DSpace

"DSpace captures your data in any format – in text, video, audio, and data. It distributes it over the web. It indexes your work, so users can search and retrieve your items. It preserves your digital work over the long term. DSpace provides a way to manage your research materials and publications in a professionally maintained repository to give them greater visibility and accessibility over time."

www.dspace.org

DSpace 6.3 - *The current stable release*

This **stable, bug fix** version was released in June 2018. This is the **recommended** version to use.

DSpace is the software of choice for academic, non-profit, and commercial organizations building open digital repositories. It is free and easy to install "out of the box" and completely customizable to fit the needs of any organization.

DSpace preserves and enables easy and open access to all types of digital content including text, images, moving images, mpegs and data sets. And with an ever-growing community of developers, committed to continuously expanding and improving the software, each DSpace installation benefits from the next.

The DSpace project was initiated in July 2000 The Massachusetts Institute of Technology (MIT) Libraries and Hewlett- Packard (HP) Labs have worked together on the development of an open source system called Dspace that functions as a repository for the digital research and educational material produced by members of a research university or institution. Dspace is a ground breaking digital repository system that captures, stores, indexes, preserves and distributes digital research materials.

VISION AND MISSION - The DSpace Project will produce the world's choice for repository software providing the means for making information openly available and easy to manage.

The DSpace Foundation was formed in 2007 as a non-profit organisation to provide support to the growing community of institutions that use DSpace. The foundation's mission is to lead the collaborative development of open source software to enable permanent access to digital works.

MEMBERS - DSpace Members are leaders from university, research, library organizations, and others, who have made a financial commitment to our open source projects. As such they have become a partner in governance, ensuring that DuraSpace projects continue to serve the global communities that depend on them into the future.

KEY FEATURES –

Application Architecture: DSpace is a full stack web application, consisting of a database, storage manager and front end web interface. The architecture includes a specific data model with configurable metadata schemas, workflows and browse/search functionality.

Modern, RESTful Web UI: DSpace 7.0 will feature a completely rewritten web user interface based on the Angular 2 javascript platform.

Built-in workflows: Originally designed for libraries, the embedded DSpace data model and approval workflows are familiar to librarians and archivists.

Built-in search engine: DSpace comes packaged with Apache Solr, an open source enterprise search platform that enables filtered (faceted) searching and browsing of all objects. The full text of common file formats is searchable, along with all metadata fields. Browse by interfaces are also configurable.

Unlimited File types: DSpace can store any type of file. In addition, it auto-recognizes files of most common formats (e.g., DOC, PDF, XLS, PPT, JPEG, MPEG, TIFF).

Metadata: By default, DSpace uses a Qualified Dublin Core (QDC) based metadata schema. Institutions can extend that base schema or add custom QDC-like schemas. DSpace can import or export metadata from other major metadata schemas such as MARC or MODS.

Tools/plugins: DSpace comes with a suite of tools (batch ingest, batch export, batch metadata editing, etc.) and plugins for translating content into DSpace objects. Additionally, commercial plugins are available through service providers.

Security: DSpace provides its own built-in authentication / authorization system, but can also integrate with existing authentication systems such as LDAP or Shibboleth.

Permissions: DSpace allows you to control read/write permissions site-wide, per community, per collection, per item and per file. You may also delegate administrative permissions per community or per collection.

Disaster Recovery: DSpace allows you to export all of your system content as AIP (Archival Information Packages) backup files. These AIPs can be used to restore your entire site, or restore individual communities, collections or items.

OAI-PMH / SWORD (v1 and v2) / OpenAIRE / Driver: DSpace complies with standard protocols and best practices for access, ingest, and export.

REST: DSpace provides RESTful APIs in accordance with modern web standards.

Configurable Database: Organizations can choose either PostgreSQL or Oracle for the database in which DSpace manages items and metadata.

Configurable File Storage: Files in DSpace can be stored either using a local filesystem (default) or a cloud-based solution, such as Amazon S3.

Data Integrity: On upload, DSpace calculates and stores a checksum for each file. Optionally, you may ask DSpace to verify those checksums to validate file integrity.

Languages: DSpace is available in over 20 languages.

DSPACE SOFTWARE WILL:

1. DSpace provides tools for management of digital assets, and is commonly used for building institutional repositories.
2. DSpace helps to create, index and retrieve various types of digital contents which include research articles, grey literature, theses, cultural materials, 3D digital scans of objects, photographs, films, audio/videos, scientific datasets, institutional records, educational materials and other forms of content.
3. The collection in DSpace is organised into communities, collections and items. The communities in DSpace include a high-level organizational structure whose only purpose is to divide collections into related groups. Each community contains one or more collections, which are containers for related items. An item is a deposited object of any type: a published article, an image, audio, or video file, notes, a presentation, etc.
4. DSpace is specially designed for digital preservation support for all the documents that are added into the repository in a simple fashion.
5. Focus on the Institutional Repository use case.
6. Be lean, agile, and flexible.
7. Be easy and simple to install and operate.
8. Include a core set of functionality that can be extended to or integrated with complementary services and tools in the larger scholarly ecosystem.

ARCHITECTURAL OVERVIEW - DSpace is a set of cooperating Java web applications and utility programs that maintain an asset store and an associated metadata store. The web applications provide interfaces for administration, deposit, ingest, search, and access. The asset store is maintained on a file system or similar storage system. The metadata, including access and configuration information, is stored in a relational database.

REQUIREMENTS FOR DSpace - DSpace is written in Java, it will therefore run on any Operating System (Linux, Windows, Mac OSX). DSpace is

built on top of free, open-source tools, such as the Apache Web server, the Tomcat Servlet engine, and the PostgreSQL relational database system. For user convenience, DSpace package the necessary JDBC and other drivers and libraries together with DSpace. This set of tools should run on any UNIX-type OS, such as Linux, HP/UX, or Solaris, and you can substitute other libraries if you need to run on another platform. The system runs on anything from a laptop to a $500K server.

MINIMAL DSPACE PRODUCTION SYSTEM REQUIREMENTS

- 2-3 GB of Random Access Memory (RAM)
- 1GB for Tomcat (e.g. "TOMCAT_OPTS=-server -Xms1024M -Xmx1024M -XX:MaxPermSize=128M -Dfile.encoding=UTF-8")
- 1GB for Database (PostgreSQL or Oracle).
- Keep in mind your Operating System also needs some memory to function. So, while DSpace may only need ~2GB of memory, you should ensure the computer itself has at least 3-4GB of RAM available overall.
- 20 GB of Storage (or roughly enough storage for all the files you wish to store in DSpace)

Currently DSpace supports only the Dublin Core metadata element set with a few qualifications conforming to the library application profile.

DSpace supports the Open Archives Initiative's Protocol for Metadata Harvesting (OAI-PMH) v2.0 as a data provider.

DATA MODELS - There are five main entities which are the hub of information structure and aggregation in DSpace

• **Communities**: an administrative/logic groping of one or more collections (and sub-communities); eg. faculties and departements in a single university, centers for geographically distributed organizations; projects/research area, etc...

• **Collections**: a grouping of items which are analogous for tipology (metadata) and workflow; currently collections are the fulcrum of archive customizations

• **Item:** a box which contains both a document metadata and one or more bitstream bundles

• **Bundle:** a groping of bistreams used to separate the original documents, those obtained from automatic process, (such as full-text extraction), archival and Creative Commons licence

• **Bitstream** (= digital content, usually a fulltext)

ADD-ON MODULES – Specific commercial add-on modules greatly enhance the functionality of the DSpace platform.

- Metadata Quality
- Listings & Reports
- Content & Usage Analysis
- Document Streaming
- Image Zoom
- Audiovisual Streaming

NOTABLE DSpace REPOSITERIES –

- The World Bank - Open Knowledge Repository
- Apollo - University of Cambridge Repository
- Digital Access to Scholarship at Harvard
- DSpace@MIT
- Spiral - Imperial College London Repository
- WHO Institutional Repository for Information Sharing
- Think Asia
- Shodhganga: a reservoir of Indian Theses

A full list of institutional repositories using DSpace software as well as others is available via the Registry of Open Access Repositories (ROAR)

SERVICE PROVIDERS - DSpace has a global, active network of registered service providers who provided commercial support, hosting, training or site customization. https://duraspace.org/dspace/resources/service-providers/

RAKHI SINGH