

Digital Repository Management System

B.TECH SEM – VII Minor PROJECT
Dept. of Computer Science & Engineering

By

Khushi Shah 19BCP067
Khushi Katariya 19BCP068

Under the Supervision
Of
Dr. Samir Patel



SCHOOL OF TECHNOLOGY
PANDIT DEENDAYAL ENERGY UNIVERSITY
GANDHINAGAR, GUJARAT, INDIA
July – December, 2022

ABSTRACT

In the digital age, libraries need to preserve and provide access to digital resources, but traditional library procedures and tools from the brick-and-mortar type library are often not suited to this task. Physical libraries and their access rely upon a publishing model that has slowly evolved over the past 500 years. In this model, each resource consists of an object in a single format, and each object remains static over time. A digital library exists within a very different framework. A single resource may consist of objects in many formats, yet each of these objects is a resource in its own right.

As the number of documents like academic projects, lesson plans, and question paper grows, it becomes difficult to store, and manage a large number of documents and find requested relevant documents by users.

- So to overcome these problems, we create Digital Repository Management System
- Chance of loss of record is lower and record searching is easy and time-saving.
- Different rights can be given as admin, submitter, reviewer, and user.
- OS Independent

The primary aim of the project is to assist librarians, library managers, and archivists by storing and preserving articles, e-books, resources used for teaching, and audio-visual items in a digital repository. The system will also provide the feature to catalog, index, and retrieve these resources. It will make the resources widely available and more convenient to use for admin as well as users.

This project is specifically aimed at creating a digital repository management system for Pandit Deendayal Energy University. An attempt has been made in creating a web application using various technologies. Currently, features such as adding files, and browsing has been implemented. The technologies used are HTML, CSS, PHP, Bootstrap, and XAMP.

TABLE OF CONTENTS

Sr. No.	Chapters	Page No.
1	Introduction	4
2	Literature Review	5 - 6
3	Proposed Work	7
4	Implementation Details	8 - 12
5	Result and Comparison with existing work	12
6	Summary and Future Directions.	13
7	References	14

INTRODUCTION

On a basic level, a digital repository is simply a collection of digital resources. These materials may have been converted from an analog format, such as paper, or they may have been born digital. Regardless, of the type of resources, it contains, one of the primary functions of digital repositories is to preserve electronic resources, though they must also provide a system for cataloging, indexing, and retrieving digital materials. Unlike physical repositories, which exist to preserve specific information artifacts, the purpose of a digital repository is to preserve access to information artifacts. In contrast to their counterparts, electronic documents do not physically exist- they are a stream of data that must be copied, transmitted across networks, and interpreted by software to be useful. Technology changes rapidly, so the data streams that represent digital resources must constantly evolve in order to remain useful.

Building a repository presents many challenges but it also brings benefits to the library as well. Aside from improving desktop access to materials, it gives libraries a chance to make unique resources widely available. This project “ Institutional Digital Repository System” has the aim of allowing every member, staff, and student of the Pandit Deendayal Energy University the use it. They can access and upload content in order to review it as and when required.

A digital repository can be rewarding for patrons and staff alike if appropriate resources can be dedicated to the project. It can help the library improve its ability to deliver the next generation of library services while making new resources available to more patrons than was possible before. And because these resources are very new, they can make the library more interesting and exciting for both users and staff.

LITERATURE REVIEW

In the following, the four open-access Open sources digital library software are compared based on the characteristics identified in the previous section. The level of support of each characteristic and specific considerations for each DL system are discussed.

1. DSPACE

DSpace is the most popular among the digital library solutions available in the open source domain. The basic entity in DSpace is *item*, which contains both metadata and digital content. Qualified Dublin Core (DC) [8] metadata fields are stored in the item, while other metadata sets and digital content are defined as bitstreams and categorized as bundles of the item. The internal structure of an item is expressed by structural metadata, which define the relationships between the constituent parts of an item. DSpace uses globally unique identifiers for items based on CNRI Handle System. Persistent identifiers are also used for the bit streams of every item. Supports collections of items and communities that hold one or more collections. An item belongs to one or more collections, but has only one owner collection. Dspace stores qualified DC metadata in a relational database (PostgreSQL or Oracle). Other metadata sets and digital content are represented as bitstreams and are stored on filesystem. Lucene supports fielded search, stemming and stop words removal. Searching can be constrained in a collection or community. Also, browsing is offered by default on title, author and date fields. It supports users (e-people) and groups that hold different rights. Authentication is provided through user passwords, X509 certificates or LDAP. Access control rights are kept for each item and define the actions that a user is able to perform. Although DSpace has a flexible object model is not so open in constructing very different objects with independent metadata sets because of its database oriented architecture. The user interface is fixed and provides only minor presentation interventions. Another disadvantage is the full support of only specific file formats as digital content.

2. FEDORA

The basic entity in Fedora is *digital object*. The internal structure of digital object is determined from the Fedora Object XML (FOXML), which is based on Metadata Encoding and Transmission Standard (METS)[9]. Digital object contains metadata and digital content (both treated as datastreams). Fedora supports collections using RELS-EXT DataStream that contains a basic relationship ontology. In this DataStream the relationships between digital objects (like isMember Of Collection or is PartOf) are expressed using RDF. Fedora does not provide a mechanism to manipulate these relations. Metadata and digital content are both considered datastreams of the digital object. Datastreams can be stored (a) internally on the digital object XML file, (b) on filesystem as managed content or (c) on an external source. One

or more metadata sets can be concurrently used, while different file formats can be stored as separate datastreams in a digital object. Default indexing is provided for the DC metadata set and digital object's system metadata (persistent identifier, creation/modification date, label, content model). Indexing and searching is managed from a relational database (MySQL, Oracle or PostgreSQL). : The web interface of Fedora provides a search environment to the end-user, where he/she may execute simple keyword or field search queries. The default view of digital objects is restricted to the presentation of the system metadata and the datastreams. It supports users and groups authorized for accessing specific digital objects using XACML policies. Authentication is provided through LDAP or for specific IP addresses.

3. GREENSTONE

Basic entity in Greenstone is *document*, which is expressed in XML format. Documents are linked with one or more resources that represent the digital content of the object. Each document contains a unique document identifier but there is no support for persistent identifiers of the resources. : A collection in Greenstone defines a set of characteristics that describe its functionality. These characteristics are: indexing, searching and browsing capabilities, file formats, conversion plugins and entry points for the digital content import. There are also some characteristics for the presentation of the collection. The representation of hierarchical structure in text documents is supported for chapters, sections and paragraphs. Both documents and resources are stored on filesystem. Metadata are user defined and are stored in documents using an internal XML format. Indexing is offered for the text documents and specific metadata fields. Searching capabilities provided for defined sections in a document (Title, chapter, paragraph) or in whole document. New collections and the contained documents are built using the Greenstone Librarian Interface or the command line building program. The default web user interface provides browsing and searching into collections, navigating into hierarchical objects (like books) using table of contents. Presentation of documents or search results may differ depending on specified XSLTs.

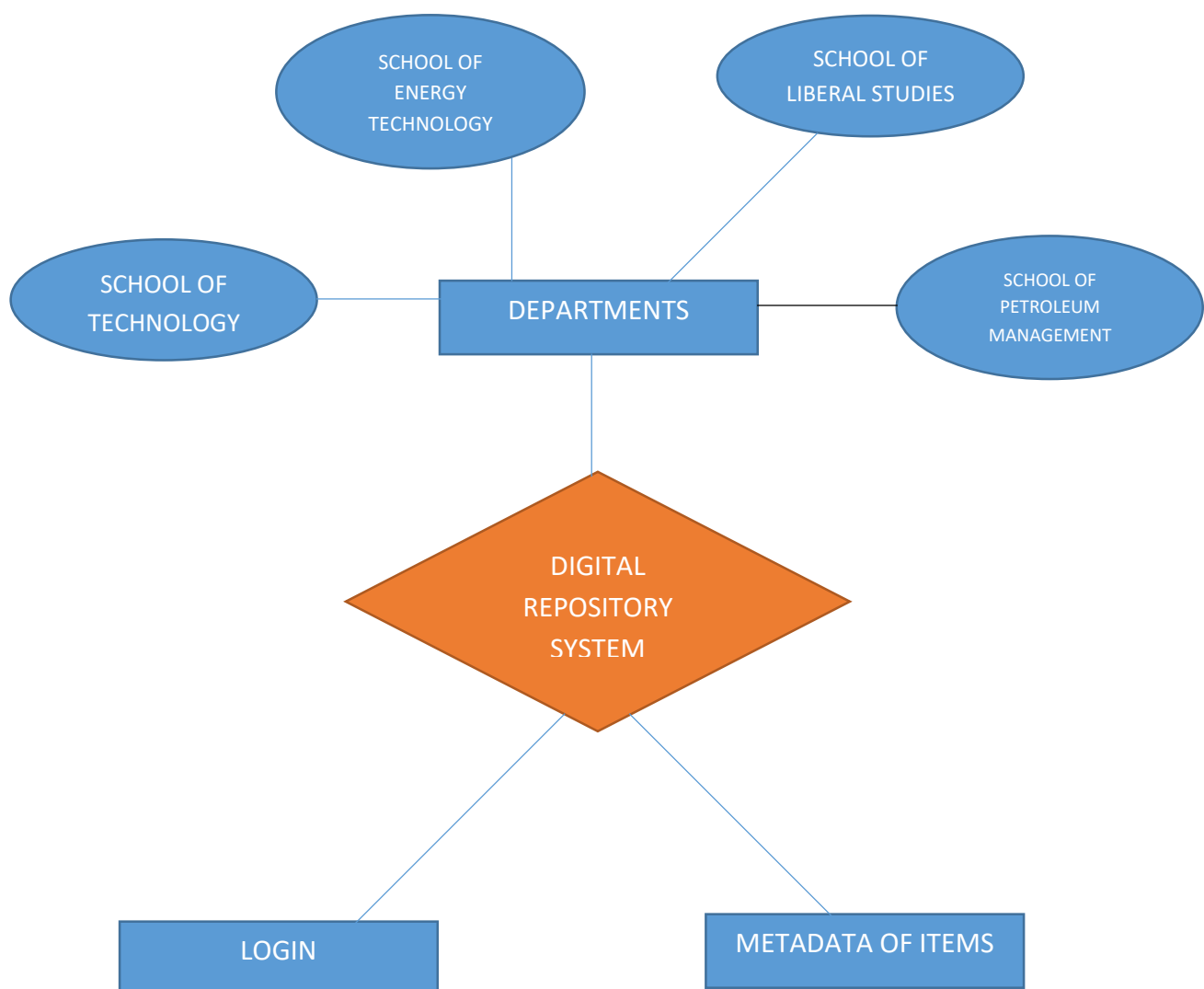
4. EPrints

Basic entity in EPrints is the *data object*, which is a record containing metadata. One or more documents (files) can be linked with the data object. Each data object has a unique identifier. There is no consideration of collections in EPrints. Data objects are grouped depending on specific fields (subject, year, title, etc). There is no definition of relations between documents, except using URLs in specific metadata fields. Metadata fields in EPrints are user-defined. The data object, containing metadata, is stored in a MySQL database and the documents (digital content) are stored on filesystem. Indexing is supported for every metadata field, using the MySQL database. Full text indexing is supported for selected fields. Combined fielded search and free text search are provided to the end-user. Browsing is provided using specified fields (e.g. title, author, subject). A default web user interface is provided for the creation and editing of objects. Authority records can be used helping the completion of specific fields (e.g. authors, title). Objects can also be imported from text files using multiple formats (METS, DC, MODS, BibTeX, EndNote). The web user interface provides browsing by selected metadata fields (usually subject, title or date). Browsing can be hierarchical for subject fields. Searching

environment allows user to restrict the search query using multiple fields and select values from lists.

PROPOSED WORK

Currently, one module has been prepared for the current project. The module of uploading the documents has been completed in this project. It is basically uploading various documents such as books, pdfs, etc. into the web application so that the users can see the uploaded documents. The project is specifically designed for the use of librarians and library users. The product will work as a complete user interface for digital repository management. Below is the flowchart of the



IMPLEMENTATION DETAILS

1. FrontEnd

For the frontend side of the project HTML, CSS, and Bootstrap was used.



Figure 1: Home Page



Home Browse Support/Help Login



DIGITAL REPOSITORY
PANDIT DEENDAYAL ENERGY
UNIVERSITY

Login to CRM

[New User? Click here to register!](#)

Please enter your e-mail address and password into the form below:


E-mail Address:

Password:

[Log in](#)

[Have you forgotten your password?](#)

Figure 2: Login Page



Home Browse Support/Help Login

Home / My Profile

Edit your Profile

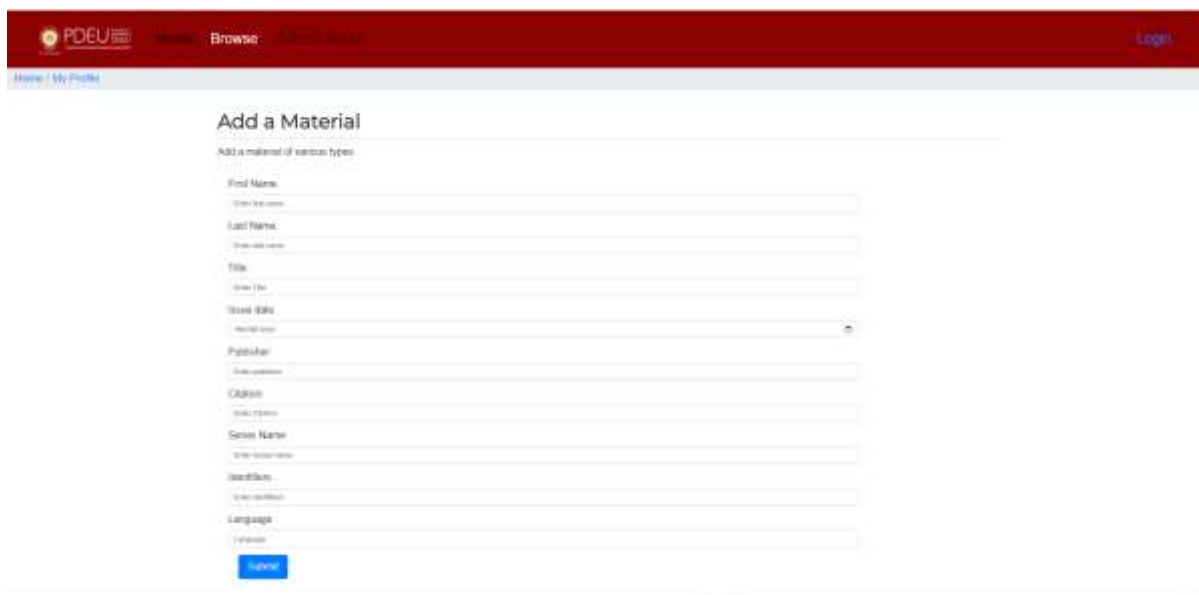
Profile: Khushi Shah

[Edit & Save Profile](#)

My Submissions

Issue Date	Title	Author	Submitted/Unfinished
01/11/22	MDI Paper	Khushi	Submitted

Figure 3: Profile Page



Add a Material

Add a material of various types

First Name: [Create New](#)

Last Name: [Create New](#)

Title: [Create New](#)

Date Added: [Create New](#)

Publisher: [Create New](#)

Author: [Create New](#)

Series Name: [Create New](#)

Location: [Create New](#)

Language: [Create New](#)

Figure 4: Metadata Page

2. Backend

XAMPP stands for cross-platform, Apache, MySQL, PHP, and Perl. XAMPP is a free and open source web server that allows you to develop, test, and build websites on a local server.

To run PHP for the web, you will need to install a web server like Apache and a database like MySQL – and both are supported by XAMPP.

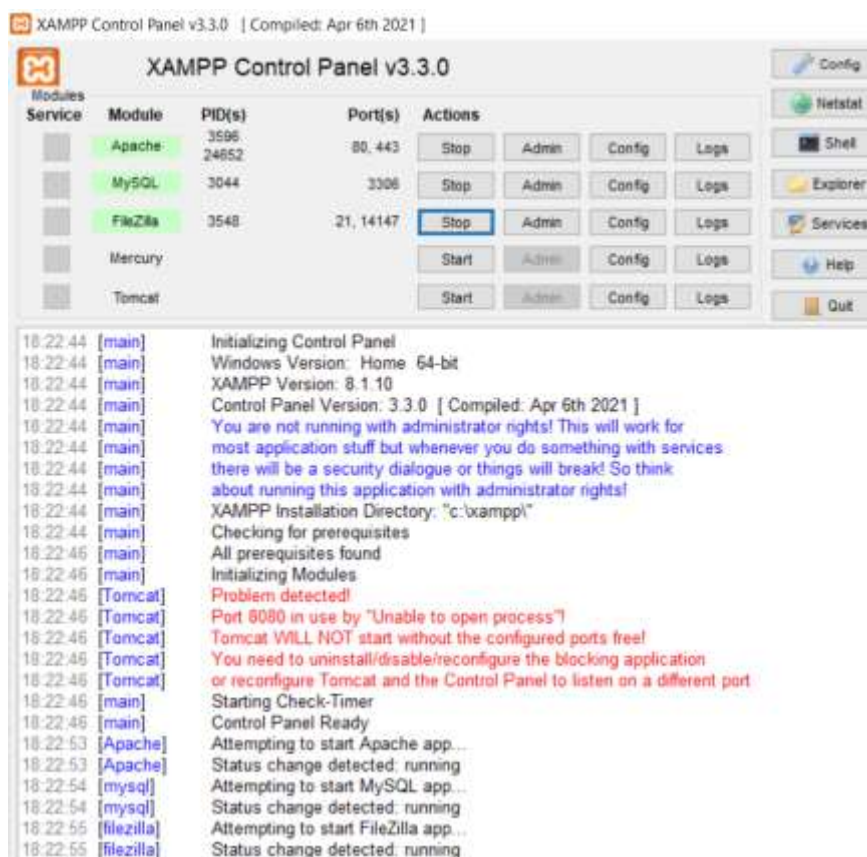


Figure 5: Control Panel XAMPP



Figure 6: XAMPP running on localhost

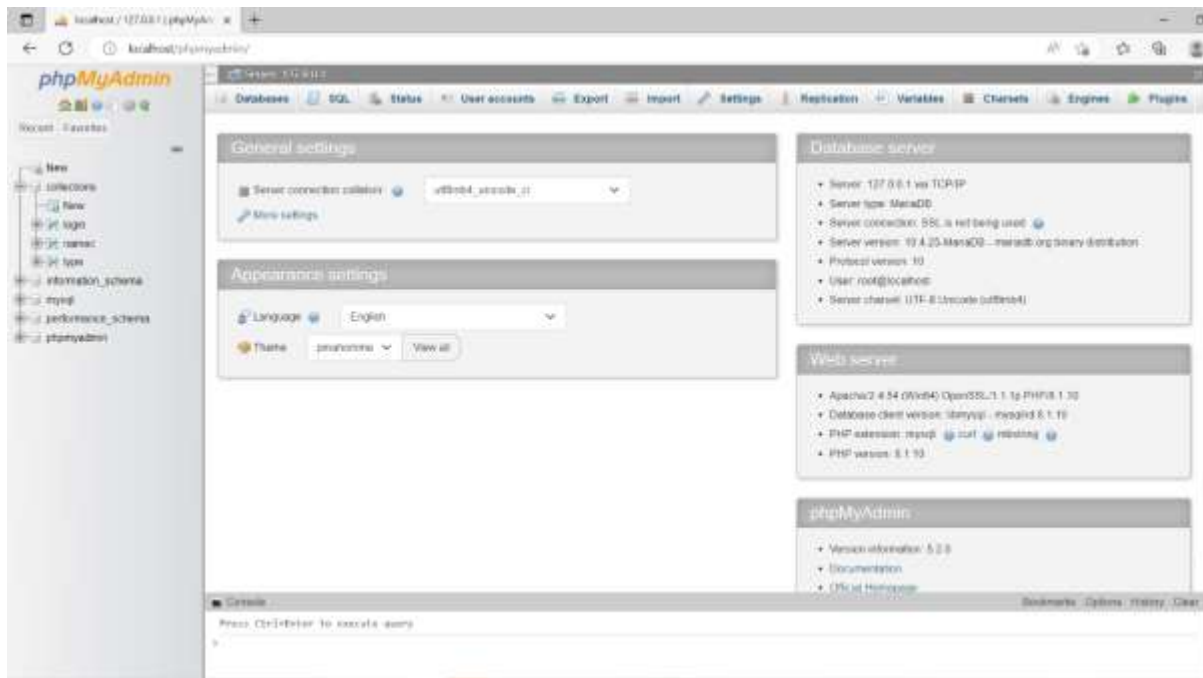


Figure 7: phpMyadmin page

In the phpMyAdmin page, we have created a database named collections that contain two tables titled Login and Type.

Login Table contains information about the username and password of the person that has registered once and wants to login again

Type table contains the metadata of the files like author name, issue date of book, series name of book, collection name where the file is to be uploaded etc. One can see this data on display.php page.

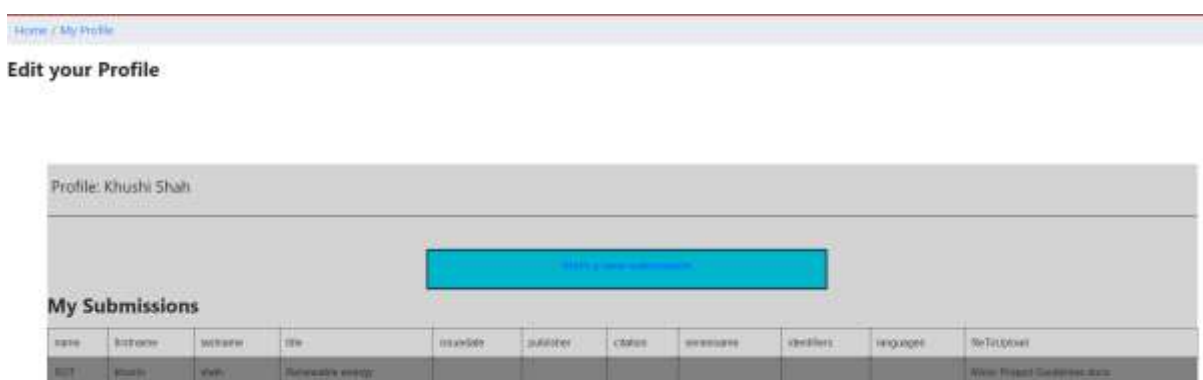


Figure 8: Display php Page

Db_conn.php contains the code to connect with database using mysqli_connect

```
<?php

$name = "localhost";
$username = "root";
$password = "";
$db_name = "collections";

$conn = mysqli_connect($name, $username, $password, $db_name);

if (!$conn) {
    echo "Connection failed!";
    exit();
}
```

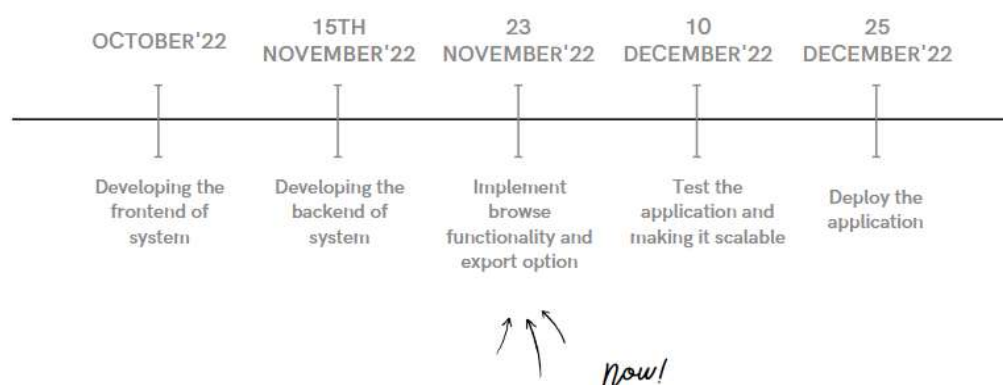
RESULTS AND COMPARISON WITH THE EXISTING WORK

- We are making the Digital Repository Management system customized for PDEU , i.e., University-specific. As Dspace that is used currently does not provide a proper hierarchy of collections.
- We are making it platform Independent and open Source unlike some softwares such as Omeka that works only on Linux.
- We are implementing precise control limits to be placed on staff and user functions. Admin, staff, and users so that unwanted materials are not uploaded by unauthorized users.

SUMMARY AND FUTURE WORK

The system provides a computerized version of library resources management system which reduces human work and need of physical space for storing resources. This project facilitates users to access book and other resources information based on technology and domains, this project is incorporated with user friendly interface which increases efficiency of repository management system.

The design and implementation of a Digital Repository Management System for the Pandit Deendayal Energy University has provided a solution to most of the challenges encountered in keeping hardcopy of books, reports, exam papers etc., and also aided in the process of digitization of different work. This repository system tool is expected to impact positively on the process of digitization of research papers, books, papers, thesis etc.. The system helps librarians and other people that are concern with the process of digitization and making items more accessible to students, faculties



REFERENCES

1. <https://upguage.com/product-details/Document-Management-System> (Image 1)
2. <https://minds.wisconsin.edu/handle/1793/38291/recent-submissions>
- 3 . <http://ir.pdpu.ac.in:8080/jspui/>
4. <https://merlot.org/merlot/index.htm>
5. <http://pub.demo.eprints-hosting.org/>
6. <https://www.greenstone.org/>
7. <https://www.figma.com/file/MLxfjopUeAgyIFY78ER9na/DRM?node-id=115%3A47>
8. Sahu, Ramani & Karadia, Alekh. (2015). COMPARATIVE STUDY OF AN OPEN SOURCE DIGITAL LIBRARY SOFTWARE: DSPACE, GREENSTONE AND EPRINT.
9. https://www.researchgate.net/publication/271647555_Institutional_Repositories_Software_for_Digital_Libraries_in_the_Digital_Environment