

Adapting Kepler Framework for Enriching Institutional Repositories: An Experimental Study

A. Ramnishath, Francis Jayakanth, Filbert Minj, and T.B. Rajashekar

National Centre for Science Information, Indian Institute of Science, Bangalore, India
ramnishath@gmail.com,
{franc, filbert, raja}@ncsi.iisc.ernet.in

1 Introduction

There is growing trend towards academic and research organizations to establish OAI-compliant institutional repositories. ePrints@IISc (<http://eprints.iisc.ernet.in/>) is the institutional repository of Indian Institute of Science (IISc), Bangalore. Though the repository is growing steadily, mediated submission by the ePrints@IISc staff is the predominant mode of enriching the repository. We have been exploring viable means of getting our researchers to contribute more actively to the repository. Observations have recently been made as to why researchers might be reluctant to contribute to central repositories [1]. It has been suggested that it might be useful to provide researchers with tools to easily create and share Personal Digital Repositories (PDR) designed to organize and facilitate their research and learning agendas. The collection in the PDR is built and managed by the scholar based on individual needs. A network of such PDRs could form the basis for a bottom up, organic approach to enrich centralized institutional repositories.

2 Implementation Details

We carried out a study to explore the bottom up approach using the Kepler Communal Digital Library Framework [2, 3]. In adapting the Kepler framework, our objectives were: a) To enhance Kepler Archivelet to serve as a PDR for the individual researcher, to provide controlled web access to publications in the PDR, and allow controlled exposure of metadata and full text for harvesting by the Kepler Group Server, and b) To enhance Kepler Group Server to enforce document type and metadata policies of the ePrints@IISc repository on the network of registered PDRs and to harvest metadata and full text from individual PDRs as per the harvesting rights permitted by the PDRs and ingest them to ePrints@IISc repository. Key requirement was to enforce the collection policies of the ePrints@IISc repository at PDR level for content publishing, and at the Kepler Group Server level for harvesting permitted content from PDRs. This was achieved by introducing three key enhancements:

- A ‘foreign.xml’ file at the Group Server level, configurable by the ePrints@IISc repository manager reflecting the collection policies of the repository, to be conformed to by PDRs. Each PDR registering with the Group Server automatically downloads this file and is used during content publishing

- Customized metadata driver at PDR level imposing the policies encoded in 'foreign.xml' file
- Customized 'NCSI-OAI-DC' metadata scheme for harvesting by the Group Server as per the collection policies, in addition to universal OAI-DC module.

Apart from carrying out these specific enhancements, we could successfully achieve the following:

- Adapted the Kepler Archivelet and Group Server framework in a centralized institutional repository setting, to support bottom up content publishing and repository level content aggregation as per the collection policy requirements of the central repository
- Enhanced the Kepler Archivelet to develop a prototype PDR supporting key features such as: richer menu of document types (e.g. conference papers, presentations, preprints) for publishing including metadata specific to these document types; option to register with the Group Server; option to produce a PDR home page based on a configurable style sheet showing personal details of the researcher and archived publications providing an additional means to reach users on the Internet; and option to manage access rights indicating which publications to appear on the PDR home page and which publications can be harvested at metadata and/or full text level
- Enhanced the Kepler Group Server to enforce key collection policies (e.g. document types, metadata and subject categories) on registered PDRs; to harvest metadata and full text as per these policies and the access rights defined by PDRs; and to export this content in XML format as per the import requirements of ePrints@IISc repository, and
- Imported the content to ePrints@IISc repository with provision for indicating possible duplicates, for review by the repository moderator

While this study was carried out in the specific institutional repository set up at IISc, which uses the EPrints.org repository software, the enhancements we have carried out could be applicable in different repository software situations.

References

1. Gandel, Paul, B, Katz, Richard, N, Metros, Susan, E.: The weariness of the flesh: Reflection on the life of the mind in an era of abundance. *EDUCAUSE Review*, Vol. 39, no. 2 (2004) 40-51
2. Liu, X, Kurt, M, Zubair, M.: Enhanced Kepler Framework for Self-Archiving. *Proceedings of the 2002 International Conference on Parallel Processing Workshops*, p.455, August 18-21 (2002)
3. Maly, K, Nelson, M, Zubair, M, Amrou, A, Kothamasa, S, Wang, L.: Kepler – A Communal Digital Library. Technical report, Old Dominion University (2004)