IR@NAL: Journey in Green Road OA Publishing
Shivaram BS and Poornima Narayana
ICAST, CSIR-NAL
{ shivaram, poornima}@nal.res.in

Abstract:
The Institutional Repository at NAL is the digital archive of the research output of the scientists and technologists. The Repository was established with objectives of increasing visibility for intellectual output of an organization and long term preservation with well established polices. GNU Eprints open source IR tool is used for archiving and managing the digital collection. This paper in detail explains the inception and the saga of six year journey of NAL-IR in green road OA publishing channel. The paper discusses the technology employed, methodology adopted for data collection and preservation, customization of NAL-IR. The process of user education regarding OA publishing with respect to green road is discussed in detail with statistics. It is summarized by providing the current status and the statistics on number of hits received.

1. Introduction

1.1 Open Access
Advent of ICT has revolutionized the process of scholarly publishing thus leading to E-Publishing. Electronic publishing of Journals, Books, Reports, Patents, etc has enhanced speedy communication among research scholars. OA (Open Access) publishing is also one such new model in ICT environment which is getting popularized across the world. Many organizations and scholarly community motivated with advantages of OA publishing started contributing towards it. Open access (OA) is a novel publishing model where scholarly literature is made available free on internet. In brief, Open Access (OA) is free, immediate, permanent online access to the full text of research articles for anyone, web wide, without the severe restrictions on use commonly imposed by publisher copyright agreements. Open Access Publishing is achieved in two ways i.e. Golden road and Green road. The "golden road" of OA journal-publishing , where journals provide OA to their articles either by charging the author-institution a publication or processing fee instead of charging a subscription fee from the user-institution, or by simply making their online edition free for all and recouping the publication and production costs from other source. The "green road" of OA is self-archiving, where authors provide OA to their own published articles, by making their own eprints (the final accepted version) freely available to all by placing them in their institutional or central repositories.

1.2 CSIR-NAL
National Aerospace Laboratories (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 major 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 Centre and the computerized full scale fatigue test facility. The laboratory has developed three major aircrafts, HANSA, the first composite trainer aircraft, SARAS, a fourteen seater multi role light transport aircraft, and NM5, a five seater aircraft.

1.3 NAL-IR (www.nal-ir.nal.res.in)
The Information Centre of NAL (ICAST), an ISO-9001:2000 Certified Centre along with its parent organization has been recognized 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. 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 Institutional Repository at NAL is the digital archive of the research output of the scientists and technologists. Since the inception of NAL during 1959, till date the R&D staffs have published more than 20,000 research publications in various forms. 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 Waikato, 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 through NAL’s intranet. During 2004, the work progressed rapidly with the adoption of the open source software GNU Eprints 2.0 for archiving and managing the digital collections. The knowledge base of NAL-IR covers Journal Articles, Conference Papers, Technical reports, Presentation/lectures, Project documents, Patents, Thesis, Images and Book chapters. In the year 2010 NAL-IR upgraded to GNU Eprints 3.0 with enhanced features. Growth of NAL-IR is depicted in the following chart.
2. Infrastructures
NAL has FDDI network of 1GB with 100mbps LAN spread in three wide campuses and 10 MBPS Internet connectivity. ICAST is well equipped with Application Servers, Terabyte storage and backup solution, high end scanners like Minolta book/microfiche scanners and good number of flatbed scanners with back to back and batch mode facility.

2.1 NAL-IR Hardware & Software configuration
NAL-IR if installed on HP Proliant ML 150 G3 series terabyte server with INTEL (R) Xeon 3.0 GHZ processor. Eprints 3.0 developed by South Hampton University is being used as IR tool on Fedora 9.0 with all dependency applications like Apache, MySql, PERL, ImageMagic, Xpdf, Unicode strings etc.

3. Documents at NAL-IR

3.1 Document Types:
NAL-IR knowledge base covers journal articles, conference papers, technical reports, presentation/lectures, preprints, thesis, images etc. Major share in NAL-IR comprises of technical reports, journal articles and conference papers. Following chart highlights the share of various documents in NAL-IR as on date.

3.2 Content acquisition:
3.2.1 Conventional Data Collection work-flow @ the ICAST
1. Recent publications: The Annual Report published at CSIR-NAL covers research output of NAL scientists, one of the master lists for collection building. Authors are contacted individually either for hard copy or soft copy preferably of their research papers.
2. Interaction with Authors – Archival team contact authors for their respective publications at regular intervals.
3. Bibliographic Database: CD-ROM/online version is used for searching the research publications of NAL scientists, especially journal articles and conference papers.
4. E-Journals: CSIR-NAL is subscribed to major e-journal publishers viz., Elsevier, Springer, AIAA, Wiley etc, these e-journal portals are scanned for its publications at regular intervals.

3.2.2 Digitization
ICAST has maintained print archive of CSIR-NAL publications like technical reports, special publication, project reports etc. These documents are identified for the digitalization process with well defined specification. All documents were scanned with 300dpi using adobe acrobat professional package and converted them to editable .pdf file using built in OCR feature.

3.2.3 Automated publication module:
CSIR-NAL publishes good number of peer reviewed research papers and high quality technical reports. To bring out these publication authors need to follow well established administrative procedure and get approval from competent authorities such as publication committee approval. Head of the department, ICAST planned to introduce fully automated publication modules to facilitate authors in publishing their research work and also acquiring all NAL publications to the repository. Following flowchart gives functionalities of automated publication module.
3.3 Deposition Process:
Scientists and technologists at CSIR-NAL are encouraged to deposit their publications at repository by registering as authenticated users at NAL-IR. Regular orientation programmes are arranged to all contributors for educating them regarding the concepts of OA, advantages, and workflow mechanism of NAL-IR. Around 500+ authors have registered and contributing their publications. Trained IR staff at ICAST are also engaged in the document deposition process. They are given detailed training on different aspects of IR and document deposition process.

3.4 Methodology Adopted

4. NAL-IR Policies

The Institutional Repository was setup with well established polices regarding metadata, content access, submission, and archiving. Any user on Internet can access the metadata free of charge and re-use in any medium without prior permission for not-for-profit purposes provided the OAI Identifier or a link to the original metadata record are given. Content of the full-text and other full data items available on public domain can be accessed free of charge for personal research or study, educational, or not-for-profit purposes without prior permission or charge with due acknowledgement to authors. Submission policies for depositing documents at NAL-IR are allowed to accredited members of the institution (CSIR-NAL), or their delegated representatives. Authors may only submit their own work for archiving and in the case of collaborative work at least one of the creators must be from NAL. In case of authors’ deputation to and from NAL, it is mandatory to have NAL affiliation in their work for deposition to NAL-IR. The validity and authenticity of the content of submissions is the sole responsibility of the depositor. Creators can also send their articles to NAL-IR via e-mail, Though IR staff will deposit on behalf of creators, still they strongly recommend creators to deposit. In case NAL-IR receives proof of copyright violation, the relevant item will be removed or blocked immediately from the repository. With the aim of long term preservation all documents at repository are retained with continued readability and accessibility. In case of issues like legal or copyright violations, national security, plagiarized/falsified research work etc documents will be removed or withdrawn.

5. Customization

5.1 Subject
Research and development activities at CSIR-NAL largely covers aerospace and allied disciplines like material science, structural engineering, physics, space, geosciences mathematics etc.
Research output published by the institute can be broadly classified under above mentioned subject headings. Default library of congress (LC) subject heading list available with Eprints3.0 was customized to NASA subject categories to meet NAL-IR requirement.

5.2 Browse and Search
Ease and effective information leading to document retrieval from digital repository is an important feature which can be achieved by browsing and searching. Various browsing options: by document type, year, subject, and author are enabled. Along with above default options browse by department/division has been added to cater the requirements of NAL. There may be chances of formation of new department/s or merging of similar existing departments over a period of time etc., provisions are made to make available the respective documents under the same department browse options. Default simple and advanced search facility available with Eprints is enabled.

5.3 Homepage customization:
Being unique is a common psychological human trait in the present world. Similar logic applies to each website designed by human being on the Internet. Each website is designed with objective of embedding maximum information regarding their area of interest and concepts. NAL-IR homepage is designed to reflect the area of research carried out at the institution. All basic information regarding the institution, objectives, polices, guidelines, etc are reflected in the home page. It has been customized to list the latest (8 titles) in the home page so the users will come to know about the latest additions. Web links to OAI-PMH aggregator which harvest NAL-IR and other related repositories who work in the similar area are also been listed as a reference sources for IR users.
6. User Education
6.1 Need for User Education:
By analyzing annual reports and administrative documents of various departments it has been observed that NAL’s scientific output has crossed over 10,000 in three major document types. They are journal articles, conference papers and project documents, the last both in classified and unclassified format. Through informal survey it is revealed that majority of the scientific staff, especially younger scientists and technologists were not aware about concepts, sources, and advantages of Open access publishing and had apprehension about copyright and other legal issues, complicated uploading procedure etc. These issues motivated the IT staff of the Information Center (ICAST) to conduct extensive training program across the existing 3 campuses of NAL.

With formal approval from competent authority all 15 core departments of NAL were informed regarding the Outreach program. Dates and timings were finalized in consultation with heads of each department ensuring maximum people to benefit from the program. Necessary handouts, training program brochure were distributed to each division well in advance to create awareness about training program. Steps were taken to impart training at their premises. All scientists and technical officers were requested to attend training program.

Training topics broadly divided in to two sessions, first covering on exhaustive resources and services of ICAST, concepts, advantages of Open access publishing, various channels of OA publishing, copy right issues related to OA and Institutional repositories followed by the theme talk which concentrated on NAL-IR. The scientists were educated about IR polices (Submission, Metadata, Archiving, Access etc), technology used, work flow, access statistics. Available records for each division were listed and top ten authors who had deposited from the division were represented in graphical mode as an encouraging measure. This exercise was repeated for all divisions which enthused and motivated many participants follow the suite. At the end of the session, 20-30 minutes of interactive discussion with participants helped training team to judge their affinity towards OA and plan accordingly to improve the repository status. Number of scientists and technologists participated in user education program is given the following chart.

7. Usage
One of the prime objectives in establishing NAL-IR was to enhance research visibility of the organization and proof is required that IR content is being accessed. Usage of NAL-IR content is significantly increased over period of time from its inception. Following chart indicated the access statistics from various countries till metadata level. Views from NAL campuses are excluded from the current analysis. NAL-IR administration has noticed good numbers of full text downloads and document requests through email for restricted documents.

8. Future plans:
NAL-IR administration plans to implement interactive statistical module to highlight most downloaded title/s, authors, departments etc., in order to motivate scientists. Integrating ISI-WoS and Google scholar plug-ins to tract citations of journal and conference titles is also being thought for implementation. Conducting awareness programmes at regular intervals to populate the repository by getting hold of more documents is also in the pipeline.

References