Open archives for Library and Information Science: an international experience...

4th EBIB Conference Internet in libraries Open Access
7th December 2007
Torun, Poland

Antonella De Robbio [1]
E-LIS Admin. team
http://eprints.rclis.org/
CAB Centro di Ateneo per le Biblioteche
University of Padua,
Via Anghinoni, 3 - 35121 PADOVA (Italy)
Tel: +39 049 8273654
Email: antonella.derobbio@unipd.it
URL:
http://www.math.unipd.it/~derobbio/home/schedaeng.htm

Paula Sequeiros [2]
Instituto de Sociologia
Universidade do Porto
R. Campo Alegre, 4150 PORTO
PORTUGAL
Tel: +351 226077190
Email: psequeiros@gmail.com
URL:
http://sigarra.up.pt/flup/unidades_geral.visualizar?p_unidade=59

Abstract

Library and information domain being highly integrated with computing science and technology, bibliometric and citational analysis, the team that launched this project felt that LIS could set an interesting reference case to other communities in providing a state-of-the-art model for the OA movement and for digital libraries, and most particularly to the open archive model, within which E-LIS is a disciplinary repository. Furthermore the evolving academic discourse surrounding the concept of Open Access (OA) provided the underlying motivation for the establishment of E-LIS Eprints for Library and Information Science in compliance with the Open Archives Initiative (OAI) protocol.

Authors who contribute to an e-print archive are participating in a global effort by universities, researchers, libraries, publishers, editors, and readers, to redefine the mechanisms of scholarly communication. E-LIS makes LIS research more visible and accessible, which in turn will increase its status and public value. The trend towards digital libraries and self-archiving is one of the more recent developments in the field of Library and Information Science (LIS). The development of an international LIS network is being stimulated by the extension of the OA concept to LIS works and facilitated by the dissemination of material within the LIS community. These are some of the reasons for the success of E-LIS as an organizational model developed within an OAI compliance

---

[1] **Antonella de Robbio** is the Coordinating Manager of the Library System of the University of Padova, the University Centre for Libraries (CAB). Since 2002 she has been responsible for the Projects Section and Digital Library: digitalization, eprints repositories (Open Access in general), bibliographical services to support e-learning. Until 2002 she was responsible for the Mathematics Library of the University of Padova, where in collaboration with other authors, she developed a set of tools for displaying and connecting classification schemes, applied to the MSC, ZDM, PACS and CCS, as well as to DDC. She is the Italian coordinator of MAI the Italian MetaOpac (since 1999). In 2003 she implemented E-LIS for which she is the Manager.

[2] **Paula Sequeiros** is a researcher with the Institute of Sociology in Porto, Portugal. She was a Documentalist in the banking industry and university librarian director, now dedicates herself to LIS teaching and training. She is an E-LIS Editor for Portugal since 2006.

framework and for this respect E-LIS can be regarded as a tool for disseminating the OA philosophy.

Built in 2003, E-LIS is an international open access archive related to librarianship, information science and technology, and related disciplines. Being the first international e-server in this area, E-LIS is a part of the RCLIS (Research in Computing, Library and Information Science) project and is organized, managed and maintained by an international team of librarians working on a voluntary basis. All editors share a common vision and mission, while contributing to E-LIS with their own experience and competence. E-LIS deals with each country specific issues to decide the best solution for technical and non-technical barriers so that international visibility can be promoted whilst national interests are served. International co-operation can facilitate debate on current issues on many levels and provides new professional experiences and expertise to editors, on a personal level.

To date E-LIS is the largest repository in library and information science and after a few years contains almost 7000 papers from 82 countries. Poland, being a very active country in the Europe continent, is also contributing, within this framework, with national papers.

This presentation describes the main technical and organizational characteristics of the archive, its configuration and customization, and states its policies, aims and mission. Main focus, however, is placed on E-LIS organizational model and on strategic issues related to OA. It also reasons on how E-LIS efforts are contributing to outline some of the challenges and opportunities resulting from a global vision of the Library and Information Science domain, giving E-LIS team the necessary motivation to engage in such a venture and to further develop international research activities.

Keywords: Library and Information Sciences; E-LIS; Open Access Repositories

Concepts underlying E-LIS

An evolving academic discourse around the concept of Open Access (OA) provided the underlying motivation for the establishment of **_E-LIS Eprints for Library and Information Science_**. Authors who contribute to an e-print archive are participating in a global effort by universities, researchers, libraries, publishers, editors, and readers, to redefine the mechanisms of scholarly communication. E-LIS will make LIS research more visible and accessible, which in turn will increase its status and public value.

One of the more recent developments in the field of Library and Information Science (LIS) is a trend towards digital libraries and self-archiving. Self-archiving can be defined as the deposit of a digital document in a public, free-access repository, for example, an e-print archive. An e-print archive is a collection of digital research documents such as articles, book chapters, conference papers and data sets. E-prints are the digital texts from research papers, before or after refereeing or publication[3].

Established in 2003, E-LIS (http://eprints.rclis.org) is an international Open Archive for Library and Information Science. Being the first international e-server in this area, E-LIS is a part of the RCLIS (Research in Computing, Library and Information Science) project and is organised, managed and
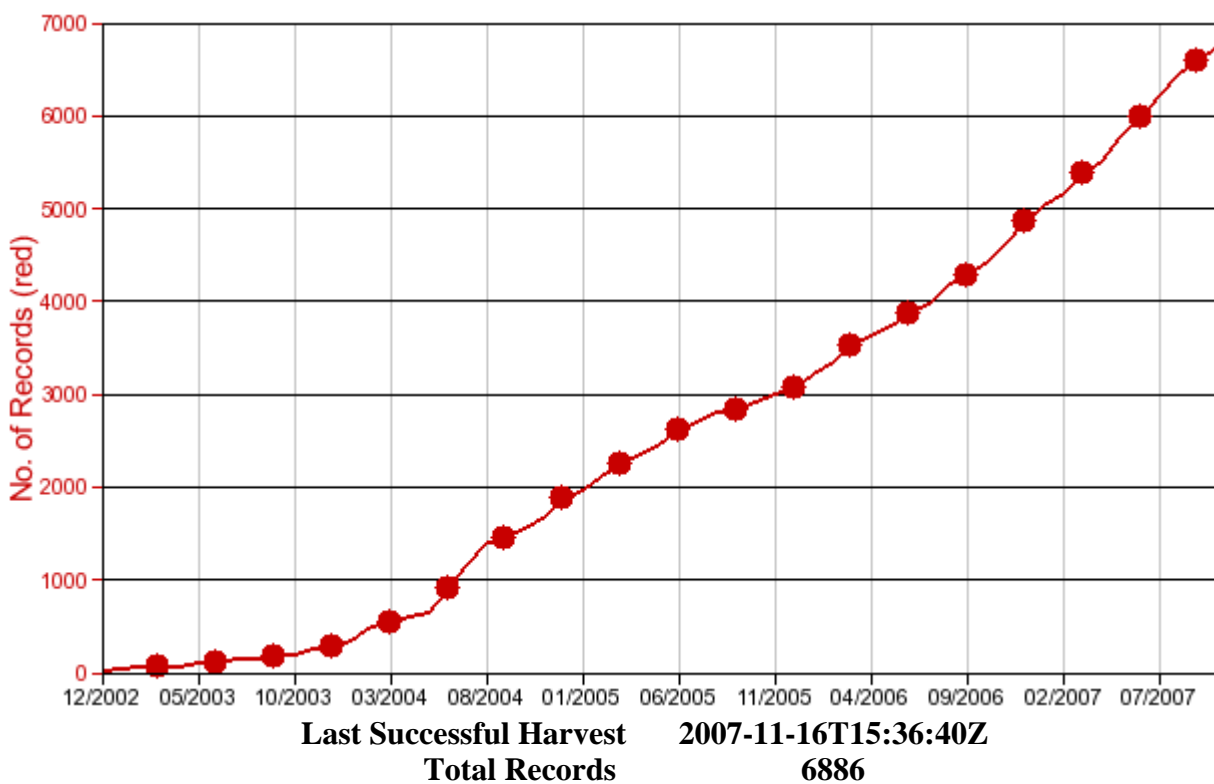
---

[3]   A draft is called a *preprint* before refereeing and publication, the refereed, accepted, final draft is called a *postprint*, so the term *e-prints* includes both.

maintained by an international team of librarians, with different backgrounds, working on a voluntary basis. Positing itself as a not-for-profit project, it is freely accessible, users being able to search and access full-text documents. It is based on a spirit similar to the Open Source Initiative in computing whereby people from all over the world co-operate in building software for the public domain.
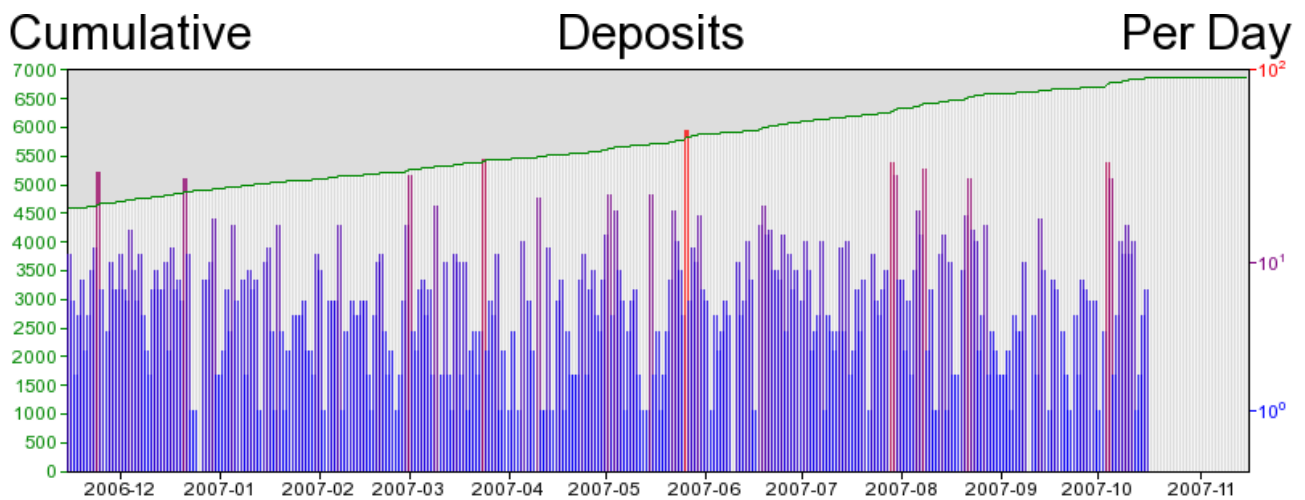
E-LIS is supported by GNU Eprints software and is based on OAI (Open Archives Initiative) protocols and standards to allow for interoperability between repository servers.

The archive: documents, authors and services

Statistics point to over 2400 post-prints and 400 pre-prints deposited documents. These consist of over 1750 articles from 200 different journals and 600 conference papers from 250 conferences.
As far as quality is concerned we can say that 62% of documents were refereed.

Last Successful Harvest   2007-11-16T15:36:40Z
Total Records   6886

There is more to create an archive than putting software into a machine: creating an eprint archive is not difficult, but filling it is a daunting task.

Cumulative   Deposits   Per Day

Both traditional LIS topics and interrelated disciplines are covered. Any LIS scientific or technical finished document, published or unpublished, in any language and format may be accepted and then accessed by LIS communities in any country.

Authors are invited to self-archive their work. A proxy service supports authors who may feel less IT-confident or concerned over editorial agreements with prior publishers. Metadata are set on the basis of a 23 document typology and then checked by an international editors committee in accordance with editorial guidelines Agreements with institutions and library associations from many countries have also been a part of the submission policy which is open to proposals from new partners. Until now 60% of submissions were mediated by E-LIS editors, following these agreements, or on behalf of authors, for some reason unable to do it themselves, while another 30% were self-archived. In the aggregate, authors amount to 2540.

Starting from its homepage, users perceive two main structures organising all contents, both displaying second levels. The first structure is the JITA classification schema for LIS, displaying main subject divisions and subdivisions; the second structure is a geographical ordering based on continent and country. Papers can be viewed on a single country basis, highlighting the internationality of the project, in addition to facilitating user access.

E-LIS value-added features provide tools for editors and for users. These include, among others, automatic alerts for editors, full metadata display of metadata records, full-text search, browsing by user, a counter in the homepage as well as statistics from Webalizer.
E-LIS also provides reference linking with ParaTools: a *seek* button parses citations and links them with online resources that hold a strong probability of providing the free full article.
An e-mail alert service is currently subscribed by 1020 users.

Organisational model: E-LIS teams

The administration team deals with strategic issues including the initiative's future direction, its policies and their impact on the user community as well as its connection with other scientific communities. Being the core of the organisational model, the administration staff that also holds responsibility in determining the archive's international and future vision and in keeping in mind an understanding of national and international needs within a world framework of Open Access strategic issues.
Forty two countries from all five continents are currently involved in E-LIS, each working towards a common vision while maintaining their own cultural individuality. And these are: Argentina,

Austria, Bosnia and Herzegovina, Brazil, Bulgaria, Canada, Chile, China, Colombia, Costa Rica, Croatia, Cuba, Czech Republic, Finland, Germany, Greece, India, Indonesia, Israel, Italy, Mexico, Moldova, Nepal, New Zealand, Nigeria, Pakistan, Peru, Poland, Portugal, Russia, Serbia and Montenegro, South Africa, Spain, Sri Lanka, Switzerland, Thailand, Turkey, Uganda, United Kingdom, United States, Zambia, Zimbabwe.

The international context

Other similar initiatives within the LIS area differ from E-LIS in terms of organisation and vision, which in E-LIS case are clearly international (for other repositories' mission, aims and activities see tables below). Nevertheless E-LIS holds a stance that it is important to collaborate with these similar, although different, repositories in order to build a service provider for the LIS context.

E-LIS singularity lies in the fact that it is not commercially-driven (there is neither funding nor interest in receiving a revenue or profit) and that it's based on voluntary work.

Repository policies

E-LIS policies determine its identity, quality and direction. An archive's organisational model is the sum of its policies and an archive without policies is like a library without rules, especially in such a case as E-LIS.
Main policy concerns are discussed and democratically agreed upon by the editorial staff. These are structured along the following axes[4]:

1. Mission: aims and objectives, what it is and where it's going, target audience and communities involved, clearly outlining its previously discussed international traits;
2. Submission policies: who can deposit and how;
3. Metadata quality;
4. Copyright policies;
5. Organisational model.

E-LIS copyright policy is clearly stated and follows the RoMEO Project recommendations concerning the self-archiving of documents.

All the editors, over sixty now, share this common vision and mission, while contributing to E-LIS with their own experience and competence. E-LIS deals with each country specific issues to decide the best solution for technical and non-technical barriers so that international visibility can be promoted whilst national interests are served. International co-operation can facilitate debate on current issues on many levels and provides new professional experiences and expertise to editors, on a personal level.
There is an undergoing debate, about the launching of a service provider to harvest and gather national metadata from national LIS archives which could become a part of the RCLIS infrastructure.

A must for E-LIS are preservation issues, specially the concern for the implementation of policies on the preservation of metadata. A key step in developing such policies is the identification of the types of material in the repository in terms of technical structure, or file formats (e.g. pdf, html). Institutional archives or cross-searched disciplinary repositories are an upstream resource for most

---

[4]   for further information refer to E-LIS website

new research and educational materials in the digital era and so «it is important to build the concept of preservation from the outset» as suggested by JISC Circular 4/04, note 10.

Metadata designed for managing digital content over a long period of time is commonly referred to as *preservation metadata.* Usually these are necessary to hold documents, inform, describe and record a range of activities related to preserving specific digital objects in a repository. To date, repositories, E-LIS included, have no «formal preservation policy»; indeed such policy is «being preceded by *de facto* rules on file formats and transformations but no provision for acquiring the source versions is being made», according to recent survey results. «The strategy for preservation should be determined by the nature and need of the repository, and should be driven by repository policy rather than the other way around» (idem).

For this reason PRESERV *(PReservation Eprint SERVices)* is born[5], a JISC project researching and developing infrastructural digital preservation services for institutional repositories, offering a tool integrated into ROAR (Registry of Open Access Repositories, http://roar.eprints.org/).
E-LIS responded to this survey and «is working to allow for a *Preserv profile* within ROAR. *Preserv profiles* were produced applying The National Archives (TNA, UK) PRONOM-DROID file format identification service to repository data collected by the Celestial harvester. The resulting graphical view of a repository is then «broken down by file formats through ROAR's user interface. The National Archives curates a database of file formats, PRONOM, aiming to identify repositories content by using TNA's DROID (Digital Record Object Identification) open source software, which may be downloaded and applied for any repository».

Moreover OpenDOAR (The Directory of Open Access Repositories, http://www.opendoar.org/)

«has produced a useful and practical tool [http://opendoar.org/tools/en/policies.php] that helps build preservation policy on top of policies for metadata, data, content and submission. Its preservation policy definition form is especially perceptive to allow for external partnerships» .
PRONOM-ROAR[6] suggests a new perspective. On one hand, because it implements an ingest service based on OAIS reference model for institutional archives built on EPrints software. On other, it adapts EPrints software to allow the collection and dissemination of preservation-oriented metadata to supplement the current bibliographic information. Following this line, GNU EPrints version 3, relapsed on January 2007, includes a number of features that help support the preservation of digital objects. It interacts with preservation services by providing features for complex-object export, recording the history of changes to a document, and by preserving rights declarations. These new features were jointly developed with the PRESERV project, with coding on the METS and Creative Commons (CC) licensing components provided by the latter.

Statistical analysis, based on an automated crawler and on the *Pronom file format identification tool*, showed the breakdown of file types contained in E-LIS: for a total of 5077 records an 86% were found to be PDF/MS-Word[7].

This summer (August 2007) Andrea Marchitelli, Antonella De Robbio and Imma Subirats Coll have been collaborating with project DRAMBORA (Digital Repository Audit Method Based on Risk Assessment). The Digital Curation Centre (DCC) and Digital PreservationEurope (DPE)

---

[5]   The project developed from February 2005 to January 2007 involving Southampton University (School of Electronics and Computer Science), The National Archives, The British Library and Oxford University (Library Services, Systems and Electronic Resources Service)
[6]   For more on PRONOM-ROAR profiling see this short illustrated guide
http://trac.eprints.org/projects/iar/wiki/Profile
[7]   for the complete table refer to Morrison, et al. 2007.

created the Digital Repository Audit Method Based on Risk Assessment (DRAMBORA) toolkit (http://www.repositoryaudit.eu/)

The toolkit is intended «to facilitate internal audit by providing repository administrators with a means to assess their capabilities, identify their weaknesses, and recognise their strengths. Digital repositories are still in their infancy and this model is designed to be responsive to the rapidly developing landscape. The development of the toolkit follows a concentrated period of repository pilot audits undertaken by the DCC, conducted at a diverse range of organisations including national libraries, scientific data centres and cultural and heritage data archives.

The purpose of the DRAMBORA toolkit is to facilitate the auditor in:
– defining the mandate and scope of functions of the repository
– identifying the activities and assets of the repository
– identifying the risks and vulnerabilities associated with the mandate, activities and assets
– assessing and calculating the risks
– defining risk management measures
– reporting on the self-audit».

Some procedures for risk identification, resulting from a deep analysis, highlighted that E-LIS needs to improve at technical management level, while at content level it showed to be well-prepared.

Conclusion

In the LIS community most published papers came traditionally from the US, UK and Europe; on the contrary, with OA encouraging a wider audience, it has been pointed out that in E-LIS high quality papers are now coming not only from the countries believed to be most innovative, but also from other countries, from India, Africa, Asia, Eastern Europe, as well as Central and South America. This an especially desirable development for the inclusiveness generated. This same trend is being reinforced by the intensive collaboration from Polish librarians, who have deposited good quality papers in E-LIS, thanks to our two excellent editors Bozena Bednarek-Michalska and Lidia Derfert-Wolf. Polish librarians have particularly evidenced very high bibliometric skills which have proved to be very stimulating for the whole E-LIS community. E-LIS in Poland now lays a path alongside more general efforts to promote Open Access initiatives in and across European countries.

Library and information domain being highly integrated with computing science and technology, bibliometric and citational analysis, the team that launched this project felt that LIS could set an interesting reference case to other communities in providing a state-of-the-art model for the OA movement and for digital libraries, and most particularly to the open archive model, within which E-LIS is a disciplinary repository.

The development of an international LIS network is being stimulated by the extension of the OA concept to LIS works and facilitated by the dissemination of material within the LIS community. These are some of the reasons for the success of E-LIS as an organizational model developed within an OAI compliance framework and for this respect E-LIS can be regarded as a tool for disseminating the OA philosophy.

The invisibility of scholarship from countries not at core of the most advanced ones, and specially from African countries or India, was regarded as a *fait accompli* in the pre-Internet era. Nowadays, discussion on the digital divide notwithstanding, this scholarship may expose their work via the networked services of digital repositories, such as E-LIS where papers from 82 countries are currently deposited in 36 languages: Afrikaans, Basque, Bosnian, Bulgarian, Catalan, Chinese,

Croatian, Czech, Danish, Dutch, English, French, German, Greek, Hungarian, Indonesian, Italian, Kannada, Malay, Malayalam, Maori, Nepali, Norwegian, Persian, Polish, Portuguese, Romanian, Russian, Serbian, Sinhala, Slovak, Slovenian, Spanish, Swedish, Turkish, Ukrainian.

Directories and digital libraries spread widely as a result of technological interventions and as a reaction to demanding and competitive economical conditions. E-LIS efforts are contributing to outline some of the challenges and opportunities resulting from a global vision of the Library and Information Science domain, giving E-LIS team the necessary motivation to engage in such a venture and to further develop international research activities.

Bibliography

BROWN, ADRIAN. 2006. Automatic Format Identification Using PRONOM and DROID. **Digital Preservation Technical Paper** [on line], no. 1, 7 March. Available from: <http://www.nationalarchives.gov.uk/aboutapps/fileformat/pdf/automatic_format_identification.pdf>.

HEY, JESSIE M. N., et al. 2005. PRESERV (PReservation Eprint SERVices). In: **JISC Joint Programmes Meeting**. Cambridge, England: JISC. Also available from: <http://eprints.soton.ac.uk/16281/>.

HITCHCOCK, STEVE, et al. 2007. **Survey of repository preservation policy and activity** [on line]. Southampton: Preserv Project, IAM Group, School of Electronics and Computer Science, University of Southampton, [accessed 2007-11-15]. Available from: <http://preserv.eprints.org/papers/survey/survey-results.html>.

JISC. 2004. **JISC Circular 4/04: Full Text: Call for Projects in Supporting Institutional Digital Preservation and Asset Management** [on line]. JISC. Disponible en:< http://www.jisc.ac.uk/uploaded_documents/04_04FullText.pdf>.

MORRISON, HEATHER, et al. 2007. E-LIS:the Open Archive for Library and information Science. **The Charleston Advisor**, no. July, pp. 56-59, 23, 26. Also available from: <http://eprints.rclis.org/archive/00011032/02/charleston_advisor.pdf>.

ROMEO PROJECT. 2002. **Project RoMEO** [on line]. Loughborough, Leics: Department of Information Science, Loughborough University, Available from: <http://www.lboro.ac.uk/departments/ls/disresearch/romeo/>.

ROSS, SEAMUS, et al. 2007. **Building Trust in Digital Repositories Using DRAMBORA** [on line]. JISC, Available from: <http://www.jisc.ac.uk/events/2007/06/repositories_conference/repos_pres_session34_notes.aspx>.

Further reading on E-LIS
Barrueco Cruz, José Manuel and Subirats Coll, Imma. 2003. RCLIS: towards a digital library for Information Science. In *Proceedings Libraries in Digital Age (LIDA),* Dubrovnik and Mljet. Available from: <http://eprints.rclis.org/archive/00000352/>

Bassi, Maria Cristina, et al. 2005. E-LIS : E-prints in Library and Information Science. In *Proceedings World Library and Information Congress : 71st IFLA General Conference and*

Council. "Libraries - A voyage of discovery", Oslo. Available from: <http://eprints.rclis.org/archive/00004497/>

Bednarek-Michalska, Bozena and Derfert-Wolf, Lidia. 2006. E-LIS - archiwum publikacji z zakresu bibliotekoznawstwa i informacji naukowej. Uwagi praktyczne dla polskich autorów. *Biuletyn EBIB* (3/73). Available from: <http://eprints.rclis.org/archive/00005878/01/ ebib_elis.pdf>

Bednarek-Michalska, Bozena and Derfert-Wolf, Lidia. 2005. E-LIS: E-prints in Library and Information Science. Available from: <http://eprints.rclis.org/5315/>

De Robbio, Antonella. 2003. E-LIS: un open archive per library and informations science. AIB Notizie 15, 2003(2): pp.12. (Italy). Available from: < http://eprints.rclis.org/archive/00002154/>.

2004. Kumar, Anil and Kalyane, V. L. Bibliographics for the 983 eprints in the live archives of E-LIS: trends and status report up to 7th July 2004, based on author-self-archiving metadata. Available from: <http://eprints.rclis.org/archive/00001927/>

De Robbio, Antonella. 2003. E-LIS: un Open Archive in Library and Information Science. *Bibliotime*, VI (1). Available from: <http://eprints.rclis.org/archive/00000201/>

Lewandowski, Dirk. 2004. E-LIS: E-prints in Library and Information Science. In *Proceedings Internationales Symposium fur Informationswissenschaft (ISI).* Available from: <http://eprints.rclis.org/archive/00002532/>

De Robbio, Antonella and Subirats Coll, Imma. 2005. E-LIS: an international Open Archive towards Open Digital Libraries building, *High Energy Physics Libraries Webzine*, issue 11, August. Available from: <http://library.cern.ch/HEPLW/11/papers/1/>

Medeiros, Norm. 2004. A repository of our own: the E-LIS e-prints archive. *OCLC Systems & Services* 20 (2): pp.58-60. Available from: <http://eprints.rclis.org/archive/00001655/>

2004. Subirats Coll, Imma and Barrueco Cruz, José Manuel Un archivo abierto en ciencias de la documentación e información. *El Profesional de la Información* 13(5): pp. 346-352. Available from: <http://eprints.rclis.org/archive/00002472/>

Melinscak Zlodi, Iva and Pavelic, Damir. 2004. Upravljanje otvorenim arhivom : primjer ELIS-a. In *Katic*, Tinka Eds. *Proceedings Arhivi, knjiznice, muzeji 7*, Porec. Available from: <http://eprints.rclis.org/archive/00003020/>

Subirats Coll, Imma; Arencibia, Jorge Ricardo and De Robbio, Antonella. 2004. E-prints for Library and Information Science (E-LIS) : la tecnología al servicio de la bibliotecología y las ciencias de la información. *ACIMED* 12(6). Available from: <http://eprints.rclis.org/archive/00002849/>

Sequeiros, Paula. 2006. Acesso Aberto a documentos sobre Documentação e Informação. *Cadernos BAD*, (1). Also vailable from: <http://eprints.rclis.org/archive/00008309/01/eLIS.pdf>

Sequeiros, Paula. 2007. E-LIS fala Português? É um prazer conhecê-lo!. In *Proceedings 9º Congresso Nacional de Bibliotecários Arquivistas e Documentalistas*, Ponta Delgada. Available from: <http://eprints.rclis.org/archive/00008770/>