Upgrading From 1.6.x to 1.7.x

In the notes below `[dspace]` refers to the install directory for your existing DSpace installation, and *[dspace-source]* to the source directory for DSpace 1.7.x. Whenever you see these path references, be sure to replace them with the actual path names on your local system.

Upgrade Steps

Before upgrading you need to check you are using the current recommended minimum versions of Java (1.6), Maven (2.0.8 or above) and ant (1.7 or above). For more details, see the current listing of Prerequisite Software

1. **Backup Your DSpace.** First, and foremost, make a complete backup of your system, including:
   - A snapshot of the database. *To have a "snapshot" of the PostgreSQL database, you need to shut it down during the backup. You should also have your regular PostgreSQL Backup output (using Postgres' pg_dump command).*
   - The asset store (`[dspace]/assetstore` by default)
   - Your configuration files and customizations to DSpace (including any customized scripts).
2. **Download DSpace 1.7.x** Retrieve the new DSpace 1.7.x source code either as a download from DSpace.org or check it out directly from the SVN code repository. If you downloaded DSpace do not unpack it on top of your existing installation. Refer to Installation Instructions, Step 3 for unpacking directives.
3. **Stop Tomcat.** Take down your servlet container. For Tomcat, use the `$CATALINA_HOME/bin/shutdown.sh` script. (Many Unix-based installations will have a startup/shutdown script in the `/etc/init.d` or `/etc/rc.d` directories).
4. **Apply any customizations.** If you have made any local customizations to your DSpace installation they will need to be migrated over to the new DSpace. These are normally housed in one of the following places:
   - JSPUI modifications: `[dspace-source]/dspace/modules/jspui/src/main/webapp/`
   - XMLUI modifications: `[dspace-source]/dspace/modules/xmlui/src/main/webapp/`
5. **Update Configuration Files.** Some parameters have changed and some are new. You can either attempt to make these changes in your current 1.6.x `dspace.cfg` file, or you can start with a new 1.7 `dspace.cfg` and re-modify it as needed. Configuration changes are noted below:
   - *__CORRECTION__* There was a missing hyphen "-" in the property key for mail character set:

     ```
     # Set the default mail character set. This may be over ridden by providing a line
     # inside the email template "charset: <encoding>", otherwise this default is used.
     #mail.charset = UTF-8
     ```

   - *__CORRECTION__* This was moved from the end of the solr configuration section to just under Logging Configurations:

     ```
     # If enabled, the logging and the solr statistics system will look for
     # an X-Forward header. If it finds it, it will use this for the user IP Address
     # useProxies = true
     ```

   - *__CHANGE__* The MediaFilter is now able to process Power Point Text Extracter

     ```
     #Names of the enabled MediaFilter or FormatFilter plugins
     filter.plugins = PDF Text Extractor, HTML Text Extractor, \
                                  PowerPoint Text Extractor, \
                                  Word Text Extractor, JPEG Thumbnail
     # [To enable Branded Preview]: remove last line above, and uncomment 2 lines below
     #                       Word Text Extractor, JPEG Thumbnail, \
     #                       Branded Preview JPEG

     #Assign 'human-understandable' names to each filter
     plugin.named.org.dspace.app.mediafilter.FormatFilter = \
       org.dspace.app.mediafilter.PDFFilter = PDF Text Extractor, \
       org.dspace.app.mediafilter.HTMLFilter = HTML Text Extractor, \
       org.dspace.app.mediafilter.WordFilter = Word Text Extractor, \
       org.dspace.app.mediafilter.PowerPointFilter = PowerPoint Text Extractor, \
       org.dspace.app.mediafilter.JPEGFilter = JPEG Thumbnail, \
       org.dspace.app.mediafilter.BrandedPreviewJPEGFilter = Branded Preview JPEG

     #Configure each filter's input format(s)
     filter.org.dspace.app.mediafilter.PDFFilter.inputFormats = Adobe PDF
     filter.org.dspace.app.mediafilter.HTMLFilter.inputFormats = HTML, Text
     filter.org.dspace.app.mediafilter.WordFilter.inputFormats = Microsoft Word
     filter.org.dspace.app.mediafilter.PowerPointFilter.inputFormats = Microsoft Powerpoint, Microsoft
     Powerpoint XML
     filter.org.dspace.app.mediafilter.JPEGFilter.inputFormats = BMP, GIF, JPEG, image/png
     filter.org.dspace.app.mediafilter.BrandedPreviewJPEGFilter.inputFormats = BMP, GIF, JPEG, image/png
     ```

   - *__CHANGE__* The Crosswalk Plugin Configuration has changed with additional lines. Edit your file accordingly:

```
# Crosswalk Plugin Configuration:
#   The purpose of Crosswalks is to translate an external metadata format to/from
#   the DSpace Internal Metadata format (DIM) or the DSpace Database.
#   Crosswalks are often used by one or more Packager plugins (see below).
plugin.named.org.dspace.content.crosswalk.IngestionCrosswalk = \
  org.dspace.content.crosswalk.AIPDIMCrosswalk = DIM, \
  org.dspace.content.crosswalk.AIPTechMDCrosswalk = AIP-TECHMD, \
  org.dspace.content.crosswalk.PREMISCrosswalk = PREMIS, \
  org.dspace.content.crosswalk.OREIngestionCrosswalk = ore, \
  org.dspace.content.crosswalk.NullIngestionCrosswalk = NIL, \
  org.dspace.content.crosswalk.OAIDCIngestionCrosswalk = dc, \
  org.dspace.content.crosswalk.DIMIngestionCrosswalk = dim, \
  org.dspace.content.crosswalk.METSRightsCrosswalk = METSRIGHTS, \
  org.dspace.content.crosswalk.RoleCrosswalk = DSPACE-ROLES

plugin.selfnamed.org.dspace.content.crosswalk.IngestionCrosswalk = \
  org.dspace.content.crosswalk.XSLTIngestionCrosswalk, \
  org.dspace.content.crosswalk.QDCCrosswalk

plugin.named.org.dspace.content.crosswalk.StreamIngestionCrosswalk = \
  org.dspace.content.crosswalk.NullStreamIngestionCrosswalk = NULLSTREAM, \
  org.dspace.content.crosswalk.CreativeCommonsRDFStreamIngestionCrosswalk = DSPACE_CCRDF, \
  org.dspace.content.crosswalk.LicenseStreamIngestionCrosswalk = DSPACE_DEPLICENSE

plugin.named.org.dspace.content.crosswalk.DisseminationCrosswalk = \
  org.dspace.content.crosswalk.AIPDIMCrosswalk = DIM, \
  org.dspace.content.crosswalk.AIPTechMDCrosswalk = AIP-TECHMD, \
  org.dspace.content.crosswalk.SimpleDCDisseminationCrosswalk = DC, \
  org.dspace.content.crosswalk.SimpleDCDisseminationCrosswalk = dc, \
  org.dspace.content.crosswalk.PREMISCrosswalk = PREMIS, \
  org.dspace.content.crosswalk.METSDisseminationCrosswalk = METS, \
  org.dspace.content.crosswalk.METSDisseminationCrosswalk = mets, \
  org.dspace.content.crosswalk.METSRightsCrosswalk = METSRIGHTS, \
  org.dspace.content.crosswalk.OREDisseminationCrosswalk = ore, \
  org.dspace.content.crosswalk.DIMDisseminationCrosswalk = dim, \
  org.dspace.content.crosswalk.RoleCrosswalk = DSPACE-ROLES
```

- *__NEW__*

```
plugin.named.org.dspace.content.crosswalk.StreamDisseminationCrosswalk = \
  org.dspace.content.crosswalk.CreativeCommonsRDFStreamDisseminationCrosswalk = DSPACE_CCRDF, \
  org.dspace.content.crosswalk.CreativeCommonsTextStreamDisseminationCrosswalk = DSPACE_CCTEXT, \
  org.dspace.content.crosswalk.LicenseStreamDisseminationCrosswalk = DSPACE_DEPLICENSE
```

- *__CHANGE__* The Packager Plugin Configuration has changed considerably. Carefully revise your configuration file:

```
Packager Plugin Configuration:
#   Configures the ingest and dissemination packages that DSpace supports.
#   These Ingester and Disseminator classes support a specific package file format
#   (e.g. METS) which DSpace understands how to import/export.  Each Packager
#   plugin often will use one (or more) Crosswalk plugins to translate metadata (see above).
plugin.named.org.dspace.content.packager.PackageDisseminator = \
  org.dspace.content.packager.DSpaceAIPDisseminator = AIP, \
  org.dspace.content.packager.DSpaceMETSDisseminator = METS, \
  org.dspace.content.packager.RoleDisseminator = DSPACE-ROLES

plugin.named.org.dspace.content.packager.PackageIngester = \
  org.dspace.content.packager.DSpaceAIPIngester = AIP, \
  org.dspace.content.packager.PDFPackager  = Adobe PDF, PDF, \
  org.dspace.content.packager.DSpaceMETSIngester = METS, \
  org.dspace.content.packager.RoleIngester = DSPACE-ROLES
```

- *__CHANGE__* The Mets Ingester configuration has change and been updated. Carefully edit:

```
#### METS ingester configuration:
```

```
# These settings configure how DSpace will ingest a METS-based package

# Configures the METS-specific package ingesters (defined above)
# 'default' settings are specified by 'default' key

# Default Option to save METS manifest in the item: (default is false)
mets.default.ingest.preserveManifest = false

# Default Option to make use of collection templates when using the METS ingester (default is
false)
mets.default.ingest.useCollectionTemplate = false

# Default crosswalk mappings
# Maps a METS 'mdtype' value to a DSpace crosswalk for processing.
# When the 'mdtype' value is same as the name of a crosswalk, that crosswalk
# will be called automatically (e.g. mdtype='PREMIS' calls the crosswalk named
# 'PREMIS', unless specified differently in below mapping)
# Format is 'mets.default.ingest.crosswalk.<mdType> = <DSpace-crosswalk-name>'
mets.default.ingest.crosswalk.DC = QDC
mets.default.ingest.crosswalk.DSpaceDepositLicense = DSPACE_DEPLICENSE
mets.default.ingest.crosswalk.Creative\ Commons = DSPACE_CCRDF
mets.default.ingest.crosswalk.CreativeCommonsRDF = DSPACE_CCRDF
mets.default.ingest.crosswalk.CreativeCommonsText = NULLSTREAM

# Locally cached copies of METS schema documents to save time on ingest.  This
# will often speed up validation & ingest significantly.  Before enabling
# these settings, you must manually cache all METS schemas in
# [dspace]/config/schemas/ (does not exist by default).  Most schema documents
# can be found on the http://www.loc.gov/ website.
# Enable the below settings to pull these *.xsd files from your local cache.
# (Setting format: mets.xsd.<abbreviation> = <namespace> <local-file-name>)
#mets.xsd.mets = http://www.loc.gov/METS/ mets.xsd
#mets.xsd.xlink = http://www.w3.org/1999/xlink xlink.xsd
#mets.xsd.mods = http://www.loc.gov/mods/v3 mods.xsd
#mets.xsd.xml = http://www.w3.org/XML/1998/namespace xml.xsd
#mets.xsd.dc = http://purl.org/dc/elements/1.1/ dc.xsd
#mets.xsd.dcterms = http://purl.org/dc/terms/ dcterms.xsd
#mets.xsd.premis = http://www.loc.gov/standards/premis PREMIS.xsd
#mets.xsd.premisObject = http://www.loc.gov/standards/premis PREMIS-Object.xsd
#mets.xsd.premisEvent = http://www.loc.gov/standards/premis PREMIS-Event.xsd
#mets.xsd.premisAgent = http://www.loc.gov/standards/premis PREMIS-Agent.xsd
#mets.xsd.premisRights = http://www.loc.gov/standards/premis PREMIS-Rights.xsd

#### AIP Ingester & Disseminator Configuration
# These settings configure how DSpace will ingest/export its own
# AIP (Archival Information Package) format for backups and restores
# (Please note, as the DSpace AIP format is also METS based, it will also
# use many of the 'METS ingester configuration' settings directly above)

# AIP-specific ingestion crosswalk mappings
# (overrides 'mets.default.ingest.crosswalk' settings)
# Format is 'mets.dspaceAIP.ingest.crosswalk.<mdType> = <DSpace-crosswalk-name>'
mets.dspaceAIP.ingest.crosswalk.DSpaceDepositLicense = NULLSTREAM
mets.dspaceAIP.ingest.crosswalk.CreativeCommonsRDF = NULLSTREAM
mets.dspaceAIP.ingest.crosswalk.CreativeCommonsText = NULLSTREAM

# Create EPerson if necessary for Submitter when ingesting AIP (default=false)
# (by default, EPerson creation is already handled by 'DSPACE-ROLES' Crosswalk)
#mets.dspaceAIP.ingest.createSubmitter = false

## AIP-specific Disseminator settings
# These settings allow you to customize which metadata formats are exported in AIPs

# Technical metadata in AIP (exported to METS <techMD> section)
# Format is <label-for-METS>:<DSpace-crosswalk-name> [, ...] (label is optional)
# If unspecfied, defaults to "PREMIS"
aip.disseminate.techMD = PREMIS, DSPACE-ROLES

# Source metadata in AIP (exported to METS <sourceMD> section)
# Format is <label-for-METS>:<DSpace-crosswalk-name> [, ...] (label is optional)
# If unspecfied, defaults to "AIP-TECHMD"
```

```
aip.disseminate.sourceMD = AIP-TECHMD

# Preservation metadata in AIP (exported to METS <digipovMD> section)
# Format is <label-for-METS>:<DSpace-crosswalk-name> [, ...] (label is optional)
# If unspecified, defaults to nothing in <digiprovMD> section
#aip.disseminate.digiprovMD =

# Rights metadata in AIP (exported to METS <rightsMD> section)
# Format is <label-for-METS>:<DSpace-crosswalk-name> [, ...] (label is optional)
# If unspecified, default to adding all Licenses (CC and Deposit licenses),
# as well as METSRights information
aip.disseminate.rightsMD = DSpaceDepositLicense:DSPACE_DEPLICENSE, \
    CreativeCommonsRDF:DSPACE_CCRDF, CreativeCommonsText:DSPACE_CCTEXT, METSRIGHTS

# Descriptive metadata in AIP (exported to METS <dmdSec> section)
# Format is <label-for-METS>:<DSpace-crosswalk-name> [, ...] (label is optional)
# If unspecfied, defaults to "MODS, DIM"
aip.disseminate.dmd = MODS, DIM
```

- *__NEW__* A new property has been added to control the discovery index for the Event System Configuration:

```
# consumer to maintain the discovery index
event.consumer.discovery.class = org.dspace.discovery.IndexEventConsumer
event.consumer.discovery.filters =
Community|Collection|Item|Bundle+Add|Create|Modify|Modify_Metadata|Delete|Remove
```

- *__NEW__* License bundle display is now configurable. You are able to either display or suppress.

```
# whether to display the contents of the licence bundle (often just the deposit
# licence in standard DSpace installation
webui.licence_bundle.show = false
```

- *__CORRECTION__* Thumbnail generation. The width and height of generated thumbnails had a missing equal sign.

```
# maximum width and height of generated thumbnails
thumbnail.maxwidth  = 80
thumbnail.maxheight = 80
```

- *__CORRECTION and ADDITION__* Authority Control Settings have changed. Formerly called ChoiceAuthority, it is now referred to as DCInputAuthority.

```
## The DCInputAuthority plugin is automatically configured with every
## value-pairs element in input-forms.xml, namely:
##   common_identifiers, common_types, common_iso_languages
##
## The DSpaceControlledVocabulary plugin is automatically configured
## with every *.xml file in [dspace]/config/controlled-vocabularies,
## and creates a plugin instance for each, using base filename as the name.
## eg: nsi, srsc.
## Each DSpaceControlledVocabulary plugin comes with three configuration options:
# vocabulary.plugin._plugin_.hierarchy.store = <true|false>    # default: true
# vocabulary.plugin._plugin_.hierarchy.suggest = <true|false>  # default: true
# vocabulary.plugin._plugin_.delimiter = "<string>"            # default: "::"
##
## An example using "srsc" can be found later in this section

#plugin.selfnamed.org.dspace.content.authority.ChoiceAuthority = \
# org.dspace.content.authority.DCInputAuthority, \
# org.dspace.content.authority.DSpaceControlledVocabulary
```

- *__NEW__* Controls autocomplete for authority control

```
## demo: subject code autocomplete, using srsc as authority
## (DSpaceControlledVocabulary plugin must be enabled)
#choices.plugin.dc.subject = srsc
#choices.presentation.dc.subject = select
#vocabulary.plugin.srsc.hierarchy.store = true
#vocabulary.plugin.srsc.hierarchy.suggest = true
#vocabulary.plugin.srsc.delimiter = "::"
```

- ***NEW*** You are now able to order your bitstreams by sequence id or file name.

```
#### Ordering of bitstreams ####

## Specify the ordering that bitstreams are listed.
##
## Bitstream field to sort on.  Values: sequence_id or name. Default: sequence_id
#webui.bitstream.order.field = "sequence_id"

## Direction of sorting order. Values: DESC or ASC. Default: ASC
#webui.bitstream.order.direction = ASC
```

- ***NEW*** DSpace now includes a metadata mapping feature that makes repository content discoverable by Google Scholar:

```
##### Google Scholar Metadata Configuration #####
google-metadata.config = ${dspace.dir}/config/crosswalks/google-metadata.properties
google-metadata.enable = true
```

- ***NEW*** XMLUI is now able to concatenate CSS, JS and JSON files:

```
# Enabling this property will concatenate CSS, JS and JSON files where possible.
# CSS files can be concatenated if multiple CSS files with the same media attribute
# are used in the same page. Links to the CSS files are automatically referring to the
# concatenated resulting CSS file.
# The theme sitemap should be updated to use the ConcatenationReader for all js, css and json
# files before enabling this property.
#xmlui.theme.enableConcatenation = false

# Enabling this property will minify CSS, JS and JSON files where possible.
# The theme sitemap should be updated to use the ConcatenationReader for all js, css and json
# files before enabling this property.
#xmlui.theme.enableMinification = false
```

- ***NEW*** XMLUI Mirage Theme. This is a new theme with it's own configuration:

```
### Setings for Item lists in Mirage theme ###
# What should the emphasis be in the display of item lists?
# Possible values : 'file', 'metadata'. If your repository is
# used mainly for scientific papers 'metadata' is probably the
# best way. If you have a lot of images and other files 'file'
# will be the best starting point
# (metdata is the default value if this option is not specified)
#xmlui.theme.mirage.item-list.emphasis = file
```

- ***NEW*** OAI Response default change.

```
# DSpace by default uses 100 records as the limit for the oai responses.
# This can be altered by enabling the oai.response.max-records parameter
# and setting the desired amount of results.
oai.response.max-records = 100
```

- *__CHANGE__* EPDCX property key has been renamed.

```
# Define the metadata type EPDCX (EPrints DC XML)
# to be handled by the SWORD crosswalk configuration
#
mets.default.ingest.crosswalk.EPDCX = SWORD
```

- *__NEW__* New SOLR Statistic Property keys:

```


# Timeout for the resolver in the dns lookup
# Time in milliseconds, defaults to 200 for backward compatibility
# Your systems default is usually set in /etc/resolv.conf and varies
# between 2 to 5 seconds, to high a value might result in solr exhausting
# your connection pool
solr.resolver.timeout = 200

----

# Enable/disable logging of spiders in solr statistics.
# If false, and IP matches an address in solr.spiderips.urls, event is not logged.
# If true, event will be logged with the 'isBot' field set to true
# (see solr.statistics.query.filter.* for query filter options)
# Default value is true.
#solr.statistics.logBots = true
```

6. **Build DSpace.** Run the following commands to compile DSpace.:

```
cd [dspace-source]/dspace/
mvn -U clean package
```

   You will find the result in `[dspace-source]/dspace/target/dspace-[version]-build.dir` . Inside this directory is the compiled binary distribution of DSpace. Before rebuilding DSpace, the above command will clean out any previously compiled code ('clean') and ensure that your local DSpace JAR files are updated from the remote maven code repository.
7. **Update DSpace.** Update the DSpace installed directory with the new code and libraries. Issue the following commands:

```
cd [dspace-source]/dspace/target/dspace-[version]-build.dir
ant -Dconfig=[dspace]/config/dspace.cfg update
```

8. **Update the Database.** You will need to run the 1.6.x to 1.7.x database upgrade script.
   - For PostgreSQL:

     ```
     psql -U [dspace-user] -f [dspace-source]/dspace/etc/postgres/database_schema_16-17.sql [database name]
     ```

     (Your database name is by default 'dspace'). Example:

     ```
     psql -U dspace -f [dspace-source]/dspace/etc/postgres/database_schema_16-17.sql dspace
     ```

   - For Oracle: Execute the upgrade script, e.g. with sqlplus, recording the output:
     a. Start SQL*Plus with `sqlplus [connect args]`
     b. Record the output: `SQL> spool 'upgrade.lst'`
     c. Run the upgrade script `SQL> @[dspace-source]/dspace/etc/oracle/database_schema_16-17.sql`
     d. Turn off recording of output: `SQL> spool off`
9. **Generate Browse and Search Indexes**. It's always good policy to rebuild your search and browse indexes when upgrading to a new release. To do this, run the following command from your DSpace install directory (as the 'dspace' user):

```
[dspace]/bin/dspace index-init
```

10. **Deploy Web Applications**. If your servlet container (e.g. Tomcat) is not configured to look for new web applications in your `[dspace]/webapps` directory, then you will need to copy the web applications files into the appropriate subdirectory of your servlet container. For example:

```
cp -R [dspace]/webapps/* [tomcat]/webapps/
```

11. **Restart servlet container.** Now restart your Tomcat/Jetty/Resin server program and test out the upgrade.
12. **Add a new crontab entry**, or add to your system's scheduler, the following, run as the DSpace user, to enable routine maintenance of your SOLR indexes. If you do not run this command daily, it is likely your production instances of DSpace will exhaust the available memory in your servlet container

```
[dspace]/bin/dspace stats-util -o
```