TechnicalFAQ

This page addresses the most commonly-asked technical questions about DSpace. See also the User FAQ page for answers to general questions about DSpace.

Table of Contents

1. What is the most current release of the DSpace software?
2. Where can I download the DSpace open source software?
3. I have technical questions about DSpace – how can I find solutions?
4. Is there install doc for DSpace?
5. How do I report bugs?
6. How do I request new features?
7. Internal System Error: I keep receiving the message “The system has experienced an internal error”. How do I diagnose the problem?
   7.1 To view the DSpace log live
8. What's the best strategy for backing up/restoring DSpace?
9. How do I run DSpace on port 80/443?
10. Do I need HTTPS (SSL)?
11. How do I run DSpace on Windows?
12. How do you append data from a test server to a production environment?
13. Are there any tips on running DSpace on RedHat/Fedora Core Linux?
14. filter-media Hangs While Creating Search Index
15. Setting logging level up to DEBUG
16. How do I add/index/display another Dublin Core field?
17. How do I add a field to the submission UI?
18. How do I add a field to the Lucene search index?
19. How do I add a field to the simple item display page?
20. I've made a lot of modifications to DSpace, and I want to bring my code up to date with a later version of DSpace. What's the easiest way to do that?
21. When trying to add content to DSpace, all of the fields filled out appear null.
22. How do I run DSpace with Apache and Tomcat using the mod_jk connector?
23. How do I run multiple instances of DSpace in one Tomcat Server?
24. How do you localise the Creative Commons license?
25. DSPACE with ISO-8859-1 and others encodings
26. Clearing Cocoon (XMLUI) cache
27. I changed access restrictions on a collection, but restrictions didn't change on its items
28. How can I change what is displayed in XMLUI for a particular collection?
29. I'm getting 'relation "bi_5_dis" does not exist' (or any other number)
30. Duplicate items in Recent Submissions in XMLUI Discovery
31. How do I update Messages.properties without rebuilding DSpace?
32. I'm getting "SolrException: Bad Request" followed by a long query or a "tooManyClauses" Exception
33. Errors connecting to Solr
34. Error starting Tomcat: SEVERE: Error listenerStart
35. Metadata values in CSV export seem to have duplicate columns

What is the most current release of the DSpace software?

The most recent stable version of DSpace, recommended for production use, is available from GitHub.

To see what's in store for the future, see the DSpace Releases planning page and the DSpace RoadMap.

Where can I download the DSpace open source software?

The most recent stable version of DSpace, recommended for production use, is freely available from GitHub.

If you are a developer, DSpace is also available to be checked out using Git/GitHub: https://github.com/DSpace/DSpace See also Development with Git.

I have technical questions about DSpace – how can I find solutions?

You’re in the right place. First, you’ll want to search this FAQ to see if your question has been asked/answered already. Next, try searching the DSpace System Documentation and the DSpace-Tech mailing list archives.

Still need help?

Post your questions to the DSpace-tech mailing list, where members of the DSpace community will offer their assistance. And once you find the answer, if you think others will be asking the same question post them here in the Technical FAQ page.
Is there install doc for DSpace?

Yes, the DSpace System Documentation includes a section on Installation. Read through this FAQ and also see Installation Guides.

If you run into any problems or have further questions, send an email to the DSpace-Tech mailing list for help.

How do I report bugs?

If you think you’ve found a bug, we recommend taking the following steps.

1. First, attempt to determine if it really is a bug by asking on the DSpace-Tech mailing list to see if others have encountered this same problem.
2. Once you’ve verified it with someone else, please submit a bug report to DSpace Issue Tracker (JIRA). Try to be as descriptive as you can, and detail the steps that someone else would need to take to encounter this same problem.
3. One of the developers will review this bug and get back to you. Don’t be surprised if he/she asks some followup questions. The developer may need to figure out exactly how you encountered the bug so that he/she can figure out the best way to fix it.

Thanks for being willing to help locate and report bugs! An active community helps us make more stable, bug-free software.

How do I request new features?

Before putting in a new feature request, you should make sure that the feature would be useful to the broader DSpace community and isn’t institutional specific. It’s also worth mentioning that since the DSpace developers are all volunteers, you may want to work to locate an interested developer who can help you scope out and begin to develop this new feature.

If you’ve come up with a new feature idea, we recommend taking the following steps.

1. First, attempt to determine if this feature will be useful to everyone using DSpace. The best way to do this is to describe this feature in an email to the DSpace-Tech mailing list in order to begin the discussion.
2. Once you’ve determined your feature would be useful to all, please submit a feature request to DSpace Issue Tracker (JIRA). Try to be as descriptive as possible. Describe the feature in as much detail as you can. Make sure to also describe why you feel this feature would be useful (i.e., what is the use case?)
3. Once your feature request is made, the DSpace developers will begin to review it and determine if a volunteer would step up to help developing it. Please be patient. Sometimes a volunteer can be found quickly. Other times it can take a while. If you can help us locate a volunteer developer, it is even more likely that the feature can be developed relatively quickly. If you don’t hear anything for a while, follow-up with the DSpace developers by posting a comment to your initial feature request in Jira.

Also see How to Contribute to DSpace

Internal System Error: I keep receiving the message "The system has experienced an internal error". How do I diagnose the problem?

See the Troubleshoot an error page for hints on which log files to check for errors, and the recommended way of sending these to the DSpace-Tech Mailing List

To view the DSpace log live

```bash
## In Linux, run the following command then reproduce the internal error state
## use the pathname prefix appropriate to your DSpace installation
tail -F /var/logs/dspace.log
```

2009-05-18 12:52:45,935 WARN org.dspace.app.webui.servlet.DSpaceServlet @ name\n@place:session_id=F32DD4066DD475B8841A2A2204961B9A:ip_addr=1.2\n.3.4:database_error;org.postgresql.util.PSQLException: ERROR: duplicate key \nviolates unique constraint "metadatafieldregistry_pkey"
org.postgresql.util.PSQLException: ERROR: duplicate key violates unique constra\nt "metadatafieldregistry_pkey"
at org.postgresql.core.v3.QueryExecutorImpl.receiveErrorResponse(QueryE\nxecutorImpl.java:1525)

What's the best strategy for backing up/restoring DSpace?

From the Documentation, see Storage Layer (which describes the basics for backing up your database and file storage). Also see AIP Backup and Restore, which describes how to backup your DSpace to a set of Archival Information Packages (AIPs).
Finally, there is a somewhat outdated summary at BackupRestore.

How do I run DSpace on port 80/443?

The official Installation Documentation provides a section on "DSpace over HTTPS" which describes how to configure Apache HTTPD (responding on ports 80 and 443) to redirect to a DSpace running on Apache Tomcat. There are many other guides out on the web for redirecting requests from Apache HTTPD (running on ports 80/443) to Apache Tomcat (usually this redirect is configured via either mod_proxy_ajp or mod_jk). Those guides should also work for DSpace, as DSpace requires no special configurations for Apache HTTPD + Apache Tomcat.

Also see Running DSpace on Standard Ports for further examples for how this could be achieved.

Do I need HTTPS (SSL)?

The official Installation Documentation provides a section on "DSpace over HTTPS" which describes enable HTTPS either in Tomcat directly, or in Apache HTTPD (and redirecting requests to Tomcat).

Also see SecuringDspace, ServletSecurity for more details.

How do I run DSpace on Windows?

See Installing DSpace on Windows

How do you append data from a test server to a production environment?

Many people have two installs of DSpace: a test server and a production server. The problem is how to take data from the test server and integrate (move, append, migrate, etc.) it into the production server. The most common example is when you are ready to move a new collection from test into production. Note that this is a somewhat different problem than the need to move all data from a test server and completely overwrite any data on the production all system, which may happen when first moving out of test and into production. RichardJones covers that scenario more completely here, along with a number of other DSpace-related topics.

The main problem in accomplishing this task easily is that some of the metadata will be duplicated if the export from your test server is imported directly and without modification into the production server. In particular, the following metadata elements will be affected:

- identifier.url
- description.provenance
- date.accessioned
- date.available
- date.issued
- format.extent
- format.mimetype
- All of these values should be deleted to avoid duplication, except for date.issued if the items have been published or publicly distributed before. However, Cody Green has written a bash shell script that will traverse a DSpace export (e.g., from your test server) and fix the metadata. Here is the link to that script: http://sunsite.utk.edu/diglib/dspace/dspace_migrate.sh (please note that the script will remove the date.issued field for all records).
- Also, the script is now included in the latest release.

Are there any tips on running DSpace on RedHat/Fedora Core Linux?

Many people have run into trouble getting DSpace running on Redhat/Fedora Core Linux, specifically, encountering CLASSPATH errors and relating to libgcj.

The easiest way to resolve this is to remove all the PM-based Java components, and install the original binary packages: Java SDK from Sun, and Ant and Tomcat from Apache.org.

To remove the PM Java stuff, do:

```
rpm -e libgcj
```

This will probably throw up a load of dependency errors. You'll need to remove the packages that depend on libgcj too.

Another useful list from RichardJones of packages he had to remove from Fedora Core 2:
Clive Gould has published a very useful blog explaining how to get DSpace working with a default install of CentOS 5.x, including Apache/Tomcat integration.

The Post Street Archives in Midland, Michigan with funding from the Herbert H. and Grace A. Dow Foundation worked with a team of students at Michigan Technological University to install DSpace for their digital archives. They produced http://www.poststreetarchives.org/DSpaceInstall.pdf this set of installation instructions for novice users on Fedora Core 4.

You also may wish to check Installation Guides for additional installation tips/suggestions relating to specific operating systems.

filter-media Hangs While Creating Search Index

If filter-media appears to hang after it creates the thumbnails and indexes text content, you may consider running filter-media as root. For some reason, certain users report that filter-media will hang after printing the following message:

Creating search index:

My personal experience shows that this happens when I run it as dspace, but not as root. Not sure yet why this happens, but this may be a quick fix until it is sorted out. Also, bear this potential "problem" in mind when setting up cron jobs for filter-media, index-all, and others.

Also note that running index-all as dspace does not hang, even though it does (apparently) the same thing.

Setting logging level up to DEBUG

If something doesn't seem to be working, you can often find out more by upping the logging level of DSpace. To do this:

1. Edit [dspace]/config/log4j.properties.
2. Change the level by changing this line:

   log4j.rootCategory=INFO, A1

   to this:

   log4j.rootCategory=DEBUG, A1

3. Restart Tomcat (or Resin, etc.).

Now you'll find there's a lot more information in [dspace]/log/dspace.log. You can switch on and off logging for particular DSpace classes. For example, if you just want to know about SQL queries going to the database, you could leave the main level above as INFO, and instead add the following line (then restart Tomcat):

   log4j.logger.org.dspace.storage.rdbms.DatabaseManager=DEBUG, A1

More information about DSpace log files is available in the DSpace System Documentation.

How do I add/index/display another Dublin Core field?

Adding a new Dublin Core field can be done from the Administrative User Interface of DSpace.

- Login as an Administrator, and look for the page called "Metadata Schema Registry".

After you've added a new field, you can add it to the Search (Lucene) Index or Browse Index by following the directions in the "Configuration and Customization" chapter of the DSpace System Documentation.

In that chapter you'll also find hints/tips on customizing your User Interface display to show this new field.

Additional help can be found in the on this wiki.
How do I add a field to the submission UI?
The DSpace System Documentation includes a chapter titled "Customizing and Configuring Submission User Interface" which details how to make this and many other Submission UI customizations.

How do I add a field to the Lucene search index?
The DSpace System Documentation includes a section on "Configuring Lucene Search Indexes" in the chapter on "Configuration and Customization".

How do I add a field to the simple item display page?
See the on this wiki for more help.

I've made a lot of modifications to DSpace, and I want to bring my code up to date with a later version of DSpace. What's the easiest way to do that?
We recommend using Git's merge features to help manage this. See Development with Git for some sample workflows.

When trying to add content to DSpace, all of the fields filled out appear null.
After setting up the DSpace database and properly setting up the application, most functionality appears normal. However, when adding content all the fields are null when it comes time to review a content submission. The Java errors thrown revolve around a null pointer exception.

* The root cause of the problem is that the database is set up using some character encoding other than UTF-8. The database must be set up using UTF-8 character encoding.

How do I run DSpace with Apache and Tomcat using the mod_jk connector?
See ModJk.

How do I run multiple instances of DSpace in one Tomcat Server?
See MultipleDspaceOneServer

How do localise the Creative Commons license?
See LocaliseCreativeCommons

DSPACE with ISO-8859-1 and others encodings
If you want to use an encoding other than UTF-8 you can edit the file src/org/dspace/storage/rdbms/DatabaseManager.java and modify the next lines
else if ((jdbcType == Types.CLOB && "oracle".equals(dbName)) || jdbcType == Types.VARCHAR)
{
    // Support CLOBs in place of TEXT columns in Oracle
    row.setColumn(name, results.getString(i));
}
/*else if (jdbcType == Types.VARCHAR)
{
    try
    {
        byte[] bytes = results.getBytes(i);
        if (bytes != null)
        {
            String mystring = new String(results.getBytes(i), "UTF-8");
            row.setColumn(name, mystring);
        }
        else
        {
            row.setColumn(name, results.getString(i));
        }
    }
    catch (UnsupportedEncodingException e)
    {
        // do nothing, UTF-8 is built in!
    }
}*/

Clearing Cocoon (XMLUI) cache

Apache Cocoon, the DSpace component XMLUI is based on, keeps a cache of visited pages to speed up access to them. This can pose a problem usually during development of XMLUI themes, when you update XSL, but an older version of the page is served from cache instead of the current version being constructed. It’s safe to clear the cache, the only impact it can have is on speed until the cache is repopulated.

In DSpace 1.8.0 and newer, there is a convenient link in XMLUI to clear the cache without restarting the servlet container:

1. log in as DSpace administrator
2. click on "Control Panel" in the sidebar menu
3. click on the "Java Information" tab
4. click on the "(Clear Cache Immediately)" link

Next to that link, you can also see how many pages are currently cached.

In DSpace versions up to 1.7.2, you have to clear the cache manually by shutting down Tomcat, deleting the cache directory and then restarting Tomcat:

1. ${tomcat6.home}/bin/shutdown.sh
2. cd ${tomcat6.home}/work/Catalina/{appropriate.domain.dir}/_
3. rm -rf cache-dir

Depending on your Tomcat version, whether you used a package and where you installed it, some common locations of the cache directory are:

/var/cache/tomcat5/work/Catalina/localhost/xmlui/cache-dir
/var/lib/tomcat6/work/Catalina/localhost/.../cache-dir
/opt/tomcat6/work/Catalina/localhost/xmlui/cache-dir
/usr/local/tomcat6/work/Catalina/localhost/xmlui/cache-dir

I changed access restrictions on a collection, but restrictions didn't change on its items
You have to remember that access restrictions are inherited from its parent collection only when items are submitted, not when access restrictions are modified. So any new items submitted to that collection will have the collection's restrictions you set, but existing items will keep their current permissions.

In other words: "When a new item gets archived in the repository, it inherits the default policies for an item, defined by the collection in which the item is archived. It's important to realize that these policies can altered for each item, bundle or bitstream, without altering other items or collection properties. Policies are really tied to items, bundles and bitstreams." (source)

So how would you do a mass-change of permissions?

- In XMLUI, use Advanced Policy Manager. Log in as an administrator, in the Administrative menu, click Authorizations (under Access Control.) Then in the main pane, under the search box, there's a link to the tool (Advanced authorizations tool: Click here to go to the item wildcard policy admin tool). See also this video guide: https://youtu.be/PFlq-yGSPY4
- You can write a curation task (see DSpace Curation System, Curation Task Cookbook)
- You can write custom SQL query on the "resourcepolicy" table.
  - In order to see this changes take effect in the UI (in Discovery search/browse), be sure to update the discovery index (using the -f flag) after running the SQL query.

How do I remove the search box from the front page in XMLUI?

This will remove the search box from the homepage body and keep the search box in the menu:

```xml
<xsl:template name="disable_front-page-search" match="dri:div[@id='aspect.discovery.SiteViewer.div.front-page-search']">
</xsl:template>
```

How can I change what is displayed in XMLUI for a particular collection?

If you want items in a particular collection to be rendered differently (add/remove/modify elements) than in the rest of the repository, you can use this recipe.

```xml

<xsl:template ...>
  ...
  <xsl:if test="$mycollection">
    <p>This item belongs to collection 123456789/1234.</p>
  </xsl:if>
  ...
</xsl:template>
```

I'm getting 'relation "bi_5_dis" does not exist' (or any other number)

These are the tables corresponding to individual browse indexes. Browse indexes are defined in dspace.cfg as webui.browse.index.1 through webui.browse.index.n and you can modify which metadata fields you want in the indexes, but you have to remember that after any such change, you need to re-create the indexes.

See ReIndexing Content (for Browse or Search) in the official documentation for your version of DSpace.

See also the relevant dspace.cfg section.

Duplicate items in Recent Submissions in XMLUI Discovery

This happens when you skip a step in the instructions for Discovery#EnablingDiscovery. In particular, forgetting to set recent.submissions.count = 0 will mean that recent items continue to be rendered by the Browse aspect in addition to the Discovery aspect, resulting in duplication.

To fix this, set recent.submissions.count = 0 in [dspace]/config/dspace.cfg and restart your servlet container. If needed, also clear the Cocoon cache: TechnicalFaq#TechnicalFaq-ClearingCocoon(XMLUI)cache

How do I update Messages.properties without rebuilding DSpace?
The Messages.properties file is used for translations both in DSpace core ([dspace]/lib/dspace-api-4.0-SNAPSHOT.jar) and in DSpace JSPUI ([dspace]/webapps/jspui/WEB-INF/lib/dspace-api-4.0-SNAPSHOT.jar). You usually want to update the JSPUI one:

1. find dspace-api-X.X.jar in [dspace]/webapps/jspui/WEB-INF/lib/
2. unzip it in an external folder
3. modify the Message.properties file
4. copy Message.properties to [dspace]/webapps/jspui/WEB-INF/classes/
5. restart your servlet container

I'm getting "SolrException: Bad Request" followed by a long query or a "tooManyClauses" Exception

From the related email thread:

The issue is related to the Access Rights Awareness feature introduced in Discovery and affect both the search and the browse system (when SolrBrowseDAO are used, optional in 3.x, by default in 4.0).

The issue here is that a single user is included in several groups - that means that the Solr query will include a filter to limit the results to only items that are visible by group1, group2, ..., groupN.

- If the number of group is high, you can reach the HTTP header limit (4kB or 8kB by default, see maxHttpHeaderSize Tomcat 7 docs) or a tooManyClauses Exception in Solr, that can be "solved" incrementing this parameter: https://github.com/DSpace/DSpace/blob/dspace-4_x/dspace/solr/search/conf/solrconfig.xml#L474
- When the number is to large you could also consider to disable the Access Rights Awareness feature, commenting this line: https://github.com/DSpace/DSpace/blob/dspace-4_x/dspace/config/spring/api/discovery.xml#L25

Solutions:

in case of exceeding header size:

- increase maxHttpHeaderSize in Tomcat

in case of tooManyClauses Exception:

- increment maxBooleanClauses; or
- disable Access Rights Awareness

Errors connecting to Solr

Due to localhost restriction, Solr is only listening on the interface associated with the "localhost" hostname (loopback interface, usually with IP address 127.0.0.1 or ::1). Trying to access it via the FQDN hostname (e.g. http://example.com:8080/solr/statistics) will result in an error (usually status code 403 Unauthorized).

You have to configure


Also make sure you have

solr.url=http://localhost:8080/solr/oai in [dspace]/config/modules/oai.cfg and

search.server=http://localhost:8080/solr/search in [dspace]/config/modules/discovery.cfg

All of these can be changed at once at build time by changing

solr.server=http://localhost:8080/solr in [dspace-source]/build.properties

Error starting Tomcat: SEVERE: Error listenerStart

Problem: When starting Tomcat, the Tomcat log contains the following error:

Jun 01, 2016 4:41:08 PM org.apache.catalina.core.StandardContext startInternal
SEVERE: Error listenerStart

Explanation: While this may indicate various errors, one possible source of problems is re-using configuration from an older version of DSpace which is incompatible with the new code. Most likely the incompatibility is in [dspace]/config/spring/.

Solution: Move away your customized configuration directory, start again with the fresh configuration corresponding to the DSpace version you're installing or updating to and then make your changes in that configuration.
Metadata values in CSV export seem to have duplicate columns

Symptom: In a CSV export from Batch Metadata Editing, you can often see that the same column seems to be listed twice, where each of these columns contains values in different rows. The two columns may look like this: dc.title, tc.title[]

Explanation: This is a harmless inconsistency in how DSpace stores metadata. The brackets designate which language the metadata value is in (each metadata field may have multiple values in multiple languages, which is appears in the CSV like this: dc.title, dc.title[de], dc.title[en_US]). In the database, an empty language may have two representations - either an empty string "" or a special NULL value. As DSpace organically developed over time, the programmers of various data ingest methods chose one of these two representations. If you use multiple ingest methods, you may end up with metadata that use both.

Fix 1: In the CSV file, move the values from one of the columns to the other. Then import the CSV file. Upon next export, the empty column won't appear in the CSV file anymore.

Fix 2: Use the following SQL statement to fix the data: UPDATE metadatavalue SET text_lang=NULL WHERE text_lang='';

Any of these fixes will only fix your current data. If you continue to use the two different ingest methods, you'll get new such duplicates.

The permanent fix would have to fix the import methods. The JIRA issue is DS-2174 - Getting issue details...