNU Eprints version 3.1.2.1.

E-LIS accepts any scholarly or technical document, published or unpublished, that is relevant to LIS and has the form of a finished document that is ready to enter into a process of communication. This results in a broad coverage both in terms of content and types of submitted material.

As of April 14, 2009, E-LIS contains 9025 documents. For the last 12 months the growth rate of the repository has been at about 2% or 140 documents per month. The steadily growing number of documents shows that E-LIS has achieved a sustainable acceptance in the LIS community.

The invisibility of scholarship from so called developing countries is regarded as a major problem in the academic community. Thanks to the existence of networked services of digital OA repositories, scholars from these countries can access relevant knowledge via these services and in turn have the opportunity to disseminate their documents to their global academic communities at practically no cost. Accordingly, the mission of E-LIS is to serve the international LIS community and provide a place where all authors can deposit their documents and contribute to the worldwide dissemination of knowledge. Authors who deposit their documents in open archives, such as E-LIS, are participating in a global endeavor by universities, researchers, libraries, (some) publishers, editors and readers to redefine and eventually transform the mechanisms of scholarly communication. In E-LIS, more than 5900 users from 90 countries have been depositing documents in 37 languages. Table 1 and Table 2 show the outstanding international contribution to E-LIS.

Table 1: Top 10 languages in E-LIS (data retrieved on April 14, 2009)

Language	Number of Documents
Spanish	3146
English	2822
Italian	937
German	402
Greek	322
Portuguese	283
Turkish	255
Serbian	205
Catalan	194
Polish	121

Table 2: Top 10 contributing countries to E-LIS (data retrieved on April 14, 2009)

Country	Number of Documents
Spain	1774
Italy	1005
India	534
USA	530
Cuba	528
Mexico	382
Greece	370
United Kingdom	368
Turkey	320
Argentina	314

The documents in E-LIS are not only scholarly papers before and after peer-review (so called preprints or post-prints), but also conference papers, conference posters, presentations, books, book chapters, technical reports, working papers, theses, newspaper and magazine articles, etc. All in all, more than 20 document types are available in E-LIS. Table 3 shows the distribution of documents to the top 5 document types and indicates how many of them have undergone peer-review.

Among all documents deposited in E-LIS, about 54% are marked as peer-reviewed. This means that roughly half of the documents are OA literature in the narrower sense of the meaning. The high number of non peer-reviewed documents is founded on the fact that E-LIS not only serves LIS scholars in the field, but also librarians and other information professionals. The latter often deposit and use document types that are by nature not subject to peer-review, e.g. tutorials, reports, library instruction materials, etc. However, when searching E-LIS it is possible to limit to a desired subset of documents, e.g. peer-reviewed articles in a certain language.

Table 3: Top 5 document types in E-LIS (data retrieved on April 14, 2009)

Document type	Number of documents	among them peer-reviewed
Journal articles / postprints	4111	3042
Conference papers	2111	1012
Presentations	921	n/a
Preprints	323	61
Thesis	260	n/a

Organization of E-LIS

Since its beginning, E-LIS has been organized, managed and maintained by an international team of librarians and information professionals with different vocational and educational backgrounds. The organization of E-LIS relies on voluntary work; however, coordinating volunteers without having a clear organizational structure and defined policies is like herding cats. Therefore, E-LIS has an elaborated organizational structure that is made up of three sections.

The administrative section consists of the administrative board (8 members) which is responsible for the overall policy and organizational development. The administrative board comprises the following roles or functions: the chief executive who is the representative of E-LIS, the coordinator of the editors who appoints all country and regional editors, and the chair of the technical board that appoints and removes members from the technical board provided that at least 2 in 3 of the existing members express agreement with the appointment or dismissal. The chair of the administrative board appoints and removes members from the administrative board, provided that at least 2 in 3 of the existing members express agreement with the appointment or dismissal. The administrative board can dismiss and replace the chair by another person among its members by majority vote.

The second section consists of 67 editors from 40 countries; the maximum number of editors per country is three. The editors' main task is to approve the documents deposited in E-LIS in terms of metadata quality. The group of editors is also the forum where important policy issues are discussed and democratically agreed upon. Moreover, the editors have the duty of promoting E-LIS in their countries and keeping close contact with local LIS communities, relevant academic institutions, and authors. 4 regional editors support the country editors in the approval process. They supervise the quality of the metadata approved by the editors and support the collaboration of editors in case of cross-country or regional projects.

The third section is the technical board (4 members) that concentrates on the software – its implementation, the development of added value functionalities, and its operation within the OAI framework. The technical board can dismiss its chair and replace it by another person among its members by majority vote. Figure 1 shows the functional organizational structure of E-LIS.

Benefits for authors

Providing the right services for authors and publishers is a vital part of repository management. So, why should authors deposit their full texts in E-LIS? There are basically two major benefits: the worldwide dissemination of papers and detailed usage statistics for each document.

The main purpose of E-LIS is to give authors in the LIS community the opportunity to self archive their papers and to disseminate them as widely as possible via the Internet. E-LIS achieves this by being part of the so called *Open Archive Initiative* (OAI) (<http://www.openarchives.org>). The OAI architecture contains two functional roles: *data provider* and *service provider*. A data provider, such as E-LIS, maintains one or more repositories, and the information objects in these repositories (e.g. full texts) are described with structured metadata (e.g. citation records). Service providers harvest the metadata (mostly in Dublin Core format) from several data providers according to the protocol *OAI-PMH* (*OAI Protocol for Metadata Harvesting*), a set of request and responses carried over HTTP. The service providers offer a search-and-retrieval interface that allows users to search metadata records from various data providers. The metadata records in the service providers contain a hyperlink to the information object in the corresponding data provider. The value of service providers lies in the aggregation of data: end-users do not have to search the relevant repositories individually, but now can rely on one resource to search information within their relevant discipline [4].

Service providers can either be general or discipline specific. Some prominent general service providers are *OAIster* (<http://www.oaister.org>), *BASE – Bielefeld Academic Search Engine* (<http://base.ub.uni-bielefeld.de>), *Scientific Commons* (<http://en.scientificcommons.org>) or *Scirus* (<http://www.scirus.com>). The discipline specific service provider in the field of LIS is *DL-Harvest* (<http://dlharvest.sir.arizona.edu>) that harvests metadata from 14 repositories that are dedicated fully or partly to LIS [5]. E-LIS is harvested by all of these service providers; this means that all documents in E-LIS can also be found easily via several other academic databases on the Internet.

Additionally to the metadata harvesting by service providers, all records and – if technically possible – full texts in data providers are indexed by Internet search engines. Academic search engines like *Google Scholar* (<http://scholar.google.com>) explicitly direct their robots

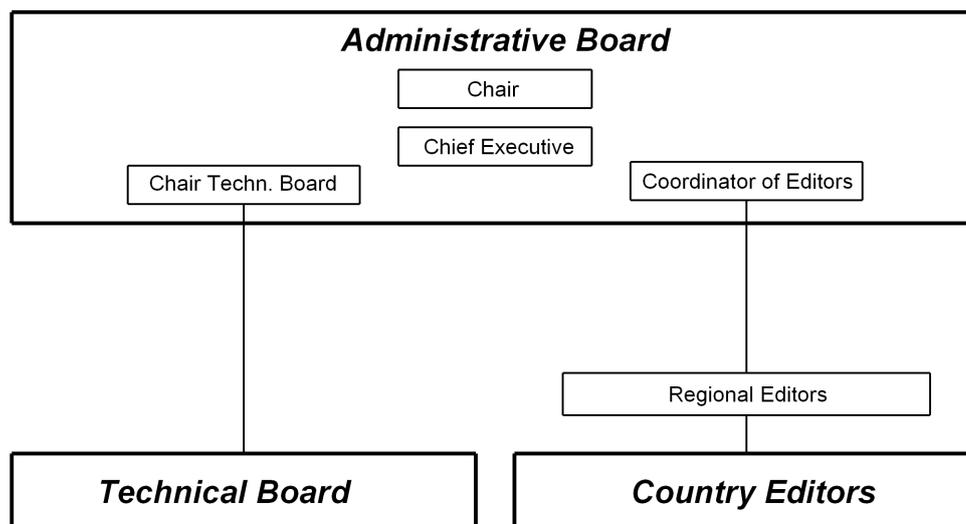


Figure 1: Functional organization chart of E-LIS

to registered institutional or subject OA repositories. Due to the fact that individual homepages of researchers are often ignored by academic search engines, putting documents somewhere on the Internet is only half way there. In order to disseminate full texts as widely as possible and make them visible to OAI service providers and academic search engines, authors should deposit their documents in OA repositories. Moreover, these repositories provide a stable and persistent URL for citation purposes.

Another advantage for authors is the availability of detailed usage statistics that provide monthly document and abstract views broken down by country. While the characteristics of the OAI architecture benefit all OAI compliant repositories, the availability of detailed usage statistics is a feature developed by the technical staff of E-LIS. Figure 2 shows the usage statistics for an item in German language deposited in November 2008.

In addition, E-LIS tries to keep pace with the evolving requirements of authors and repository users. Due to the recent migration to Eprints version 3 users have gained advantage from new features, e.g.:

- it is possible to download the metadata of records in various formats like BiBTeX, RefMan RIS (for EndNote, Reference Manager), etc., and
- authors can define an embargo period for their documents. This means that the full-text of such documents can be used only after a defined period due to copyright agreements with publishers. However, the citation will be provided immediately after depositing.

Improvements and add-ons related to multimedia content and Web 2.0 functionalities are goals of further developments.

Benefits for LIS publishers and associations

Not-for-profit publishers of LIS journals can use E-LIS as a platform for the secondary publication of their content. For example, all articles published in *ACIMED*, a Cuban journal dedicated to medical information and librarianship published by the *National Center for Scientific Research*, have been deposited in E-LIS from the first issue in 1993 up to now.

Professional or academic associations that organize conferences and publish conference proceedings can make a similar use of E-LIS. This can be exemplified by the complete inclusion of conference papers of the *Pan-Hellenic Conferences of Greek Academic Libraries*: due to a cooperation of the *University of Cyprus Library* (which had created a digital database with the proceedings) with Greek and Cypriot E-LIS editors it was possible to transfer more than 200 records with full-texts into E-LIS, and it is planned to continue this collaboration.

These two representative examples show how publishers and associations can communicate and promote their efforts in a highly effective and cost-efficient way. In general, such cooperations work best with mediated archiving, because it is easier to achieve a high metadata quality, e.g., when entering multipart names of conference proceedings. Moreover, mediated archiving is also a special service to the authors. If publishers or associations need assistance with mediated archiving, the country editors of E-LIS will give support or hand the issue over to the technical team (in case of bulk deposits). It is recommended that the publishers or associations obtain the right to store the conference papers, presentations, book chapters, articles, etc. in E-LIS from the authors in advance. In order to avoid conflicts, an opt-out policy should exist for those authors who do not want to share their documents open access.

For this eprint: [[Past four weeks](#)] [[This year](#)] [[Last year](#)] [[All years](#)]

Abstract views and document downloads for all years

	Abstracts	Downloads
Views	284	142

Views by country (derived from IP address of query) for all years

Country	Abstracts	Downloads
Germany	193	99
Austria	14	12
Switzerland	8	11
United States	38	5
Spain	7	3
France	4	2
Latvia	2	2
Hungary	3	1
Italy	3	1
Netherlands	3	1
Croatia	1	1
Greece	1	1
United Kingdom	0	1
Mexico	1	1
Czech Republic	0	1
Luxembourg	2	0
Denmark	1	0
Bosnia	1	0
Canada	1	0
Brazil	1	0

History of views for this ePrint

Period	Abstracts	Downloads
2009 Feb	7	7
2009 Jan	32	20
2008 Dec	71	43
2008 Nov	174	72

Figure 2: Example of detailed data for document usage

Currently, E-LIS collaborates with the following institutions and associations:

- AAB – *Andalusian Librarians Association* (Spain)
- AHDI – *Asociación Hispana de Documentalistas en Internet* (Spanish speaking world)
- AIDA - *Italian Association for Advanced Documentation* (Italy)
- AVEI – *Associació Valenciana d'Especialistes en Informació* (Spain)
- *Centre for Digital Library Research, University of Strathclyde* (United Kingdom)
- CNBA – *Coordinamento Nazionale Biblioteche Architetture* (Italy)
- CNIC – *National Center for Scientific Research* (Cuba)
- IBICT – *Brazilian Institute on Science and Technology* (Brazil)
- *Polish Librarian Association* (Poland)
- SEDIC – *Spanish Society of Documentation and Scientific Information* (Spain)
- TLA – *Turkish Librarians Association* (Turkey)
- UNAK – *University and Research Librarians Association* (Turkey)
- *Universidad Politécnica de Valencia* (Spain)

New partners are always welcome. Institutions or associations interested in a cooperation are invited to contact the responsible country editor.

Policy management

Additionally to an appropriate organizational structure, clear and defined policies are a prerequisite for successfully running a repository. There are various challenges in everyday repository management that require clear policies, e.g.:

- an author wants to withdraw one or more full texts,
- an author suggests fundamental changes in metadata or wants to modify the full text,
- an author deposits a paper that seems to be questionable in terms of content,
- a publisher contacts the repository in terms of copyright infringement,
- a library association intends to archive the proceedings of its conferences in the repository, but wonders if long term digital preservation of the full texts will be guaranteed.

Many E-LIS editors have already been confronted with one or more of these issues. Experience in the repository community has shown: the more detailed and specific the policies, the easier such issues can be handled. Policies should be clear and binding, but they never can be seen as carved in stone. They have to keep pace and evolve with changes in the repository environment. Due to this E-LIS has been engaged in three policy related activities: policy self assessment using the DRAMBORA audit scheme, reflection of its preservation management,

and formulating and communicating its policies in a standardized and structured way.

Policy self-assessment

In 2007 E-LIS completed the *DRAMBORA* audit (<http://www.repositoryaudit.eu>). *DRAMBORA* (Digital Repository Audit Method Based On Risk Assessment) can be implemented as self-assessment and covers all kinds of activities that are likely to have an impact on digital objects, e.g. missions and objectives of the repository, organizational issues, technical means, staff, funding, etc. The *DRAMBORA* audit facilitates [6]:

- documentation of organizational and regulatory frameworks,
- identification of activities, assets and their owners,
- identification and assessment of risks associated with managing digital information,
- management of risks to ensure business continuity and future use of information.

The 10 core requirements are [6]:

- mandate and commitment to digital object maintenance,
- organizational fitness,
- legal and regulatory legitimacy,
- efficient and effective policies,
- adequate technical infrastructure,
- acquisition and ingest,
- preservation of digital object integrity, authenticity and usability,
- metadata management and audit trails,
- dissemination, and
- preservation planning and action.

E-LIS has been an early adopter of the *DRAMBORA* audit. In a nutshell, the result was that E-LIS in general seems to be well prepared; however, some shortcomings were discovered, e.g. the documentation of procedures, technical routines, and preservation issues could be improved. Accordingly, the main advantages of completing the audit process were a detailed reflection on the policies, to get information which further documentation is needed, and obtaining an understanding of the severity of various risks that constitute a risk-ranking.

Preservation management

Preservation issues are of growing importance in the repository community. With regard to E-LIS, it is rather a *presentation* than a *preservation repository*. It therefore primarily aims at the broad and quick dissemination of LIS materials. However, keeping abreast of preservation issues and complying with conventional preservation practices and standards is a must for all OA repositories. To date, most of the repositories, including E-LIS, have no formal preservation policy, except of the recommendation of file formats that are likely to facilitate long-term preservation. The first step in defining a preservation

strategy is to identify the file formats of the information objects in the repository, because an accurate knowledge of file formats is a prerequisite for preservation planning and active preservation. Therefore E-LIS provides a profile that shows the breakdown of file formats contained in the repository that is provided in the context of the project *Preserv* (<http://preserv.eprints.org>) that aims at enabling long-term access to materials in repositories [3]. The *Preserv* Profile is available within the *Registry of Open Access Repositories ROAR* (<http://roar.eprints.org>) (Figure 3).

This histogram shows that at about 86% of the files are in one of the various PDF formats. At about 9% of the files are already available in PDF/A format (Portable Document Format – Archival (1)) that is considered as suitable for long-term preservation. Each bar in the histogram is a hyperlink to the list of the records that contain the defined file type.

The software used for E-LIS, EPrints v3, helps to support the preservation of digital objects by the means of various features that were jointly developed with the *Preserv* project. For example, the new history module of EPrints allows the recording of the history of changes to a repository object by updating its preservation metadata, i.e. the metadata designed for managing digital content over a long period of time.

Although all types of files can be deposited in E-LIS, authors are recommended to deposit PDF or HTML files. Plain text files are also supported. Now that the types of materials in the repository are known and the applied software provides the relevant technical features, a preservation policy can be elaborated.

Policy communication

Not only the development of clear policies is important for repository management, but also the communication of the policies to users and other stakeholders. Thus E-LIS intends to define its policies in a standardized way in order to achieve more clarity for authors and repository staff. The Directory of Open Access Repositories *OpenDOAR* provides on its website an *Policies Tool* (<http://www.opendoar.org/tools/en/policies.php>) that helps repository administrators to formulate and communicate their policies [7]. The tool recommends two types of policies:

- minimum options that allow the achieving of OA goals and
- optimum options for refinements for more use or better quality.

The overall policy is made up of 5 specific policies:

- metadata policy for information describing items in the repository
- data policy for full-text and other full data items
- content policy for types of document and data set held
- submission policy concerning depositors, quality and copyright
- preservation policy

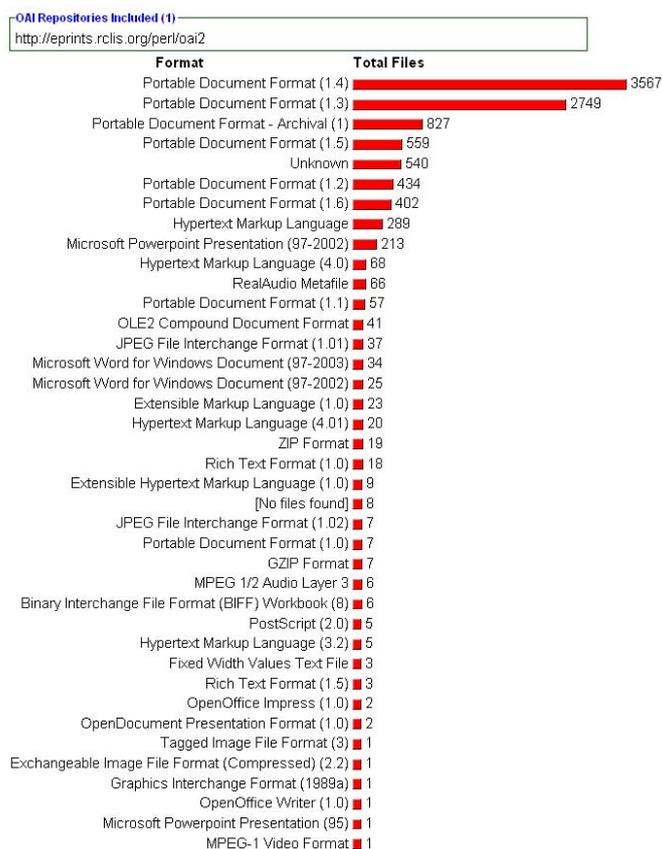


Figure 3: Preserv Profile of E-LIS (profile retrieved on April 16, 2009)

In the time of writing this article, the administration of E-LIS is defining its policy using this tool.

Conclusion

The development of E-LIS, especially the steadily growing number of deposits, shows that this repository is on the right track. Therefore, E-LIS is not only a successful disciplinary archive for the LIS community, but also can act as a model for other communities or disciplines that want to set up a centralized disciplinary OA repository. Moreover, the organizational model is an example of a very effective global collaboration.

However, this does not mean that the E-LIS team can rest on its laurels. A repository is a growing and evolving organism and has to adapt to the changing repository environment and user requirements. Especially, the policy development is of importance – defining, revising and communicating policies are a key issue in the management of repositories.

References

- Bailey CW. What is open access? In: Jacobs N, editor. *Open Access: Key Strategic, Technical and Economic Aspects*. Oxford: Chandos Publishing; 2006. p. 13-26. Preprint available from: <http://www.digital-scholarship.com/cwb/WhatsOA.pdf>
- De Robbio A, Subirats Coll I. E-LIS: an international open archive towards building open digital libraries. *High Energy Physics Library Webzine*. 2005;(11). Available from: <http://eprints.rclis.org/4476/>
- Morrison H, Subirats Coll I, Medeiros N, De Robbio A. E-LIS: The open archive for library and information science. *The Charleston Advisor*. 2007;9(1):56-61. Available from: <http://eprints.rclis.org/11032/>
- Breeding M. Understanding the Protocol for Metadata Harvesting of the Open Archives Initiative. *Computers in Libraries*. 2002;22(8):24-9.
- Coleman A, Roback J. Open access federation for library and information science: dLIST and DL-Harvest. *D-Lib Magazine*. 2005;11(12). Available from: <http://www.dlib.org/dlib/december05/coleman/12coleman.html>
- DRAMBORA. DRAMBORA: digital repository audit method based on risk assessment. 2007. Available from: http://www.repositoryaudit.eu/img/drambora_flyer.pdf
- Millington P. Moving Forward with the OpenDOAR Directory. Presented at: 8th International Conference on Current Research Information Systems, Bergen, 11th-13th May 2006. Available from: <http://www.opendoar.org/documents/BergenPresentation20060512Handouts.ppt>

Corresponding author:

Antonella De Robbio
Università degli Studi di Padova - Center for University Libraries
Via Anghinoni 3, 35121 Padua, Italy
antonella.derobbio@unipd.it

Please cite as

De Robbio A, Katzmayr M. The management of an international open access repository: the case of E-LIS. *GMS Med Bibl Inf.* 2009;9(1):Doc09.

This article is freely available from

<http://www.egms.de/en/journals/mbi/2009-9/mbi000137.shtml>

Published: 2009-06-16

Copyright

©2009 De Robbio et al. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by-nc-nd/3.0/deed.en>). You are free: to Share – to copy, distribute and transmit the work, provided the original author and source are credited.