



Open Source Software Packages for E-Resource Management

DR. MEHULKUMAR S. PATEL

Librarian, Sankalchand Patel College of Engineering, Visnagar

&

MR. RAJANIKANT P. PATEL

Research Scholar, HNGU, Patan, Gujarat &

Librarian, Narsinhbhai Patel Dental College and Hospital, Visnagar
Gujarat (India)

Abstract:

This paper describes the detailed information about E-Resource Management Software Packages. The software packages involve DSpace, EPrints, Greenstone, Fedora etc. Main points included in this paper are, features of ERM software packages, authentication and authorization detail, accessibility, supported software/operating systems, web support, customization detail, administrative functions, searching facility etc. This may be useful for the library personnel who deal with the E-Resources and facing many problems in E-Resource Management activities.

Keywords: Apache Ant, DSpace EPrints, E-Resource, Open Source, Fedora Greenstone, Metadata, Open Archives Initiative, Tomcat and PostgreSQL,

1. Preface

In present days publication work of information resources in Electronic form is increasing day by day. So, librarian's duty became more challenging towards the E-Resource Collection development, E-Resource selection, Licensing process, E-Resource acquisition process, E-Resource access management etc. Now a day's some readymade E-Resource management software systems are available in market which can release the librarian's burden & help in smooth functioning of E-Resource Management process.

2. Definition & Meaning of E-Resource Management:

Definition of E-Resource Management given below,

“Systems developed to assist librarians in the control of licensed third-party resources published electronically (database, e-books, e-journals etc.) including license management, renewal, legal use, access management, and collection development.”

- *Encyclopedic Dictionary of Library and Information Science, Volume-2, 2010.*

In 2001, a small group of academic librarians began to create metadata specifications for managing electronic subscriptions and their associated titles. The group was later asked by the Digital Library Federation (DLF) to deliver formal specification for vendors, as a replacement for various home grown systems. Innovative interfaces inc. was the first library automation vendor to market and ERM software module based on the DLF specifications, and other vendors have followed suit.

- *Encyclopedic Dictionary of Library and Information Science, Volume-2, 2010.*

3. E-Resources Management Software Packages

For smooth running of E-Resource Management some readymade Open Source Software packages are available in the market. Some examples are as follow.

3.1 DSpace

DSpace is a joint venture of “Massachusetts Institute of Technology’s libraries (MIT)” and “Hewlett-Packard (HP) Labs” and it is freely available as an open source system that can be customized and improved further. DSpace runs on UNIX or LINUX operating system and also successfully installed on Window XP and Window 2000 professional. Other open source tools like Apache Ant, Apache Tomcat and PostgreSQL are also needed to install the DSpace.

Some characteristics of DSpace are as:

1. It is a service model for open access and/or digital archiving for continuing access.
2. Provide a platform to frame an Industrial Repository and the collections are searchable and retrievable by the web.
3. Helps to make available institute-based scholarly material in digital formats. The collections will be open and interoperable.

The features of DSpace as Digital Management Software are as follows:

3.1.1 Authentication

DSpace allows contributors to limit access to items in DSpace, at both the collection and the individual item level. The mechanism whereby the system securely identified its users.

3.1.2 Authorization

The mechanism by which a DSpace determines what level of access a particular authenticated user should have to secure resources controlled by the system is done by keeping access control policies that allow it to understand what credentials are required (if any) to undertake particular actions upon particular resources. Authentication is provided through user Passwords X509 certificates or LDAP.

Some other features are, Not-dynamic HTML document support, OAI-PMH support, object management, import & export, usage statistics, Handle system. Customization & types of document supported.

3.2 EPrints

The EPrints software is the largest and most broadly distributed or installed software as any other open repository system. The University of Soluthampton (UK) developed it and first version of the system was publicly released in late 2000. Installation of this software is relatively easier; its installation needs support from the other open source software like LINUX, Apache and MySQL.

Listed below are some of special features possessed by the EPrints:

(a) Accessibility via web browser

EPrints provides web based interface that makes it easy to use and administer.

(b) Full Text and Field Search

Searching is based on metadata not full text based search is supported by EPrints. Searching in EPrints allows scanning each of the metadata field types in the database by using simple or advanced search. Any metadata field can be searched with fine granularity by SQL querying the database.

(c) Administrative function provided

EPrints archive can use any metadata schema as being provided by the administrator. The administrator decides what metadata fields are held about each EPrints item. This is specified in three or four stages: (1) Definition of a maximal set of metadata fields that should be stored (e.g. authors, title, journal, journal volume, etc.), (2) Definition of different types of EPrints (e.g. refereed journal article, thesis, technical report, unpublished preprint, etc.) (3) Specification for each type which metadata fields should be stored, and which of those fields are mandatory, (4) Decide how these metadata fields should be projected into the Open Archives world. (If necessary, interoperability can be switched off, but this is strongly discouraged.).

(d) Open Source Software

EPrints uses traditional technologies and runs on pure Open Source systems. It uses MySQL, Apache database and web server. MySQL is the world's most popular open source database, recognized for its speed and reliability and Apache has been the most popular web server on the Internet since April of 1996. EPrints is programmed by using the script language "Perl", that is low level but powerful.

(e) Three user roles:(administrator, editor and author)

1. Administrator role controls all back-end options such as organization of records, web interface appearance and functionality, and all other server-side settings.
2. Editor role reviews submissions before they are published online and may edit metadata on submissions to maintain consistency or correct errors.
3. (iii) Author role allows submission of documents and management of previously submitted documents.

(f) OAI-PMH Support

EAS is fully interoperable with OAI (Open Archives Initiative) Protocol for Metadata Harvesting. Open Archives protocol allows sites to programmatically retrieve or 'harvest' the metadata from several sources, and offer services using that metadata, such as indexing or linking services. Such a service allows e-prints servers create the potential for a global network of cross-searchable research information, by allowing the contents of servers around the world searched simultaneously by using the OAI (Open Archives Initiative) protocol.

(g) Multilingual Support:

Unicode is used throughout the software, allowing any language to be processed in a consistent manner.

(h) File formats supported:

Functions with many file types, including: PDF, HTML, JPEG, TIFF, MP3, and AVI etc. Metadata schema can be tailored to meet the requirements.

(i) Statistics

Statistics are provided for administrative usage. Statistical reports/summary can be used for performing analysis on repository.

(j) Customization

The EPrints data modal consist of user defined metadata. In order to export data in other formats plugins can be written. For developers who wish to access the core Digital Library functionality Core API in Perl language is provided.

(k) Item preview in EPrints

Thumbnail preview of documents and images is generated automatically upon file upload.

3.3 Greenstone

Greenstone Digital Library Software, well known as GDLS, allows users to build their own digital libraries. It is multilingual software which supports various languages and is available in English, French, Spanish, Russian, Kazakh and Vietnamese. Greenstone runs on Windows, UNIX and MacOSX with the help of Java environment. It can be installed in two ways, for a local library (Intranet) it requires Java virtual environment and for Web library it requires Java virtual environment and Apache or IIS web server running and it can be customized.

Some of special features of Greenstone are,

(a) Accessible via web browser

Collections are accessed through a standard web browser (Netscape or Internet Explorer) and combine easy-to-use browsing with powerful search facilities.

(b) Full Text and Field Search

The user can search the full text of the documents, or choose between indexes built from different parts of the documents.

(c) Flexible browsing facilities

The user can browse lists of authors, lists of titles, lists of dates, classification structures, and so on. Different collections may offer different browsing facilities and even within a collection, a broad

variety of browsing interfaces are available. Browsing and searching interfaces are constructed during the building process, according to collection configuration information.

(d) Create access structures automatically

The Greenstone software creates information collections that are very easy to maintain. All searching and browsing structures are built directly from the documents themselves. No links are inserted by hand, but existing links in originals are maintained. This means that if new documents in the same format become available, they can be merged into the collection automatically. Indeed, for some collections this is done by processes that wake up regularly, scout for new material, and rebuild the indexes—all without manual intervention.

(e) Make use of available metadata

Metadata, which is descriptive information such as author, title, date, keywords, and so on, may be associated with each document, or with individual sections within documents. Metadata is used as the raw material for browsing indexes. It must be either provided explicitly or derivable automatically from the source documents. The Dublin Core metadata scheme is used for most electronic documents; however, provision is made for other schemes.

(f) Plug-in extends system's capabilities

In order to accommodate different kinds of source document, the software is organized in such a way that “plug-in” can be written for new document types. Plug-in currently exist for plain text, html, Word, PDF, PostScript, E-mail, some proprietary formats, and for recursively traversing directory structures and compressed archives containing such documents.

(g) Customization

The Greenstone allows customization of presentation of collection that are based on Extensible Style sheet Language transformation (XSLT) and other agents that govern the definite functions of Digital library.

(h) Designed for Multi-gigabyte collection

Collections can contain millions of documents, making the Greenstone system suitable for collections up to several gigabytes.

(i) Multilingual Support

Unicode is used throughout the software, allowing any language to be processed in a consistent manner. To date, collections have been built containing French, Spanish, Maori, Chinese, Arabic and English. On-the-fly conversion is used to convert from Unicode to an alphabet supported by the user's web browser.

(j) Collections support multiple formats:

Greenstone collections can contain text, pictures, audio and video clips. Most non-textual material is either linked in to the textual documents or accompanied by textual descriptions (such as figure captions) to allow full-text searching and browsing.

(k) Administrative function provided:

An “administrative” function enables specified users to authorize new users to build collections, protect documents so that they can only be accessed by registered users on presentation of a password, examine the composition of all collections, and so on. Logs of user activity can record all queries made to every Greenstone collection.

(l) Collections can be published on the Internet or on CD-ROM:

The software can be used to serve collections over the World-Wide Web. Greenstone collections can be made available, in precisely the same form, on CDROM . The user interface is through a standard web browser (Netscape is provided on each disk), and the interaction is identical to accessing the collection on the web—except that response times are more predictable. The CD-ROMs run under all versions of the Windows operating system.

3.4 Fedora

Fedora is the open source software jointly development by the University of Virginia and Cornell University. Fedora is a digital object repository management system, it is designed to be a foundation

upon which full-featured institutional repositories and other interoperable web-based digital libraries can be built.

3.4.1 Key features of Fedora Repository Software

- Store all types of content and its metadata
- Digital content of any type can be managed and maintained
- Metadata about content in any format can be managed and maintained
- Scale to millions of objects
- Access data via Web APIs (REST/SOAP)
- Provide RDF search (SPARQL)
- Rebuilder Utility (for disaster recovery and data migration)
- The entire repository can be rebuilt from the digital object and content files.
- Content Model Architecture (define "types" of objects by their content)
- Many storage options (database and file systems)
- JMS messaging (your apps can "listen" to repository events)
- Web-based Administrator GUI (low-level object editing)
- OAI-PMH Provider Service
- GSearch (fulltext) Search Service
- Multiple, customer driven front-ends.

4. Conclusion

Many academic and special libraries are acquiring various E-Resources to enhance their library and information services. U.G.C. and AICTE, New Delhi are also encouraging the University and College libraries to purchase such E-Resource packages. Because of that librarians are facing lots of challenges in E-Resource related activities. This paper can be helpful to library professionals in choosing best option of E-Resource Management Software Packages.

References

1. Encyclopaedic Dictionary of Library and Information Science, Volume-2, 2010.
2. Raghavendra, M. & Shivarama, J. (2008). Building and Managing Electronic Resources in a Networked Environment: A Modular Approach. International Conference on Knowledge For All Role of Libraries and Information Centres (Proceedings),375-377.
3. Thoudam, Suleta Devi. Management of E-Resources in the Modern Library Information System: An Outlook. p.362-363.
4. Yu, Holly and Breivold, Scott. (2008). Electronic Resource Management in Libraries Research and Practice. New York: Information Science Reference. p.73-88.
5. www.dspace.org
6. www.eprints.org
7. www.fedora-commons.org
8. www.greenstone.org