

Installation of AgriOcean Dspace on Linux.

1. Prerequisite software:

can be downloaded through the software install procedures of the different Linux distributions (tested with Fedora and Ubuntu)

- Oracle Java JDK 6 or later (standard SDK is fine, you don't need J2EE).
 - Also available at: <http://java.sun.com/javase/downloads/index.jsp>
 - Only Oracle's Java has been tested with each Dspace release and is known to work correctly. Other flavors of Java may pose problems.
- Apache Tomcat 5.5 or later.
 - Also available at: <http://tomcat.apache.org>.
 - Note that DSpace will need to run as the same user as Tomcat. Fedora and Ubuntu installation creates a tomcat user. The AgriOcean Dspace must have the same user.
 - You need to ensure that Tomcat has a) enough memory to run DSpace and b) uses UTF-8 as its default file encoding for international character support. So ensure in your startup scripts (etc) that the following environment variable is set:
`JAVA_OPTS="-Xmx512M -Xms64M -Dfile.encoding=UTF-8"`
 - Modifications in [tomcat-conf]/server.xml: You also need to alter Tomcat's default configuration to support searching and browsing of multi-byte UTF-8 correctly. You need to add a configuration option to the <Connector> element in [tomcat-conf]/server.xml: `URIEncoding="UTF-8"`
e.g. if you're using the default Tomcat config, it should read:
<!-- Define a non-SSL HTTP/1.1 Connector on port 8080 -->
<Connector port="8080"
 maxThreads="150"
 minSpareThreads="25"
 maxSpareThreads="75"
 enableLookups="false"
 redirectPort="8443"
 acceptCount="100"
 connectionTimeout="20000"
 disableUploadTimeout="true"
 URIEncoding="UTF-8"/>
- PostgreSQL 8.4 or higher
 - Also available at: <http://www.postgresql.org/> .
 - Unicode (specifically UTF-8) support must be enabled. This is enabled by default in 8.0+.
 - Download also the pgAdmin III client for Postgresql, a handy tool to manage the database.

2. AgriOcean Dspace 1.1

- Create a main directory repos at command line:
 - `sudo mkdir /repos`
 - `chown -R <user>:<tomcat user> /repos`
 - `chmod -R 775 /repos`

- Download from <http://code.google.com/p/agrioccean/downloads> the files in the directory /repos
 - Installation_AOD_Linux_beta.pdf
 - AOD-Setup-1-1_Linux_Beta.7z
- Unpack AOD-Setup-1-1_Linux_Beta.7z. If 7z is not yet installed, a modern Linux distribution will download the necessary packages automatically (with ROOT authentication requested)
 - `mv aod /repos/aod`
 - `sudo chown -R <user>:<tomcat user> /repos/aod`
 - `sudo chmod -R 775 /repos/aod`

3. Installation procedure

- PostgreSQL database: From the command line.
 - Create Postgres user
 - `sudo su postgres`
 - `createuser -U postgres -d -A -P dspace`
 - `exit`
 - Allow user to connect to the database
 - `sudo vi <postgresql directory>/pg_hba.conf`
 - `# Add this line to the configuration: local all dspace md5`
 - `sudo service postgresql restart`
 - Create database: name = aod
 - `createdb -U dspace -E UNICODE aod`
 - `psql -U dspace aod < /repos/aod/aod.sql`
- Tomcat6 server – configuration (see also under 1. prerequisite software):
 - Open <tomcat-conf>/server.xml in a text editor:
Insert the following to let tomcat know about our dspace webapps. This should be added before the closing "</Host>" tag.
 - `<Context path="/dspace" docBase="/repos/aod/webapps/dspace" allowLinking="true"/>`
 - `<Context path="/oai" docBase="/repos/aod/webapps/oai" allowLinking="true"/>`
 - `<Context path="/solr" docBase="/repos/aod/webapps/solr" allowLinking="true"/>`
- Check Dspace configuration at /repos/aod/config/dspace.cfg

| Configuration option in dspace.cfg: | Comment |
|--|---|
| !!! Touch only the configuration marked in red !!! | #: used to uncomment a whole line. |
| ##### Basic information ##### # DSpace installation directory dspace.dir = c:/repos/aod | If you want to install AgriOcean DSpace in another directory, the reference to the base directory has to be changed to in the next files: <code><repos_directory></code> 1. <code>\config\dspace.cfg</code> <code><repos_directory>\webapps</code> 2. <code>\dspace\WEB-INF\web.xml</code> 3. <code>\oai\WEB-INF\web.xml</code> 4. <code>\solr\WEB-INF\web.xml</code> |

| | |
|---|--|
| <p># DSpace host name - should match base URL. Do not include port number dspace.hostname = localhost</p> <p># DSpace base host URL. Include port number etc. dspace.baseUrl = http://localhost:8080</p> <p># DSpace base URL. Include port number etc., but NOT trailing slash # Change to xmlui if you wish to use the xmlui as the default, or remove # "/jspui" and set webapp of your choice as the "ROOT" webapp in # the servlet engine. dspace.url = \${dspace.baseUrl}/dspace</p> <p># The base URL of the OAI webapp (do not include /request). dspace.oai.url = \${dspace.baseUrl}/oai</p> <p># Name of the site dspace.name = AgriOcean</p> | <p>AgriOcean DSpace is installed on your local machine, to make it visible for external users you need to give it the name of your server: A URL or an IP E.g. www.yourinstitute.org or IP (ask your network manager). If tomcat is installed on another port you have to change the port number.</p> <p><i>When changing the port number, change it also in file <repos_directory>\config\dspace-solr-search.cfg</i></p> <p>The url of your institute will be www.yourinstitute.org:8080/dspace If you want a url without extension, you can rename the link in the context path to ROOT (in /<tomcat-config dir>/server.xml – see above)</p> <p>Here you can adapt the name, which will be used on the website</p> |
| <p>##### Database settings #####</p> <p># URL for connecting to database #db.url = \${default.db.url} db.url = jdbc:postgresql://localhost:5432/aod</p> <p># JDBC Driver #db.driver = \${default.db.driver} db.driver = org.postgresql.Driver</p> <p># Database username and password #db.username = \${default.db.username} #db.password = \${default.db.password} db.username = dspace db.password = ****</p> | <p>If you have defined another database, user, you can change it at resp. db.url and db.username. Fill out the password for the postgresql user who manages the aod database.</p> |
| <p>##### Email settings #####</p> <p># SMTP mail server mail.server=smtgmail.com</p> <p># SMTP mail server authentication username and password (if required) mail.server.username = aaa@gmail.com mail.server.password =*****</p> <p># SMTP mail server alternate port (defaults to 25) mail.server.port = 25</p> <p># From address for mail mail.from.address = oceandocs@gmail.com</p> | <p>It is possible to use a gmail as a mail server.</p> <p>Every user should create their own gmail account and fill out the mail and password where required.</p> <p>If you use a local email server, no authentication is required..</p> |

| | |
|---|--|
| <pre># Currently limited to one recipient! feedback.recipient = oceandocs@gmail.com # General site administration (Webmaster) e-mail mail.admin = oceandocs@gmail.com # Recipient for server errors and alerts alert.recipient = oceandocs@gmail.com # Recipient for new user registration emails registration.notify = oceandocs@gmail.com mail.server.disabled = true</pre> | <p>An option is added to disable the mailserver. By default, this property is set to false. By setting mail.server.disabled = true, DSpace will not send out emails. This is <u>especially useful for development and test environments</u> where production data is used when testing functionality.</p> |
| <pre># Default language for metadata values default.language = en</pre> | <p>If you use another language as default for the submitted information, you can change it here. Use the ISO693-1 definition (e.g. fr=français, es=Espanol, ...)</p> |
| <pre>##### Usage Logging ##### solr.log.server = http://localhost:8080/solr/statistics</pre> | <p>set this to be the port you run the dspace "solr" webapp on, by default, we are assuming a test configuration with tomcat still running on port 8080</p> |

- Final Steps:

- Initialize indexing script:
 - `sudo -[tomcat user] /repos/aod/bin/dspace index-init`
- create new administrator (optional) – an administrator already exist: agr@ocean
password: agri
 - `sudo -[tomcat user] /repos/aod/bin/dspace create-administrator`
- Setup crons: `sudo -[tomcat user] crontab-e`
Timings of the crons are for illustration purpose only. Actual timings on the server may be different.

Send out subscription e-mails at 01:00 every day

0 1 * * * [dspace]/bin/sub-daily

Run the media filter at 02:00 every day

0 2 * * * [dspace]/bin/filter-media

Run the checksum checker at 03:00

0 3 * * * [dspace]/bin/checker -lp

Mail the results to the sysadmin at 04:00

0 4 * * * [dspace]/bin/dsrun org.dspace.checker.DailyReportEmailer -c

Clean up the database nightly at 4.20am

20 4 * * * vacuumdb --analyze dspace > /dev/null 2>&1

Run stat analysis

0 1 * * * [dspace]/bin/stat-general

0 1 * * * [dspace]/bin/stat-monthly

0 2 * * * [dspace]/bin/stat-report-general

0 2 * * * [dspace]/bin/stat-report-monthly

- Restart tomcat
- Open URL in web browser
 - Fire up you favourite web browser and go to:
 - <http://localhost:8080/dspace>
- You should now be able to see AgriOceanDspace website up and running

For running Dspace over port 80: See

<https://wiki.duraspace.org/display/DSPACE/DspaceOnStandardPorts>

For Advanced Installation options: go to the Dspace pages at

http://www.dspace.org/1_7_1Documentation/Installation.html#Installation-AdvancedInstallation.

- 'cron' Jobs
- Multilingual Installation
- **Dspace over HTTPS:**
- The Handle Server (not installed for AgriOcean Dspace)
- Updating Existing Handle Prefixes
- **Google and HTML sitemaps**
- DSpace Statistics

Marc Goovaerts

Hasselt University Library,

Diepenbeek, February 13, 2012.

Marc.goovaerts@uhasselt.be