Institutional Repositories in India: A Case Study of National Aerospace Laboratories

Poornima Narayana*, B S Biradar** and I R N Goudar*
* Information Center for Aerospace Science & Technology
National Aerospace Laboratories
Bangalore –560017 India
** Chairman, Department of Library and Information Science
Kuvempu University, Shankarghatta
Shimoga – 577 451 Karnataka, India

1. Introduction

The emergence of Internet has brought enormous opportunity to bring the results of research primarily to academicians, scientists and scholars through digital communication to anyone at any time, anywhere in the world. Although it is now possible to have free access to exhaustive information on the web, still significant amount of research is not available freely. While the delivery technique for scientific publications has changed rapidly, the economic ramifications have hardly changed. During the 1990s several e-print archives as well as a few hundred peer-reviewed, electronic, scholarly journals emerged. The common denominator for most of these is that they offer free access to the electronic product. This has become known as "open access publishing".

2. Open Access Movement

Open access may be defined as a philosophy to achieve the goal of accessing and making available the digital material free of charge which may or may not be free from copyright and licensing restrictions. 'Open access' (OA) means that a reader of a scientific publication can read it over the Internet, print it out and even further distribute it for non-commercial purposes without any payments or restrictions. At the most the reader is in some cases required to register with the service in question, which for instance can be useful for the service providers in view of the production of readership statistics. The use of the content by third parties for commercial purposes is, however, as a rule prohibited. Thanks to the open availability the linking from reference lists to OA publications is substantially facilitated, since the reader does not encounter barriers such as use licenses, and each reference is only a mouse-click away. In general, the author keeps almost complete copyright and can also publish the material elsewhere. The concept of OA came into existence sometime in 1991 due to the necessity of facilitating scholarly communication. According to Berlin Declaration act "open access is a comprehensive source of human knowledge and cultural heritage approved by the scientific community". Budapest Initiative defines Open access as 'freely available on Internet for the public, permitting to read, download, copy,

distribute, print, search or link to the full text, crawl them for indexing, pass them as data to software or use them for any other lawful purpose, without financial, legal or technical barriers other than those inseparable from gaining access to the Internet itself". The four most important OA channels are electronic, refereed, scientific periodicals, research-area-specific archive (e-print) servers (in this paper called subject-specific repositories), institutional repositories of individual universities/institutions, and self-posting on authors' home pages. The e-prints archives. Open access journals and self archiving institutional repositories are the initiatives towards open access. E prints archive is an online repository of materials, freely available on the web for widest possible dissemination of knowledge. The Networked Digital Library of Theses and Dissertations (NDLTD), an international organization is dedicated to promoting the adoption, creation, use, dissemination and preservation of electronic analogues to the traditional paper-based theses and dissertations.

2.1. OA Software Packages

A number of software packages both proprietary and free on net have been developed for archiving and managing digital collections. However open source software packages such as Dspace developed by MIT and HP (*http://www.dspace.org/*), E-prints developed by University of Southampton (*http://www.eprints.org/*), Fedora developed by University of Virginia and Cornel University (*http://www.fedora.info/)* are driving the OA movement especially the development of IRs in the world.

3. Institutional Repositories (IRs)

Institutional repositories represent an important OA-channel and are relatively new developments in scholarly communication process compared to open journals and subject-specific repositories. The development of institutional repositories emerged as a new strategy that allows universities to apply serious, systematic leverage to accelerate changes taking place in scholarship and scholarly communication, both moving beyond their historic relatively passive role of supporting established publishers in modernizing scholarly publishing through the licensing of digital content, and also scaling up beyond ad-hoc alliances, partnerships, and support arrangements with a few select faculty pioneers exploring more transformative new uses of the digital medium. The characteristics of IR include - Institution-based; Scholarly material in digital formats; Cumulative and perpetual; Open and Interoperable. Institutions and their libraries are in a better position than individual researcher to guarantee that the material is available even after decades and that the collection is systematically maintained, for instance, to take account of changing file formats and media. Institutional repositories represent an integral part of the long-term strategies of the universities in question, in particular as these have to redesign their publishing and library policies to take into account the totally new conditions

created by the Internet. The institution's own production of theses and working papers can easily be put up on such repositories, but in the long run the posting of the central production of the university's researchers, that is, their conference and, in particular, journal papers, is crucial. Although institutional repositories can be seen as useful marketing channels for individual universities their most significant impact on the global scale can only be achieved via co-operation via open access indexing services.

Institutional initiatives to promote open access to the research work carried out by them has been made possible through establishing Institutional Repositories. These are "digital archives of intellectual products created by the faculty, staff and students of an institution or group of institutions accessible to end users both within and without the institution. The IR may hold various kinds of publications, such as pre-prints and post-prints of journal articles, conference papers, research reports, theses, dissertations, seminar presentations, working papers and other scholarly items. This way, intellectual contributions of researchers are made accessible free of charge to the whole community of researchers across the world. Thus, the open access which was evolved out of the necessity of wider access to scholarly publication relies on the initiatives of individuals (self archives), institutions. It is more of a philosophy of facilitating wider communication, feedback and use.

3.1. Metadata Harvesting

Another service that gained momentum after establishment of different institutional repositories or open access archives is the metadata harvester. A metadata harvesting service harvests or indexes metadata from OAI- compliant archives or repositories and journals through harvesting software that supports a protocol known as OAI-PMH (Open Access Initiative - Protocol for Metadata Harvesting). The standard protocol used is OAI PMH (Open Archives Initiative-Protocol for metadata harvesting) . Some of the well known Metadata harvesters are: OAIster set up at University of Michigan which harvests and indexes from 634 institutions., ARC at Old Dominion University, ArXIV of physics, Mathematics and computer science, UIUC registry of cultural heritage materials etc.

4. Indian Scenario

4.1. Research and Scholarly Communication

India has prospered through it's strong academic and research establishments. The R&D organizations have also developed expertise in their respective areas that are now recognized worldwide. Leading Indian scientific research institutions, such as Indian Institute of Science (IISC), Indian Institutes of Technology (IITs), Indian Statistical Institute (ISI), laboratories under the Council of Scientific and Industrial Research (CSIR), Indian Council of Medical Research

(ICMR), Indian Council Agricultural Research (ICAR), Indian Space Research Organisation (ISRO), Department of Atomic Energy (DAE), traditional universities, deemed universities and Corporate R & Ds have been playing crucial role towards national development. The high quality research accompanied by innumerable scholarly communications to various national and internationals journals and conferences has put India in the forefront in the developing world and leader of South Asian countries. Only the elite institutions have reasonably good information provision facilities that support scholarly communications. On one hand the paucity of funds for the subscription based scholarly journals and on the other the shrinking budget discourage both the access to vast scholarly publications and publication process itself. The open access literature plays a vital role, both in terms of research communication and access, provided, of course, the benefits in terms of economic and social recognitions are assured by this system. While the ICT infrastructure necessary to take advantage of the open access is not adequate in developing countries, the situation in India, is the other way round. The situation has improved to a considerable extent. The number of Internet subscribers in were 140,000 in 1998 and now the number has crossed 5 million. There is a big leap in the telecommunication facility and Internet bandwidth available.

4.2. Open Access

There is a healthy sign of adopting the open access much faster in India compared to many developing countries. A good number of high quality, peer-reviewed open access journals are being published covering a wide spectrum of subjects. While there are many publishers in this category, six major publishers need special mention. They are - Indian Academy of Sciences (IAS, 11 journals); Indian National Science Academy (INSA, 4 journals); Indian Medlars Center of NIC (MedInd, 38 journals); Medknow publications (28 journals); Indian journals.com (8 journals) and Kamala-Raj enterprises (5 journals). Libraries and information centers in India attached to various types of institutions are now taking part in open access movement, by establishing institutional digital repositories to provide worldwide access to their research documents.

4.3. Institutional Repositories

In India, some institutions, like Indian Institute of Science; Indian Institute Management, Kozhikode; Indian Statistical Institute, Bangalore; Indian Institute of Technology, Delhi; National Institute of Technology Rourkela; National Aerospace Laboratories (NAL); National Chemical Laboratory; Information and Library Network (INFLIBNET); National Institute of Oceanography; Raman Research Institute; etc. have set up open access institutional repositories (IRs) that archive and disseminate research outputs of respective institutions. While self-archiving by authors is slowly picking up, mostly it is IR staff that collect and archive the documents to respective IRs on behalf of authors. There are some subject specific IRs as well, Librarian's Digital Library (LDL) of Documentation

Research and Training Center (DRTC), Bangalore and OpenMed@NIC of National Informatics Centre, New Delhi are examples giving access to LIS and biomedical literature respectively. Vidyanidhi of University of Mysore is an example of document type specific collection that archives and provides access to theses and dissertations of Indian Universities. Some archives like OpenMed@NIC offer RSS (Really Simple Syndication or Rich Site Summary) feeds to the registered users, who also get notification on addition of new items in regular intervals. Search engines and metadata harvesters index most of the Indian operational repositories. Registry of Open Access Repositories (ROAR) from University of Southampton lists out 22 IRs of Academic and Research Institutions in India as depicted in Table 1. These IRs have adopted self-archiving model and have been using two most popular open source software – Dspace and GNU Eprints.   Somehow ROAR has not listed few Indian IRs such as that of Indian Council for Research on International Economic Relations (*www.icrier.org/publications.html*),   IITB   (*www.library.iitb.ac.in/~mnj/gsdl/cgi-bin/library*), IIT Kanpur (*library.iitk.ac.in:8080/examples/thesis/index.html*), INSA Digital Library (*drtc.isibang.ac.in/insa*), Vidhyanidhi of University of Mysore (*www.vidhyanidhi.org.in*). A number of new initiatives still being in the initial stages are yet to be listed by ROAR.

Table 1
Indian IRs Listed by ROAR

Name	Host Institution	URL	Items Jun 2006	Soft. Used
Eprints@IISc	Indian Institute of Science, (IISc) Bangalore.	http://eprints.iisc.ernet.in/	5106	EPrints
Dspace@IIMK	Indian Institute of Management, Kozhikode (IIMK)	http://dspace.iimk.ac.in/	146	EPrints
Dspace@IIA	Indian Institute of Astrophysics (IIA)	http://prints.iiap.res.in/	998	DSpace
Dspace@NITR	National Institute of Technology, Rourkela (IITR)	http://dspace.nitrkl.ac.in/dspace/	269	DSpace
ETD@IISc	Indian Institute of Science (IISc)	http://etd.ncsi.iisc.ernet.in/	166	DSpace
Dspac@ INFLIBNET	INFLIBNET	http://dspace.inflibnet.ac.in	428	DSpace
Librarian's Digital Library (LDL)	Docum. Res. & Training Centre (DRTC)	https://drtc.isibang.ac.in/	249	DSpace
NAL Institutional Repository	Nat. Aerospace Laboratories (NAL)	http://nal-ir.nal.res.in/	886	EPrints
EPrints at NCL	Nat. Chemical	http://dspace.ncl.res.in/	354	DSpace

	Laboratory (NCL)			
Digital Repository Service of NIO	National Institute of Oceanography (NIO)	http://drs.nio.org/drs/	141	DSpace
Dspace@NITR	National Institute of Technology, Rourkela	http://dspace.nitrkl.ac.in/dspace/	269	DSpace
Eprints at OWSA	OneWorld South Asia Open Archive Initiative	http://open.ekduniya.net/	89	EPrints
OpenMED at NIC	National Informatics Centre (NIC)	http://openmed.nic.in/	1125	EPrints
Eprint@DU	University of Delhi	http://www.du.ac.in/	104	EPrints
Digital Repository of RRI	Raman Research Institute (RRI)	http://dspace.rri.res.in/	1296	DSpace
Eprints@Bioinf-ormation	Bioinformation	http://www.bioinformation.net/	#	EPrints
DSpace@UOH	University of Hyderabad	http://igmlnet.uohyd.ernet.in	#	DSPace
EPrints@IIMK	Indian Institute of Management, Kozhikode (IIMK)	http://eprints.iimk.ac.in/	146	EPrints
EPrints@IIITA	Indian Institute of Inf. Technology, Allahabad (IIITA)	http://eprints.iiita.ac.in/	#	EPrints
ISI Library, Bangalore	Indian Statistical Institute, Bangalore	http://library.isibang.ac.in:8080/dspace/	#	DSpace
Eprints at RGCFB	Rajiv Ghandi Center For Biotechnology	http://www.rgcb.res.in/	#	EPrints
Eprints at Medknow	MedknowEprints	http://eprints.medknow.com/	#	Eprints

 # Number of items not available in ROAR and could not be ascertained through respective IR sites as some of them were not accessible.

4.4. Metadata Harvesters

Few Indian institutions have been experimenting metadata harvesting. Search Digital Libraries (SDL) of DRTC is one such service harvesting library information science subject-specific open access archives and repositories. The 'Knowledge Harvester@INSA', is an experimental initiative from Indian National Science Academy that harvests metadata from 3 archives. "SJPI Cross Journal Search Service" is a recent initiative from NCSI at IISc that harvests metadata from 13

Indian open access journals. Indian Institute of Technology, Delhi has initiated a metadata harvesting service called SEED that indexes 4 archives. NAL has initiated harvesting OAI compliant IRs of CSIR Laboratories in India and making them accessible through a unified search interface. All these metadata harvesters in India use PKP Harvester, developed by Public Knowledge Project UBC, Canada. Open J-Gate (*www.openj-gate.org*), a free service on net is an open access journals indexing service initiated by Informatics India Private Limited. It covers more than 3,500 open access academic, research and industry journals world over of which > 1,500 are peer-reviewed scholarly journals. More than 0.3 million new articles added every year to this service.

5. Institutional Repository at NAL

5.1. Information Centre at NAL (*www.icast.org.in*)

NAL, a constituent of CSIR is India's pre-eminent civil R & D establishment in aeronautics and allied discipline and has made very significant contributions to all Indian aerospace programmes. It is spearheading the effort to design and develop small and medium size civil aircrafts in India. NAL is well equipped with modern and sophisticated national facilities like the Wind Tunnel Center and the computerized full scale fatigue test facility. The laboratory has developed two major aircrafts, HANSA, the first composite trainer aircraft and the SARAS, a fourteen seater multi role light transport aircraft. The Information Centre of NAL (ICAST), an ISO-9001:2000 Certified Centre along with its parent organisation has been recognised as the National Information Centre for Aerospace Science and Technology by the National Information System for Science and Technology (NISSAT/DSIR), UNESCO and AR&DB. The ICAST with its state-of-the-art expertise, infrastructure and services caters to the information requirements of the Indian aerospace community in particular and the engineering and technical personnel in general. The centre is well known for its aerospace collections of books, journals and specifically technical reports from NASA, DLR, ONERA, NLR, ARL and UTIAS. It offers Online, Web and CD-ROM based literature search services including access to e-journals, News Clipping Service, Online Public Access Catalogue, Union Catalogue of Serials. The centre has setup a portal 'AeroInfo' (www.aeroinfo.org.in), first of its kind in the country, which serves as one window information search facility for aerospace Web sources in the world in general and India in particular.

5.2. NAL-IR (*www.nal-ir.nal.res.in*)

ICAST initiated setting up of its own repository during 2003 using, the then most popular open source software Greenstone Digital Library (GSDL), developed at University of Waikatoo, NZ. More than 300 papers at abstract level along with few full text contributed by the scientists at NAL were uploaded, but were made accessible on NAL's intranet. It was in the year 2004, the work progressed

rapidly with the adoption of the open source software GNU E Prints for archiving and managing the digital collections. This knowledge base covers journal articles, conference papers, technical reports, presentation/lectures, preprints, thesis, images etc. One can browse the documents by author, division, subject, date and document type. The screen shot of NAL-IR home page has been given in Figure 1.

5.2.1. Technologies Employed and Supported

NAL-IR runs on GNU EPrints open archive software (revision EPrints 2.3.6), developed by the University of Southampton, England and freely distributable archive system available from *software.eprints.org*. This IR is powered by MySQL for database, Apache Web server, scripting language PERL5, XML standard for retrieving, the Document Object Model (DOM) platform- and language-neutral interface that allow programs and scripts to dynamically access and update the content, structure and style of documents, ParaCite for both reference parsing and location. The IR Supports OAI-PMH for effective interoperability, VLit for electronic locations (URLs), Valid XHTML and Valid CSS for checking Cascading Style Sheets for display.

5.2.2. Collection Organisation and Search Features

While the scientific staff are trained to contribute their publications and other documents directly to IR, but mostly the staff of information center collect the documents available both in soft copy and hard copy form and upload them to IR. To facilitate the digitization job, especially for documents published earlier, high end scanners like Minolta book scanner, microfiche scanners and flatbed scanners with back to back and batch mode facility have been deployed. The bibliographic staff of center add the metadata based on relevant tools including NASA's subject index to maintain consistency. Access and retention policies have been adopted covering issues like - Who can deposit; What can they deposit; Can items submitted once be withdrawn; Implications of staff movements; Publications brought out when one is on deputation abroad or to another institution; Interdivisional publications; Publications based on R & D work concerned to sponsored projects; Patentable research output; Classified documents; etc. The level of accessibility include fully open on Internet with or without full text, only on institute's Intranet with or without full text, accessible only to concerned heads of the departments and project leaders. The IR facilitates browsing by year of publication, author, concerned department, subject category and document type. One can also search documents by author, title, subject, publisher using simple and advanced search features with many options.

5.2.3. Full Text Downloads

5.2.3.1. Country-wise Downloads

More than 13,500 downloads to the full text of NAL-IR has been recorded between 15th Jan to 25th June, 2006 (Table 2). This statistics does not include the downloads within NAL. The number of documents downloaded varied depending upon the date of uploading and thereby this statistics is basically an indicator of download trend of IR documents. Notably, downloads from United States is highest with 33.6% followed by India 23.61%, United Kingdom 9.48%, Canada 3.63%, Japan 3.2%, China 1.65%, Germany 1.37%, Netherlands 1.21% and so on.

5.2.3.2. Document-wise Downloads

Table – 3 gives the range of statistics of downloads to the documents held by NAL-IR. It is difficult to mention the percentage of documents receiving downloads as the documents have been uploaded during the entire period specified. However, it is satisfying to note that >33% of the documents uploaded have been downloaded, although the number of downloads vary from document to document. Three documents have been downloaded more than 200 times followed by 8 documents between 100-200; 13 documents between 50-100 and so on.

Table 2
Country-wise Full Text Downloads
(January 15 to June 25)

Sl. No.	Country	Total Access	Percentage
1	United States	4550	33.63
2	India	3195	23.61
3	United Kingdom	1283	9.48
4	Canada	491	3.63
5	Japan	433	3.2
6	China	224	1.65
7	Germany	186	1.37
8	Netherlands	164	1.21
9	Sweden	146	1.08
10	France	138	1.02
11	South Korea	104	0.77
12	Pakistan	93	0.69
13	Italy	89	0.66
14	Malaysia	88	0.65
15	EU	80	0.59
16	Russia	78	0.58

Table 3
Document-wise Downloads
(From January 15 to June 25)

Range of downloads	No. of Documents
200 - 379	3
100 - 199	8
50 - 99	13
25 - 49	23
20 - 24	21
15 - 19	18
> 15	175

| 17 | Taiwan | 75 | 0.55 |
| 18 | 58 countries & others | 753 | |

5.2.4. Immediate and Future Plan

The Center plans to adopt Dspace for archiving and maintaining the NAL-IR. Once the job concerned to usual document types like journal articles, conference papers and technical reports is completed, other documents types like presentations, images, patents, theses would be taken up. The metadata harvesting facility is still in experimental stage. Shortly the same would be taken up on regular basis covering IRs of all 40 CSIR chain of laboratories and also those of aerospace institutions in the country. The center is in the process of conducting workshops for training staff of both CSIR and aerospace set ups.

6. References:

1. Arunachalam, Subbaiah: India's March Towards Open Access. In: SciDev.Net 5 march 2004.

2. Bo-Christer Bjork: Open Access to Scientific Publications: An Analysis of the Barriers to Change? In:Information Research, Vol. 9 No. 2, January, 2004.

3. Branin, J; Smith, M; Staples T, Van de Velde (2002): Institutional Repositories:Pt1: Overview and Organizational Issues. Pt 2: Functionality and technical issues. In*: www.cni.org/tfms/2002b.fall/projects.html*

4. Coleman, Anita; Roback, Joseph: Open Access Federation for Library and Information Science. In: D Lib magazine, 11 (12), Dec 2005.

5. Das, Anup Kumar; Sen, B K; Dutta, Chaitali (2005): Collection Developments in Digital Repositories in India. In: *http://tdil@mit.gov.in/apr_2005.htm*

6. David Prosser: Information revolution: Can Institutional Repositories and Open Access Transform Scholarly Communication In: ELS gazette V15, 2003.

7. Lynch, Clifford A: Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age. In: ARL Bimonthly Report 226, February 2003.

8. Rajashekar T B: Improving the Visibility of Indian Research: An Institutional, Open Access Publishing Model. In: Indo-US Workshop on Open Digital Libraries and Interoperability, June 23-5, 2003.

9. Suber, Peter (2002): Open Access to Scientific Journal Literature. In: Journal of Biology, 1(1), p. 1-3.

10. Suber, Peter; Arunachalam, Subbaiah: Open Access to Science in the Developing World. In: Open Access, Unrestricted Access to Published Research. In: 93rd Indian Science Congress, 3-7 January 2006, Hyderabad, India.

11. Yeats, Robin: Institutional Repositories. In: Vine, 33 (2), 2003, p 96-101.

*Poornima Narayana: *poornima@css.nal.res.in*
*I R N Goudar: *goudar@css.nal.res.in*
 ** B S Biradar: *bsbiradar53@rediffmail.com*

Figure 1: NAL Institutional Repository (NAL-IR)

National Aerospace Laboratories
Institutional Repository

Home   About   Browse   Search   Register   User Area   Help

Welcome to NAL Institutional Repository

NAL's Institutional Repository is the digital archive of the research output of our scientists. This knowledge base covers journal articles, conference papers, technical reports, presentation/lectures, preprints, Thesis, images etc. One can browse the documents by author, division, subject, date and document type. Both simple and advanced search facilities have been provided. ICAST hopes to make this archive exhaustive very soon. NAL scientists are welcome to submit their publications on their own by registering.

Search Records:
(Title/Abstract/Authors/Creators/Date)

Search Fulltext: **Search**

Number of items in the Repository: **876**

Browse	Search	Deposit
• By Year • By Subject • By Department • By Author • By Type	• Simple Search • Advanced Search	• Register With Us • Sign in to Deposit • Forgot Password ? Click Here • Latest Additions

Contact Information   Powered by EPrints   A service from ICAST