

GYH d'UbX'WcbZ][i fUh]cb'cZ a digital library based on Ubuntu and DSpace

Author

José Vicente García Martínez

Josev.garcia@gmail.com

Dissertation Supervisor

Ismael Ripoll

iripoll@disca.upv.es

Escuela Técnica Superior de Ingeniería Informática

www.etsinf.upv.es

Polytechnic University of Valencia (Spain)

www.upv.es



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA



Fundación
Globalidad y
Microeconomía

Table of contents

1.	Introduction.....	5
1.1.	Overview of the Organization.....	5
1.2.	Objectives	6
1.3.	Dissertation structure	7
2.	Analysis	9
2.1.	Choosing Open Source digital library software	9
2.2.	Choosing a Linux distribution	11
2.3.	Integration needs with the current web platform	12
3.	Installation and Configuration	15
3.1.	Linux distribution setup.....	15
3.2.	Software prerequisites setup	16
3.2.1.	Java	16
3.2.2.	Apache Tomcat.....	17
3.2.3.	Maven	20
3.2.4.	Postgresql	21
3.2.5.	Xpdf	22
3.2.6.	Postfix	22
3.2.7.	Security configuration.....	24
3.2.8.	Configure Unattended Updates:	25
3.3.	DSpace installation.....	27
3.3.1.	Configure Folder Permissions	28
3.3.2.	Hide Community or Collection from list	28
3.3.3.	Localization	31
3.4.	DSpace Interface Configuration.....	32
3.4.1.	Set Mirage as the default search theme	32
3.4.2.	Customize the Mirage theme	32
3.5.	Searches and Results.....	33
3.5.1.	Enable Discovery as the default search engine.....	33
3.5.2.	Select the indexed fields for the search	36
3.5.3.	Hide metadata listed in results from the user.....	36
3.5.4.	Define search indexes.....	36
3.5.5.	Update thumbnails and full text sources.....	37
3.5.6.	Enable thumbnails in the search results	37
3.5.7.	XPDF MediaFilter Configuration	37
3.6.	Metadata Configuration	39

3.6.1.	Change item metadata style	39
3.6.2.	Change item description metadata	40
3.6.3.	Change visibility of items through RSS, OAI and Subscriptions...41	
3.7.	Import and Export Information	41
3.7.1.	Export a collection.....	41
3.7.2.	Import a collection	41
3.8.	Schedule 'cron' Jobs	41
3.9.	Version Update	43
3.10.	Results	47
4.	Conclusions	51
4.1.	Technologic framework	52
5.	Bibliografy	54

1. Introduction

1. Introduction

The purpose of this project is to analyze the requirements, prepare the technologic environment and perform the deployment of a Digital Repository to improve the information processes of an organization.

A previous project has been conducted to determine the more suitable tool for this matter, which has concluded that the software that best meets the requirements is the DSpace Digital Repository.

The organization that will use this repository is the "Globalidad y Microeconomía" foundation, sited in the innovation campus of the Polytechnic University of Valencia.

This repository is going to be used to store the internal documentation that the organization generates as well as the great amount of external documentation that feeds the foundation for the proposal of achieving their functions.

By means of this tool will be possible to execute full text searches in the contents of the documents, as well as storing the appropriate metadata for each document. In this way a series of documents will be published in the repository's public interface so the users can download and see their contents.

This document describes the steps taken to choose the technologic environment, to install the linux server, to install the required software and to configure and customize each piece of the solution.

In addition this project aims to contribute with the open software community, disclosing the acquired knowledge with the installation, configuration and customization of this software, to facilitate the task of implementing this software.

1.1. Overview of the Organization

The organization beneficiary of the results of this project is The "*Globalidad y Microeconomía*" Foundation. Is Managed by Mr. Justo Nieto Nieto, who is an authority in the divulgation of knowledge as having hold the position of rector of the Polytechnic University of Valencia as well as Culture Adviser of the Valencian Province Government. This foundation is placed in the Innovation Polytechnic City of the Polytechnic University of Valencia, Camino de Vera without number. Building 8B, Access N, 5th Plant.

The foundation births in 2008 in the Polytechnic University of Valencia. Is an without commercial reasons and is promoted from the UPV and its council, which is the maximum authority of the foundation and is totally independent in their decisions.

The Foundation was created with the main objective of realizing proposals that can become opportunities for business success, generally from

unconventional Innovation. The foundation aims to be a benchmark for the innovation in Valencia. Providing entrepreneurs, institutions, organizations and society with material and relevant activities related to innovation, among which we find:

- Research and knowledge generation on Innovation.
- Organizing outreach Innovation.
- Design of Innovation Policy for Institutions.
- Training for Innovation.
- Innovation initiatives.
- Advising in The Innovation Lab, choosing an unconventional initiative and performing a quick assessment of the goodness of that opportunity by a team capable of such analysis.

The foundation council is composed by 15 members:

- *Asociación Española de Fabricantes de Azulejos y pavimentos Cerámicos (ASCER).*
- *Associació d' Empreses Innovadores Valencianes (AVANT).*
- *Asociación para el cuidado de la calidad de vida (CUIDA).*
- *Grupo para el Desarrollo y la Innovación (INDEHOLD II).*
- *Universidad Miguel Hernández de Elche.*
- *Universidad Politécnica de Cartagena.*
- *Universidad Politécnica de Valencia.*
- *Rural caja.*
- *Banco Santander, S.A.*
- *Productos Editoriales Periódicos, S.A. (Economía 3).*
- *Colegios Oficiales de Ingenieros Agrónomos de Levante.*
- *Colegio Oficial de Arquitectos de la Comunidad Valenciana.*
- *Colegio Oficial de Ingenieros Técnicos Industriales de Valencia.*
- *Consejo de Colegios Oficiales Aparejadores y Arquitectos Técnicos de la Comunidad Valenciana.*
- *Fundación Instituto Valenciano de Tecnología (INVATE).*

The foundation also spreads innovation knowledge through ideas, published books, its specialized library, the internet site, seminars, conferences, journeys, etc...

1.2. Objectives

- Chose the appropriate Linux distribution to accommodate the digital repository.
- Administer the security requirements for the Linux server.
- Configure the required software for the installation of the chosen digital repository.

- Deployment of the chosen tool.
- Customize the user interface to facilitate the integration with the existent foundation's internet site.
- Customize the operational aspects of the software by editing the source code when there is no possibility to do it through the settings.
- Publish the results of this project to help to the free software community.

1.3. Dissertation structure

The dissertation follows the next structure:

- **Chapter 2: Analysis**

This section describes the process conducted to choose the software for the digital repository and the Linux distribution to be installed in the server. As well remarks the integration needs with the current web site.

- **Chapter 3: Installation and configuration**

This chapter lists the steps carried out to install the Linux server, to install the required software and to configure and customize each piece of the solution.

- **Chapter 4: Conclusions**

This chapter describes a short reflexion about the conclusions reached by the author during the fulfilment of this Project. Likewise this section lists the technologies used in this work.

- **Chapter 5: Bibliography**

List of the references to the documents, books or web pages consulted by the author.

2. Analysis

2. Analysis

2.1. Choosing Open Source digital library software

The aim of this process is to obtain a detailed specification of the information system, so that information meets the needs of users and provides the basis for the subsequent design of the system.

A previous project [Cuellar, 2011] carried out by an Information Science professional specialized in libraries, separately analyzed different software related to digital libraries: EPrints, DSpace, Fedora and Zentity, with the aim of comparing these software packages and determine their strengths and weaknesses when it comes to the election.

The requirements marked by the organization determine that the system must have the characteristics are of a digital library and a document management system, but shall prevail the document management system over the library functions.

The list of requirements of the organization can be resumed as follows:

- Must be developed using free software to reduce costs and contribute to the expansion of free software.
- The system should have features that facilitate the customization of the user interface. Will be appreciated the improvements offered to the user. Must have the ability to restrict access to system information to external users.
- Must have a full-text search Engine.
- Loan options.
- You should be able to register documents using a web interface with possibility of attaching files.
- Possibility to query the database from the internet and retrieve documents.
- The system should be able to retrieve the documents by category.
- Will be assessed the support of the community of users and the number of software updates.
- It will be appreciated the number of organizations that have implemented the system (Degree of adoption)
- Must have features that facilitate the customization of search options.
- Batch Import Capacity: Requires that the software has selected the option to do a mass import of documents, especially at the beginning to reduce implementation time.

Having assessed all analyzed characteristics and completion tests performed with different web pages each, can be said that the most appropriate software for the Foundation "Globalidad y Microeconomía" is DSpace. This is an information system with digital repository architecture that captures, stores, sorts, preserves, and distributes digital research material in order to ensure, preserve and distribute all intellectual production.

The DSpace software fulfils most of the initial requirements as described in the following list:

- You can customize the user interface very easily have powerful tools to configure the user interface using XML files in case XMLUI based interface. And you can customize the search fields through configuration files and metadata can be customized through the web interface. has automatic generation of thumbnails (small images to preview your document) of uploaded documents.
- At the same time you can also customize the search options, to list and search for communities.
- Ability to restrict access to system information to external users via a username and password.
- Possesses extraction tools text automatically imported documents and allows searches on it using a search engine implemented internally.
- Loan options not considered but are enlisted some metadata fields with loan information and user.
- One of the requirements was that he could enlist documents using web interface and possibility of including files. It has also proven to be very simple and fast.
- You can check the web interface to query from the internet without any problems and without being inside the Polytechnic University of Valencia.
- DSpace has the ability to retrieve documents by category. Since it has advanced search and there you can do the categories of most interest to the user.
- Updates from DSpace was born in 2002 the organization has published a major version (increase functionality almost a year) and several minor versions (correct bugs and add functionality less important) every few months.
- DSpace has over 80 developers worldwide to contribute code and 15 committees working together to develop new community updates sent by The community has strength when software development and project continuity commitment.
- DSpace excels by far in the number of organizations that have implemented over the other.
- Is developed in an open source platform so that any organization can use it without paying license taxes.

It has an indexing engine and is able to search into metadata and full text (optional).

DSpace is a project of the libraries at MIT (Massachusetts Institute of Technology) Hewlett-Packard Co. His initial goal was to create a scalable and sustainable, capable of hosting more than 100,000 digital content units produced each year by the MIT faculty and researchers: articles, reports, communications, also databases, computer programs, video recordings, presentations used in class, etc.

This software has been developed using existing standards and standards allowing you to easily integrate with other information systems. Standards such as OAI-PMH, OAI-ORE, SWORD, WebDAV, OpenSearch, OpenURL, RSS, ATOM

When local authentication mechanisms, using plugins for most authentication methods of the university, including: LDAP (LDAP and hierarchical), Shibboleth, X.509, based on IP. Additionally, DSpace has its own internal authentication method.

It is available in more than twenty languages.

Its rate is high and regular updating and correcting programming errors that arise, rather quickly.

Create permanent URLs stored materials. Allows backing up files automatically.

It currently has more than 900 organizations using the software. In 2007, a census of institutional repositories in the United States and found that CLIR DSpace Repository was preferred by 446 participants in the survey. On the website you can read DSpace long list and a conceptual map of all registered users who have implemented the DSpace, its most common use is in academic and research libraries.

You can manage and preserve any document format (DOC, PPT, XLS, ODT, PDF, Word, JPEG, MPEG, TIFF files ...).

The end user interface is user friendly and supports search and view documents. These can be opened in a Web browser or by means of conventional software.

PostgreSQL database and Oracle are the database engines supported which are widely used in information systems and have demonstrated their robustness and reliability.

2.2. Choosing a Linux distribution

During the analysis phase, a previous study of the needs of the organization has been carried out. As well the chief and workers of the "Fundación Globalidad y Microeconomía" have been interviewed to gather the requirements of the Foundation.

The premises that I'm going to take into account, according to the organization requirements are the following:

- Use as a server
- Freeware
- Supported by the community
- Frequent releases
- Commitment with security updates
- Long time support
- Stability
- Existence of prebuilt and tested software packages
- 64bit architecture support

After excluding a great number of linux distributions that doesn't fulfill the initial requirements, the ones to be evaluated are the following:

Fedora
Ubuntu
Mandriva
OpenSuSE
Debian

All those distributions achieve to some extent the requirements. The ones to be evaluated are the versions oriented to act as a server if available. Between all those Ubuntu LTS shines in long time support, which is one of the most appreciated characteristics, because the less maintenance effort, the better option.

Ubuntu also excels in community support and this distribution has superb commitment with stability and security updates so in the end this is the chosen distribution.

2.3. Integration needs with the current web platform

The digital repository has to integrate with the current web platform by connecting both web pages smoothly so that the users not notice the swap during navigation.

This is going to be achieved by coping the existing HTML design to the DSpace template, and keeping the DSpace content inside of a frame the same way the main web page of the Foundation does.

Screenshots of the main page and the DSpace page can be found in the section "*3.4. Results*".

3. Installation and Configuration

3. Installation and Configuration

3.1. Linux distribution setup

The Linux distribution chosen in the section "2.2. *Choosing a Linux distribution*" has been Ubuntu Server LTS. At the time of the development of this project, the last version was Ubuntu Server 12.04 LTS.

Steps to be carried out:

1. Download an Ubuntu Server 12.04 LTS iso image, burn it and follow the steps.
2. Change keyboard configuration if necessary:

```
# sudo dpkg-reconfigure console-setup
```

3. Update packages to last version:

```
# sudo aptitude upgrade
```

To move to the last release of Ubuntu (only if a newer version has been released since having installed the operating system)

```
# sudo do-release-upgrade
```

4. If you want to install the graphic user interface:

```
# sudo aptitude install ubuntu-desktop
```

to start the graphic user interface:

```
# startx
```

5. Install build-essential to be able of compile programs:

```
# sudo aptitude install build-essential
```

3.2. Software prerequisites setup

3.2.1. Java

An optional step that can be done after installation would be to switch to the Sun/Oracle Java JDK. The tasksel task to install Tomcat installs the default OpenJDK which is a viable form of Java, however the official recommendation of DSpace is to use the Sun/Oracle Java JDK which offers better performance and other proprietary enhancements.

Enable the Canonical Partners repository.

The Sun Java is available in the partners repository which makes for an easy installation. From the GUI this can be changed by going to Software Sources.

```
# sudo vi /etc/apt/sources.list
```

Uncomment the line:

```
deb http://archive.canonical.com/ubuntu maverick partner
```

Update the catalog of packages:

```
# sudo apt-get update
```

Install Sun Java

```
# sudo apt-get install sun-java6-jdk sun-java6-plugin
```

Change the in-use Java to Sun Java, as opposed to OpenJDK

First we list the available jdk's installed on the system, then we set the sun java to be the new default.

```
# sudo update-java-alternatives -l
# java-6-openjdk 1061 /usr/lib/jvm/java-6-openjdk
# java-6-sun 63 /usr/lib/jvm/java-6-sun
# sudo update-java-alternatives -s java-6-sun
```


3.2.2. Apache Tomcat

Download the latest version of tomcat. For example *tomcat.7.tar.gz*

```
# tar xvzf apache-tomcat-7.0.0.tar.gz
# sudo mv apache-tomcat-7.0.0/ /usr/share/tomcat7
# sudo useradd -g tomcat -d /usr/share/tomcat7/ tomcat
# sudo usermod -G www-data tomcat
# sudo chown -R tomcat:tomcat /usr/share/tomcat7
```

Configure permissions so webapps folder to be a socket:

```
# sudo chmod -R 2755 /dspace/webapps
```

If the previous command fails is possible you have to do this:

```
# sudo chmod 0440 /etc/sudoers
```

Edit the *catalina.sh* script:

```
# sudo gedit /usr/share/tomcat7/bin/catalina.sh
```

Insert the *JAVA_HOME* and *JRE_HOME* after the first line, so the file is as follows:

```
#!/bin/sh
JAVA_HOME="/usr/lib/jvm/java-6-sun"
JRE_HOME="/usr/lib/jvm/java-6-sun/jre"
TOMCAT_HOME="/usr/share/tomcat7"
TOMCAT_USER="dspace"
CATALINA_HOME="/usr/share/tomcat7"
CATALINA_BASE="/usr/share/tomcat7"
# Licensed to the Apache Software Foundation (ASF)...
#...
```

Edit the *tomcat-users.xml* script:

```
# sudo gedit /usr/share/tomcat7/conf/tomcat-users.xml
```

```
<?xml version='1.0' encoding='utf-8'?>

<tomcat-users>
  <role rolename="tomcat"/>
  <role rolename="manager-gui"/>
  <role rolename="manager-script"/>
  <role rolename="manager"/>
  <role rolename="admin-gui"/>
  <role rolename="admin-script"/>
```

```
<role rolename="admin"/>
<user username="exampleuser" password="examplepassword" roles="manager-
gui"/>
<user username="dspace" password="examplepassword"
roles="dspace,manager"/>
</tomcat-users>
```

Start tomcat server:

```
# sudo /usr/share/tomcat7/bin/startup.sh
```

We obtain the following in the console:

```
Using CATALINA_BASE:   /usr/share/tomcat7
Using CATALINA_HOME:   /usr/share/tomcat7
Using JRE_HOME:        /usr/lib/jvm/java-6-sun/jre
Using
CLASSPATH:              /usr/share/tomcat7/bin/bootstrap.jar:/usr/share/tomca
t7/bin/tomcat-juli.jar
```

Verify that `JRE_HOME` is the one that we defined.

After having run tomcat we check that the server is working:

```
# netstat -tapn
```

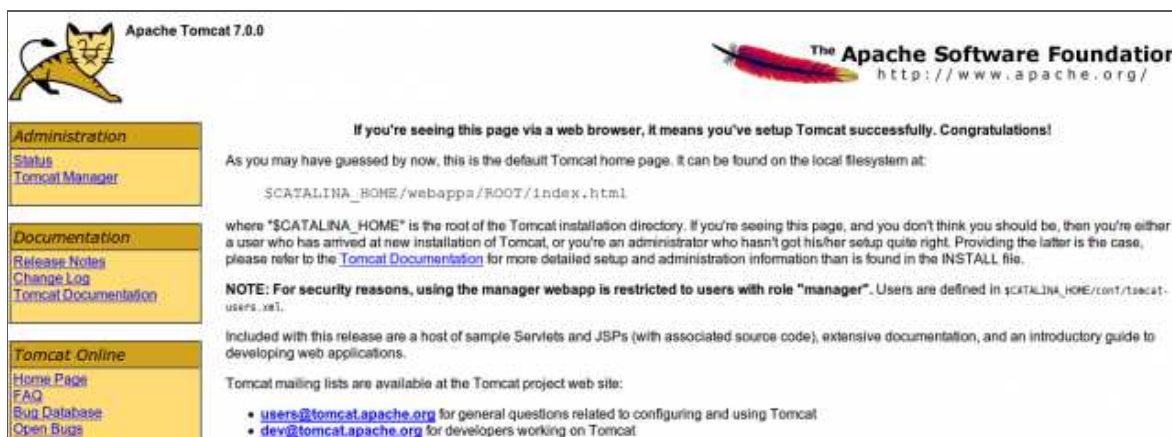
You should see something like this

```
tcp6 0 0 :::8080 :::* LISTEN 2848/java
```

After that open a web browser and type the following URL:

```
http://127.0.0.1:8080/6
```

A page like the next one should appear.



Apache Tomcat 7.0.0

The Apache Software Foundation
http://www.apache.org/

If you're seeing this page via a web browser, it means you've setup Tomcat successfully. Congratulations!

As you may have guessed by now, this is the default Tomcat home page. It can be found on the local filesystem at:

```
$CATALINA_HOME/webapps/ROOT/index.html
```

where `$CATALINA_HOME` is the root of the Tomcat installation directory. If you're seeing this page, and you don't think you should be, then you're either a user who has arrived at new installation of Tomcat, or you're an administrator who hasn't got his/her setup quite right. Providing the latter is the case, please refer to the [Tomcat Documentation](#) for more detailed setup and administration information than is found in the `INSTALL` file.

NOTE: For security reasons, using the manager webapp is restricted to users with role "manager". Users are defined in `$CATALINA_HOME/conf/tomcat-users.xml`.

Included with this release are a host of sample Servlets and JSPs (with associated source code), extensive documentation, and an introductory guide to developing web applications.

Tomcat mailing lists are available at the Tomcat project web site:

- users@tomcat.apache.org for general questions related to configuring and using Tomcat
- dev@tomcat.apache.org for developers working on Tomcat

Administration

- Status
- Tomcat Manager

Documentation

- Release Notes
- Change Log
- Tomcat Documentation

Tomcat Online

- Home Page
- FAQ
- Bug Database
- Open Bugs

Useful commands:

Init the server:

```
# sudo /usr/share/tomcat7/bin/startup.sh
```

Stop the server:

```
# sudo /usr/share/tomcat7/bin/shutdown.sh
```

Configure tomcat automatic start:

To make tomcat automatically start when we boot up the computer, you can add a script to make it auto-start and shutdown.

```
# sudo gedit /etc/init.d/tomcat7
```

Now paste in the following:

```
#!/bin/sh
#
# Tomcat auto-start
#
case $1 in
start)
sh /usr/share/tomcat7/bin/startup.sh
;;
stop)
sh /usr/share/tomcat7/bin/shutdown.sh
;;
restart)
sh /usr/share/tomcat7/bin/shutdown.sh
sh /usr/share/tomcat7/bin/startup.sh
;;
esac
exit 0
```

You'll need to make the script executable by running the chmod command:

```
# sudo chmod 755 /etc/init.d/tomcat7
```

The last step is actually linking this script to the startup folders with a symbolic link. Execute these two commands and we should be on our way.

We create a soft link of the script found in */etc/init.d* pointing to folders that are executed when entering / leaving a runlevel */etc/rcN.d*, where N is

the runlevel. When we want to stop the service in runlevel N, a softlink beginning with K is created. If we want it to get up, it will begin with S. The S and K will be followed by a number, indicating the order in which they will rise / stop services. You can create the appropriate links.

The rc2.d softlink in the startup script of the application will start only with runlevel 2

```
# sudo ln -s /etc/init.d/tomcat7 /etc/rc1.d/K99tomcat7
# sudo ln -s /etc/init.d/tomcat7 /etc/rc2.d/S99tomcat7
# sudo /etc/init.d/tomcat7 restart
```

To check that tomcat is running in levels 1 and 2 as we have configured, you can run the tool:

```
# sysv-rc-conf
```

If the tool is not present in your system use the following command to install it.

```
# sudo apt-get install sysv-rc-conf
```

This is a good opportunity to remove unwanted services. If you want to delete any service entry you can do it with the following command:

```
# sudo update-rc.d -f tomcat remove
```

Tomcat configuration:

Append the following lines to `/etc/default/tomcat7` to set the preferences necessary for DSpace:

```
TOMCAT7_USER=dspace
TOMCAT7_SECURITY=no
```

3.2.3. Maven

Maven allows a project to build using its project object model (POM) and a set of plugins that are shared by all projects using Maven, providing a uniform build system.

Install the package for the maven build utility:

```
# sudo aptitude install maven2
```

3.2.4. Postgresql

```
# sudo aptitude install postgresql
```

Postgresql configuration:

The first task is to reset the "*postgres*" user password, who has administrator permissions for the DB server. Open a terminal and type:

```
# sudo su postgres -c psql template1
template1=# ALTER USER postgres WITH PASSWORD 'password';
template1=# \q
```

After that change the "*postgres*" user password. This will be the user suited to run the service.

```
# sudo passwd -d postgres
# sudo su postgres -c passwd
```

To reset the postgres service type:

```
# sudo /etc/init.d/postgresql-8.4 restart
```

After resetting the system check postgres be up:

```
# netstat -tapn
```

Something like the following should appear:

```
tcp6 0 0 :::5433 :::* LISTEN
```

Setup pgadmin (optional)

```
# sudo aptitude install pgadmin
```

To open pgadmin3 from Ubuntu menu go to:

Applications/Programming/Pgadmin III

Connect to Postgresql indicating the name, server and port:

```
Name: MyConnectionWithPostgresql
Server: localhost
Port: 5433
```

Install libpg-java package for the Postgres JDBC driver:

```
# sudo aptitude install libpg-java
```

3.2.5. Xpdf

XPDF is a suite of tools for Portable Document Format (PDF) files. To enable support for Foreign Languages including Chinese to view PDF Files you need to install the language support. Here are the instructions.

The tools include xpdf, a PDF viewer (in the package xpdf-reader), and PDF converters (including to/from PostScript) (in the package xpdf-utils).

To install Xpdf type: aptitude install xpdf

```
# sudo aptitude install xpdf
```

3.2.6. Postfix

Postfix is a mail server widely used in unix and Linux servers. Postfix is required to send automatic e-mails to users of DSpace.

To install Postfix in Ubuntu type:

```
# sudo aptitude install postfix
```

In the configuration Windows that will appear select:

- Only local e-mail
- Server name: myservername.com

If the assistant does not appear and Postfix is installed with the default options, the configuration can be changed later with the following command:

```
# dpkg-reconfigure postfix-config
```

The Postfix configuration file can be tweaked to be able to send mail through gmail.com, edit the file:

```
# sudo gedit /etc/postfix/main.cf
```

Add:

```
# relay
relayhost = [smtp.gmail.com]
smtp_use_tls = yes
smtp_sasl_auth_enable = yes
smtp_sasl_password_maps = hash:/etc/postfix/sasl/sasl_passwd
smtp_sasl_security_options = noanonymous
smtp_sasl_tls_security_options = noanonymous
smtp_generic_maps = hash:/etc/postfix/generic
```

Change the file */etc/postfix/sasl/sasl_passwd*:

```
[smtp.gmail.com] mygmailuser:mypassword
```

Set the required permissions:

```
# sudo chmod 600 /etc/postfix/sasl/sasl_passwd
```

Update search tables:

```
# postmap /etc/postfix/sasl/sasl_passwd
```

Add the following line to */etc/postfix/main.cf*

```
# smtp_generic_maps = hash:/etc/postfix/generic
```

Add the following line to */etc/postfix/generic*

```
# sudo gedit /etc/postfix/generic
```

Set

```
root@mydomain.com myemail@gmail.com
```

Reset the postfix server:

```
# sudo /etc/init.d/postfix restart
```

Send a test e-mail from the command line to check the configuration:

```
# telnet 127.0.0.1 25
HELO 127.0.0.1
MAIL FROM: <examplemail@gmail.com>
RCPT TO:<examplemail@gmail.com>
DATA
Subject: este es mi asunto
Este es el cuerpo
<INTRO>
.
<INTRO>
QUIT
```

3.2.7. Security configuration

Create the UNIX 'dspace' user, update the password, create the directory in which you will install dSpace, and ensure that the UNIX 'dspace' user has write privileges on that directory:

```
# sudo useradd -m dspace
# sudo passwd dspace
# sudo mkdir /dspace
# sudo chown dspace /dspace
```

Create the PostgreSQL 'dspace' user and the 'dspace' database. Using sudo as the Unix '*postgres*' user, authorize the 'dspace' user. You will need to select a password and specify 'n' in the "create new roles" prompt. Then, as the '*dspace*' user, create the database.

```
# sudo -u postgres createuser -U postgres -d -A -P dspace
# sudo -u dspace createdb -U dspace -E UNICODE dspace
```

Change ownership of the tomcat directories to the dspace user:

```
# sudo chown -R dspace /var/cache/tomcat6
# sudo chown -R dspace /var/lib/tomcat6
# sudo chown -R dspace /var/log/tomcat6
# sudo chown -R dspace /etc/tomcat6
```


3.2.8. Configure Unattended Updates:

The `unattended-upgrades` package can be used to automatically install updated packages, and can be configured to update all packages or just install security updates. First, install the package by entering the following in a terminal:

```
# sudo apt-get install unattended-upgrades
```

To configure `unattended-upgrades`, edit `/etc/apt/apt.conf.d/50unattended-upgrades` and adjust the following to fit your needs:

```
Unattended-Upgrade::Allowed-Origins {  
    "Ubuntu lucid-security";  
    // "Ubuntu lucid-updates";  
};
```

Certain packages can also be *blacklisted* and therefore will not be automatically updated. To blacklist a package, add it to the list:

```
Unattended-Upgrade::Package-Blacklist {  
    // "vim";  
    // "libc6";  
    // "libc6-dev";  
    // "libc6-i686";  
};
```

The double `"/"` serve as comments, so whatever follows `"/"` will not be evaluated.

To enable automatic updates, edit `/etc/apt/apt.conf.d/10periodic` and set the appropriate **apt** configuration options:

```
APT::Periodic::Update-Package-Lists "1";  
APT::Periodic::Download-Upgradeable-Packages "1";  
APT::Periodic::AutocleanInterval "7";  
APT::Periodic::Unattended-Upgrade "1";  
APT::Periodic::RandomSleep "6000";
```

The above configuration updates the package list, downloads, and installs available upgrades every day. The local download archive is cleaned every week.

You can read more about **apt** Periodic configuration options in the `/etc/cron.daily/apt script header`.

The results of **unattended-upgrades** will be logged to
/var/log/unattended-upgrades

Source: Automatic Updates,
<https://help.ubuntu.com/10.04/serverguide/C/automatic-updates.html>

3.3. DSpace installation

Download and Install DSpace

Create the [dspace] directory. The [dspace] directory is where the running dspace code will reside.

```
# sudo mkdir /dspace
```

Download the Source Release

The source release allows you to customize every aspect of DSpace. This step downloads the compressed archive from SourceForge, and unpacks it in your current directory. The dspace-1.x.x-src-release directory is typically referred to as [dspace-src].

```
# wget
  http://sourceforge.net/projects/dspace/files/DSpace%20Stable/1.7.2/dspace-1.7.2-src-release.tar.bz2
# tar -xvjf dspace-1.7.2-src-release.tar.bz2
```

Compile and Build DSpace

The source release that has been obtained is human readable source code, and must be compiled to machine code for the server to run it. "mvn package" compiles the source code, and "ant" will do all the work necessary to initialize the database with the DSpace schema, and copy all of the compiled machine code to a location where the web server can serve it.

```
cd into
  dspace-1.7.2-src-release
```

```
# sudo mvn -U package
```

Your first run of Maven downloads a lot of dependencies. Be prepared for several minutes of download activity, followed by several minutes of build activity. Note that if you accidentally run Maven using gcj instead of Sun Java, and it fails, you should remove the #/.m2 directory (`rm -rf /home/dspace/.m2` before proceeding with the correct java).

```
cd into
  [dspace-src]/dspace/target/dspace-[version]-build.dir/
```

Initialize the database and install the software:

```
# sudo ant fresh_install
```

Remark: If the build fails two things are necessary to do before a new attempt: 1. Remove the remains of the failed build (*execute ant clean*), 2. Remove the dspace tables from the database by dropping and recreating it (execute *dropdb -U dspace dspace; createdb -U dspace -E UNICODE dspace*). Of course the reason for the failing must be cured too.

Create the initial DSpace administrator:

```
# sudo /dspace/bin/dspace create-administrator
```

Start Tomcat:

```
# sudo /etc/init.d/tomcat6 start
```

Open the new URL in your Web browser:

http://hostname:8080/jspui or *http://hostname:8080/xmlui*

(Adjust for your hostname and port number, accordingly maybe 8180)

Source: Installing DSpace 1.7.2 on Ubuntu

<https://wiki.duraspace.org/display/DSPACE/Installing+DSpace+1.7+on+Ubuntu#InstallingDSpace1.7onUbuntu-DownloadandInstallDSpace>

3.3.1. Configure Folder Permissions

At the time when DSpace is up and running we can step to its customization.

Set the appropriate permissions to dspace user to be able to access the main folder:

```
# sudo chmod 775 -R /dspace/
```

3.3.2. Hide Community or Collection from list

As we want that documents in our internal repositories to be hidden from general public (anonymous user group), we'll have to make some changes to some sources for this to be accomplished. We'll use a piece of code from Lucas van Schaik written for the Leiden University repository.

Step 1: Change file

`/dspace-api/src/main/java/org/dspace/content /Community.java`

Near line 276 (v1.3.2) / line 303 (v1.6) change the code in function `findAllTop` from (in v1.6 this looks a bit different):

```
// First check the cache
Community fromCache = (Community) context.fromCache(
    Community.class, row.getIntColumn("community_id"));
if (fromCache != null)
\{
    topCommunities.add(fromCache);
\}
else
\{
    topCommunities.add(new Community(context, row));
\}
```

to this:

```
// First check the cache
Community aCommunity = (Community) context.fromCache(
    Community.class, row.getIntColumn("community_id"));
if (aCommunity == null) \{
    aCommunity = new Community(context, row);
\}
if(AuthorizeManager.authorizeActionBoolean(context,aCommunity,Constants.R
EAD)) \{
    top.add(aCommunity);
\}
```

Near line 455 (v1.3.2) / 467 (v1.4.2) / 628 (v1.6) change the code in function `getCollections` from (in v1.6 this looks a bit different):

```
// First check the cache
Collection fromCache = (Collection) ourContext.fromCache(
    Collection.class, row.getIntColumn("collection_id"));
if (fromCache != null)
\{
    collections.add(fromCache);
\}
else
\{
    collections.add(new Collection(ourContext, row));
\}
```

to this:

```
// First check the cache
Collection aCollection = (Collection) ourContext.fromCache(
    Collection.class, row.getIntColumn("collection_id"));
if (aCollection == null) \{
    aCollection = new Collection(ourContext, row);
\}
if(AuthorizeManager.authorizeActionBoolean(ourContext,aCollection,Constan
ts.READ)) \{
    collections.add(aCollection);
\}
```

Near line 504 (v1.3.2) / 515 (v1.4.2) / 687 (v1.6) change the code in function `getSubcommunities` from:

```
// First check the cache
Community fromCache = (Community) ourContext.fromCache(
    Community.class, row.getIntColumn("community_id"));
if (fromCache != null)
\{
    subcommunities.add(fromCache);
\}
else
\{
    subcommunities.add(new Community(ourContext, row));
\}
```

to this:

```
// First check the cache
Community aCommunity = (Community) ourContext.fromCache(
    Community.class, row.getIntColumn("community_id"));
if (aCommunity == null) \{
    aCommunity = new Community(ourContext, row);
\}

if(AuthorizeManager.authorizeActionBoolean(ourContext, aCommunity, Constant
s.READ)) \{
    subcommunities.add(aCommunity);
\}
```

Step 2: Change community and/or collection authorizations

- log on as an administrator
- navigate to the community or collection you want to hide
- click on the edit button near "Community's Authorizations:" or "Collection's Authorizations:"
- change or delete the anonymous READ policy

Step 3: build / install / restart

Do the stuff you normally do when deploying a new version of DSpace as described in the "Build DSpace" part of the "3.9. Version Update" section.

Step 4: test it

- make sure you are not logged on
- go to your Communities & Collections list page (*/community-list*)
- your community or collection you just changed, should not be visible
- log on as someone who should be able to see your community or collection (administrators can always see it)
- go to your Communities & Collections list page (*/community-list*)

- your community or collection should not be visible

Source: [<https://wiki.duraspace.org/display/DSPACE/Hide+Community+or+Collection+from+list>]

3.3.3. Localization

In order to deploy a multilingual version of DSpace you have to configure two parameters in `[dspace-source]/config/dspace.cfg`:

- Set `default.locale`, e.g.

```
default.locale = es
```

- Set `webui.supported.locales`, e.g.

```
webui.supported.locales = es, en
```

Enable Spanish language

In `[dspace]/config/dspace.cfg` set:

```
default.locale = es
webui.supported.locales = es
xmlui.supported.locales = es
```

In the following path:
`[dspace]/webapps/xmlui/i18n`

Copy the file:
`messages_es.xml`

Obtain the Spanish translation of DSpace from the following Source:
<http://dspace-sandbox.googlecode.com/svn/modules/dspace-xmlui-lang/trunk/src/main/webapp/i18n/>

Move the file `messages.xml` to `messages_en.xml`

```
# sudo mv messages.xml messages_en.xml
```

Copy `messages_es.xml` to `messages.xml`

```
# sudo mv messages-es.xml messages.xml
```

3.4. DSpace Interface Configuration

3.4.1. Set Mirage as the default search theme

The Mirage theme is an advanced faceted search interface similar to the one used by Amazon when searching for products. It shows the information contained in the database classified by facets or types. As the user selects more facets the results are more specific.

In `[dspace]/config/dspace.cfg` change:

```
<theme name="Default Reference Theme" regex=".*" path="Reference/" />
```

To this:

```
<!-- <theme name="Default Reference Theme" regex=".*" path="Reference/" /> -->
```

Change:

```
<!-- <theme name="Atmire Mirage Theme" regex=".*" path="Mirage/" /> -->
```

To this:

```
<theme name="Atmire Mirage Theme" regex=".*" path="Mirage/" />
```

In `[dspace]/config/dspace-solr-search.cfg`

Change the port in:

```
solr.search.server = http://localhost:8080/solr/search
```

To this:

```
solr.search.server = http://localhost:8180/solr/search
```

3.4.2. Customize the Mirage theme

First edit the file:

```
[dspace-source]/dspace-xmlui/dspace-xmlui-webapp/src/main/webapp/themes/Mirage/lib/xsl/core/page-structure.xml
```

Here you can change the structure of the generated xhtml.

Likewise you can customize the aspect changing the css files found in:


```
[dspace-source]/dspace-xmlui/dspace-xmlui-  
webapp/src/main/webapp/themes/Mirage/lib/css/style.css and base.css
```

Moreover you can change some design elements in the file:

```
[dspace-source]/dspace-xmlui/dspace-xmlui-  
webapp/src/main/webapp/themes/Mirage/lib/xsl/core/navigation.xsl
```

1. Rebuild the DSpace installation package by running the following command from your `[dspace-source]/dspace/` directory:

```
# sudo mvn package
```

1. Update all DSpace webapps to `[dspace]/webapps` by running the following command from your `[dspacesource]/dspace/target/dspace-[version]-build.dir` directory:

```
# sudo ant -Dconfig=[dspace]/config/dspace.cfg update
```

1. Deploy the the new webapps: (not necessary if our tomcat points to our directory: `[dspace]/webapps` to serve the pages)

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

2. Restart Tomcat

```
# sudo /etc/init.d/tomcat1 restart
```

For more information consult:

[<http://www.slideshare.net/tDonohue/making-dspace-xmlui-your-own>] where the inner working and configuration of xmlui via overlays is explained in-depth.

3.5. Searches and Results

3.5.1. Enable Discovery as the default search engine

Enable the Discovery Aspects in the XMLUI by changing the following settings in `config/xmlui.xconf`

Comment out: SearchArtifacts

Uncomment: Discovery

```

<xmlui>
  <aspects>
    <aspect name="Artifact Browser" path="resource://aspects/ArtifactBrowser/" />
    <aspect name="Browsing Artifacts" path="resource://aspects/BrowseArtifacts/" />
    <!--<aspect name="Searching Artifacts" path="resource://aspects/SearchArtifacts/"
/>-->
    <aspect name="Administration" path="resource://aspects/Administrative/" />
    <aspect name="E-Person" path="resource://aspects/EPerson/" />
    <aspect name="Submission and Workflow" path="resource://aspects/Submission/" />
    <aspect name="Statistics" path="resource://aspects/Statistics/" />

    <!--
      To enable Discovery, uncomment this Aspect that will enable it
      within your existing XMLUI
      Also make sure to comment the SearchArtifacts aspect
      as leaving it on together with discovery will cause UI overlap issues-->
    <aspect name="Discovery" path="resource://aspects/Discovery/" />

    <!--
      This aspect tests the various possible DRI features,
      it helps a theme developer create themes
    -->
    <!-- <aspect name="XML Tests" path="resource://aspects/XMLTest/" /> -->
  </aspects>

```

Enable the Discovery Indexing Consumer that will update Discovery Indexes on changes to content in XMLUI, JSPUI, SWORD, and LNI in `config/dspace.cfg`

Add discovery to the list of `event.dispatcher.default.consumers`

Change `recent.submissions.count` to zero

```

#### Event System Configuration ####

# default synchronous dispatcher (same behavior as traditional DSpace)

event.dispatcher.default.class = org.dspace.event.BasicDispatcher
event.dispatcher.default.consumers = search, browse, discovery, eperson, harvester

#Put the recent submissions count to 0 so that discovery can use it's recent submissions,
#not doing this when discovery is enabled will cause UI overlap issues
#How many recent submissions should be displayed at any one time

recent.submissions.count = 0

```

Check that the port is correct for `solr.search.server` in `[dspace]/config/dspace-solr-search.cfg`

If all of your traffic runs over port 80, then you need to remove the port from the URL

```

##### Search Indexing #####
solr.search.server = http://localhost/solr/search

```

From the command line, navigate to the `dspace` directory and run the command below to index the content of your DSpace instance into Discovery.

```
# ./bin/dspace update-discovery-index
```

In `[dspace]/config/xmlui.xconf`

Change

```
<aspect name="Searching Artifacts"  
  path="resource://aspects/SearchArtifacts/" />
```

To this:

```
<!-- <aspect name="Searching Artifacts"  
  path="resource://aspects/SearchArtifacts/" /> -->
```

In `[dspace]/config/dspace.cfg`

Change

```
event.dispatcher.default.consumers = search, browse, eperson, harvester
```

To this:

```
event.dispatcher.default.consumers = search, browse, discovery, eperson,  
  harvester
```

Change:

```
recent.submissions.count = 5
```

To this:

```
recent.submissions.count = 0
```

Check that the port is correct for `solr.search.server` in `config/dspace-solr-search.cfg`

If all of your traffic runs over port 80, then you need to remove the port from the URL

```
##### Search Indexing #####  
solr.search.server = http://localhost/solr/search
```

From the command line, navigate to the `dspace` directory and run the command below to index the content of your DSpace instance into Discovery.

```
# ./bin/dspace update-discovery-index
```

Source:

<https://wiki.duraspace.org/display/DSPACE/Discovery+Configuration>

3.5.2. Select the indexed fields for the search

In `/dspace/config/dspace.cfg`

```
search.index.1 = author:dc.contributor.*
search.index.2 = author:dc.creator.*
search.index.3 = title:dc.title.*
search.index.4 = keyword:dc.subject.*
search.index.5 = abstract:dc.description.abstract
search.index.6 = author:dc.description.statementofresponsibility
search.index.7 = series:dc.relation.ispartofseries
search.index.8 = abstract:dc.description.tableofcontents
search.index.9 = mime:dc.format.mimetype
search.index.10 = sponsor:dc.description.sponsorship
search.index.11 = identifier:dc.identifier.*
search.index.12 = language:dc.language.iso
```

After changing the configuration run `[dspace]/bin/index-init` to regenerate the indexes.

3.5.3. Hide metadata listed in results from the user

In `[dspace]/config/dspace.cfg` is needed to establish the following values:

```
metadata.hide.dc.description.provenance = true
metadata.hide.dc.prestamo.email = true
metadata.hide.dc.prestamo.fechadevolucion = true
metadata.hide.dc.prestamo.fechainicio = true
metadata.hide.dc.prestamo.nombre = true
metadata.hide.dc.prestamo.telefono = true
```

3.5.4. Define search indexes

To select the fields that the user will use when searching do the following.

In `[dspace]/config/dspace.cfg` is needed to the following values:

```
webui.browse.index.1 = dateissued:item:dateissued
webui.browse.index.2 = author:metadata:dc.contributor.*,dc.creator:text
webui.browse.index.3 = title:item:title
webui.browse.index.4 = subject:metadata:dc.subject.*:text
```

3.5.5. Update thumbnails and full text sources

In the command line exec the following command:

```
# sudo /dspace/bin/dspace filter-media
```

3.5.6. Enable thumbnails in the search results

In the file `[dspace]/config/dspace.cfg` set:

```
webui.browse.thumbnail.show = true
webui.browse.thumbnail.maxheight = <maxheight in pixels>
webui.browse.thumbnail.maxwidth = <maxwidth in pixels>

webui.item.thumbnail.show = true

xmlui.theme.mirage.item-list.emphasis = file
```

Settings for Item Preview:

```
webui.preview.enabled = true
# max dimensions of the preview image
webui.preview.maxwidth = 600
webui.preview.maxheight = 600
# the brand text
webui.preview.brand = Fundacion Globalidad y Microeconomia
# an abbreviated form of the above text, this will be used
# when the preview image cannot fit the normal text
webui.preview.brand.abbrev = FGyM
# the height of the brand
webui.preview.brand.height = 20
# font settings for the brand text
webui.preview.brand.font = SansSerif
webui.preview.brand.fontpoint = 12
#webui.preview.dc = rights
```

3.5.7. XPDF MediaFilter Configuration

This filter extracts better the pdf text and is able to create thumbnails for pdf, but not enabled by default to avoid complicating the installation.

Install xpdf. Package to manage pdfs:

```
# sudo aptitude install xpdf
```

Install curl if it is not already installed. This package is useful to download from command line using http:

```
# sudo aptitude install curl
```

Download Java Advanced Imaging Image I/O Tools:

```
# curl -O http://download.java.net/media/jai-imageio/builds/release/1.1/jai_imageio-1_1-lib-linux-i586.tar.gz
```

Uncompress:

```
# tar xzf jai_imageio-1_1-lib-linux-i586.tar.gz
```

Install it in the Maven repository. From the same folder where's uncompressed the tar file, execute:

```
# sudo mvn install:install-file -Dfile=jai_imageio-1_1/lib/jai_imageio.jar -DgroupId=com.sun.media -DartifactId=jai_imageio -Dversion=1.0_01 -Dpackaging=jar -DgeneratePom=true

# curl -O https://maven.nuxeo.org/nexus/content/repositories/public/javax/media/jai_core/1.1.2_01/jai_core-1.1.2_01.jar
```

Install it in the Maven repository. From the same folder where's uncompressed the tar file, execute:

```
# sudo mvn install:install-file -Dfile=jai_core/jai_core-1.1.2_01.jar -DgroupId=javax.media -DartifactId=jai_core -Dversion=1.1.2_01 -Dpackaging=jar -DgeneratePom=true
```

Edit the file `[dspace]/config/dspace.cfg` and add or replace the following lines:

```
xpdf.path.pdftotext = /usr/bin/pdftotext
xpdf.path.pdfppm = /usr/bin/pdfppm
xpdf.path.pdfinfo = /usr/bin/pdfinfo

filter.plugins = \
    PDF Text Extractor, \
    PDF Thumbnail, \
    HTML Text Extractor, \
    Word Text Extractor, \
    JPEG Thumbnail

plugin.named.org.dspace.app.mediafilter.FormatFilter = \
    org.dspace.app.mediafilter.XPDF2Text = PDF Text Extractor, \
    org.dspace.app.mediafilter.XPDF2Thumbnail = PDF Thumbnail, \
    org.dspace.app.mediafilter.HTMLFilter = HTML Text Extractor, \
    org.dspace.app.mediafilter.WordFilter = Word Text Extractor, \
    org.dspace.app.mediafilter.JPEGFilter = JPEG Thumbnail, \
    org.dspace.app.mediafilter.BranchedPreviewJPEGFilter = Branded Preview JPEG

filter.org.dspace.app.mediafilter.XPDF2Thumbnail.inputFormats = Adobe PDF
```

```
PDFfilter.org.dspace.app.mediafilter.XPDF2Text.inputFormats = Adobe PDF
```

Build and install following the "*Build DSpace*" of the "*3.9. Version Update*" section.

cd into the `[dspace-src]/dspace` directory.

```
# sudo mvn -Ppdf-mediafilter-support package
```

```
cd into [dspace-src]/dspace/target/dspace-[version]-build.dir/
```

```
# sudo ant -Dconfig=[dspace]/config/dspace.cfg update
```

3.6. Metadata Configuration

3.6.1. Change item metadata style

Files to edit:

```
[dspace]/config/dspace.cfg
```

```
[dspace-source]/config/language-packs/Messages.properties
```

1. To change the metadata shown, change the order of the Dublin Core after the tag: `webui.itemdisplay.default` (in `Dspace.cfg`), or add new elements:

```
webui.itemdisplay.default = dc.title, dc.title.alternative,  
dc.contributor.*, dc.subject, dc.date.issued(date), dc.publisher
```

2. To change the name of the field shown for an specific metadata, find its key name (should begin with `'metadata'`) in `message.properties` file and change it by:

```
metadata.dc.title.alternative = Title (French)
```

3. Follow the steps listed in the part "*Buld DSpace*" of the "*3.9. Version Update*" section.

4. Reindex DSpace

This process rebuilds DSpace search indexes.

You have to execute it after a task which modifies the content of these indexes (eg manual changes of the metadata, withdrawing items) In addition, is necessary running it if the search indexes are modified.

If desired you can schedule the process to reindex DSpace daily if you make many small changes for a while, without reindexing, search function of DSpace may become slow.

```
# sudo [DSpace]/bin/dspace index-all
```

3.6.2. Change item description metadata

Edit the following file:

```
[dspace-src]/dspace/dspace-xmlui/dspace-xmlui-  
webapp/src/main/webapp/themes/Mirage/lib/xsl/aspect/artifactbrowse  
r/item-view.xsl
```

To add new metadata is needed to duplicate this xml structure:

```
<!-- publisher row -->  
<xsl:when test="$clause = 3 and (dim:field[@element='publisher' and  
not(@qualifier)])">  
  <div class="simple-item-view-other">  
    <span class="bold">  
      <i18n:text>xmlui.dri2xhtml.METS-1.0.item-publisher  
</i18n:text></span>  
    <span>  
      <xsl:for-each select="dim:field[@element='publisher' and  
not(@qualifier)]">  
        <xsl:copy-of select="."/node()"/>  
        <xsl:if test="count(following-sibling::dim:field[@element='publisher'  
and not(@qualifier)]) != 0">  
          <br/>  
        </xsl:if>  
      </xsl:for-each>  
    </span>  
  </div>  
<xsl:call-template name="itemSummaryView-DIM-fields">  
  <xsl:with-param name="clause" select="($clause + 1)"/>  
  <xsl:with-param name="phase" select="$otherPhase"/>  
</xsl:call-template>  
</xsl:when>
```

In the beginning of the file we have several examples:

Title row
Author(s) row
Publisher row

Using these as a template, add the metadata that we are interested to show from our Dublin Core collection.

Keep in mind that in '\$clause = x' x has to be increasing with integers beginning with 1 in the order that you want to give to the metadata.

Likewise if the Dublin Core metadata does not have qualifier, use the form: not(@ qualifier), using publisher row column as pattern.

3.6.3. Change visibility of items through RSS, OAI and Subscriptions

Visibility in Browse and Search Indexes in *DSpace.cfg*

```
harvest.includerestricted.rss = false
harvest.includerestricted.oai = false
harvest.includerestricted.subscription = false
```

3.7. Import and Export Information

3.7.1. Export a collection

```
# sudo /dspace/bin/dspace export -t COLLECTION -i 123456789/4 -d
/home/root/exports -n 1
```

3.7.2. Import a collection

```
# sudo /dspace/bin/dspace import -a -e examplemail@gmail.com -c
123456789/4 -s /home/root/exports -m /home/root/mapfile
```

3.8. Schedule 'cron' Jobs

A couple of DSpace features require that a script is run regularly - the email subscription feature that alerts users of new items being deposited, and the new 'media filter' tool, that generates thumbnails of images and extracts the full-text of documents for indexing.

To set these up, you just need to run the following command as the *dspace* UNIX user:

```
# crontab -e
```

Then add the following lines:

```
# Send out subscription e-mails at 01:00 every day
0 1 * * * [dspace]/bin/dspace sub-daily
# Run the media filter at 02:00 every day
0 2 * * * [dspace]/bin/dspace filter-media
# Run the checksum checker at 03:00
0 3 * * * [dspace]/bin/dspace checker -lp
# Mail the results to the sysadmin at 04:00
0 4 * * * [dspace]/bin/dspace checker-emailer -c
```

Naturally you should change the frequencies to suit your environment.

PostgreSQL also benefits from regular 'vacuuming', which optimizes the indexes and clears out any deleted data.

Become the *postgres* UNIX user, run *crontab -e* and add (for example):

```
# Clean up the database nightly at 4.20am
20 4 * * * vacuumdb --analyze dspace > /dev/null 2>&1
```

In order that statistical reports are generated regularly and thus kept up to date you should set up the following cron jobs:

```
# Run stat analysis
0 1 * * * [dspace]/bin/dspace stat-general
0 1 * * * [dspace]/bin/dspace stat-monthly
0 2 * * * [dspace]/bin/dspace stat-report-general
0 2 * * * [dspace]/bin/dspace stat-report-monthly
```

Obviously, you should choose execution times which best fulfill your needs, and you should ensure that the *report* scripts run after the analysis scripts to give them time to complete (8 months of logs can take around 25 seconds to complete).

Source: Chapter 3. DSpace System Documentation: Installation
[http://www.dspace.org/1_6_0Documentation/ch03.html]

3.9. Version Update

First download the latest version of DSpace from the repository. For this install subversion:

```
# sudo aptitude install subversion
```

Download source code from the repository:

```
# sudo svn export --force
http://scm.dspace.org/svn/repo/dspace/tags/dspace-1.7.2
home/dspace/dspace-1.7.2-src/
```

Source: Building DSpace From Source

[<https://wiki.duraspace.org/display/DSPACE/Building+DSpace+From+Source>]

In case you have to upgrade your version of DSpace, it would be necessary to move the changes that have been made. To do this you should copy the changes we have made to the files to the new version. Keep in mind that some of the files that have been customized may have changed in the next version, so you will have to check the differences with some program like Beyond Compare (which graphically shows the changes between two files and allows incorporate the significant changes from one version to the other in a simple and intuitive way).

List of changed files to the "Globalidad y Microeconomía" Foundation installation of Dspace v1.7.2

[dspace-src]/dspace/config/dspace.cfg

Main DSpace configuration file.

[dspace-src]/dspace/config/dspace-solr-search.cfg

Solr configuration file.

[dspace-src]/dspace/config/news-xmlui.xml

File containing the description that appears in the main page of DSpace.

[dspace-src]/dspace/config/xmlui.xconf

XMLUI interface configuration file.

[dspace-src]/dspace/config/oaicat.properties

Change this only if OAI-PMH is going to be used to export metadata. Here is the configuration of the crosswalk plugins used to disseminate content.

[dspace-src]/dspace-api/src/main/java/org/dspace/app/mediafilter/XPDF2Thumbnail.java

This fixes a bug in the code for which the temporary thumbnail intended to be read was `outPrefix + "-000001.ppm` while actually xpdf version we have installed in our system (Version: 3.02-2ubuntu1.1) that generates it to `outPrefix + "-1.ppm."` so we have replaced the path in the code.

```
[dspace-src]/dspace-xmlui/dspace-xmlui-  
webapp/src/main/webapp/i18n/messages.xml  
[dspace-src]/dspace-xmlui/dspace-xmlui-  
webapp/src/main/webapp/i18n/messages_es.xml  
[dspace-src]/dspace-xmlui/dspace-xmlui-  
webapp/src/main/webapp/i18n/messages_?? .xml
```

This is the localization file. There is a messages file for each language configured in DSpace. As we only support Spanish, this file corresponds to the interface descriptions in Spanish.

```
[dspace-src]/dspace-xmlui/dspace-xmlui-  
webapp/src/main/webapp/themes/Mirage/images/cabecerahome.jpg
```

This is the background image of the main page.

```
[dspace-src]/dspace-xmlui/dspace-xmlui-  
webapp/src/main/webapp/themes/Mirage/images/logo.gif
```

Organization logo used in the main page.

```
[dspace-src]/dspace/dspace-xmlui/dspace-xmlui-  
webapp/src/main/webapp/themes/Mirage/lib/css/style.css
```

Styles needed by the html design.

```
[dspace-src]/dspace/dspace-xmlui/dspace-xmlui-  
webapp/src/main/webapp/themes/Mirage/lib/xsl/core/page-  
structure.xsl
```

Definition of the web in XHTML. Has been customized to integrate with the existent organization main page.

```
[dspace-source]/dspace-xmlui/dspace-xmlui-  
webapp/src/main/webapp/themes/Mirage/lib/xsl/core/navigation.xsl
```

A link to the help has been added in this file.

```
[dspace-src]/dspace/dspace-xmlui/dspace-xmlui-  
webapp/src/main/webapp/themes/Mirage/lib/xsl/aspect/artifactbrow  
se/item-view.xsl
```

Metadata description.

Build DSpace.

Run the following commands to compile DSpace:

```
# cd [dspace-source]/dspace/  
# sudo mvn -U clean -Ppdf-mediafilter-support 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 ('package'), 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 repository.

When recompiling be sure to use the option "*-Pxpdf-mediafilter-support*", otherwise when executing the filter-media application to extract text and thumbnails from PDFs you will get the following error:

```
ERROR filtering, skipping bitstream:

Item Handle: 123456789/675
Bundle Name: ORIGINAL
File Size: 490404
Checksum: 3425e8df23bf424a01f2fc409703f881 (MD5)
Asset Store: 0
javax.imageio.IIOException: Can't read input file!
javax.imageio.IIOException: Can't read input file!
  at javax.imageio.ImageIO.read(ImageIO.java:1275)
  at
  org.dspace.app.mediafilter.XPDF2Thumbnail.getDestinationStream(XPDF2Thumbnail.java:244)
  at
  org.dspace.app.mediafilter.MediaFilterManager.processBitstream(MediaFilterManager.java:737)
  at
  org.dspace.app.mediafilter.MediaFilterManager.filterBitstream(MediaFilterManager.java:561)
  at
  org.dspace.app.mediafilter.MediaFilterManager.filterItem(MediaFilterManager.java:511)
  at
  org.dspace.app.mediafilter.MediaFilterManager.applyFiltersItem(MediaFilterManager.java:479)
  at
  org.dspace.app.mediafilter.MediaFilterManager.applyFiltersAllItems(MediaFilterManager.java:414)
  at
  org.dspace.app.mediafilter.MediaFilterManager.main(MediaFilterManager.java:333)
  at sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method)
  at sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:39)
  at
  sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:25)
  at java.lang.reflect.Method.invoke(Method.java:597)
  at org.dspace.app.launcher.ScriptLauncher.main(ScriptLauncher.java:183)
```

Stop Tomcat. Take down your servlet container. For Tomcat, use the `$CATALINA/shutdown.sh` script. (Many Unix-based installations will have a startup/shutdown script in the `/etc/init.d` or `/etc/rc.d` directories.)

Backup Your DSpace. Make a complete backup of your system

Have a snapshot of the PostgreSQL database. Use `pg_dump` command:

```
#!/usr/bin/pg_dump --host localhost --port 5432 --username postgres
--format custom --blobs --column-inserts --verbose
--file "/tmp/dspacePostgres.Backup" dspace pg_dump: reading schemas
```

Backup The asset store (`[dspace]/assetstore` by default, and any other assetstores configured in the `[dspace]/config/dspace.cfg` "assetstore.dir" and "assetstore.dir.#" settings)

Backup your configuration files and customizations to DSpace (including any customized scripts).

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
#sudo ant -Dconfig=[dspace]/config/dspace.cfg update
```

Generate Browse and Search Indexes.

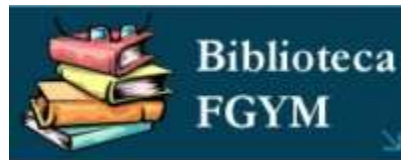
Though there are not any database changes between 1.7 and 1.7.1 release, it makes good policy to rebuild your search and browse indexes when upgrading to a new release. To do this, run the following command from your DSpace install directory (as the dspace user):

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

Source: Duraspace, Upgrading a DSpace Installation,
<https://wiki.duraspace.org/display/DSDOC/Upgrading+a+DSpace+Installation>

3.10. Results

The integration with the current web site has been carried out by inserting a button in the main page of the foundation to link to the digital library.



The following screenshot shows the final result of the button embedded into the right column of the main page of the foundation.

La **Fundación Globalidad y Microeconomía** nace para intentar ofrecer al emprendedurismo, oportunidades de la innovación no convencional susceptibles de ser éxito empresarial.

NOTA: No es que tengamos más preferencia por los nubarrones que presagian tormenta que por el cielo azul y luminoso del Mediterráneo, pero el gran ideólogo de la Innovación, Schumpeter, se refería a ésta como un *vendaval de destrucción creativa*. A él le dedicamos esta iconografía.

Seminario sobre Oportunidades de Negocio
Fundación Globalidad y Microeconomía

14 de Junio de 2012

Ya tiene a su disposición todo el material del **SEMINARIO SOBRE OPORTUNIDADES DE NEGOCIO FGYM** que tuvo lugar el día 14 de Junio en el cubo rojo de la CPI de la Universidad Politécnica de Valencia.

calendario de actividades
Julio 2012

					1
2	3	4	5	6	7
8	9	10	11	12	13
14	15	16	17	18	19
20	21	22	23	24	25
26	27	28	29	30	31

Biblioteca FGYM

seminario oportunidades

seminario innovación FGYM

blog

Main web page of the Foundation.

The DSpace welcome page has been customized to adopt the same design, as described in the section "2.3. Integration needs with the current web platform"



Main web page of the digital library

4. Conclusions

4. Conclusions

This project has provided the "*Globalidad y Microeconomía*" Foundation with a digital repository to facilitate their daily work.

The hardware infrastructure consists of a server sited in the headquarters of the foundation which has hosted the software required for the project.

The operating system Ubuntu Server LTS has been installed and configured in the server.

Having has been installed the software prerequisites of the digital repository.

Finally has been installed the DSpace tool and has been configured and customized to fulfill the organization's initially proposed requirements.

This has been the hardest and longest stage due to the inherent difficulty of editing the source code, given that some required features were not included in the DSpace set of functionalities.

After the installation and configuration process, was started the importation of the metadata and the documents into the DSpace repository.

The result has been very positive and is currently in use by the internal workers of the foundation and by the web visitors. The index and cover of the books of the library of the foundation can be consulted in the DSpace repository as well as a great amount of digital documents related to the innovation.

The foundation's digital library can be accessed from the main page of the foundation in the URL: <http://www.fundaciongym.es>

4.1. Technologic framework

The software versions used to accomplish the project have been the following:

- Ubuntu 12.04 LTS
- Java 6 SDK
- Apache Tomcat 7
- Maven 2
- Postgresql 8.4
- Xpdf
- Postfix
- DSpace 1.7.2

5. Bibliography

5. Bibliografy

[Cuellar, 2011] *Empar Cuéllar Alejandro, Diseño e implantación de un Sistema de Gestión de Información y Documentación (2011).*

Ubuntu, Automatic Updates,

<https://help.ubuntu.com/10.04/serverguide/C/automatic-updates.html>

Lukas van Schaik, Hide Community or Collection from list,

<https://wiki.duraspace.org/display/DSPACE/Hide+Community+or+Collection+from+list>

Duraspace, Discovery Configuration,

<https://wiki.duraspace.org/display/DSPACE/Discovery+Configuration>

Chapter 3. DSpace System Documentation: Installation,

http://www.dspace.org/1_6_0Documentation/ch03.html

Duraspace, Installing DSpace 1.7 on Ubuntu

<https://wiki.duraspace.org/display/DSPACE/Installing+DSpace+1.7+on+Ubuntu#InstallingDSpace1.7onUbuntu-DownloadandInstallDSpace>

Duraspace, Building DSpace From Source

<https://wiki.duraspace.org/display/DSPACE/Building+DSpace+From+Source>

Duraspace, Rebuild DSpace,

<https://wiki.duraspace.org/display/DSPACE/Rebuild+DSpace>

Jenson Taylor, Manually Installing Tomcat 7 on Ubuntu,

<http://jensontaylor.blogspot.com/2010/09/manually-installing-tomcat-7-on-ubuntu.html>

Duraspace, Upgrading a DSpace Installation,

<https://wiki.duraspace.org/display/DSDOC/Upgrading+a+DSpace+Installation>

Ahmed Ban, Install Ubuntu,

<http://tazjelapp.blogspot.com/2011/01/install-part-1.html>

Texas digital Library, Learning to use Manakin,

<http://www.tdl.org/wp-content/uploads/2009/04/LearningToUseManakin.pdf>

Fundación Globalidad y Microeconomía,

<http://www.fundaciongym.es>