DSpace—A Base for Developing Repositories in India

Dr. Geetanjali Rana Kanwar*
Assistant Librarian, Govind Ballabh Pant University of Agriculture and Technology, Pantnagar, Uttarakhand, India

Abstract

DSpace is an open source dynamic digital repository. Web-based interface makes it easy for a submitter to create an archival item by depositing files. It was designed to handle any format from simple text document to datasets and digital video. Data files also called bitstreams are organized together into related sets. Each bitstream has a technical format and other technical information. This technical information is kept with bitstreams to assist with preservation over time. An item is an archival atom consisting of group related to content and associated descriptions (metadata). An item exposed metadata is indexed for browsing and searching. Items are organized into collections logically related material. A community is the highest level of DSpace content hierarchy. They correspond to parts of the organization such as department, labs, research centres/schools. The modular architecture allows for creation of large repositories that can be extended across institutional boundaries. It is committed to reliable file preservation to offer functional preservation where files are kept accessible as technology formats, media and paradigms evolve over time for many types of files as possible. The end-user interface supports browsing and searching the archives. Once the item is located, web-native formatted files can be displayed in a web-browser while other formats can be downloaded and opened with a suitable application programme.

Keywords: Metadata, digital, preservation, DSpace, e-theses, Digital Library, JAVA, HTML, institutional repository

*Author for Correspondence E-mail: geetanjali.rk@gmail.com

INTRODUCTION

DSpace is a platform that allows to capture items in any format whether in text, video, audio and data. It distributes it over the web. It indexes work for the users to search and retrieve items. It preserves digital work over long term. It is typically used as institutional repository which captures and ingest of materials including metadata about materials. It has easy access to the materials both by listing and searching. It supports long term preservation of the materials. It provides a way to manage research materials and publications in a professionally maintained repository to give them greater visibility and accessibility over time.

DSpace supports submission of, management of and access to digital content, format text, images, audio, video organised on organizational needs of large universities which includes communities and collections. The features of DSpace are:

- Digital Preservation persistent ID’s, support levels for different file formats.
- Access Control
- Versioning
- Search and Retrieval
- Based on Qualified Dublin Core metadata
- OAI-PMH data provider and to support metadata harvesters.

FOUNDATION OF DSPACE

DSpace introduced in July, 2000 by HP-MIT alliance. It was formed in 2007 as a non-profit organization to provide support to develop use of digital technology. It leads to development of open source software to enable permanent access to the digital works. It has a strong network of service providers and training resources. It also supports in using the open
standards and software. DSpace was developed as open source so that institutions and organizations with minimal resources could run it. It is designed to run on UNIX platform with other open source middleware, tools and programmes written by DSpace team. All the original codes are used in JAVA programming language.

Other pieces of the technology stack include a relational database management system (PostgreSQL), a Web Server and JAVA Servlet engine (Apache and Tomcat both from Apache Foundation), Jena (an RDF toolkit from HP Labs), OAIC at from OCLC and several other useful libraries. The system is available on SourceForge linked from both the DSpace informational website and HP Labs site.

The different software releases of DSpace are:
Version 1.0—8th November, 2002
Version 1.1—8th May, 2003
Version 1.2—13th August, 2004
Version 1.3—3rd August 2005
Version 1.4—26th July, 2006
Version 1.5—25th March, 2008

Description of DSpace are:
Year of creation: 2002
Development Organization: DSpace Foundation (MIT/HP)
Programming Language: JAVA, Database; PostgreSQL, Oracle
Metadata format: Dublin Core
Website: http://dspace.org/

DEVELOPMENT MODEL OF DSPACE
The code for DSpace is kept within a source code control system (http://dspace.svn.sourceforge.net/viewer/dspace/). The system allows code to be added or modified over time while maintaining a track of all changes and a note of why the change was made and by whom. It assist with development of software and ensures the quality and traceability of the code. Any past version of DSpace can also be downloaded from system in an individual state as originally distributed. Control of source code repository is delegated to a small group of committers (http://wiki.dspace.org/index.php/DSpace_contributors). Only contributors can change the code and release new versions. The committers work with wider community of DSpace users. Anyone can submit big fixes, new features or feature requests. It can all be done through Source Forge administrative system (http://sourceforge.net/projects/dspace/). Support is provided on an informal basis via email in three different lists:
1. DSpace-Tech for Technical Support
2. DSpace General for general questions and announcements
3. DSpace-Dev for discussing development issues

COMMUNITY DEVELOPMENT MODEL OF DSPACE
It is an open source software with BSD licence. The community development model shown in Figure 1. DSpace Look has:
1. Source code control repository (SVN)
2. Committers
3. Community welcome to submit bug reports, patches, feature requests
4. E-mail lists for support

Use of DSpace
DSpace can be used to store any type of digital medium which includes:
1. Journal Papers
2. Data sets
3. Electronic Theses
4. Reports
5. Conference posters
6. Videos
7. Images

Benefits of Using DSpace
1. Research results out quick
2. Reaching worldwide audience through search engines such as google
3. Storing reusable teaching materials that can use with course management systems
4. Archiving and distributing material that currently put on personal website
5. Storing students projects/theses
6. Keeping track of own publications/bibliography
7. Having persistent network identifier for work that never changes or break
8. No more page charges for images. Images can be persistent identifiers in published articles
REASONS FOR DESIGNING DSPACE

1. Designed for specific in-house scholars: These institutions established open access institutional repositories that disseminate research outputs. These are self-archived. The administrator of the repositories collects the research documents from different sources and submits the documents to institutional repository on behalf of persons concerned.

2. Designed for specific subjects: Some institutional repositories are designed to store and provide access to specific subject collections of documents. It is not affiliated to any specific institutes providing institutional repository facility but cannot research articles to respective subject field of interest. E.g. LDL (Librarian’s Digital Library) developed by DRTC, Bangalore which is a subject-specific repository for library and information professionals. Openmed@NTC, maintained by NIC, New Delhi provides access to bio-medical literature.

3. Designed for specific documents: Institutional repository is designed for storing and providing access to the documents for specific type of collections, e.g., Vidyanidhi of University of Mysore which provides storage and access to theses and dissertations.

4. Digital Repository Software: IBM introduced Digital Library Software in 1991 for managing collection of digital files. The technology helps in storage, maintenance, retrieval and display digital content. DSpace with website http://www.dspace.org was developed jointly by MIT library and HP. It captures, stores, indexes, preserves and redistributes for research material formats of the organizations. It supports institutional repositories and electronic records management also. It can be used worldwide for the digital archiving needs.

The government of India is very keen towards creating institutional repositories as a national movement under research and development activities. It doesn’t require huge amount of finance but only require co-operation and technical knowledge sharing among professionals. UGC, CSIR, ICMR, AICTE supports financially for designing institutional repositories. DSpace as a repository software provides information openly available and easy to manage.

DSPACE TECHNOLOGY

The operating system for DSpace is Unix or Linux. It is written in JAVA. It has postgresQL relational database. It provides complete Web User Interface but Java API’s available. DSpace search uses an index of the content of submitted metadata for both local and federated catalogs. DSpace configuration allows full text searching of certain document types including PDF if text-based or OCR-ed, MS Word, plain text and HTML indexing for full text searching is not instantaneous. The index is prepared periodically on daily basis as per configuration and configured to transform and extract text from certain other document types such as power point. The images types like JPG can be transformed and extracted to thumbnail images for efficient searching. The compressed files, odt, etc. are limited in searching corresponding metadata. From user’s perspective, DSpace search is effective where quality metadata has been provided for a document and for text-based documents once the full-text index is updated. DSpace is open source software that can be downloaded, use and modify for free. The software is shared under BSD (Berkeley Software Distribution) licence [1–3].

India started institutional repository movement in 2004 and growing fast. UGC 2005 Act supported and recommended to set up institutional and educational repositories all over India. National Knowledge Commission in 2007 strongly open access to public funded research literature and initiative for building nationwide Institutional Repository. DELNET, INFLIBNET are also supporting in modernization of library and setting up the institutional repositories. The institutional repositories owned by R&D Institutes/Academic institutes in India based on secondary sources, metadata harvesting services, directions, etc. for study. The data related to R & D institutional repositories have been collected from respective institutions co-
websites and other secondary sources. The Case study methods are followed to develop the institutional repositories. The data’s are analyzed based on certain parameters such as number of documents, software used, growth of collection, etc. [4–6].

In India, DSpace registered service providers made investment in DSpace technology with a commitment to work co-operatively with DuraSpace organization to best serve community of users. The repository needs server, modern automated scanner, high pixel camera, stable network connection and necessary open source software as per requirement. Table-1- Registered Service Providers of DSpace in India shows the different DSpace service providers of India.

![DSpace Diagram](graphic_url)

**Fig. 1: DSpace Look.**

<table>
<thead>
<tr>
<th>S.No.</th>
<th>Name of DSpace Service Provider</th>
<th>Services Provided</th>
<th>Language</th>
<th>Contact</th>
<th>E-mail id</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>D Technologies, New Delhi</td>
<td>Add-on development, Analysis, Consultancy, Customization, Design, Development, Hosting, Installation/configuration, Integration, Training</td>
<td>English, Hindi</td>
<td>+91-9717396363</td>
<td><a href="mailto:sales@d2t.co.in">sales@d2t.co.in</a></td>
</tr>
<tr>
<td>2.</td>
<td>Informatics, Bangalore</td>
<td>Consultancy, Development, Installation/configuration, Integration, Migration, Training</td>
<td>English, Hindi, Kannada, Tamil, Telugu, Bangla, Marathi, Malayalam</td>
<td>+91-80-40387777</td>
<td><a href="mailto:info@informaticsglobal.com">info@informaticsglobal.com</a></td>
</tr>
<tr>
<td>3.</td>
<td>Vir Softech Pvt. Ltd., Noida, Uttar Pradesh</td>
<td>Add-on development, Analysis, Customization, Data conversion, Development, Installation/configuration, Integration, Migration, Training</td>
<td>English, Hindi, Bengali</td>
<td>+91-9871258989</td>
<td><a href="mailto:deepak@virsoftech.com">deepak@virsoftech.com</a></td>
</tr>
</tbody>
</table>
OBJECTIVES FOR DEVELOPING INSTITUTIONAL REPOSITORY

1. To create global visibility for an institution's scholarship
2. To collect content in a single location
3. To provide open access to institutional research output by self-archiving
4. To store and preserve other institutional digital assets including unpublished or otherwise easily lost like theses or technical reports
5. Providing new opportunities for archiving and preservation of valuable digital works
6. Reduce duplication of records and inconsistencies in multiple instances of the same works
7. Offer greater flexibility and integration than other mechanisms of disseminating scholarly works with more coherent security and preservation of digital materials
8. Assists research collaboration through facilitating free exchange of scholarly information.
9. Increase visibility, reputation and prestige of institution

The different repositories using DSpace in India mentioned in Table 2.

There are total 54 repositories in different states of India mentioned in Table 3 who are using DSpace in India.

Table 2: Different Repositories using DSpace in India.

<table>
<thead>
<tr>
<th>S.No.</th>
<th>Name of Institution</th>
<th>DSpace Link</th>
<th>State</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>IGNOU</td>
<td><a href="http://www.egyankosh.ac.in/">http://www.egyankosh.ac.in/</a></td>
<td>Delhi</td>
</tr>
<tr>
<td>3</td>
<td>Indian Institute of Technology, Delhi</td>
<td>eprint.iitd.ac.in/</td>
<td>Delhi</td>
</tr>
<tr>
<td>4</td>
<td>Indraprastha Institute of Information Technology</td>
<td><a href="http://repository.iitd.edu.in/jspui/">http://repository.iitd.edu.in/jspui/</a></td>
<td>Delhi</td>
</tr>
<tr>
<td>5</td>
<td>NICSAIR</td>
<td><a href="http://nisl.niscair.res.in/handle/123456789939">http://nisl.niscair.res.in/handle/123456789939</a></td>
<td>Delhi</td>
</tr>
<tr>
<td>7</td>
<td>Parliament of India, Official debates of Rajyasahha</td>
<td><a href="http://rsdebate.nic.in/">http://rsdebate.nic.in/</a></td>
<td>Delhi</td>
</tr>
<tr>
<td>8</td>
<td>World Health Organization Southeast Asian Region Digital Repository</td>
<td><a href="http://repository.searo.who.int/">http://repository.searo.who.int/</a></td>
<td>Delhi</td>
</tr>
<tr>
<td>9</td>
<td>National Centre for Antarctic and Ocean Research</td>
<td><a href="http://dspace.ncaor.org:8080/dspace/index.jsp">http://dspace.ncaor.org:8080/dspace/index.jsp</a></td>
<td>Goa</td>
</tr>
<tr>
<td>10</td>
<td>National Institute of Oceanography</td>
<td><a href="http://dhrs.nio.org/drs/index.jsp">http://dhrs.nio.org/drs/index.jsp</a></td>
<td>Goa</td>
</tr>
<tr>
<td>11</td>
<td>INFLIBNET—Institutional Repository</td>
<td><a href="http://ir.inflibnet.ac.in/">http://ir.inflibnet.ac.in/</a></td>
<td>Gujarat</td>
</tr>
<tr>
<td>12</td>
<td>INFLIBNET—Institutional Shodhanga Indian ETD</td>
<td><a href="http://shodhganga.inflibnet.ac.in/">http://shodhganga.inflibnet.ac.in/</a></td>
<td>Gujarat</td>
</tr>
<tr>
<td>13</td>
<td>M.S. University of Baroda</td>
<td><a href="http://14.139.121.106:8080/jspui/">http://14.139.121.106:8080/jspui/</a></td>
<td>Gujarat</td>
</tr>
<tr>
<td>14</td>
<td>Pandit Deendayal University, Gandhinagar</td>
<td><a href="http://203.77.192.116:8080/xmlui">http://203.77.192.116:8080/xmlui</a></td>
<td>Gujarat</td>
</tr>
<tr>
<td>15</td>
<td>University of Kashmir</td>
<td><a href="http://dspaces.uok.edu.in/jspui/#.Uf4RptLTWng">http://dspaces.uok.edu.in/jspui/#.Uf4RptLTWng</a></td>
<td>Jammu &amp; Kashmir</td>
</tr>
<tr>
<td>16</td>
<td>Indian Institute of Astrophysics, Bangalore</td>
<td><a href="http://prints.iiap.res.in/">http://prints.iiap.res.in/</a></td>
<td>Karnataka</td>
</tr>
<tr>
<td>17</td>
<td>Indian Institute of Horticulture Research</td>
<td><a href="http://www.erepo.iihr.ernet.in/">http://www.erepo.iihr.ernet.in/</a></td>
<td>Karnataka</td>
</tr>
<tr>
<td>18</td>
<td>Indian Institute of Science, Bangalore</td>
<td><a href="http://etd.nesi.iisc.ernet.in/">http://etd.nesi.iisc.ernet.in/</a></td>
<td>Karnataka</td>
</tr>
<tr>
<td>19</td>
<td>Indian Statistical Institute, Bangalore</td>
<td><a href="http://drtc.isibang.ac.in/">http://drtc.isibang.ac.in/</a></td>
<td>Karnataka</td>
</tr>
<tr>
<td>20</td>
<td>Institute for Social and Economic Change</td>
<td><a href="http://203.200.22.249:8080/jspui/">http://203.200.22.249:8080/jspui/</a></td>
<td>Karnataka</td>
</tr>
<tr>
<td>21</td>
<td>KLE University</td>
<td><a href="http://182.48.228.18:8080/jspui/">http://182.48.228.18:8080/jspui/</a></td>
<td>Karnataka</td>
</tr>
<tr>
<td>22</td>
<td>Raman Research Institute, Bangalore</td>
<td><a href="http://dspace.rrri.res.in/">http://dspace.rrri.res.in/</a></td>
<td>Karnataka</td>
</tr>
<tr>
<td>23</td>
<td>T. John Institute of Technology</td>
<td><a href="http://122.181.173.88/jspui/">http://122.181.173.88/jspui/</a></td>
<td>Karnataka</td>
</tr>
<tr>
<td>24</td>
<td>University of Mysore, Vidyanidhi Project</td>
<td><a href="http://dspace.vidyaniidhi.org.in:8080/dspace/">http://dspace.vidyaniidhi.org.in:8080/dspace/</a></td>
<td>Karnataka</td>
</tr>
<tr>
<td>25</td>
<td>Vivekanand Institute of Technology—Dyuthi</td>
<td><a href="http://dyuthi.cusat.ac.in/xmlui/">http://dyuthi.cusat.ac.in/xmlui/</a></td>
<td>Kerala</td>
</tr>
<tr>
<td>26</td>
<td>Cochin University of Science and Technology—Dyuthi</td>
<td><a href="http://dyuthi.cusat.ac.in/xmlui/">http://dyuthi.cusat.ac.in/xmlui/</a></td>
<td>Kerala</td>
</tr>
<tr>
<td>27</td>
<td>Cochin University of Science and Technology—Digital Library</td>
<td><a href="http://dspace.cusat.ac.in/jspui/">http://dspace.cusat.ac.in/jspui/</a></td>
<td>Kerala</td>
</tr>
</tbody>
</table>
Table 3: Number of DSpace Repositories in States of India.

<table>
<thead>
<tr>
<th>S.No.</th>
<th>State of India using DSpace Repository</th>
<th>Number of Repositories using DSpace in Indian States</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Karnataka</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>Delhi</td>
<td>08</td>
</tr>
<tr>
<td>3</td>
<td>Kerala</td>
<td>07</td>
</tr>
<tr>
<td>4</td>
<td>Maharashtra</td>
<td>07</td>
</tr>
<tr>
<td>5</td>
<td>West Bengal</td>
<td>05</td>
</tr>
<tr>
<td>6</td>
<td>Gujarat</td>
<td>04</td>
</tr>
<tr>
<td>7</td>
<td>Uttaranchal</td>
<td>02</td>
</tr>
<tr>
<td>8</td>
<td>Goa</td>
<td>02</td>
</tr>
<tr>
<td>9</td>
<td>Orissa</td>
<td>02</td>
</tr>
<tr>
<td>10</td>
<td>Tamil Nadu</td>
<td>02</td>
</tr>
<tr>
<td>11</td>
<td>Meghalaya</td>
<td>01</td>
</tr>
<tr>
<td>12</td>
<td>Pondicherry</td>
<td>01</td>
</tr>
<tr>
<td>13</td>
<td>Punjab</td>
<td>01</td>
</tr>
<tr>
<td>14</td>
<td>Jammu &amp; Kashmir</td>
<td>01</td>
</tr>
<tr>
<td>15</td>
<td>Uttar Pradesh</td>
<td>01</td>
</tr>
<tr>
<td></td>
<td>Total Repositories using DSpace in Indian States</td>
<td>54</td>
</tr>
</tbody>
</table>
Fig. 2: States of India having different number of DSpace Repositories.

The number of digital repositories in states of India having DSpace digital repositories can also be seen in different zones of pie-charts of Figure 2.

**BENEFITS OF DEVELOPING INSTITUTIONAL REPOSITORY**

The Institutional Repositories develop with a framework to provide better services to the research and development of any organization and archive the works in digital forms [7–10]. The benefits of creating institutional repositories are:

1. Digital Preservation of documents
2. Centralization and storage of all types of institutional output including unpublished literature
3. Support for learning and teaching
4. Standardization of institutional records
5. Ability to keep track and analyze research performance
6. Breaking down of publisher costs and permission barrier
7. Alleviation of requirement to trust publishers to maintain information in long term without commercial benefit of authors
8. Promotion of philosophy of wider communications
9. Online access for institutional research publications
10. Opportunities for new forms of scholarly communications.

**STEPS FOR BUILDING INSTITUTIONAL REPOSITORY**

1. Go to local DSpace homepage and click on My DSpace icon. The log-in screen appears.
2. Log-in-name created a dedicated e-mail account for uploading publications.
3. After log-in as an administrator on the top of the right there is an option to Create Top-Level Community by using feature community/ collection listing page. This option helps to create new top level community and take one to the Edit Community page. Enter name of top level community and then click on Create.
4. Creating a sub-community after creation of top level community go to community home page of the community that is to be parent of new sub-community. Click on Create sub-community in the Admin Tool Box at top right corner of page. Insert sub-community name then click on create.
5. After clicking sub-community, create collection under sub community, various main headings whichever relevant.
6. Describe collection checkboxes. Choose the relevant options then click on next. Write collection name, book chapters, etc. After entering the collection name then click on update.

7. Start uploading documents to DSpace collection. On clicking the name of collection it will take to uploading a document into DSpace collection. The collection in journal articles is being created.

8. Click on submit to collection. Once logged in DSpace for submitting the documents, the submission steps will cover seven steps to upload documents into DSpace which are:
   a. Describe
   b. Describe
   c. Describe
   d. Upload
   e. Verify
   f. License
   g. Complete

CONCLUSION
The institutional repository is a very powerful tool that can serve and support institutions of higher education, and scholars. If institutional repository is properly developed and maintained, it can achieve number of goals, and needs of higher education system. The joint efforts from the faculty and the library working in partnership with a full understanding can develop a beneficial repository. DSpace is the best advance software for academic, non-profit, and commercial organizations to build up open digital repositories. It is free and easy to install and can be completely customizable according to the needs of any organization/institutions. It preserves and enables easy and open access to all types of digital content including text, images, mpegs and data sets, etc. The developers of DSpace continuously working on expanding and improving the DSpace software in advance forms.

REFERENCES
1. Duraspase, Discovery Configuration. Available from Wikipedia.
8. Sharma PL. Institutional Repositories: An essential tool for information and knowledge sharing. drtc.isibang.ac.in, 9 july,2018

Cite this Article