**Establishment of an Open Access Repository in a university:
The Goa University saga**

Dr Gopakumar V
librarian@unigoa.ac.in
Dr Anuradha V
anuradhav@goa.bits-pilani.ac.in

Abstract

The scholarly communication scenario was totally monopolized by the publishers hindering the access to scholarly literature. Costs for subscription and restricted access alienated researchers from getting access to such literature. Open Access and Institutional repositories have come up to save the situation to a great extent. Some people believe that Open Access IRs are alternatives to the traditional publishing where as others believes that IRs only supplement the traditional publishing. However, Academic institutions all over the world have set up IRs and allowing faculty and researchers to self-archive or deposit their research output. This article explains in detail the story of the establishment of an IR at Goa University and the issues and impediments the team faced while establishing it.

Keywords: Open Access, Institutional Repositories, Traditional publishing, Scholarly communication.

Open Access: an introduction

The traditional scholarly communication cycle is characterized by the monopoly of journals. This monopoly has remained an impediment in access to research output. Academics always wanted to assert their control over the scholarship. But their efforts remained pointless for want of a clear cut strategy to overcome the impediment created by the journals. Some of them tried to publish their research outcome in their websites. The cost of journals escalated every year when the budgets to libraries were cut drastically creating a situation commonly called in library and information science parlance as "Serials Crisis".

Journal publishers reigned the scenario not allowing their monopoly in the scholarly communication process to ease a bit.

But the 2002 Budapest Open Access Initiative was a significant change in this whole situation. Academicians declared that they have the pragmatic solution to evade the monopoly of publishers – they have the Internet as a medium to publish. The declared that they will publish their research output in the public domain on the Internet for free and open access to any one needs it, discarding the monopoly of publishers. This concept was termed as "Open Access".

In the initial stages strategies were not clear, but with the 2003 Bethesda statement of Open Access and the Berlin Declaration of Open Access in October 2003, clarity in strategies were evolved. Open Access publications are to be freely and permanently available online on the Internet for unrestricted use with permission to distribute and reproduce in any medium providing due recognition to the original author/editor.

Two major models evolved in Open Access publication. They were named as "Golden Open Access publishing" and "Green Open Access publishing", sometimes called as Green and Gold routes for Open Access publishing. Golden Route for Open Access denotes the publication of a research article in an Open Access Journal. Presently there are hundreds of such Open Access Journals getting published from various corners of the world and dominating the scholarly publishing scenario. These are collectively listed in the website www.doaj.org. The funny part of this story is that even the proprietary journal publishers were also forced to start Open Access organs for their publications to sustain their position in the scholarly publishing cycle.

In Green route to Open Access the scholars will publish with commercial publishers and deposit their research papers in "Open Access Institutional Repositories" as per the copy right agreement. Clifford Lynch (2003), executive director of the Coalition for Networked Information, describes an IR as "a set of services that a university offers to the members of its community for the management and dissemination of digital materials created by the institution and its community members." These repositories being Open Access and remains in the public domain, serves as a tangible indicator of the quality of an academic institute. It also helps to demonstrate scientific, societal and economic relevance of the research activities of that institute to which it belongs. This in turn increases the visibility, status and public value of the institute. "Most repository platforms are full-text indexed in search engines, which offers scholars as well as the publishing institution high discoverability and wide dissemination opportunities."(Tennant, 2002) Hence within the first decade of the 21st century itself academic institutions have come out with institutional repositories on a global basis. Many Indian academic institutes and research centers also have followed the suite. The best repositories all over the world have been listed in the directory www.open doar.org.

Goa University and it's IR

The Goa University was established in the year 1985 and presently it has 6 faculties working in the campus with 24 departments and 3 research centers. Ever since it's inception lot of research activities are happening in the university and is visible from the graphical representation of the research output in numbers.
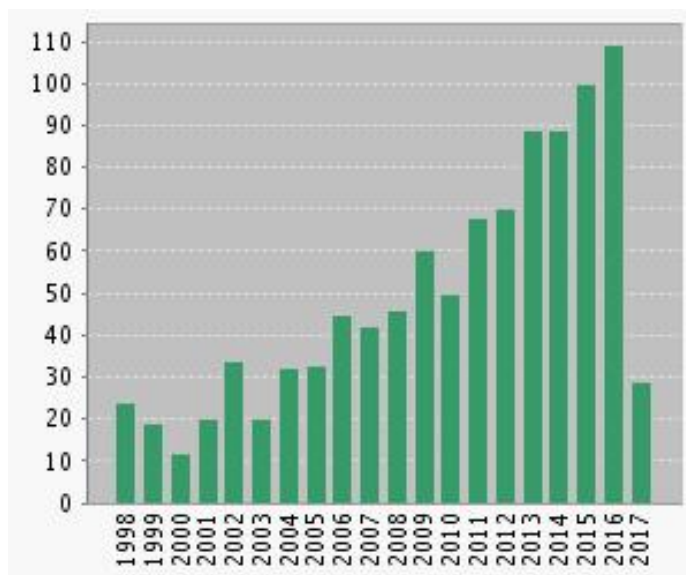
Figure 1. Published Items of Goa University in Each Year (Web of Science output)

Genesis

The establishment of an institutional repository for an institution existing for the last 30 years posed several issues and challenges. The story would not be complete without describing these issues and challenges. As such it is explained in seven steps.

They are

1. Digitization of available documents of the pre-digital era

2. Collection of metadata of publications related to Goa University

3. Collection of full-text of the publications related to Goa University

4. Establishment of the IR using a standard software

5. Establishing communities, sub-communities and collections.

6. Uploading of Metadata and full-text documents to the IR

7. Continuous monitoring of usage and maintenance of statistics.

Step1. Digitization of documents of the pre-digital era

The Goa University Library started its initiative of establishing the digital repository with the impetus it received from the award of a special grant from the INFLIBNET Centre under the Sodhganga project for establishing a digitization Centre at Goa University. The 516 theses held by the library at that time were subjected to scanning and converted to pdf file with editable text through OCR process. The work was carried out through outsourcing to an external agency as the library lacked necessary infrastructure. All theses in book form were opened from their binding and scanned. The images were then stored in three formats viz. .jpeg, .pdf and .tiff. The pdf files had editable text with them to help in indexing when they are uploaded in to the repository. This was possible through the OCR process. The scanned theses included theses from local languages also. They could not be subjected to OCR for want of required software.

Step 2. Collection of Metadata

As mentioned earlier, Goa University is in its third decade and a lot of research publications have emerged from this institution. The task of collecting information of past publications is a massive mission but was easy because of the availability of the copies of the annual reports maintained by the library which carried list of faculty publications on a regular basis. In some situations the same publications appeared in subsequent years twice but the team could get 70% of the details of publications through this method itself. Faculty members also came forward with the details of publications.

Step 3. Collection of full-text documents

The availability of metadata alone was not enough for a standard Institutional Repository. As far as possible full-text surrogates of the

articles also have to be placed along with the metadata. E-mails and consecutive notices were sent to faculty members to submit copies of their publications. Majority of the faculty members volunteered in submitting their publications in full-text hard copies. They were digitized later. The faculty members who had apprehensions about the copy right issues were told about the policy that will be followed by the IR and convinced them. This helped the team to collect the maximum possible metadata and full-text of faculty publications. We have decided to keep the metadata of an item in the repository, even if we don't have the full-text available with us. This will help in making users aware of the availability of an item.

As discussed earlier faculty members and research scholar were requested to submit copies of their publications or at least metadata. Simultaneously the team collected copies from databases where the university library has access to and also from the sister institutes and from the journal archives of the University. The team was monitoring the acquisition of original documents to the maximum extent possible.

Step 4 Installation of IR

Once we paved way to the gathering of faculty publications, we decided to install the software for the IR. First decision to take was about the software to be used. We were aware of software like DSpace, E-Prints and Greenstone. There were several rounds of discussions and we came to the conclusion to use DSpace due to the simplicity of it's interface. We made comparative study with the repository of the nearby research institute, the National Institute of Oceanography. The software was installed on a standalone server purchased out of Sodhganga grants.

Step 6 Repository structure

Once installing DSpace on the server was done, the team had to think of the structure of the repository. As known to us the Institutional Repository software allows us to create communities, sub-communities and collections. This should suite the institutional structure and should help in easy retrieval of items. The following structure was adopted after discussion with faculty and university authorities.

Community	Faculties of the University
Collection	Departments under vision faculties
Document types	Books, Books Chapter Conf. Article, Journal article, Technical Reports, Theses

Step 7 Populating the IR with metadata and documents

Once the structure was established the documents and metadata were uploaded to the repository. We used the technique of populating a .csv file with metadata and importing it into the system. The use of .CSV file help in easy data entry, as massive data entry was required for the entire documents collected from the faculty. In case of single documents received later metadata was entered individually.

Care was taken to maintain a standard file name, and in the using of names of authors. We followed the naming conventions adopted by the CSIR Central, which was available at www.csircentral.net. This helped us to maintain standard in the metadata, which was appreciated.

Step 8 Backing up and monitoring

Once it was fully functional the repository was given wide publicity through university website and e-mails. Continuous monitoring of the use of the repository and backing up of the data is

the last step which the Computer Center of the university is managing on a regular basis

Present status

At present the IR of Goa University has 9 communities and has 4666 full-text or metadata information under them. The communities as well as the number of documents /metadata available in them are shown in Table 1. below.

Table 1. Communities in Goa University IR

Communities	Number of documents / metadata
Faculty of Commerce and Management Studies	350
Faculty of Education	8
Faculty of Languages and Literature	1144
Faculty of Law	8
Faculty of Life Sciences and Environment	1242
Faculty of Medicine	0
Faculty of Natural Sciences	1080
Faculty of Social Sciences	814
Non-faculty Groups	10

The above list very clearly shows that some faculty is having very high output from their departments.

Table no 2. Comparison of research output from various departments

Document types	Number of documents / metadata
Dept. of Chemistry	513
Dept. of Computer Science & Technology	92
Dept. of Earth Science	95
Dept. of Electronics	68

Dept. of Mathematics	55
Dept. of Physics	257
Dept. of Biotechnology	74
Dept. of Botany [438]	438
Dept. of Marine Sciences [166]	166
Dept. of Microbiology [205]	205
Dept. of Zoology [193]	193

If we further check under departments under a faculty we will see that some departments are going very high and some lagging behind. This will detract departments and the faculty of them becoming exposed and negatively affecting the purpose of an IR. In the Table no. 2 it can be seen that the research output from Chemistry Botany and Physics are very high in comparison to other departments.
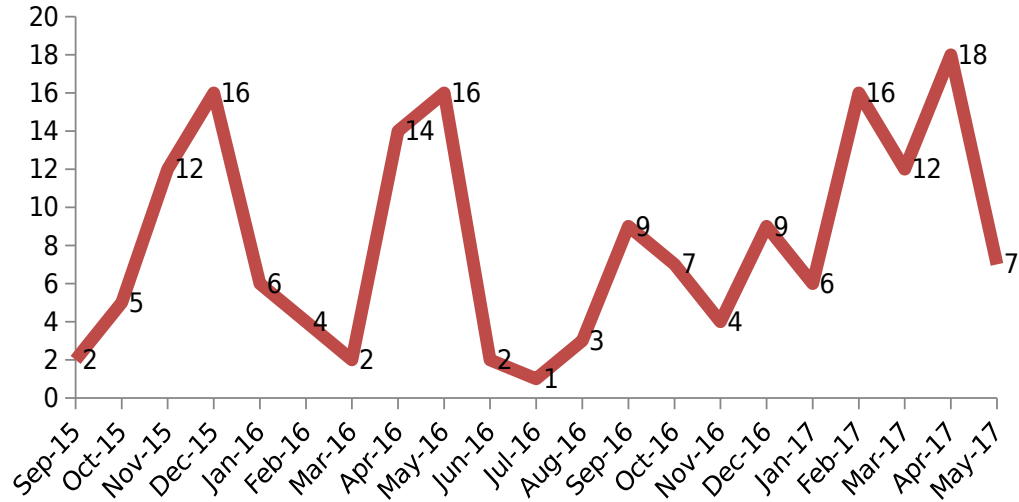
The item wise list of various document types available in the IR is shown in Table no. 2

Table 2. Document types in Goa University IR

Document types	Number of documents / metadata
Book	251
Book chapter	619
Conference article	274
Journal article	2888
Technical report	10
Thesis	591

The documents which have copy right issues are not allowed to be openly downloaded from the Institutional repository. But a needy person can make a request to the admin and can send a copy of the document over e-mail. We had received several requests and the extent of such a service is depicted in the Figure 2.

Figure 2. Documents supplied from IR on request (August 2015 to May 2017)

Conclusion

The establishment of the IR for Goa University has helped to get more visibility to the scholarly output of the faculty and research scholars. This is visible from the enthusiasm of the faculty to deposit their latest articles and they opine that they are getting more citations and requests from other researchers. The IRs are meant for self archiving by the researchers and faculty, but at present we are unable to allow this as we are skeptical about the metadata that will be provided by them. However the maintenance of the IR with international standards is the responsibility of the institute and its library which in turn will give more credibility to the research work undertaken in the university.

References

Xia, J., & Sun, L. (2007). Factors to Assess Self-Archiving in Institutional Repositories. Serials Review, 33(2), 73–80. https://doi.org/10.1016/j.serrev.2006.09.002

Dunsire, G. (2008). Collecting metadata from institutional repositories. OCLC Systems & Services: International Digital Library

Perspectives, 24(1), 51–58.
https://doi.org/10.1108/10650750810847251

van Westrienen, G., & Lynch, C. A. (2005). Academic Institutional
Repositories. D-Lib Magazine, 11(9), 1–11.
https://doi.org/10.1045/september2005-westrienen

Bankier, J.-G., Foster, C., & Wiley, G. (2009). Institutional Repositories—
Strategies for the Present and Future. The Serials Librarian, 56(1–
4), 109–115. https://doi.org/10.1080/03615260802665423

Puplett, D. (2010). The Economists Online Subject Repository—Using
Institutional Repositories as the Foundation For International Open
Access Growth. New Review of Academic Librarianship, 16(S1), 65–
76. https://doi.org/10.1080/13614533.2010.509490

Drake, M. A. (2004). Institutional Repositories. Hidden Treasures.
Searcher: The Magazine for Database Professionals, 12(5), 41–45.
Retrieved from http://search.ebscohost.com/login.aspx?
direct=true&db=bth&AN=13164121&site=ehost-live

Giesecke, J. (2011). Institutional repositories: keys to success. Journal
of Library Administration, 51(5–6), 529–542.
https://doi.org/10.1080/01930826.2011.589340

Bailey, Jr., C. W., Coombs, K., Emery, J., Mitchell, A., Morris, C., Simons,
S., & Wright, R. (2006). Institutional Repositories. SPEC Kits, 48(6),
22. https://doi.org/10.1108/03055720310509064

Kim, J. (2011). Motivations of Faculty Self-archiving in Institutional
Repositories. Journal of Academic Librarianship, 37(3), 246–254.
https://doi.org/10.1016/j.acalib.2011.02.017

Lynch, C. a. (2003). Institutional Repositories: Essential Infrastructure
For Scholarship In The Digital Age. Portal: Libraries and the
Academy, 3(2), 327–336. https://doi.org/10.1353/pla.2003.0039

Tennant, R. (2002). Institutional repositories. Library Journal, 127(15),
28, 30. https://doi.org/10.1080/03615260802665423

●