NTU-IR: An Institutional Repository for Nanyang Technological University using DSpace

Jayan C Kurian[1], Dion Hoe-Lian Goh[1], Joy Lynn Wheeler[2]
[1]Division of Information Studies, Wee Kim Wee School of Communication & Information
[2]Digital Resources Division, Library
Nanyang Technological University, Singapore
{jayan, ashlgoh,jwheeler}@ntu.edu.sg

Abstract

In today's digital networked environment, institutional repositories play a significant role in scholarly communication. This is achieved by offering a practical and cost-effective means of preserving and disseminating an organization's intellectual output. The open and interoperable nature of institutional repositories has received widespread considerations from stakeholders of academic and research institutions world-wide. In addition to digital content preservation and dissemination, an institutional repository reflects an organization's visibility and reputation. By carefully customizing repository functionality, an institutional repository can cater to a wide variety of information needs in a scholarly environment. This paper discusses the initial experience of designing and implementing an institutional repository at Nanyang Technological University (Singapore). The paper discusses institutional repositories and describes repository features associated with content acquisition, interface design, metadata descriptors, and open-access authorization policies. We will discuss why DSpace was chosen for implementing our institutional repository, Nanyang Technological University – Institutional Repository (NTU-IR) and the unique features we plan to implement to meet the needs of our users.

Keywords: Institutional repositories; Content acquisition; Interface design; Metadata descriptors; Open-access authorization.

1. Introduction

An institutional repository is defined as a central reference base for an organization's intellectual output. Individual knowledge assets contributed by academic communities to an institutional repository helps to build a knowledge-base that allows information sharing and reuse. To deliver a richer set of scholarly assets in a meaningful way to repository consumers and for the timely preservation of digital resources, an institutional repository offers a viable solution that can be integrated across networked environments. The need for Nanyang Technological University – Institutional Repository (NTU-IR) is to house collections from the Asian Media Information & Communication Centre (AMIC) as the first stage. AMIC, a leading centre in the Asia-Pacific region, geared towards the development of media and communication expertise for academic and industrial audience, has wide variety of collections in digitized format (AMIC, 1971). To make this information accessible to an online community, it has to be stored in a repository that supports long-term preservation and reuse. In subsequent stages, our institutional repository will act as a repository for scholarly collections from research centers organized under various academic divisions. DSpace (DSpace Federation, 2002), an open source digital repository system was chosen based on the checklist provided in (Goh, 2006) to evaluate open source digital library software. From the checklist preliminary evaluation criteria included are content acquisition of a variety of file types, use of metadata standards, user interface, search and browse functions, and the ability of a single sign-on and access control system.

The following paper is structured as follows. Section 2 describes content acquisition; section 3 describes metadata standards and customization and search functions. Section 4

describes the NTU-IR user interface while section 5 describes our single sign-on and access control policies. An implementation overview is given in section 6, and section 7 concludes this paper with a description of the experience gained and considers future enhancements.

2. Content Acquisition

One of the major prerequisites for an institutional repository to be a successful implementation is its content acquisition strategies. An efficient content acquisition strategy makes it easier to acquire scholarly information made available by knowledge communities into a repository. DSpace supports content acquisition through an Item Import procedure. Digital assets annotated with metadata elements are imported into NTU-IR using this facility. Digital assets can take the form of text formats, image formats, presentation formats, structured formats, and audio and video formats. The Item Import procedure can also be applied to a batch of digital assets organized in a directory. This directory includes an XML formatted Dublin Core file (DCMI, 1995), a contents file that specifies resource identification, and the digital asset to be imported into the specified repository collection. A sample directory structure is shown in Figure1 and a Dublin Core file is shown in Figure2.

PDF — AMIC_1980_10_24
Adobe Acrobat Document
1,204 KB

contents
File
1 KB

dublin_core
XML Document
1 KB

Figure1: Batch Import Directory Structure

```xml
<?xml version="1.0" encoding="iso-8859-1" ?>
<!-- title of pdf AMIC_1980_10_24.pdf -->
- <dublin_core>
    <dcvalue element="creator" qualifier="conference">AMIC-ICC-HHS
      Seminar on Communication and Child/Family Welfare :
      Singapore, Oct 27-31, 1980.</dcvalue>
    <dcvalue element="title" qualifier="none">Child and Family
      Welfare Programmes and Role of Communication in
      Thailand.</dcvalue>
    <dcvalue element="contributor" qualifier="author">Bumrongsook,
      Siha-Umphai.</dcvalue>
    <dcvalue element="date" qualifier="issued">1980-10-27</dcvalue>
</dublin_core>
```
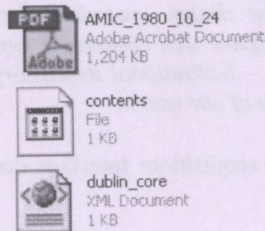
Figure 2: Dublin Core file

A batch Item Import procedure for digital assets uses the NTU-IR administration account (e.g. dspaceadmin@ntu.edu.sg), collection identification number (e.g. 123456789/257), and a map file (e.g. mapfile1212) that handles references to imported items. A persistent handler is then automatically generated for each imported item.

3. Metadata Customization and Search Functions

For effective digital asset retrieval, rich metadata descriptors have become a necessity. In NTU-IR, we customized the simple item record display (e.g. Conference Title) and enabled customized metadata elements for enhanced search. We followed a strict Dublin Core to Marc crosswalk (MARC, 2001) in order to maintain a high degree of compatibility with current cataloging standards. Full text search feature as well as advanced search based on multiple metadata elements is also enabled in NTU-IR. Individual collections in the repository as well as NTU-IR as a whole can be searched for relevant digital resources. A simple item record display for a digital resource is shown in Figure3.
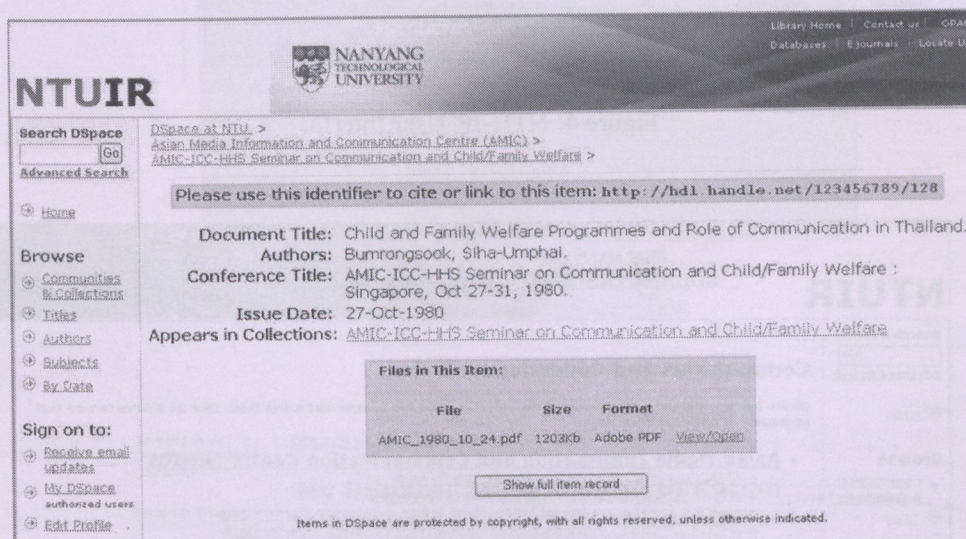

Figure 3: Simple Item Record Display

4. User Interface

The user interface is an important component that provides an easy navigation for library patrons and library administrators. The user interface of NTU-IR is given in Figure4. The default DSpace interface has been customized to give an institutional look and feel by applying customized header pages. The customized banner of NTU-IR integrates with other online library resources to support users' information seeking. Collections in NTU-IR that are named based on AMIC conference proceedings are shown in Figure5. To further customize the user interface at the community and collection level, and to support multi-lingual features, we plan to employ Manakin, an XML-based UI framework (Manakin, 2005).
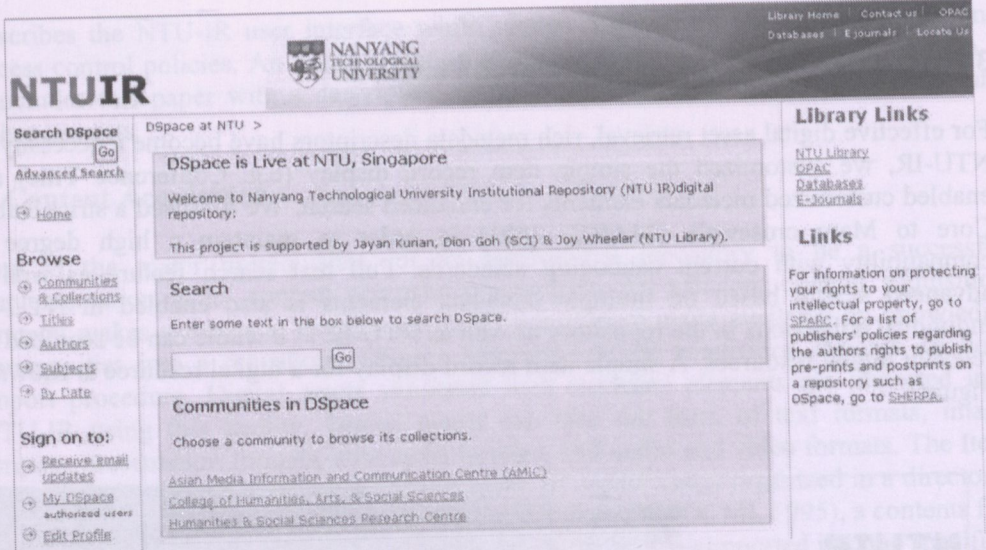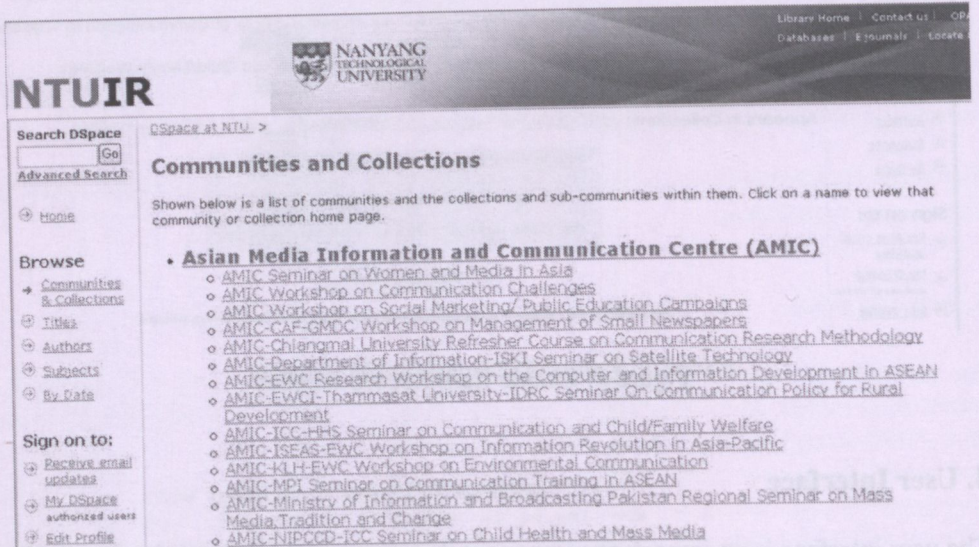
Figure 4: NTU-IR User Interface

Figure 5: Communities and Collections in NTU-IR

5. Single Sign-on and Access Control

An application is widely used and rated user friendly if it can authenticate users based on an existing authentication key enabled for domain resources. For NTU-IR, we decided on the single sign-on (SSO) authentication methodology based on NTU's hierarchical Active Directory structure. This allows all NTU staff and students to log into NTU-IR with their existing user accounts registered within NTU. Upon successful authentication, users can be automatically classified into a default group specified in the DSpace configuration file (i.e. dspace.cfg). This allows a DSpace administrator to grant relevant group privileges to be inherited by group roles. We considered three types of resource access: full text access, privileged access, and restricted access. For full text access, users have to be members of the default anonymous group. By enabling privileged access, only metadata descriptors are made searchable. Restricted access can be enabled for digital resources to be shared

among a group of authorized users. Access control policies can be enabled by the repository authorization module that allows access at the item level and at the bitstream level. This allows copyright policies for digital assets in an institutional repository. A single sign-on authentication and an individual workspace is shown below in Figure 6 and Figure 7.

Log In to NTU-Institutional Repository   Help...

Please enter your NTU Login Account and password into the form below.

NTU Login Account: [   ]

Password: [   ]

[Log In]

Have you forgotten your password?

Figure 6: Single Sign-on Authentication

Logged in as
jayan@ntu.edu.sg
(Logout)

Search DSpace
[   ] [Go]
Advanced Search

⊕ Home

Browse
⊕ Communities
  & Collections
⊕ Titles

DSpace at NTU. >

My DSpace: Jayan Chirayath Kurian   Help...

[Start a New Submission]   [View Accepted Submissions]

See Your Subscriptions

Authorization Groups I'm a Member Of

- COLLECTION_5_SUBMIT
- MYORG1_USERS

Figure 7: User Authentication and Workspace

6. Implementation

The NTU-IR project is divided into three phases with each phase spanning a period of one year. The initial phase of this project included digitizing AMIC conference proceedings and setting up a development environment. The NTU-IR database server along with a development server and a web server have been configured for a multi-tier architecture in a Windows environment. In the second phase, a soft launch will be done for evaluation and feedback. The third phase includes continued service development to accommodate feedback from users and the actual launch of the NTU-IR for public access. The supporting software requirement for DSpace 1.4 includes Postgresql 8.1, Apache Ant 1.6, and JDK 1.5 along with Tomcat 5.5 on a Windows 2003 server platform. The current status of the NTU-IR project at the writing this paper is in the early part of the second phase of development.

Kurian, Jayan C., Goh, Dion Hoe-Lian & Wheeler, Joy Lynn

7. Conclusion

The NTU-IR project aims to create an institutional repository to store and make accessible online works owned, published and created by NTU staff and students. The stakeholders of this project are Wee Kim Wee School of Communication and Information and the NTU Library. Institutional Repository features discussed in this paper include content acquisition, interface design, metadata descriptors, and open-access authorization policies. In our ongoing work, we will be considering customization of the content submission interface, additional metadata schema support, and proximity search that will add value to our institutional repository. In addition we will be considering archival standards and procedures (Tansley, 2003), security and scalability issues and standards for continued maintenance.

Acknowledgements

We wish to thank DSpace committers and DSpace mailing-list members for their valuable advice and suggestions.

References

AMIC. 1971. *The Asian Media Information & Communication Centre*. Available at: http://www.amic.org.sg/

DSpace. 2002. *The DSpace Federation*. Available at: http://dspace.org/index.html

DCMI. 1995. *The Dublin Core Metadata Initiative*. Available at: http://dublincore.org/

Goh, D.H., Chua, A., Khoo D.A., Khoo, E.B., Mak, E.B., and Ng, M.W. 2006. A checklist for evaluating open source digital library software. *Online Information Review*, Vol. 30, no.4: 360 - 379.

MARC. 2001. *Machine-Readable Cataloging to Dublin Core Crosswalk*. Available at: http:// www.loc.gov/marc/marc 2dc.html

Manakin. 2005. *The DSpace XML UI*. Available at: http://di.tamu.edu/projects/xmlui/manakin/

NTU-IR. 2006. *Nanyang Technological University - Institutional Repository*. Available at: ·http://155.69.24.193:8080/dspace/

Tansley, R. Bass, M. and Smith, M. 2003. *DSpace as an open archival information system: current status and future directions*. In: Koch, T., Solvberg, I.T. (eds.): *Research and Advanced Technology for Digital Libraries. Lecture Notes in Computer Science*, Vol. 2769. Springer, New York (2003) 446-460.