DLIST: Distributed Digital Management of the Scholarly Publication

Kamalendu Majumdar   U N Singh

Abstract

In present scenario publishers of freely available electronic journals have the potential for positively transforming the highly problematic economics of scholarly publishing. But actual use of electronic scholarly publications shows that they present serious utilization barriers to would-be readers. Practical awareness, indexing, and archiving in this new literature can overcome these barriers and is essential if such promising publications are to transform scholarly communication. DLIST (Digital Library for Information Science and Technology) can meet these needs by integrating electronic publications into existing information systems. Thus disintermediation of the scholarly publishing process will help the clientele of the low budget library in all respects.

Keywords : Digital Libraries, D-Space

0. Introduction

The vision of the digital library is not new. This is a field in which progress has been achieved by the incremental efforts of numerous people over a long period of time. However, a few authors stand out because their writings have inspired future generations. Two of them are Vannevar Bush and J.C.R.Licklider. In July of 1945, Bush, then director of the U.S.Office of Scientific Research and development, published an article titled "As We May Think" in the Atlantic Monthly This article is an elegantly written exposition of the potential that technology offers the scientist to gather, store, find and retrieve information. Much of his analysis rings as true today as it did 50 years ago. The Atlantic Monthly has placed a copy of "As We May Think" on its web site. Anyone interested in libraries or in scientific information should read it.

In the 1960s, Licklider was one of the several people at Massachusetts Institute of Technology who studied how digital computing could transform libraries. Like Bush, Licklider was most interested in the literature of science; however, he foresaw may developments that have occurred in modern computing.

1. New trend

Library gurus invariably have long lists of difficult issues to confront. These days, high on my list is the future of our university libraries. Although libraries form the basic infrastructure of the academic endeavor, I have come face to face with an unhappy fact: University librarians are now being forced to work with faculty members to choose more of the publications they can do without. The ballooning costs of academic publications are preventing faculty members and researchers from gaining access to the world's scholarship and knowledge.

2. Hurdles

Institutions are facing unprecedented budget crises just as expanding faculties and student bodies are increasing the demand for scholarly information. Even in the best of economic times, university libraries cannot hope to keep pace with the 6- to 12-percent annual inflation rate in the price of scholarly journals. And the fiscal environment today is particularly difficult. Without proper user of the huge bandwidth and the classical internet architecture, neither university librarians nor faculty members can deal with the challenges of preserving access to scholarly resources.

3.   Worldwide efforts

In this regard, the librarians should consider several strategies, including development and support to new models of scholarly publishing that can cut the costs of distributing and retrieving information. Here it can be mentioned that several organizations are experimenting with less-expensive ways to disseminate faculty research. Some of them are already well known, like
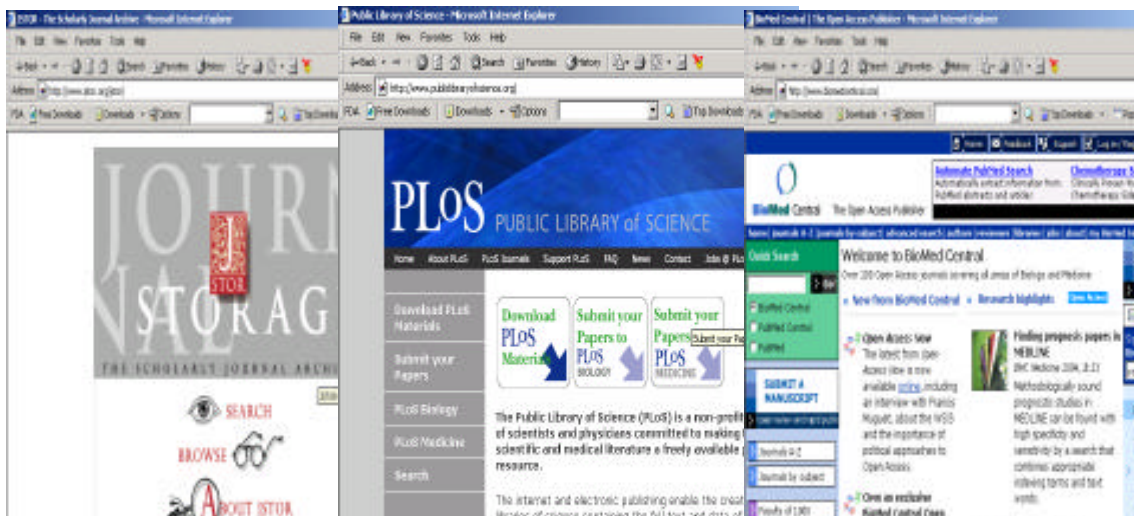

Fig: I Some online portal for fair use.

JSTOR, which digitally archives more than 300 journals in various disciplines, and Stanford University's HighWire, which stores online several hundred journals in biology, physics, and other sciences. Others, like BioMed Central and the Public Library of Science in both biology and medicine, are only just emerging. The same strategies should be applicable to our country.

4.   DLIST: Present Status

IIT was established with the objective of taking the country to technological excellence that will catapult the economy to the big league. Half a century has passed and we can proudly say that it is a goal fully accomplished. However, we have miles to go. Online publication is perhaps such an area**.** We anticipated to have a digital library which is able to deal with the e-collections e-persons and e-groups. At the same time it should enable with OAI (Open Archives Initiative) along with the support of DC (Dublin Core). Here we want to mention that an OAI is a protocol for metadata harvesting. This allows sites to programmatically retrieve or harvest the metadata from several sources, and tender services using that metadata, such as indexing or relating services. Such a service could allow users to access information from a large number of sites from one portal.

5.   Motivation

During 2002 we received a project from MHRD with a objective of developing a Real time Multimedia for distance education with an objective of creating a virtual classroom by setting a multimedia digital library server named Ekalavya at IIT, Kharagpur and to provide linkage to intellectual contents through a high-end real time multimedia server.
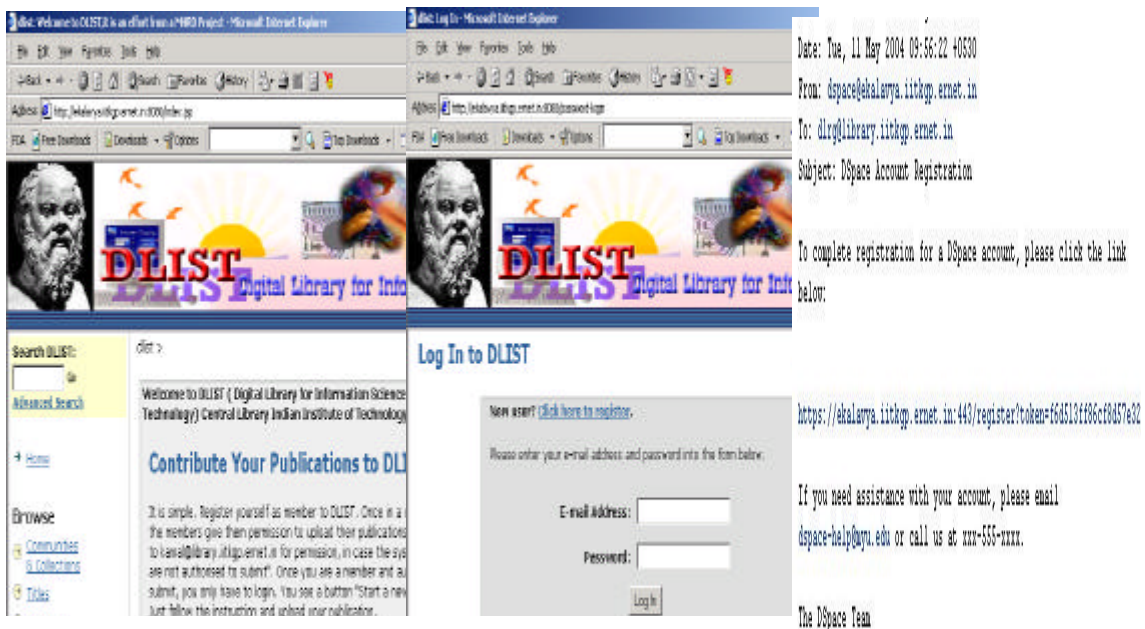
Fig: 2 Some snap sort of the IIT portal .

We are always fond of open source software. In the mean time we heard of Dspace, a digital library software platform jointly developed by Hewlett-Packard Company and MIT. In this article the authors intent to discuss the plan and success of the implementation of DSpace at Central Library, IIT, Kharagpur.

The following open sources s/w were used

•j2sdk-1_4_2_02-linux-i586	•httpd-2.0.43	•Redhat Linux 8.0
•jakarta-tomcat-4.1.27	• jk-1.2-src_connector	•openssl-0.9.7c
•apache-ant-1.5.4-bin	•dspace-1.1.1	•postgresql-7.3.4

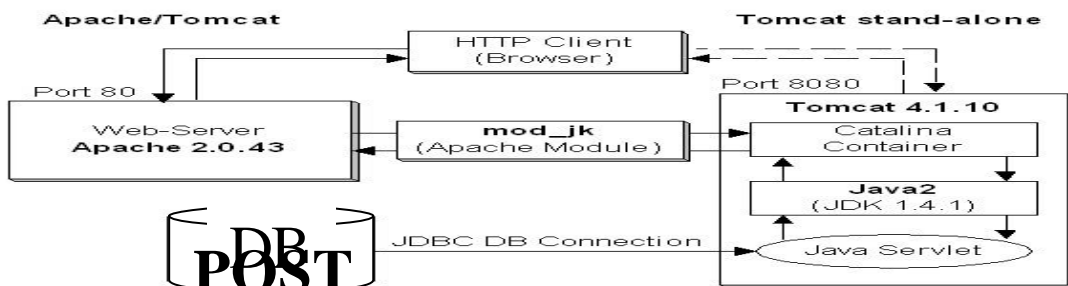How the portal actually works

Fig: 3 How the portal works .

A complete portal has been designed, developed and deployed by a small team of IIT ,Kharagpur Library as a prototype.Document has been uploaded for test run. It has been demonstrated to the fullest satisfaction that it works like any other electronic scholarly publication. The portal at IIT, Kharagpur is capable of

- ✍ Creating e-collections, e-persons and e-groups.

- ✍ Supporting the digitization of in house resources (e.g. institute theses)

- ✍ OAI (Open Archives Initiative) with the support of DC (Dublin Core).

- ✍ Programmatically retrieving or harvesting the metadata from several sources.

- ✍ Providing various indexing services using those metadata
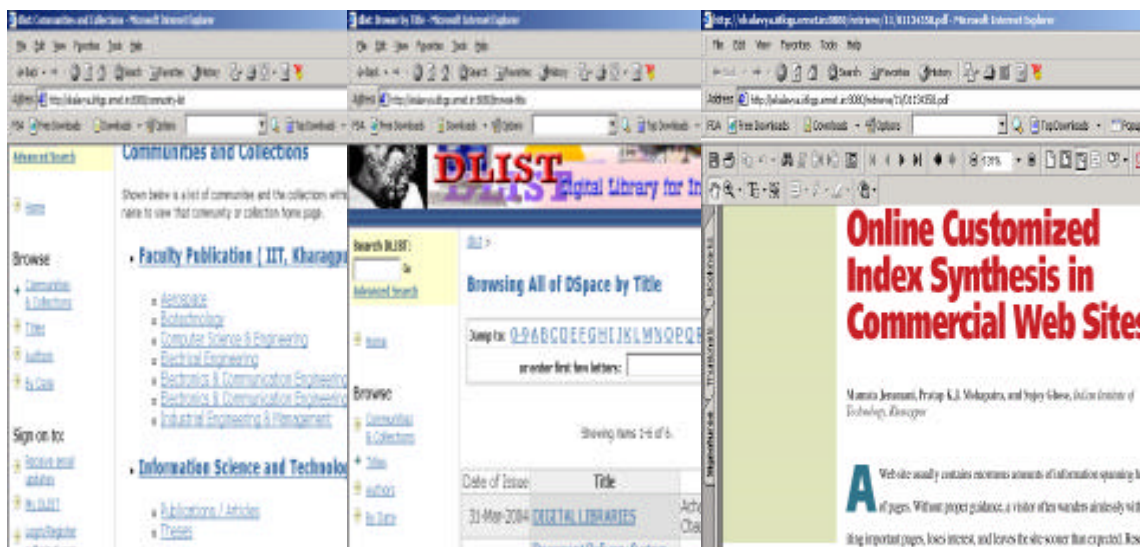
Creation of Community   Metadata search   Full text

Fig: 4 Creation of community/metadata search / access of full test document from the portal.

We have intention to help those libraries who have vary small budget and are not capable to have the costly online database and the Scholarly Community who also wants to discuss the professional issues and want to disseminate their scholarly writings to the users from various schools or colleges . We propose to publish a series of publications to give the total idea from installation of DSpace to the Organization of DSpace. In this regards this is the first effort.

6.   Methodology

Fortunately we came across the IBM software products the "Quick Start" manual. This manual introduced the product through a series of common tasks, from installation to basic configuration. The manual did not go deep into the details of the software — that's what the User Guides and other documentation were

for — but it did help new users get idea about the product quickly, and without scaring them too much. The Jakarta group produces a lot of documentation for Tomcat, but none of it seemed to be as soothingly simple to read as the "Quick Starts" and in case of DSpace, most likely the developers have an impression that the user of DSpace would have a certain echelon so they did not focus about the detailed installation of DSpace . Hence this article will prove useful. Hopefully, we have addressed the common issues facing new installation in such a way that their learning curve becomes easier to negotiate.

First we install the operating system Redhat 8.0 and on top of that we want to install Apache2

As the home of apache2 is in our case /usr/local/apache2

#tar xvfz httpd-2.0.47.tar.Z

#cd httpd-2.0.47

#./configure —prefix=/usr/local/apache2

#make

make install

It creates httpd directory at /usr/local/apache2/conf. We made the change in httpd.conf file variables Now question is how the system can know that instead of usual position (/etc/httpd/conf) httpd.conf file now we have the httpd.conf in a separate location i.e. /usr/local/apache2/conf/httpd.conf. To face these problems we have to change path of the domain at /etc/init.d/httpd. At this juncture we are also encountering error message

#/etc service httpd start

      service httpd service **failed**

we have struggle a lot at last we delete the directory " httpd" under /etc

#rm –R httpd

here also we are getting error message then we delete the .exe file of httpd under /sbin/httpd and copy the same from /usr/local/apache2 to /sbin/ and   now apache2 is working fine.

Now here we want to depicted the necessity of another webserver like program called tomcat. Tomcat, sometimes known as Apache Tomcat, sometimes as Jakarta Tomcat, is a Java servlet engine that is the reference implementation used by Sun for its Java servlet and JSP specifications. If you want to know details about tomcat please logon to http://jakarta.apache.org/tomcat,

7. Installation of Tomcat

tar xvfz jakarta-tomcat-4.1.27.tar.gz

mv jakarta-tomcat-4.1.27 /usr/local/

cd /usr/local

ln -s jakarta-tomcat-4.1.27 tomcat -> (sysmlink for tomcat)

Now problem is how apache and tomcat will interact with each other. Easiest way is to use a connector which will act as an agent between this two web servers. All the clients will send their requests to Apache and in process Apache will send these requests to Tomcat through the connector . There are several connectors developed by many professionals . We want to use mod-jk as connector. Before proceeding further you have to stop Apache and Tomcat .

 Installation of mod_jk from rpm

rpm -ivh mod_jk-ap13-1.2.2-1jpp.i386.rpm

now you have to edit three configuration files viz. httpd.conf, workers.properties and server.xml (see Appendix 1)

httpd.conf

After mod_jk has been successfully installed, a set of entries will be appended to the end of Apache's configuration file, httpd.conf. Look for the following lines at the end of the file :

```
<IfDefine HAVE_JK>
     LoadModule jk_module modules/mod_jk.so
     AddModule  mod_jk.c
     Include /etc/httpd/conf/mod_jk.conf
</IfDefine>
Include /usr/local/apache2/conf/mod_jk.conf
RedirectMatch ^/$ https://ekalavya.iitkgp.ernet.in/
RedirectMatch ^(/[^o].*) https://ekalavya.iitkgp.ernet.in$1
RedirectMatch ^(/.[^a].*) https://ekalavya.iitkgp.ernet.in$1

RedirectMatch ^(/..[^i].*) https://ekalavya.iitkgp.ernet.in$1
```

mod_jk.conf

This file is crucial to the integration, but before we go into it, we must first understand a few things about it. In mod_jk.conf we need to do the following :

- l   Define where the workers.properties file is, and what it is called
- l   Define where the log file is, and what it is called
- l   What kinds of messages to record - errors only, or more ?
- l   Whether to enable SSL or not
- l   Define the web application contexts

At this point, it is a good idea to open up the mod_jk.conf that comes with the mod_jk rpm. We will examine each of the directives in the file, relating them to the tasks outlined above.

JkWorkersFile

Value: /etc/httpd/conf/workers.properties

The location of workers.properties, a file that we will examine next.

JkLogFile

Value: /var/log/httpd/mod_jk.log

The location of the log file. Depending on the error level you specify, the file will contain either very verbose information or just the critical errors.

JkLogLevel

Value: error

There are 3 available options here :

- l   info - logs information about mod_jk's activities
- l   error - in addition to activity information, errors will also be logged
- l   debug - logs anything and everything. As its name implies, this option is ideal when troubleshooting.

Following the first 3 lines of mod_jk.conf, we encounter a series of entries that relate to SSL. We will not go into this at the moment, because we want to get a "plain-vanilla" Tomcat-Apache integration going successfully first.

What follows after the SSL section are the "contexts". We have already seen earlier that "contexts" refers to web applications that are deployed inside Tomcat. We specified a <Context> element inside Tomcat's server.xml for every web application deployed, and we have to do the same here, because Apache needs to know how to hand-off requests for web applications to Tomcat.

For every web application, we must define the contexts to Apache, and we do this by supplying the following information :

- l   Where the web application is in the file system and how do we map it to a URL ?
- l   What additional options do we want to enable for the web application ?
- l   How do we "mount" it in Apache ?

This is the web application and servlet which is working properly.

dlist Web Application

Web Application Name   dlist
Location $CATALINA_HOME/webapps/dlist
URL we want to map to   http://hostname.domain.com/ dlist/
 mod_jk Worker Name (defined inside workers.properties)   ajp13

How we express this as a context in mod_jk.conf as shown below :

#
The following line makes Apache aware of the location of the /dlist context
#
Alias /dlist "/usr/local//tomcat/webapps/dlist "
<Directory "/usr/local/tomcat/webapps/dlist ">
Options Indexes FollowSymLinks
</Directory>
#
The following line mounts all JSP files, the /servlet/ uri, and all files to Tomcat
#
<VirtualHost>

```
  JkMount /dlist/*.jsp ajp13
  JkMount /dlist/* ajp13
  JkMount /*.jsp ajp13
  JkMount /* ajp13
  JkExtractSSL On
   Jk HTTPS Indicator HTTPS
   JkSESSIONIndicator SSL_SESSION_ID
   JkCIPHERIndicator SSL_CIPHER
   JkCERTSIndicator SSL_CLIENT_CERT
 # End of Tomcat mod_jk directives
#Include /usr/local/apache2/conf/mod_jk.conf

</VirtualHost>
```

That completes our configuration for mod_jk.conf. We have two more files to edit before we are ready to begin testing.

workers.properties file will be like this :

```
worker.list=ajp13
worker.ajp13.port=8009
worker.ajp13.host=ekalava.iitkgp.ernet.in
worker.ajp13.type=ajp13
```

Now the most critical part of the installation is D Space installation and this has been described in Appendix 1.

8.   Conclusion

The faculty member who chooses alternative ways to disseminate his research should be recognized and rewarded by fellow colleagues. The rapid emergence of scholarly electronic publishing challenges our traditional methods of assessing professors' work for tenure and promotion purposes. We should take steps to guarantee that our evaluation practices keep pace with the adoption of new communication technologies. At the University of California, for instance, the Academic Senate supports consideration of electronic publications in academic peer review.

At the same time, we must not jeopardize the health or well-being of the scholarly societies and university presses that play so critical a role in academic life. Faculty members should continue to manage their intellectual property and copyright. They should decide which publishing organizations they will review, edit, and write for. When signing a publishing contract, they should determine whether to assign the publisher copyright and whether to seek a nonexclusive right to disseminate their work freely in an electronic form.

9.   Reference

1.   http://www.openarchives.org/OAI/2.0/openarchivesprotocol.htm#DefinitionsConcepts

About Authors

Mr. Kamalendu Majumdar Presently he is working as Assistant Librarian at Central Library , Indian Institute of Technology, Kharagpur, India. He holds MLibSc and BLibSc from IGNOU in the year 1994 and 1996. Currently he is going on under research program from Guru Ghasidas University. Bilaspur. His major research include multimedia information retrieval, wireless information retrieval, query processing and optimization in multiprocessor and distributed systems, and database management system and specialy interested in computer network . He has published over 35 papers in conferences/ journals in international and national level. He has also edited IMeL2002 conference proceeding as a member of editorial committee. **Email :** kamal@library.iitkgp.ernet.in

Prof. U N Singh Presently he is the Head, Dept. of Library & Information Science and Dept. of Information Technology, G G University Bilaspur, CG. He holds M.Sc and BLibSc from BHU, Varanasi, UP in the year 1979 abd 1980, Associateship in Information Science from INSDOC IN 1983, and Ph.D in Information Science in 1992 from B I T Mesra, Ranchi. His field of interest includes information science, Information storage and retrieval, and query processing , Computer network, Data and file structure, Bibliometrics, Scientometrics and Library Automation. He has published over 29 Conference/ Journal papers in international and national conferences and journals. He has guided Six Ph.D students and many MLibSc students for their project works. He is also responsible for starting many new PG level courses in Information Technology in BIT Mesra and GG University, CG. He is also member of Board of Studies and Ph. D Committee in several Indian universities. **Email :** unsingh03@yahoo.com

Appendix 1

Now the most critical part of the installation is to install the D Space. Now we have to home page Tomcat and Apache separately and both are running
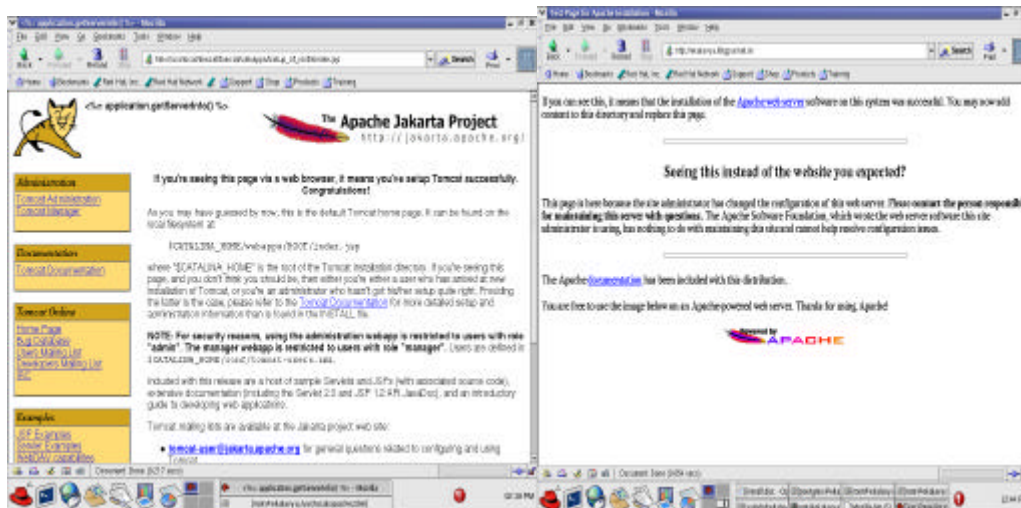
Installation of Dspace

Download the binary package—> **dspace-1.1.1.tar.gz** to /usr/local as root user

]# mkdir /dspace

]# chown dspace:dspace /dspace

]# cp /usr/local/dspace-1.1.1.tar.gz /dspace

]# chown dspace:dspace /dspace/dspace-1.1.1.tar.g

Log in as dspace user

]$ cd /dspace

]$ tar xvfz dspace **dspace-1.1.1.tar.gz** (You will get **dspace-1.1.1-source** directory)

Customise the dspace.cfg file or leave this step here...

Build & install The dspace source:

Log in as dspace user

]$ cd /dspace/dspace-1.1.1-source

]$ ant

]$ ant fresh_install

Connect Dspace webapps to tomcat

]# cd /usr/local/tomcat/tomcat/webapps/

]# mv ROOT ROOT.bak

]# ln -s /dspace/jsp ROOT

]# ln -s /dspace/jsp dlist

]# ln -s /dspace/oai dspace-oai

]# chown -R dspace:dspace ROOT dlist dspace-oai

Install the Config files:

Log in as dspace

]$ cd /dspace/bin

]$ ./install_configs

Create The administrator account.

]$ cd /dspace/bin

]$ ./create_administrator

Initialise Lucence search indicies:

]$ ./index_all

Set up email subscription feature. In the dspace user's crontab insert the following:

]# su – dspace

]$ vi emailsub

Send out subscription emails at 01:00 every day

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

]$ crontab emailsub

Customise the dspace.cfg file :

```
]# vi /dspace/config/dspace.cfg
dspace.url = https://ekalavya.iitkgp.ernet.in:443
          = http://ekalavya.iitkgp.ernet.in:80
]# vi /dspace/config/dspace.cfg
dspace.hostname = ekalavya.iitkgp.ernet.in
dspace.name = dlist
```

Note :When ever u write any thing in /dspace/config/dspace.cfg

```
U have to run install-configs from /dspace/bin
]# cd /dspace/bin
]# ./install-configs
```

Edition in server.xml in tomcat

```
]# cd /usr/local/tomcat/tomcat/conf
]# vi server.xml
add the following    lines between Host tags i.e   <Host>   </Host>
 <Context path="/dlist" docBase="dlist" debug="0">
     <Resources className="org.apache.naming.resources.FileDirContext"
    allowLinking="true" docBase="" />
 </Context>
 <Context path="" docBase="ROOT" debug="0">
     <Resources className="org.apache.naming.resources.FileDirContext"
     allowLinking="true" docBase=""/>
 </Context>
 <Context path="/dspace-oai" docBase="dspace-oai" debug="0">
     <Resources className="org.apache.naming.resources.FileDirContext"
     allowLinking="true" docBase="" />
</Context>
```

Edition in httpd.conf file in apache.

```
Add the following lines between the <VirtualHost 10.17.200.1:443>  ...</VirtualHost>
]# vi /usr/local/apache2/conf
]# vi httpd.conf
Alias /dlist "/usr/local/tomcat/jakarta-tomcat-4.1.30/webapps/dlist"
<Directory "/usr/local/tomcat/jakarta-tomcat-4.1.30/webapps/dlist">
   Options Indexes FollowSymLinks
DirectoryIndex index.html index.html index.jsp
</Directory>
JkMount  /dlist/*.jsp  ajp13
JkMount  /dlist/*   ajp13
JkMount /*.jsp ajp13
JkMount /*  ajp13
```

]# vi httpd.conf

U also need to set up apache to understand the mime-type:

add the line :

AddType   text/jsp   .jsp

Note: -> At this point make sure that ur postgresql is running

login as postgres

]# su – postgres

]$ cd /usr/local/postgres/post/bin

]$ ./postmaster -i -D /usr/local/postgres/post/data

Stop the tomcat , apache & Start the tomcat ,apache

]# /usr/local/tomcat/tomcat/bin/startup.sh

wait for 30 seconds

]# /usr/local/apache2/bin/apachectl startssl

Now u can access the dspace home page on the following URLs

https://ekalavya.iitkgp.ernet.in

https://ekalavya.iitkgp.ernet.in/dlist

Congratulations..... Dlist home page