THE USE OF DSPACE SOFTWARE IN PRESERVATION AND MANAGEMENT OF SPECIFIC LIBRARY RESOURCES.

By

ALIYU ABDULKADIR

Being a paper presented at the Kashim Ibrahim Library Senior Staff Seminar on 6th June 2011

Abstract

This paper examines Dspace software and how it is use in preserving and managing scholarly documents in Kashim Ibrahim Library, Ahmadu Bello University Zaria. The paper highlights what an Institutional repository is, what it is use for, open source software and how Dspace software runs. The paper concludes that librarians should keep abreast of the current developments in ICT and choose appropriate software that is geared toward a digital revolution.

Introduction

Digitization can remedy and ensure the global visibility of information resources in our institutions, many universities want to digitalize their intellectual wealth. Once an institution has a clear vision and defines what the repository will do for its community, subsequent decisions such as a choice of software are an essential factor to be considered. A feasibility study of the software is very important in this choice. Soft wares are initially chosen with great promises, but as the project progresses, challenges which were not identified ab-initio arise to demoralize the initial enthusiasm, therefore software with large community users and developers should be considered in order to achieve the objective of the repository.

Institutional Repository

According to Barton and Waters (2004) "institutional repositories is a database with a set of services to capture, store, index, preserve and redistribute a university's scholarly research in digital formats". The Institutional Repositories can be regarded as follows:

*Institutionally defined

*Scholarly

*Cumulative and perpetual

*Open and Interoperable (OAI-compliant and allow open access to scholarly research).

Institutional repository is used in the following ways:

*Scholarly communication

*Storing learning materials and courseware

*Electronic publishing

*Managing collections of research documents

*Preserving digital materials for the long term

*Adding to the university's prestige by showcasing its academic research

*Institutional leadership role for the Library

*Knowledge management

*Research assessment

*Encouraging open access to scholarly research

*Housing digitized collections

Most institutions and libraries are now considering Open Source Technology for building and developing their institutional repositories, because of its advantages.

Open Source Software

Open source software is software that has been released under an Open Source Initiative (OSI) Open source provides broad rights to use, modify, and distribute software. The source code is accessible and flexible to manipulate and this allows user of the software to contribute to its further development and fixing of bugs. Open source softwares are available free of cost and its license authorize users to run and distribute the software without any restriction.

Dspace

Dspace is a digital repository software platform with broad functionality for the capture, management, preservation and redistribution of digital scholarly research materials in a variety of formats, for a variety of purposes, Walker (2010). Pennock (2006) views Dspace as a digital repository system design to capture, store, index, preserve, and provide access to institutional digital research materials. Morris (2009) argues that it is an out-of-the -box repository software package for creating open repositories focused on delivering digital content to end users, and providing a full set of tools for managing and preserving content within the application. Dspace, however, is an open source software designed to capture, manage, preserve and disseminate digital scholarly research material in all forms: Dspace captures, distributes and preserves digital research products. Here you can find articles, working papers, preprints, technical reports, conference papers and data sets in various digital formats. Content grows daily as new communities and collections are added to Dspace.

Dspace was developed to solve the problem of providing a long-term home for digital materials, thus Dspace focused on open knowledge sharing and long term preservation concept. Dspace was originally designed and developed by Hewlett-Packard (HP) and Massachusetts Institute of Technology (MIT). The software development started in November 2000 as part of the HP and MIT alliance and in November 2002 they released Dspace version 1.0. In July 2007, Dspace foundation was formed. The foundation merged with Fedora Commons to create the Dura space Organization in July 2009. The design supports community participation and customization to serve the content and intellectual property of each user community. Dspace is written in Java and runs on Linux or UNIX system and Windows.

Dspace Collection Building

Dspace organizes contents into communities and collections. Collections contain items that are *similar* in some ways such as source, purpose, existing series or audience research topic, subject matter. Each collection has an approval process that includes combination of approval roles. Users *choose* a collection to submit items which could be conference presentation, research

paper, an annotated series of images etc. We at Kashim Ibrahim Library archive three collections for now, these include theses and dissertation, seminar and conference papers. The Dspace Communities correspond to administrative entities such as schools, departments, labs and research centers. Within each community there can be an unlimited number sub communities and an unlimited number of collections. Each collection may contain an unlimited number of items. When an item is submitted to a collection, the system routes the submission through the approval process; which reviews the content to ensure that it is submitted to the appropriate collection, check the submission for completeness and errors and then check or augment the submission's metadata Bass et al, (2002). All submitted and approval items enters the archive, which provides storage for the individual bit stream(s) associated with the submitted digital file. It provides users the capabilities to search, browse content of the Dspace and also capability to download authorized item. Dspace is designed to make submission easy and to provide long-term preservation of digital material stored in the repository.

Advantages of Dspace

1. Dspace has the largest community users and developers that provides frequent upgrades
2. It is completely customizable
3. It is effective for long term preservation of digital material.
4. Friendly user interface and it possesses basic and advanced search capabilities
5. Supports all file formats.

Disadvantages of Dspace

1. Dspace installation and customization require enough technical skill.
2. The authorization system is powerful but difficult and cumbersome to use.

Submission & Uploading Steps

BROWSE

Browse allows one to go through a list of items in some specified order:

Browse by Community/Collection takes one through the communities in alphabetical order and allows him to see the sub communities and collections within each community.e.g In Ahmadu Bello University Zaria there are twelve faculties which constitutes the communities and under each faculty the departments make up the sub communities in that order.

Browse by Title allows one to move through an alphabetical list of all titles of items in Dspace.

Browse by Author allows one to move through an alphabetical list of all authors of items in Dspace.

Browse by Subject allows one to move through an alphabetical list of subjects assigned to items in Dspace.

Browse by Date allows one to move through a list of all items in Dspace in reverse chronological order.

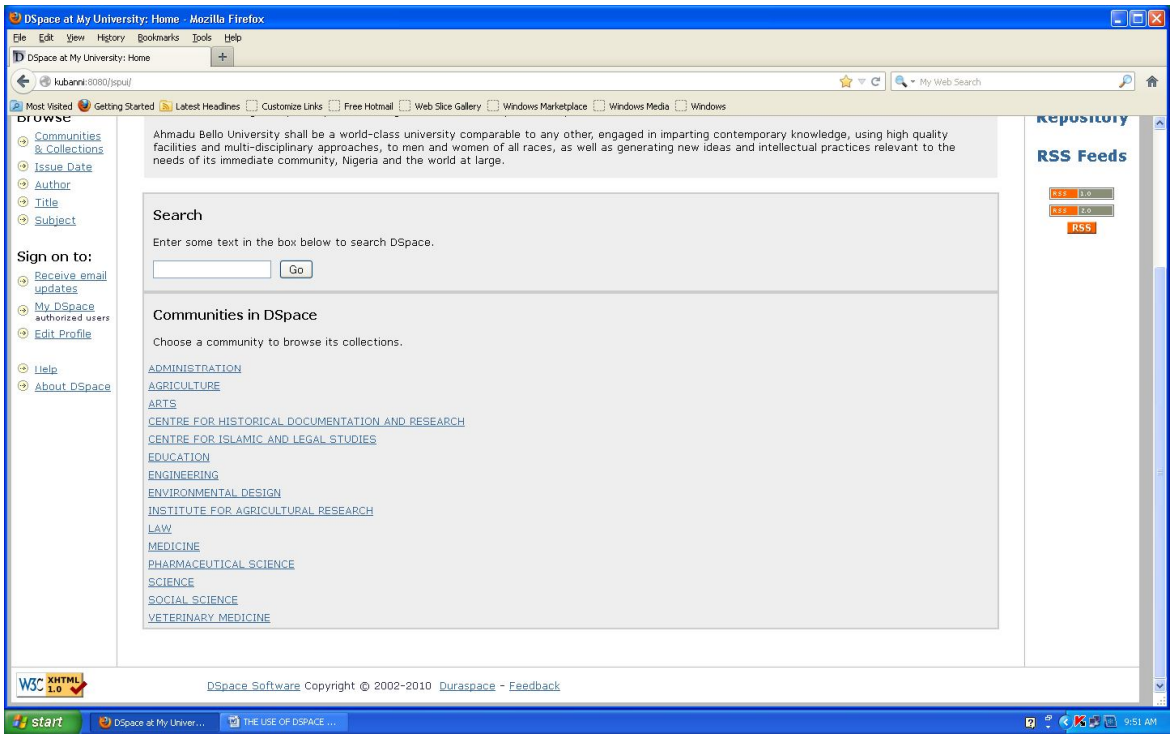
Submit is the Dspace function that enables users to add an item to Dspace. The process of submission includes filling out information about the item on a metadata form and uploading the file(s) comprising the digital item.

SEARCH

To search all of Dspace a yellow search box at the top of the navigation bar on the left is provided or the search box in the middle of the home page.



To limit one search to a specific community or collection, navigate to that community or collection and use the search bar on that page.



Communities

The Dspace content is organized around Communities which can correspond to administrative entities such as schools, departments, labs and research centers. Within each community there can be an unlimited number sub communities and an unlimited number of collections. Each collection may contain an unlimited number of items. This organization gives Dspace the flexibility to accommodate differing needs of communities by allowing them to

- Decide on policies such as: who contributes content, whether there will be a review process , who will have access
- Determine workflow - reviewing, editing, metadata
- Manage collections

Each community has its own entry page displaying information, news and links reflecting the interests of that community, as well as a descriptive list of collections within the community.

Collections

Communities can maintain an unlimited number of collections in Dspace. Collections can be organized around a topic or by type of information (such as working papers or datasets) or by any other sorting method a community finds useful in organizing its digital items. Collections can have different policies and workflows. Each Dspace collection has its own entry page displaying information, news and links reflecting the interests of users of that collection.

Sign On to Dspace

When one accesses an area of Dspace that requires authorization, the system will require him to log in. All users can register to become subscribers. Some restricted functions, such as content submission, require authorization from the community. Before one log in for the first time, he will need to click on "register with Dspace" and follow the instructions. After that, he will need to enter his e-mail address and password in the log-in form that appears. His e-mail address should include his username and domain name. It is not case sensitive.

Example: abdulaliyu74@yahoo.com

Type your password exactly as you entered it originally. It **is** case sensitive. Be sure to click on the "log in" button to continue.

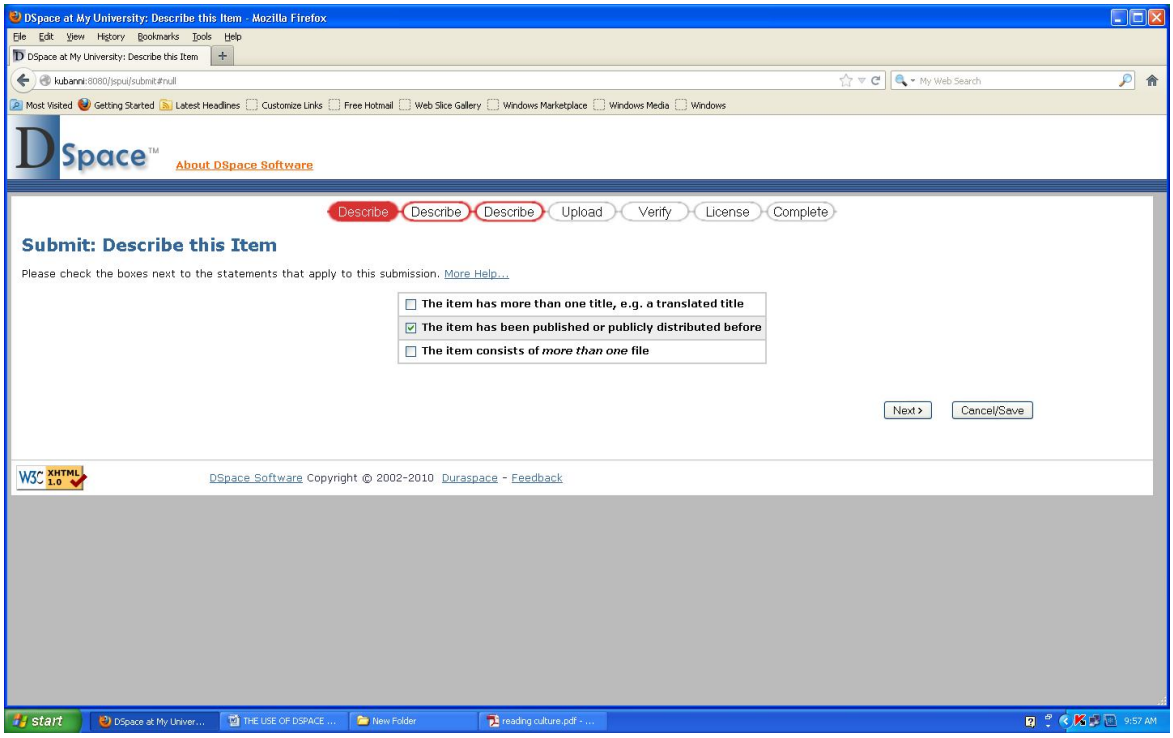
Submit

At any point in the submission process one can stop and save his work for a later date by clicking on the "cancel/save" button at the bottom of the page. The data one have already entered will be stored until he comes back to the submission, and he will be reminded on "My Dspace" page that he has a submission in process. If somehow he accidentally exit from the submit process, he can always resume from "My Dspace" page. He can also cancel his submission at any point.

Choose Collection

Progress Bar - Oval Buttons at Top of Page:

At the top of the submit pages one will find 7 oval buttons representing each step in the submission process. As one move through the process these ovals will change color. Once you have started you can also use these buttons to move back and forth within the submission process by clicking on them. You will not lose data by moving back and forth.



Select Collection:

Click on the arrow at the right of the drop-down box to see a list of Collections. Move your mouse to the collection into which you wish to add your item and click. You must be authorized

by a community to submit items to a collection. Click on the "next" button to proceed, or "cancel/save" button to stop and save or cancel your submission.

Submit: Describe Your Item -Page 1

If one responds "yes" to any of the questions on this page, he will be presented with a modified input form tailored to capture extra information. Otherwise he will get the "regular" input form.

* More than one title - Sometimes an item has more than one title, perhaps an abbreviation, acronym, or a title in another language. If this is the case, and he wants this information captured, click in the "yes" box.

* Previously issued - New items that have NOT been previously published or distributed will be assigned an issue date by the system upon Dspace distribution. If one is entering older items that have already been distributed or published, click in the "yes" box. One will receive a form prompting him for several pieces of information relating to publication.

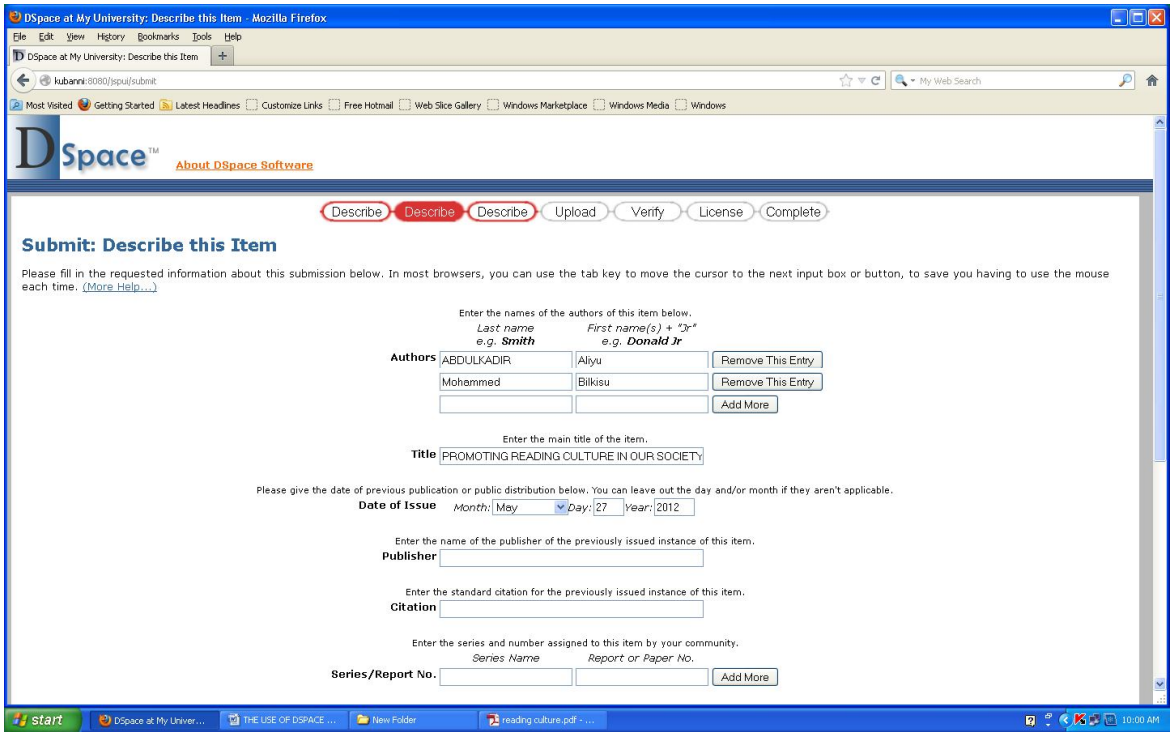
* Multiple files - An item can consist of more than one file in Dspace. A common example of this would be an HTML file with references to image files (such as JPG or GIF files). Another example of this would be an article supplemented with a video simulation and a data file. If one is submitting more than one file for this item, click in the "yes" box. Click on the "next" button to proceed, or "cancel/save" button to stop and save or cancel the submission.

Submit: Describe Your Item - Page 2

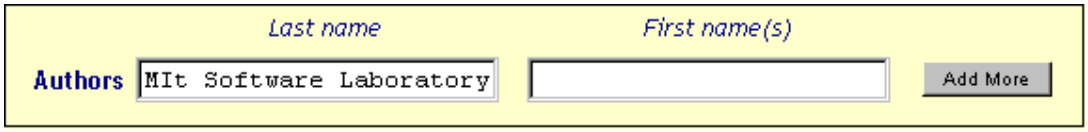
The information one fill in on these two screens will form the metadata record that will enable users to retrieve the item using search engines. The richer the metadata, the more "findable" the item will be.

Author:

This can be a person, organization or service responsible for creating or contributing to the content of the item. By clicking on the "Add More" button one can add as many authors as needed. Examples:



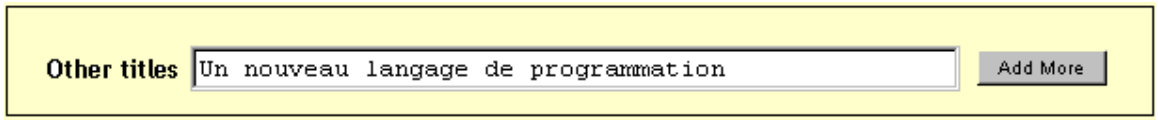
If the author is an organization, use the last name input box for the organization name:



Title:

Enter the full and proper name by which this item should be known. All Dspace items must have a title! Without a title provided one cannot proceed.

Other Title: Note - this input box appears only if one indicated on the first page that the item has more than one title. If the item has a valid alternative title, for instance, a title in another language or an abbreviation, then enter it here. Example:



Date of Issue: Note - this input box appears only if one indicated on the first page that the item has been previously published or distributed. If Dspace is the first means of distribution of this item, a date will be assigned by the system when the item becomes a part of the repository. If the item was previously published or made public, enter the date of that event here. If you don't

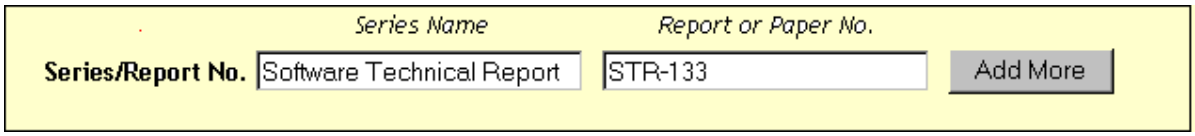
know the month, leave the default "no month"; otherwise select a month from the drop-down box. If you don't know the exact day, leave that box empty.

Publisher: Note - this input box appears only if one indicated on the first page that the item has been previously published or distributed. Enter the name of the publisher of this item.

Citation: Note - this input box appears only if one indicated on the first page that the item has been previously published or distributed. Enter citation information for this item if it was a journal article or part of a larger work, such as a book chapter. For **journal articles**, include the journal title, volume number, date and paging. For **book chapters**, include the book title, place of publication, publisher name, date and paging.

Series/Report No.:

Some of the collections in Dspace are numbered series such as technical reports or working papers. If this collection falls into that category, then there should be a default value in the **Series Name** box which one should not change, but you will have to fill in the assigned number in the **Report or Paper No.** input box. Examples:



Identifiers:

If a unique number or code that identifies this item in some system is known, enter it here. Click on the arrow to the right of the input box, and select from one of the choices in the drop down menu. The choices refer to:

Govt.doc # - Government Document Number - e.g. NASA SP 8084
ISBN - International Standard Book Number - e.g. 0-1234-5678-9
ISSN - International Standard Serial Number - e.g. 1234-5678
ISMN - International Standard Music Number - e.g. M-53001-001-3
URI - Universal Resource Identifier - e.g.. http://www.dspace.org/help/submit.html
other - A unique identifier assigned to the item using a system other than the above.

Type: Select the type of work or genre that best fits the item. To select more than one value in the list, one may have to hold down the "ctrl" or "shift" key.

Language:

Select the language of the intellectual content of the item. If the default English - United States is not appropriate, click on the arrow on the right of the drop down box to see a list of languages commonly used for publications, e.g.

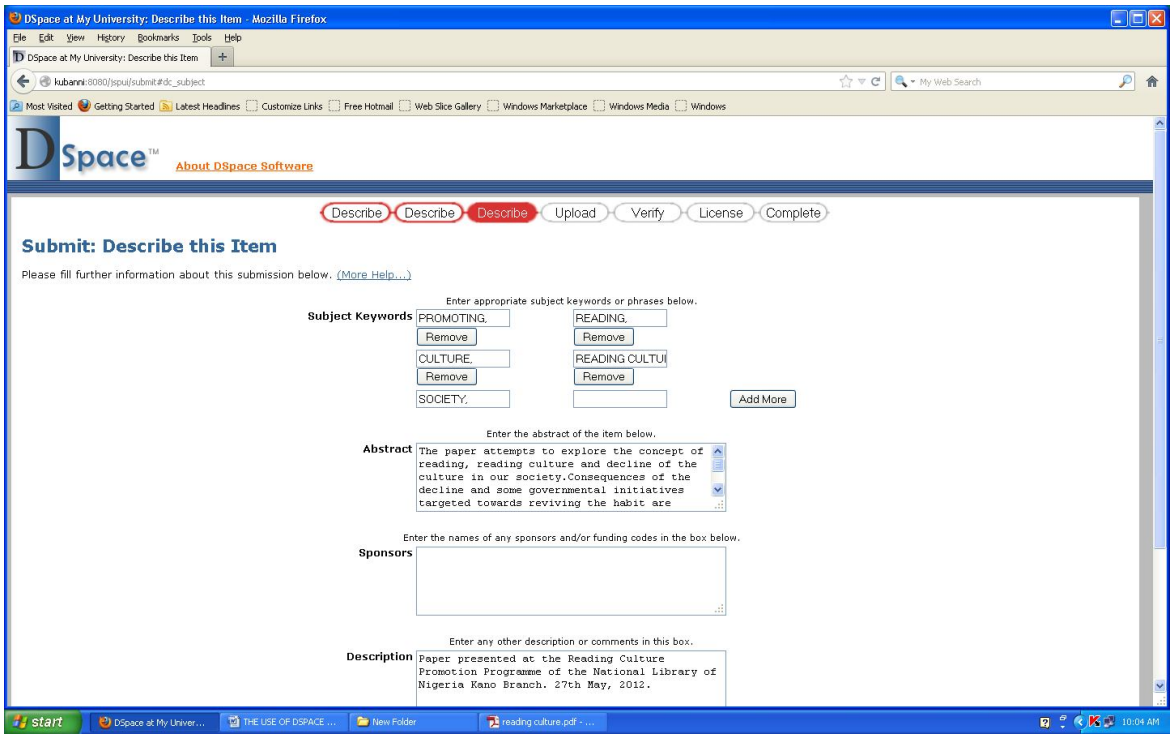


If the item is not a text document and language is not applicable as description, then select the N/A choice. Click on the "next" button to proceed, or "cancel/save" button to stop and save or cancel your submission.

Submit: Describe Your Item - Page 3

Subject/Keywords:

Please enter as many subject keywords as are appropriate to describe this item, from the general to the specific. The more words one provide, the more likely it is that users will find this item in their searches. Use one input box for each subject word or phrase. One can get more input boxes by clicking on the "add more" button. Examples:

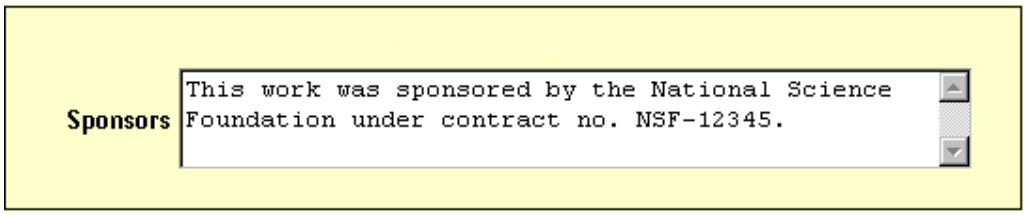


Abstract:

One can either cut or paste an abstract into this box, or you can type in the abstract. There is no limit to the length of the abstract.

Sponsors:

If the item is a product of sponsored research, one can provide information about the sponsor(s) here. This is a freeform field where you can enter any note you like. Example:



Description:

Here one can enter any other information describing the item he is submitting or comments that may be of interest to users of the item. Click on the "next" button to proceed, or "cancel/save" button to stop and save or cancel your submission.

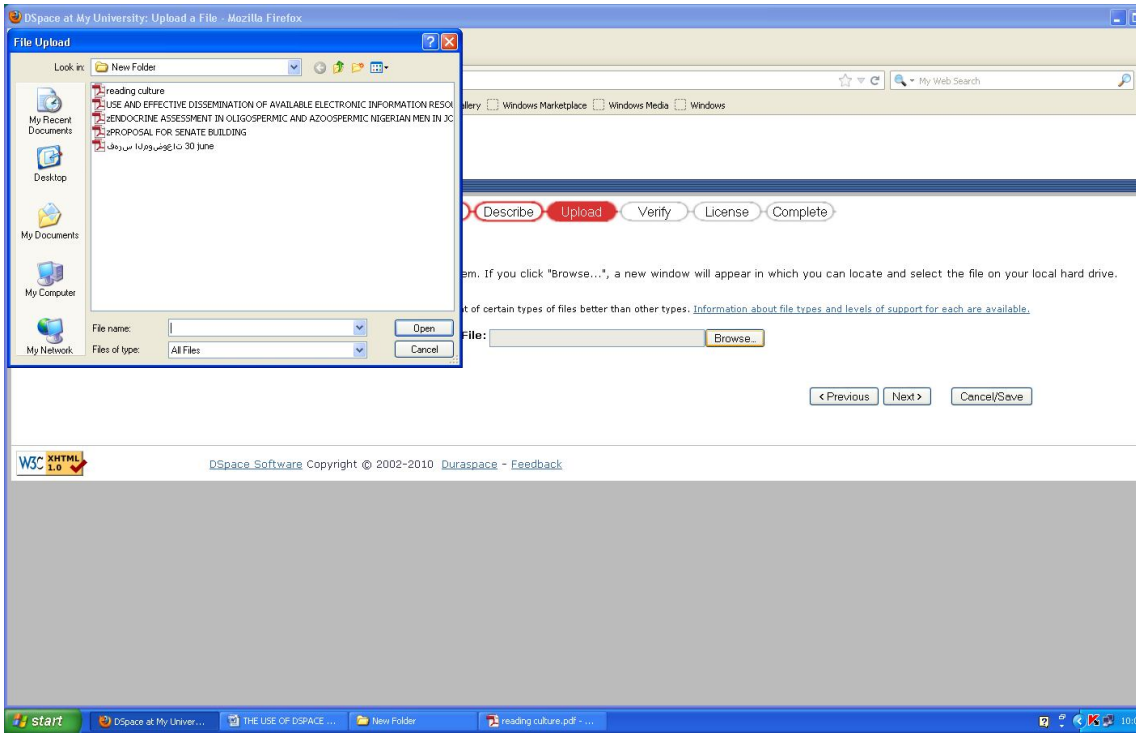
Submit: Upload a File

There are two methods of entering the name of the file one wish to upload:

1. Type the full path and file name into the input box and then click on the "next" button in the lower right hand corner of the screen.

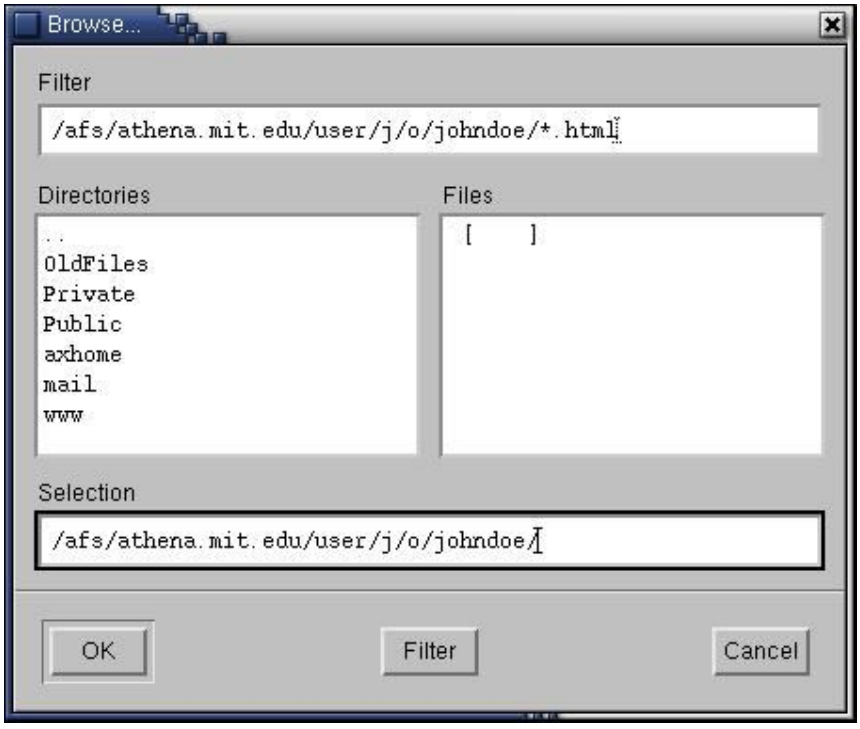
2. Click on the "browse" button and a window showing your files will appear. One can navigate through his directories and folders until he find the correct file to upload. Double-click on the file name you wish to upload, and the name will be entered into the input box.

On Windows computers, the window initially looks like this:

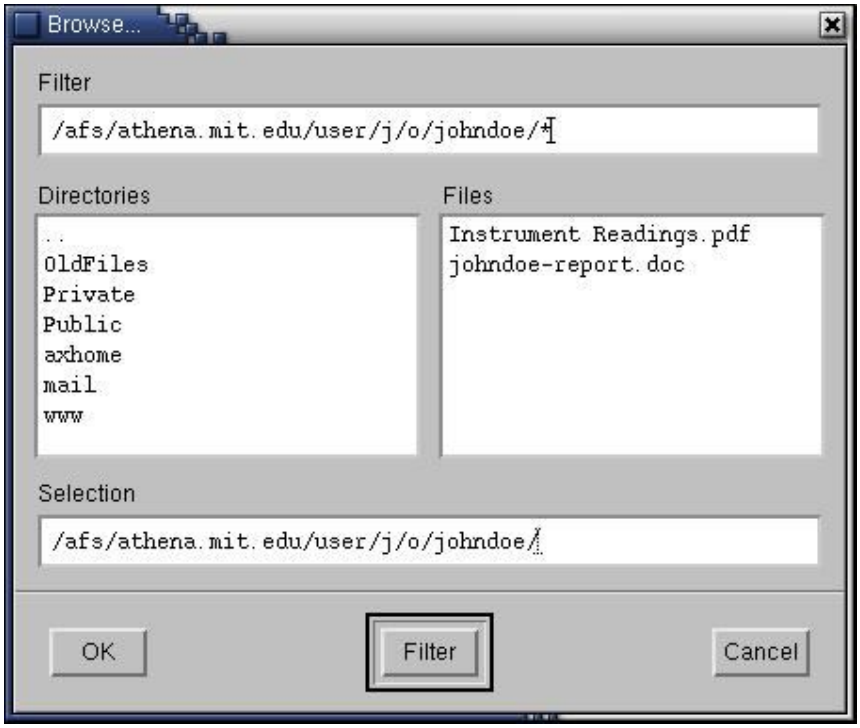


Click on the drop down arrow next to the "HTML Files" and select "All Files

On UNIX systems, the window initially looks like this:



Click to the right of the "html" in the top text box, and remove the ".html" with the backspace key, but leave the "*". Then, click on the "Filter" button at the bottom of the window. You will then see all of your files, like this:



Once the correct file name is in the input box, click on the "next" button to proceed.

File Description

If one specified at the beginning of the submit process that he has more than one file to upload, he will see an input box marked "**File Description**". The information he provide here will help users to understand what information is in each file, for instance, "main article" or "images" or "computer program" or "data set". Enter file descriptions for each item, and click on the "next" button to proceed.

Submit: File Formats

To properly archive and give access to a file, we need to know what *format* it is, for example "PDF", "HTML", or "Microsoft Word". If the system does not automatically recognize the format of the file one have uploaded, he will be asked to describe it. If the format of the file appears in the list offered, click on it and then on "Submit". If the format is not in the list, click on "format not in list" and describe the format in the text box lower down on the page.

Uploaded File

After one has uploaded the file, then check the information in the table to make sure it is correct. There are two further ways to verify that your files have been uploaded correctly:

- Click on the filename. This will download the file in a new browser window, so that you can check the contents.
- Compare the file checksum displayed here with the checksum you calculate.

If you're only uploading one file, click on "Next" when you're happy that the file has been uploaded correctly.

If you're uploading more than one file, click on the "Add another File" button this will appear if you checked "The item consists of more than one file" on the "Submit: Describe Your Item" page. When you are satisfied that all files for this item have been successfully uploaded, click on the "Next" button.

Submit: Verify Submission

This page lets one review the information he has entered to describe the item. To correct or edit information, click on the corresponding button on the right, or use the oval buttons in the progress bar at the top of the page to move around the submission pages. When you are satisfied that the submission is in order, click on the "Next" button to continue.

Click on the "Cancel/Save" button to stop and save your data, or to cancel your submission.

Submit: License

Dspace requires agreement to this non-exclusive distribution license before your item can appear on Dspace.

Submit: Submission Complete

Now that one submission has been successfully entered into the Dspace system, it will go through the workflow process designated for the collection to which he is submitting. Some collections require the submission to go through editing or review steps, while others may immediately accept the submission. One will receive e-mail notification as soon as his item has become a part of the collection or if for some reason there is a problem with his submission.

My Dspace

Is a personal page that is maintained for each member? This page can contain a list of items that are in the submission process for a particular member, or a task list of items that need attention such as editing, reviewing, or checking.

Conclusion

It is necessary for librarians to keep a watch on the developments in ICT and to choose appropriate software that is geared toward the digital age. Open source software is appropriate for libraries because librarians and open source software developers share values of freedom of information and equal access.

REFERENCES

Barton, M. R. and Waters M. M. (2004). *Creating an Institutional Repository*: LEADIRS Workbook. Retrieved from http://dspace.mit.edu/bitstream/handle/1721 accessed on 10/05/2011

Bass, M. J. et.al (2002). *Dspace-Functionality*, MIT. Available at http://Dspace.org/technology/functionality.pdf accessed on 10/05/2011

Morris, M.C. (2009). *Taking Dspace Out-of-the-Box*. Retrieved from http://expertvoices.nsdl.org/duraspace/2009/09/18/taking-dspace-out-of-the-box accessed on 10/05/2011

Pennock, M. (2006). *Dspace Digital repository Software*. Retrieved from http://www.dcc.ac.uk/resources/briefing-papers/technology.../Dspace accessed on 10/05/2011

Walker, J. H. (2010). *Dspace: A Case Study in Sustainability. MIT Libraries*. Retrieved from http://www.oss-watch.ac.uk/resources/cs-dspace.xml accessed on 10/05/201.