

DSpace

<https://duraspace.org/dspace/>

*Reviewed by Diana Dulek, Metadata Specialist at the University of Houston Libraries,
MLS student at Texas Woman's University*

DSpace is available as free, open-source software which describes itself as “the software of choice for academic, non-profit, and commercial organizations building open digital repositories.”¹

DSpace began as a collaboration between Hewlett-Packard Company and the Massachusetts Institute of Technology (MIT) Libraries in 2000. The duo then “released the system worldwide on November 4, 2002 . . . one month after its introduction as a new service of the MIT Libraries.”² Seventeen years later, an impressive community supports DSpace, with more than [1,000 instances](#) of DSpace running worldwide.

DSpace can function as a digital repository for special collections. DSpace lets users upload items individually, or through a batch import via CSV. Archivists can extend the required Dublin Core schema or add new schemas through the metadata registry (Figure 1). Once published, users can browse items through facets such as community, collection, author, date, title, and subject. Administrators can edit the metadata records of items individually, or again, through a batch edit using a CSV. DSpace lacks a true image viewer, instead providing thumbnails and a download option for each item.

¹ "About DSpace," DSpace, accessed August 14, 2019, <https://duraspace.org/dspace/about/>.

² MacKenzie Smith, Mary Barton, Mick Bass, Margret Branschofsky, Greg McClellan, Dave Stuve, Robert Tansley, and Julie Harford Walker, "An Open Source Dynamic Digital Repository," *D-Lib Magazine* 9, no. 1 (2003): 1082–9873, <http://doi.org/10.1045/january2003-smith>.

Add a new schema

Namespace: *

Namespace should be an established URI location for the new schema.

Name: *

Shorthand notation for the schema. This will be used to prefix a field's name (e.g. dc.element.qualifier). The name must be less than 32 characters and cannot include spaces, periods or underscores.

Add new schema

Figure 1: Adding a new metadata schema in DSpace

As a metadata specialist at the University of Houston (UH) Libraries, I have used DSpace for the past three years primarily to clean up metadata for our electronic theses and dissertations (ETDs). We have also launched a service to ingest faculty works into the institutional repository (IR). However an institution chooses to utilize DSpace, the repository can be divided into communities that subdivide into subcommunities and collections.

I always enjoy when we get in a new batch of ETDs to clean up because it means I get to work with DSpace. It means a respite from (until recently, as we are migrating to a Hyrax-based system) the murky waters of CONTENTdm. As I recently told our new student workers, DSpace is excellent because it is very forgiving. For example, if a user enters the incorrect date on five-hundred ETDs, the changes can be reviewed before they are approved (Figure 2). However, if a user approved the changes and does not notice the mistake until a week later, one can import the correct metadata to overwrite the mistake. This simple process shows that DSpace understands there are humans behind the data in its repositories.

Notice
Upload successful

IMPORT METADATA

Pending changes are listed below for review

Changes pending for item c92ad09d-8c22-4ad9-b4f9-99bf9605fed4 (10657/3948)	
Add: (dc.type.dcmi)	text
Remove: (dc.type.dcmi)	Text

Apply changes

Return

Figure 2: Applying changes to the metadata

As with any healthy long-term relationship, there are flaws. DSpace, as an institutional repository, seems like it should be a robust platform for student, staff, and faculty works of all kinds. However, as we recently learned at the University of Houston, MOV files do not produce thumbnails in DSpace 6, nor do they play in-browser. The file shows up, and users are expected to download it. Unless the file has substantial metadata, the user might not know what they are about to download.

DSpace 6 lacks reliable usage metrics (Figure 3). Without a way to track traffic to the university's outputs, UH has been using Google Analytics as a workaround.

General Overview

Items Archived
Bitstream Views
Item Views
Collection Views
Community Views
User Logins
Searches Performed
Licence Rejections
OAI Requests

Archive Information

Content Type	Number of items
All Items	
Research Paper	
Thesis or Dissertation	
Preprint	

Items Viewed

(more than 20 times)

Item/Handle	Number of views

Figure 3: Statistics offered by DSpace (numbers edited out)

DSpace 6 has been the most significant upgrade in the past three years, mainly because of the way the record IDs have changed from being sequential numbers to a long string of hyphenated letters and numbers that do not seem to be in any human-readable sequence. This change made it harder to sort exported metadata and find records using the ID field. I have not been able to find clear documentation on this change. The majority of DSpace's documentation is for users working through the command line, instead of a hosted instance like we use at the University of Houston Libraries.³

Additionally, the DSpace 6 upgrade resulted in several unexpected errors and strange happenings. One day after uploading a batch of ETDs, I noticed that a field was missing in the metadata records on the IR, although the CSV I imported into DSpace included the field. The next day, the missing metadata appeared in the IR. When we come across bugs like this, we put in a ticket with the Texas Digital Library (TDL) Help Desk because our instance of DSpace at UH is hosted through TDL. The developers at TDL are very responsive, and typically send a response email within the day. DSpace also uses JIRA where users can report issues and check to see if their issues are known bugs.⁴

For institutions looking at free, open-source software for their special collections, DSpace offers archivists a large community with responsive tech support, user groups, and continual software updates. However, lacking an image viewer, DSpace may lose out to other freely available software with similar capabilities, such as Omeka. DSpace 7 is tentatively slated to come out in the fourth quarter of 2019, with a preview release currently available for download and testing.⁵

³ "DSpace Documentation," DuraSpace, accessed August 14, 2019, <https://wiki.duraspace.org/display/DSDOC6x>.

⁴ "DSpace JIRA," DuraSpace, accessed August 14, 2019, <https://jira.duraspace.org/browse/DS-4331>.

⁵ "DSpace 7," DuraSpace, accessed August 14, 2019, <https://duraspace.org/dspace/dspace-7/>.