The Study of Open Source CMSs

By

CHETAN GOPILAL JAIN

A thesis submitted to the

Graduate School-New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Master of Science

Graduate Program in Electrical and Computer Engineering

written under the direction of

Prof Deborah Silver

and approved by

_____

_____

_____

_____

New Brunswick, New Jersey

May, 2010

ABSTRACT OF THE THESIS

The Study of Open Source CMSs

By CHETAN JAIN

Thesis Director:
Professor Deborah Silver

In this thesis, we evaluate different open source content management systems (CMSs) and determine their appropriateness for scientific research laboratories' website content management. We describe different CMSs and evaluate them based on the following criteria: ease of installation, usability, maintenance and updates, scalability, community strength and support, user roles and workflow, security, and Web 2.0 features. We then choose of these system, Drupal, and demonstrate its effectiveness for two different scientific websites, Bio-1 and Vizlab. Drupal allows integrating new features using community contributed modules and easy future up-gradation. Successful implementation of both projects using Drupal highlights the importance of Open Source CMSs.

Acknowledgement

I would like to thank my advisor, Prof. Deborah Silver, for her support and encouragement while writing this thesis.Also, I would like to thank my parents and family who provided me with a strong educational foundation and supported me in all my academic pursuits. I also acknowledge the help of VIZLAB at Rutgers.

Table of Contents

List of Figures

Chapter 1

Introduction

In early 2000, websites were built using various languages such as XHTML, CSS, PHP, ASP, and JavaScript, which lead to several problems and limitations such as unmanageable code, increased risk of data loss, and making it harder to update and implement basic required Search Engine Optimization (SEO) techniques [27].For small frequent changes, these tasks required hiring web developers: an expensive labor. The complexities created the need for a system to mask all the technical details and to provide a user friendly platform that organizations could use to implement basic changes. Out of this need, Content Management Systems (CMSs)were invented.

CMS platforms are frameworks which build custom content applications based on an organization's needs. The content management system provides the ability to maintain and navigate the structure of the site to the users, and allows the developer to develop the website more easily. The functionality of widely used open source CMSs are explained through the structure of files.

One of the crucial limitations of a management system is that there are numerous CMSs available as open source software and they are free to download and install. Due to this, it is becoming more and more difficult to choose and implement one particular content management system procedure for a company or an organization. The challenge in correctly implementing a CMS is both the author’s utilization of the appropriate

management system platform to create the content and the successful access by the end-user to the published website. This research is a comparative study of the most widely used CMSs such as WordPress [23], Joomla [13], and Drupal [5]. It is quite difficult to draw a conclusion as to which CMS is better as there are several pros and cons associated with each CMS. This study attempts to analyze each CMS on the basis of usage, design, performance, scalability, compatibility with different platforms, security purpose, and search engine visibility. It also highlights the statistics for different CMSs such as number of users, cost of set up, contribution by community, and its popularity. We also present an installation guide for each CMS. After the analysis we chose Drupal for implementing two science based websites called “Bio-1” and “Vizlab”.

Bio1 was developed to demonstrate Drupal’s scalability and how its framework can be used to develop complex web applications. Basically, Bio-1 is a job portal application which connects the employer and the employee from Bio Industries [3]. Approximately twenty-five Drupal modules were implemented in this project. Bio-1’s project reflects the detailed approach and how the problems were addressed later in this thesis.

Vizlab is a website, developed for Visualization Lab at Rutgers University which gives information about the lab, its research, publication, professor and students. This laboratory website demonstrates how easy it is to organize the lab’s research-related data for an administrator with just a few clicks and to allow users to search for publications using keywords.

From the user’s perspective, Drupal has great features such as taxonomy, module system, node system, CCK, and views. Drupal Access Control System was an excellent

framework for these particular websites and appropriately met the needs of the project. It allowed full control over user access permissions by creating roles and assigning what users could do. From the developer's perspective, Drupal is more of a framework and it can be easily scaled for complex sites.

The website development and testing was done on a local server, eliminating the need for security and File Transfer Protocol (FTP) file uploads via the Internet. XAMPP, a local server, is a development tool that allows website programmers and designers to test the work on computers without needing Internet access.

1.1Features of CMSs

Drupal, Joomla and WordPress are three widely used CMSs. They are analyzed in detail in the below areas.

1. **Installation:** For every CMS, one needs to download the code package of that system and install it on server where the website is hosted. For someone who is not a programmer or coder, ease of installation is a concern. There are many web hosting service providers which provide and maintain servers as well as connect the website safely and securely to the web. Costs will vary depending upon the specific hosting needs of the website. As each CMS has different technical requirements, a company's budget can impact which webhost is chosen.
2. **Maintenance and Update:** An administrator needs to look for upgrades for the CMS for various added features, security updates, and fixed bugs. The major releases of CMSs are planned, but they can occasionally break the page template and some plug-ins. It is beneficial to have a CMS that supports and provides security updates for

older versions along with newer version so those who do not want to upgrade the system immediately can continue with older version for time being. WordPress and Drupal release upgrades more frequently and support previous versions with security updates and bug fixes [1].

3. **Community Strength and Contribution:** Open Source systems are created through the efforts of community developers and users, although more recently, they are being supported by consultants and organizations. The CMS's community strength plays a vital role in choosing a CMS as one can get answers easily to his questions through active community forums and discussions. There are a number of community contributed plug-ins and themes available which are ready to use. With such a vast community, users can be sure that the system will continue to be supported in the near future. There are also numerous books and tutorials available as a reference to learners and developers.
4. **Usability:** The ease of changinga website's content is the key factor for a company or organization's preference to a specific CMS. For non-technical users, the simplicity of the admin area of CMS is of the utmost important so the staff members can be trained to add some section or page to the site, upload images, or send newsletters very easily and fast [19]. Formatting and styling the content is done through content editors. Many CMS distributions also provide support for Microsoft Word format, so users can easily copy and paste content.
5. **Scalability:** Scalability is the property of a CMS that describes how efficiently it can handle growing traffic on websites and how quick it can be expanded [6]. Websites

with tens of thousands of visitors a day will require a highly scalable CMS to meet their needs. Almost all CMSs cache the data so pages can be loaded quickly.

6. **Web 2.0 Features:** Content Management Systems provide many features for a website to interact with its users and visitors. For example, visitors can post a comment on the website, create blogs, and request RSS feeds to get content. Even site administrators can display the whole content or a part of another website through RSS feeds [1]. Moreover these CMS features can be very useful for any social networking site.
7. **Security:** To have a secure system is very important. There are many types of attacks on website such as SQL injections, link hacks, and denial of services. The ideal system finds these type of attacks very quickly and stop them from spreading [21].
8. **User Roles and Workflow:** For complex websites, the content is categorized and maintained by more than one user. If a CMS can provide variable permissions to a hierarchy of users to create, edit, or publish data, it will be a useful feature for a website administrator. For many companies, it is beneficial if the site administrators can oversee the work flow. For example, an administrator could need to see what needs to be done by a user and have the ability to post a "to do list" that can remind that user of any work that may need to be reviewed or created. Reversion of content is also a useful feature as it can save lot of time in case an administrator needs to revert back to the prior settings of a website. Drupal provides the highest control of user roles and permissions so site administrators can be very specific about roles and permissions. Joomla and WordPress have limited user types and can't limit users' specific permissions [1].

1.2 Overview of Material

This thesis continues with a general overview of Open Source CMSs and its importance (Chapter 2). In Chapter 3, we describe widely used Open Source CMS, Drupal and its key features. At the end of the chapter we present a quick installation guide of Drupal on local server. In Chapter 4, we discuss about the Joomla CMS, its key features and area of improvement in detail, followed by a description of the installation guide for Joomla on local server. The key features of WordPress CMS are discussed in detail in Chapter 5. Chapter 6 is devoted to the summary of all the three CMSs' key features based on survey of small set of corporate professionals from different industries. The two interviewers were from media industry, three from education industry and one from non-profit organization. Finally, in chapter 7 and 8 we describe the implementation of two science based websites called “Bio-1” and “Vizlab”. We conclude the thesis in Chapter 9.

Chapter 2

Open Source Content Management Systems

CMSs are software packages that let the user build and maintain a website easily in less time. An Open Source CMS is free to download and use, meaning its source code can be modified to meet project requirements with no obligations. CMSs store the data in databases to make maintaining a website less cumbersome as there are no separate files for each website page. As the content is pulled dynamically from the database it is very easy to control the placement of it on the website by setting visibility rules in advance.

A CMS separates the content from the UI (User Interface) designs by using a theme layer which is responsible for rendering page layout and CSS styles associated with it. Users can use either one of the default themes that come along with the package or a custom theme for the website. Generally, a page run by a content management system is divided into different regions. The most common regions are header, footer, right-sidebar, left-sidebar and main content. Additional regions can be added by modifying the template related files. By default, every theme comes with a single template for all pages. But in most of the CMSs one can have a custom template for a particular page. For example, it allows a user to create a 3-column layout for his home page while using a 2-column layout for all other pages by adding extra template files for home page.

CMSs are designed to be scalable and modular. Every Open Source CMS comes with set of default modules which gives the site its basic functionality. The base functionality can

be further enhanced by integrating community contributed modules or by developing custom modules.

Despite its limitations, Open Source CMSs are widely used by different industries due to its low-cost and its useful savings in time and resources.

2.1 Importance of Content Management Systems

The most important factor of CMSs is that it allows the non-technical staff member the ability to update the website content in less time. Apart from this, there are a number of other benefits in using a CMS:

1. The task of organizing data becomes much easier for larger websites.
2. For large websites, it alleviates the requirement of having to back-up all the .html files as the data is usually stored in databases.
3. CMSs eliminate the need to know how to code either in HTML or CSS. The installation process might require some technical knowledge and coding to customize a theme or template but in the long term using a CMS requires less coding with HTML, CSS, and server side scripting. This saves time as well as resources.
4. The ability to work on server with a CMS saves lot of time as there is no need to upload a website. There are fewer chances of losing or overwriting documents or files accidentally because the entire database is on the server. The only thing which needs to be considered is the backup of databases before upgrading the CMS.

5. Search engine optimization (SEO) is a practice of the webmaster that helps the website appears in relevant searches with high rankings. CMSs also provide features that help SEO. Previously, the general practice for SEO was done by getting links on external sites but the other key for a high ranking in search engine is to have an internal link that points to the landing page of the website. There are plug-ins and modules available for SEO that can optimize URLs and title tags [11].
6. The major benefit of using aCMS is one can access or edit the site from any place. Because they are entirely on web server, there is no need to depend on computers where files are stored or having to carry the site's files. By logging into the website admin area, one can edit the content in the back end and publish the pages promptly. This way, any change in the web management system can be viewed on the website in real time.
7. CMSs helps to save money by reducing the amount of time web developers have to change data such as correcting a spelling mistake, sending newsletter, or adding an article or event to a calendar. The admin area is easy, so user can add pages and images, do spell checks, send a newsletter, etc. Moreover, it saves time that is required to communicate the task to the web developer.
8. Compared to a custom CMS, which is very costly depending upon the requirements of an industry, there are Open Source CMSs available which are free to download and use.

2.2 Popular Open Source CMSs

The popularity of any Open Source CMS depends on the following important factors:

- How large and active is the community?
- How easy it is for user to install and update the package?
- How scalable and modular is the CMS?
- How user friendly it is for non-technical person?

Considering the above mentioned points, this research outlines the most widely used open source systems. All of the CMSs discussed below, are written in PHP and use MySQL as the database.

WordPress [23]:

WordPress is a very popular Open Source CMS for blogging applications. It is very easy to install and understand, and is easy to maintain and update the content. It has many predefined themes which can be used or it can be modified with some knowledge of HTML and CSS. WordPress has a very robust community which has contributed many add-on modules to extend its basic functionality. Like others, WordPress is also SEO friendly and has add-on plug-ins to make it more effective in that area. However, WordPress is not large enough to support complex sites as compared to other Open Source CMSs described here. It is a very good tool for blogging and displaying pages with an easy to use administrative area.

Drupal [5]:

As compared to WordPress, Drupal is more scalable and is a great choice for a complex website. Its huge community has contributed more than six thousand modules. Anyone can use these modules to make complex sites such as fully functional e-commerce system or a job portal. Drupal has strong support for Web 2.0 and community functionality, including groups, blogging, forums, and social networking. Also, creating and publishing a page is very easy. It is also very easy to create different roles with different set of permissions. Compared to others, Drupal has very powerful visibility settings through which admin can control what content to display on what page. Drupal also comes with default themes as part of core package and contributed themes which can be tailored to meet an individual need. Installing Drupal is not as easy as WordPress or Joomla.

Joomla[13]:

Joomla is also backed by a large community and is relatively easy to install and setup. Components of Joomla 1.5 follow the design pattern of Model-View-Controller (MVC) [16]. This layout has three distinctive zones: data collecting, representation, and interaction. These zones work independently and interact with each other, so expansion or modification of properties in any section can be done without changing other sections. One can easily develop a wide variety of functionalities, from shopping carts to community features, providing a solid base for many different kinds of sites. Joomla is good in terms of scalability to support complex sites but not as well as Drupal. Compared to Drupal, Joomla is not as flexible and has a very advanced structure. Also, it cannot be edited by internal content authors.

Chapter 3

Drupal CMS

Drupal is an Open Source CMS written in PHP and uses MySQL, PostgreSQL or MS SQL for database. Drupal can be setup on Linux, Windows or Macintosh OS. It is distributed under GPL ("GNU General Public License") [5] and is free to download. It is maintained by a large community of thousands of developers and users all over the world. Drupal has around 700 contributed themes and 6000 contributed modules which lets the user make a variety of sites ranging from a personal blog to a complex e-commerce system [5].

The architecture of Drupal is designed in such a way that the three different layers work independently and correlate with each other to give the final output. These three layers are the content which forms the website, the application algorithm that organizes this content for presentation, and the representation layer which is incorporated by the Drupal theme system. The webpage that comes to a viewer's browser goes through a sequential process in which Drupal modules take all the relevant content from the databases and then the theme gets ready for the final presentation. Un-styled HTML code is passed through template file to theme the content. Unlike Joomla, Drupal's architecture does not follow the design pattern of MVC but instead follows the Presentation-Abstraction-Control (PAC) [6].

Drupal CMS provides basic features in its package known as Drupal core [2]. These contain the ability to create user accounts and assign a hierarchy of permissions, including creating menus, customizing the themes of pages, and creating an RSS feed. Drupal provides a platform to create various types of websites such as brochure ware websites, blogs, forums, and social networking websites. Drupal's core functionalities and capabilities can be enhanced through add-on modules. These add-on modules are also known as plug-in modules which can be installed in the core package of Drupal. As Drupal adds new features, it becomes more similar to the feature requirements of a web application framework, and can also be called content management framework.

For basic website installation and administration in Drupal no coding skill is required. For developers and coders, Drupal gives a programming interface and environment. The Drupal core package is designed to be extended by a community of users. In Drupal's default set up, the website content can be supplied by either authentic or anonymous users and displayed to web visitors by a multiple of selectable criteria including by date, category, searches, etc. Drupal core also incorporates a hierarchical taxonomy system, which permits content to be classified or labeled with key words for easier access. Drupal conserves a detailed change log of core feature updates by each release.

The default Drupal package that comes with the Drupal installation includes core modules which can be used to enhance the features and functionality of the website. The default Drupal installation package includes various features such as:

- Blog and Forum
- Clean URL

- User Registration feature with different roles
- Search functionality
- Multi-level Navigation System
- Support for Open ID
- Feed Aggregator and RSS FEED
- Language Translation tool
- Workflow tools like Actions and Triggers
- Auto alert about security and new release update

Along with the core modules, the Drupal default package [5] also provides some core themes. Administrator can select these themes through a menu and customize the look and feel of the website. There are several modules added to Drupal core for non programmers to customize the themes to a certain level. For example, the Color Module helps the non-technical administrator to customize the color of theme through the browser interface.

3.1 Drupal File Structure

It is very essential to understand the basics of how the directories and files are structured on a Drupal site. When Drupal is installed, there will be a default file structure either on the local machine or on the server. Below is an example showing how each folder has all the important documentation structured and organized.

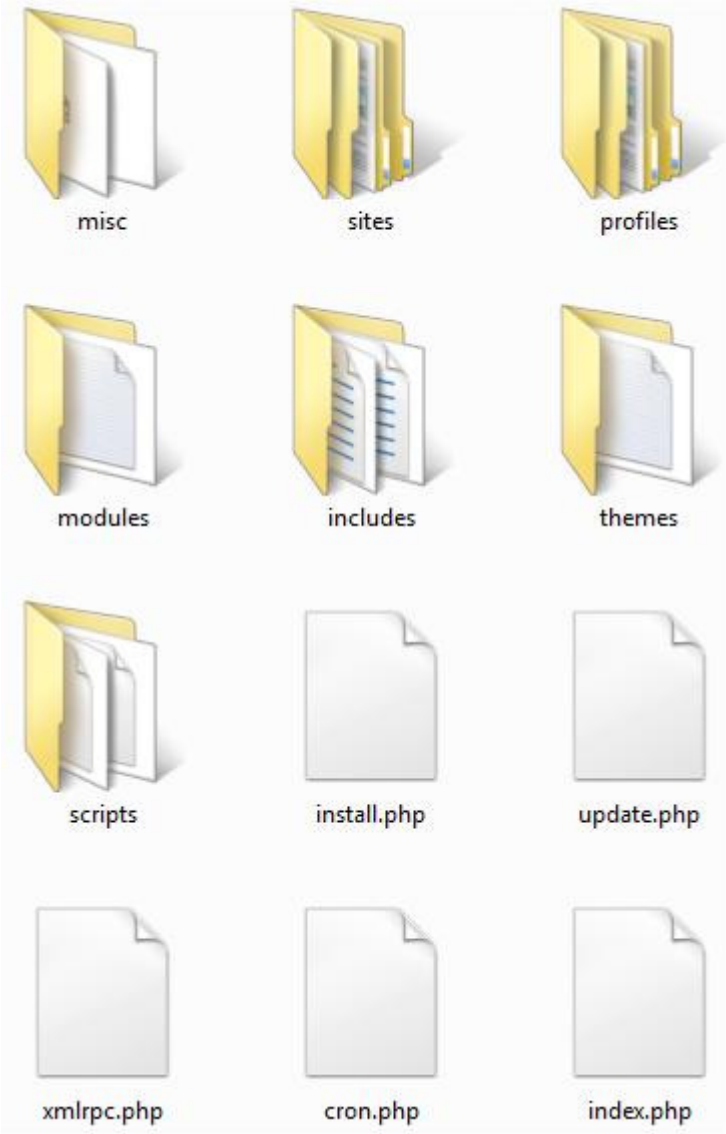

Figure 3.1: Drupal File Structure

The files are described as below:

Misc folder: It contains the core modules and each module is in its own folder.

Sites folder: This folder contains settings, modules, and themes as user's modified Drupal. Contributed modules or custom modules are stored in sites/all/modules. This folder keeps all the modification for that particular site. Inside the site directory, there is a

subdirectory called “default” which contains the default configuration file for Drupal site. Based on the information provided the Drupal installer modifies the default directory and write a setting.php. So the final settings file would be at sites/www.yoursitename.com/settings.php.

Profiles folder: It contains various installation profiles for a website. If there are other profiles besides the default profile in this subdirectory, Drupal will ask which profile user you would like to install.

Modules folder: This folder contains all the core modules.

Includes folder: This directory includes libraries of common functions used by Drupal.

Themes folder: This folder contains default themes and template engines. Additional themes which are downloaded or created should go to sites/all/themes.

Scripts folder: This folder contains scripts for several tasks such as cleaning up code, syntax check, to run Drupal from command line and handling special cases with cron.

Themes folder: This folder contains default themes and template engines. Additional themes which are downloaded or created should go to sites/all/themes.

cron.php file: This file executes periodic tasks, such as pruning database table and calculating statistics.

index.php: This file is the first entry point for Drupal.

Install.php: This file is the first entry point for the Drupal installer.

update.php: When a Drupal version is upgraded, this file updates the database schema.

xmlrpc.php: This file receives XML-RPC requests. If deployment does not intend to receive this XML-RPC request, it can be deleted safely.

Robots.txt: This file is the default implementation of the robot exclusion standard.

3.2Drupal's Key Features

Drupal's key features are as follow:

1. **Advanced Control of URL**: Drupal provides a precise control over URL structure. Each content item which is called node in Drupal can be given a custom URL. The path auto module can automate custom URL structure for each content type.
2. **Custom Content Types and Views**: Using the Views module and the Content Construction Kit (CCK), we can create new content type without writing any code. There can be created any number of custom content types and can be displayed in many different ways. Some examples of content types are forum posts, tutorials, blog spots, news stories, classified ads, and podcasts.
3. **Revision Control**: Drupal can be configured to save a new version of a page every time a page gets edited. So a user can revert back to old version if needed.
4. **Taxonomy:** Because of the built-in taxonomy feature, users can easily organize and tag content. This allows cross categorization and tagging which is very unique [9].

5. **Access Control System:** Drupal comes with very powerful user roles and access functionality. It allows the user to create different roles with different levels of permissions.
6. **Page Titles and Meta Tags**: One of the unique features of Drupal is that it gives the admin total control over the *"Page Title"* and the individual meta tag for each page.
7. **Documentation**: Drupal has excellent documentation. There are numerous tutorials to help the novice start quickly.
8. **Themeing and PHP Template**: Themeing in Drupal can be done without any PHP knowledge. Drupal uses PHP template theme engine by default.
9. **Community strength and support**: There are Drupal Forums[5] and discussions which are highly active and one can ask questions and get answers. These communities are driven by the nonprofit Drupal Foundation.
10. **Hook System**: This system in Drupal enables the user to hook in new modules easily. This hook system is invoked when any activity is done in Drupal. That action sends information to other modules which instruct them to perform a task [9].
11. **Filter System**: This system provides control over what content should be allowed to be viewed by anonymous user and admin users. One example would be if one wants to provide HTML control to admin user and filter that from an anonymous user. This feature helps to secure the website.

3.3 Areas for improvement:

Drupal has some limitations.

Performance

Sometimes Drupal's performance can affect the website especially on shared hosting. More resources are required from the server when Drupal has more modules added and hence this can slow down the website. But by including the JS aggregation module and block cache in Drupal the performance has been improved a lot.

3.4 Drupal Users:

Apart from small and midsized websites, Drupal can be scaled for high traffic websites. Some of the big companies and organizations which take advantage of Drupal are Sony, Warner Brothers Record, The United Nations, The Discovery Channel, The NHL, Forbes, Oxfam, Yahoo, AOL, and The Grateful Dead [5].

3.5Analysis of Drupal

1. **Installation:** For Drupal installation, the user does not require more technical knowledge rather than just how to connect through FTP and install databases. The installation time for the new user with knowledge of general installation of other systems might be less than half an hour.
2. **Maintenance and Update:** In the Drupal CMS, maintaining and updating site is easy. The procedure for updating the website includes the backup of the website, and then replacing the files using a web update interface. By downloading only

one files which contains the assets of website; administrators can back up the whole website. The database can be stored in PHPMyAdmin. There have been only two major upgrades in Drupal to date. Drupal gives a notice to the site admin when ever any upgrades are required. A major upgrade may affect the current template or plug-ins. Drupal also provides security updates for previous versions in case the administrator doesn't want to change the version.

3. **Community Strength and Contribution:** The community of Drupal is very large in terms of its users and developers. There are more than six hundred fifty thousand users and two thousand developers have signed up on Drupal.org. The Drupal conference is held as "DrupalCon" twice a year in North America and Europe. The community supports many active forums and discussion groups where anyone can post a question and get answers. There are many books and tutorials available for Drupal learners through the community. Apart from these, Drupal also maintains IRC channels for live chat support.
4. **Usability:** An administrator can easily access the page or section in visitor mode by just by clicking the edit button. Accessibility to an admin area requires some learning, as its default setting does not have a refined look. The core package does not include any editor like WYIWYG (What You See is What You Get) but one can be installed as a plug-in. In Drupal editing pages or sections are the same, just while adding new page one may need to link it by hand. It doesn't support Microsoft Word format in the core package, so module one can't just copy-paste the text without adding support. According to community, the new version of Drupal, Drupal 7 will be created with an easy to use admin area [5].

5. **Scalability:** Drupal is highly scalable with high traffic handling capabilities. Its WebPages are cached indefinitely as the default setting configuration but can also be manually cached for a specific time. Moreover, functionality area blocks can be cached.
6. **Web 2.0 Features:** Drupal is an excellent community platform provider. It outperforms all other options in this area. A website administrator can set permissions for site visitors to comment on any content of website. This feature facilitates social networking website which allows visitors to create a group. As per their set permissions, they can post content such as article, pictures, and videos which can be managed by the admin. This CMS also supports multi-author blogs and subscribed visitors can create their own blog in the blogging community. Outgoing RSS feeds are fully supported and add-on plug-ins can help in displaying others RSS feeds to user website. Moreover this RSS feed can be modified as per the requirement.
7. **Security:** Security updates are published on drupal.org. It gives a notice through update status plug-in to its users whenever new update is released. Drupal's active community is very active and any security holes are remedied very quickly. There are references available to guide in making a site more secure.
8. **User Roles and Workflow:** Drupal's core includes two default set of roles, anonymous user and authenticated user. Apart from these, any number of user roles can be created and assigned different permissions depending upon the content type. Add-on modules can be used to give more specific permissions to users based on content section using taxonomy function.

9. Drupal Usage Statistics

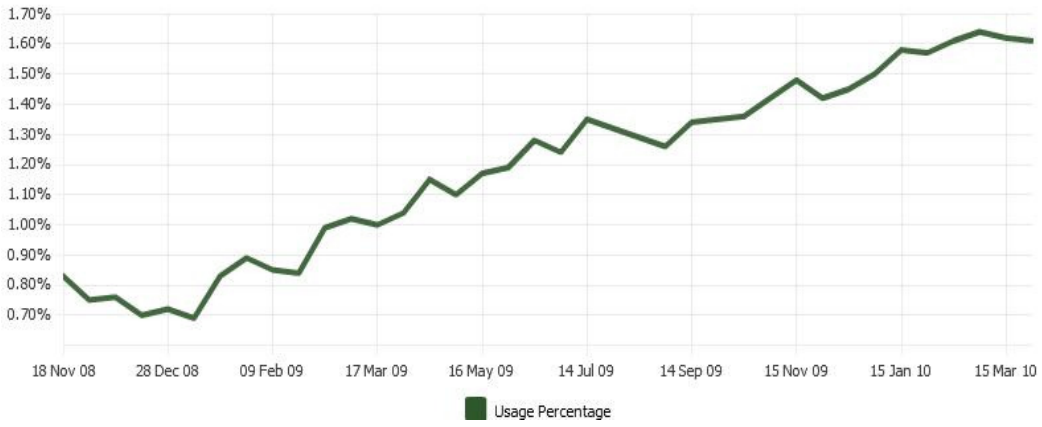

Figure 3.2 Drupal Usage Statistics [10]

This chart represents Drupal penetration over a historical time period on a large selection of website homepages queried by BuiltWith.com [10].

3.6Drupal Installation

Installing and configuring the Drupal is relatively easy. This thesis also demonstrates how to install Drupal on local server in 5 easy steps.

Step 1: Create a local server.

To setup a local server user can download the XAMPP executable file for the Windows Operating System from http://www.apachefriends.org/en/xampp-windows.html. After installation this executable will create a directory call ‘htdocs’ along with others under XAMPP directory.

Step2: Create a Database

After installing the XAMPP a user can easily manage the database through phpMyadmin. phpMyadmin is also an Open Source GUI for managing MySQL databases. phpMyadmin client can be access by visiting http://localhost/phpmyadmin . Through phpMyadmin client a user can easily create a new database and assign a user with all necessary privileges to access the database. User must keep note of username, password and database name as this will be required during installation.

Step3: Download Drupal Package.

A user can download the most stable version of Drupal (6.16 as of writing this thesis) from drupal.org and extract all the files in the htdocs folder. Under sites/default/all rename the default.settings.php file to settings.php file. Optionally, the user can create theme and module directory under /site/all for contributed themes and modules.

Step 4: Installing Drupal

Drupal will show the installation screen on running install.php script.

Figure 3.3 Drupal installation: Choose Language

User can choose the language for installation and on the next page installer will ask for database username and password.

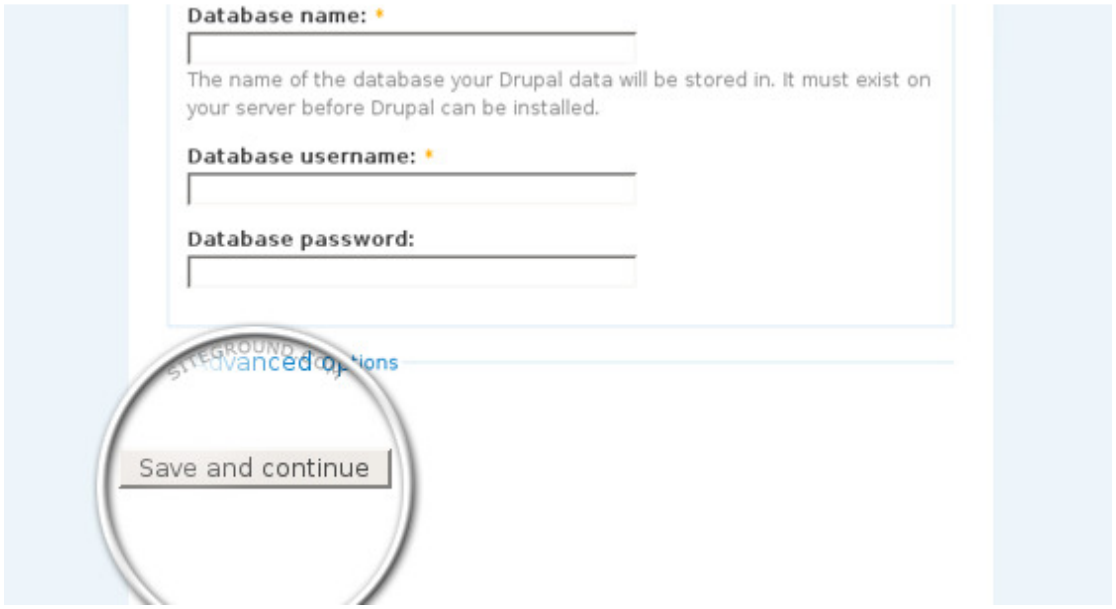

Figure 3.3 Drupal installation: Database Configuration

Step 5: Configuration of Drupal Website: After a successful connection to a database installer will install Drupal by creating tables in the database. On the next screen installer will ask a user to fill in some details of the website and create password for the admin.

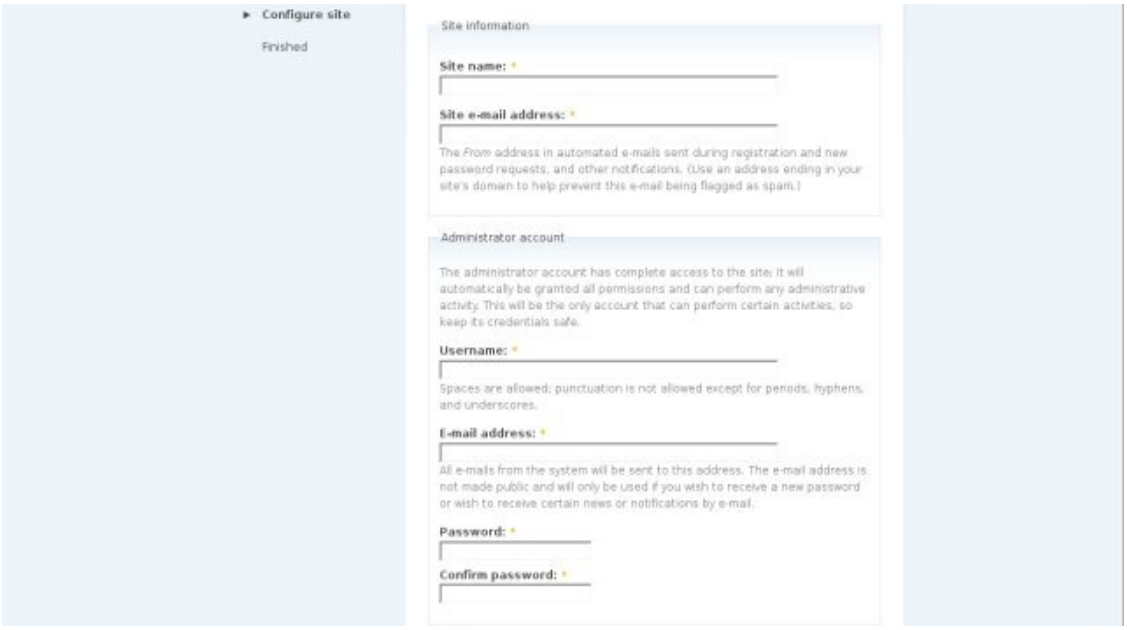

Figure 3.5 Drupal Installation: Site Configuration

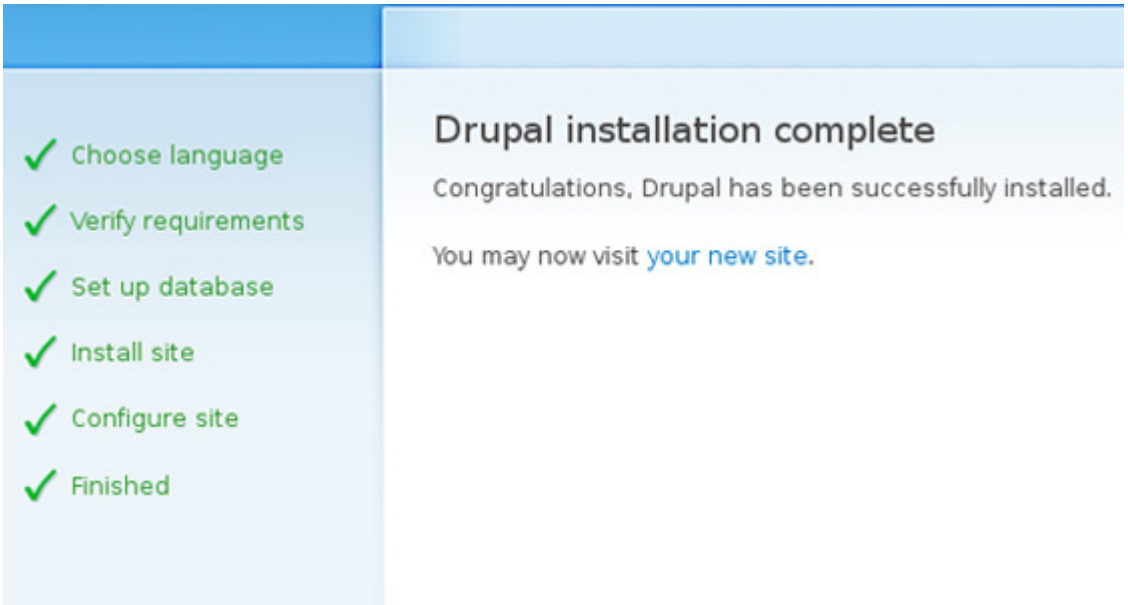

Figure 3.6 Drupal installation: Successful Installation

After saving the site information on the next screen, the administrator will see message similar to Figure 3.5 with a congratulatory message for successfully installing Drupal.

Chapter 4

Joomla CMS

Joomla is a class of Open Source CMSs written in PHP scripting language and uses MySQL database for the backend. Compared to Drupal and WordPress, Joomla is fairly new and is gaining popularity among users because of many aspects, including ease of usability and extensibility. Joomla is backed by a robust community of enthusiastic developers and users. There are around 4500 extensions and modules available to enhance the functionality of the core Joomla package. The community has also contributed around 2000+ free templates. Joomla can be installed and run on Linux, Windows or Macintosh OS. It is distributed under GPL [13] and is free to use.

Advanced components of Joomla 1.5 are built using Model-View-Controller (MVC) design pattern [16]. This model separates the data collection (Model), representation (View) and user interaction (Controller) of components. Because of such separation in three distinct parts- a hierarchy of data handlers, the theme manager, and the menu manager - the properties and methods of one section can be modified without changing the other sections [13]. Each core part of the package is assigned to an individual role and task in providing data organization and presentation.

The standard release of Joomla contains the basic features such as blogs, RSS feeds, caching, search functionality, printable versions of pages, create and manage menus, administer the system and support for language internationalization [14]. Joomla is

designed to be modular so that its core capabilities can be extended. This is done using community contributed extensions, which allow the user to make websites with different functionalities, from shopping cart to social networking, providing a solid foundation for many different kinds of sites. Joomla also support Web 2.0 features but not as efficiently as Drupal.

Joomla keeps content in its database to provide dynamic formatting. Web pages can be presented in unique format preferred by different visitors and different computers as they are not static files. Joomla templates are composed of XHTML block and in line tagged element. The theme manager interacts with data collectors and Menu manager in particular pattern.

The whole system is made up of three types of pages: *Sections*, *Categories* and *Articles*[14].The content is called *Article* in Joomla. It can be data, an image, or any type of content visible on a website. The *Article* is a basic form of content and it is not related to other properties of content such as color and font style. *Section* and *Category* can be defined as article organizers. *Articles* are stored under *Category* and *Section* can have one or more categories. One Article can only be under one category and one section. So to display the same content in different web-pages, one needs to create two different *Articles* of that content in different *Categories*.

4.1 Joomla File structure [14]:

The understanding of the basics of the directories and files structure in Joomla site is essential. When Joomla is installed, there will be a default file structure either on the local machine or on the server. Below is an example showing how each folder has all the important documentation structured and organized.

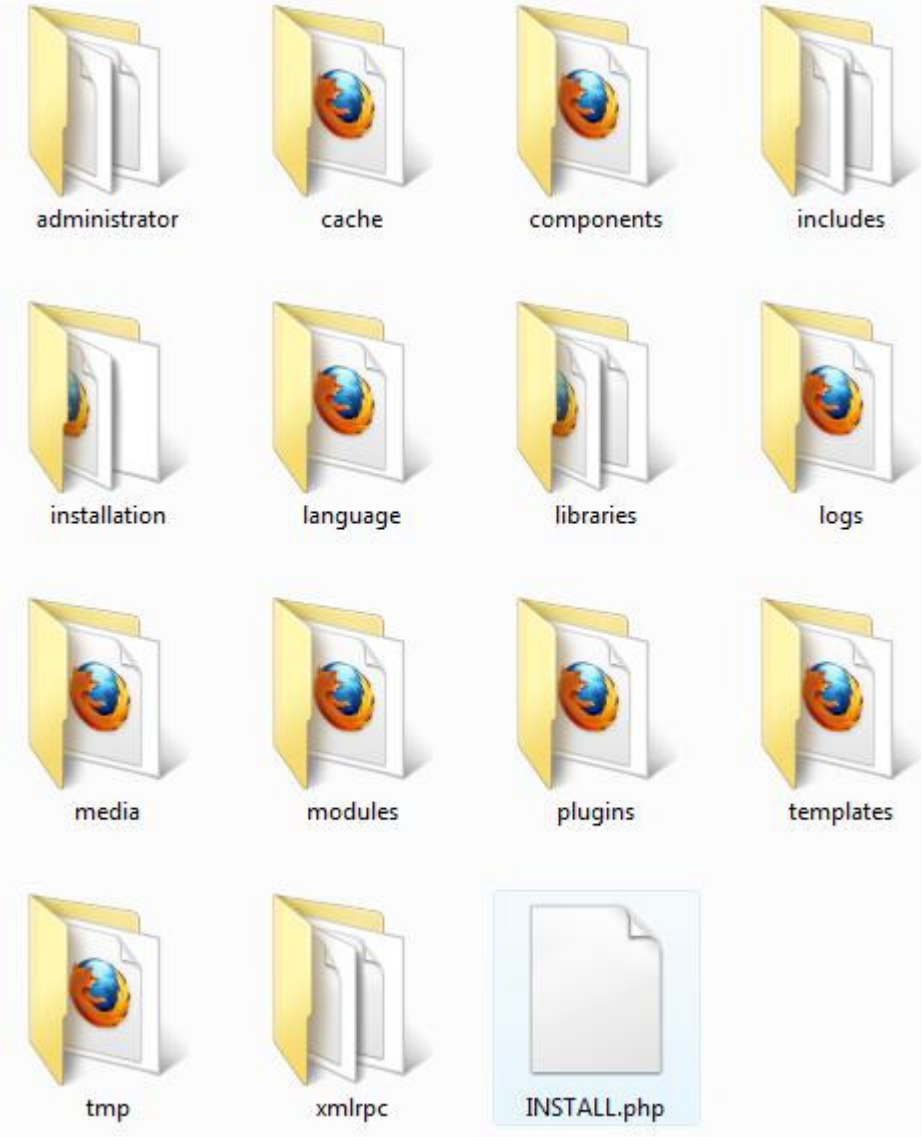

Figure 4.1 Joomla File Structure

The files are described below.

Administrator: This folder consists of all the back end files required by site administrator.

Cache: This folder is used to store cache files for Joomla as well as for some extensions.

Component: All the installed components will be stored in this folder.

Images: To store all the images, this folder will be used.

Includes: Libraries of all common functions are kept here.

Installation: For initial Joomla installation purpose this folder is used, but once it is installed, it must be removed

Language: This folder is used to store language translation files.

Libraries: Joomla API code and some other packages are kept in this folder.

Logs: This folder contains log files.

Media: All images, CSS, Javascript, flash files are stored in this directory.

Modules: All the installed modules must be stored here.

Plug-in: The entire installed plug-ins are stored here.

Templates: This folder is used to store the templates.

Tmp: This folder provides space to temporary files.

Xmlrpc: This directory is used to support web services through XML-RPC Protocols.

4.2 Joomla's key features:

Graphical Flexibility: Many packages of graphical themes are available. It is easy to create custom themes using template files, a style sheet, a file containing information and a theme screenshot. Creating a theme from HTML page is also easy.

WYSIWYG Editors: Joomla provides default editor tinyMCE which allows default video and podcast extension.

Document Manager: Joomla provides reliable document managers i.e., DocMan and RokDownload.

4.3Areas of improvement:

Structural Flexibility: The Joomla core system is developed to support only three hierarchy levels. They are *Section*, *Categories* and *Article*. If any additional level of hierarchy is required, it can be achieved by adding an external module.

Offline Access: To work on Joomla site internet connection is required. The site can be copied to a database but it is quite complicated and requires mastery over SQL database commands.

Multisite Management: Multisite management allows admin to run more than one website from a single installation package. The websites shares the one common database for storing and rendering the data. Joomla is very weak in multisite management and some available options are either unstable or expensive.

4.4 Joomla Users:

Some of the big organizations who are using Joomla arethe United Nations, Online Publication L.A Weekly, Social networking MTV Network Quizilla, IHOP, Harvard University, Senso Interiors etc [13].

4.5 Analysis of Joomla

1. **Installation:** Installation of Joomla is very easy and can be installed in less than half an hour. For installation, a user doesn't require much technical knowledge - they just have to know how to connect through FTP and install databases. It can be hosted on standard shared Linux, Apache, MySQL and PHP environment.

2. **Maintenance and Update:** In Joomla, maintaining and updating the site is easy. Like Drupal, Joomla users also need to backup the website before updating it. The

administrator can back-up the whole website by downloading a file which contains the assets of website, and replacing the web directories using a web update interface. The database can be backed up in PHPsMyAdmin. There has been only one major upgrade for Joomla to date. Joomla does not give any notice to site admin when ever any upgrade is required. Major upgrades may affect the current template or plug-ins. Joomla does not provide security updates for previous versions so administrator has to change the version for security updates.

3. Community Strength and Contribution: This CMS is supported by many independent consultants and organizations in the USA. There are several books published to help learners. Forums and discussion have been started by groups where one can ask question and get answers.

4. Usability: The user interface is very friendly with extensive use of images. Creating a new page is easy and it can be published by assigning it to the appropriate section and category. From a visitor's view of website, administrator can easily find content page and article he wants to edit but to find in an administrative view, one need to know whether it is in section, article or category as Joomla don't support cross categorization. And finding this requires learning through documentation. The Joomla core package does not support Microsoft Word format so text can't be copied and pasted.

5. Scalability: The website built on Joomla CMS can be grown as the demand for a larger site increases. Joomla is scalable and can support tens of thousands of visitors a day. Joomla's WebPages are cached for faster downloading of page and they are cached indefinitely as default but one can change the setting as needed.

6. **Web 2.0 Features:** In Joomla unlike Drupal, a website administrator cannot set permissions for site visitors to comment on any content of website through core extensions, but it can be done through a plug-in. This CMS supports a simple blog but multi-author blogs are not supported. The simple blogs can be created through content administrator on the website but cannot be created through the front-end site. Therefore, blogging communities are not supported. Also, for social networking sites, a popular extension exists to create groups and content such as articles or images. It is then managed by site administrator. Outgoing RSS feeds are fully supported and add-on plug-ins can help in displaying other RSS feeds to the user website. However users cannot modify those feeds.

7. **Security:** The updates on security are released on joomla.org and these updates occur frequently. Joomla has had only one major upgrade and it doesn't support the legacy version with security updates and fixed bugs. Therefore, it is recommended to upgrade Joomla's older version.

8. **User Roles and Workflow:** The Joomla CMS supports three administrative roles: to create new content, to edit existing content, and to publish content. In this CMS, one cannot give permission to particular users to edit or publish based on section of website or type of content. There is no notification system in which to forward or flag content. The site based on Joomla CMS is not very flexible for many different content editors as it does not give an easy way to set up a workflow based on type of content.

9. Joomla Usage Statistics

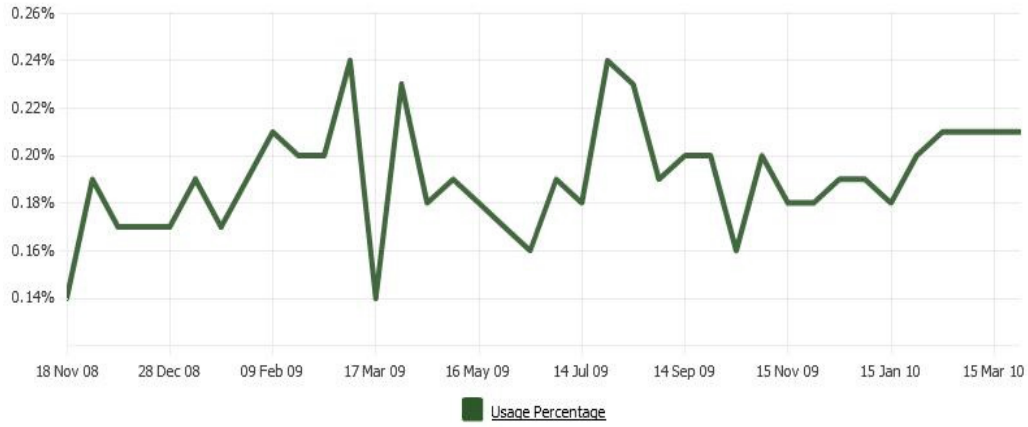

Figure 4.2: Joomla Usage Statistics[15]

This chart represents Joomla penetration over a historical time period on a large selection of website homepages queried by BuiltWith [15].

4.6 Joomla Installation

Installing and configuring the Joomla is comparatively easy. Most of the hosting companies provide a one click installation of Joomla and other Open Source CMSs using the Fatanstico tool in Control Panel (Cpanel) [13]. Below are the three easy steps to install and setup the Joomla on local server running on Windows Operating System.

Step 1: Create a local server.

To setup a local server user can download the XAMPP executable file for the Windows Operating System from http://www.apachefriends.org/en/xampp-windows.html. After installation this executable will create a directory call 'htdocs' along with others under XAMPP directory.

Step2: Create a Database

After installing the XAMPP a user can easily manage the database through phpMyadmin. phpMyadmin is also an Open Source GUI for managing MySQL databases. phpMyadmin client can be access by visiting http://localhost/phpmyadmin . Through phpMyadmin client a user can easily create a new database and assign a user with all necessary privileges to access the database. User must keep note of username, password and database name as this will be required during installation.

Step3: Download Joomla Package.

A user can download the most stable version of Joomla (1.5 as of writing this thesis)from joomla.org and extract all the files in the htdocs directory.

Step 4: Installing Joomla

To begin the installation process user should navigate to URL http://localhost/index.php

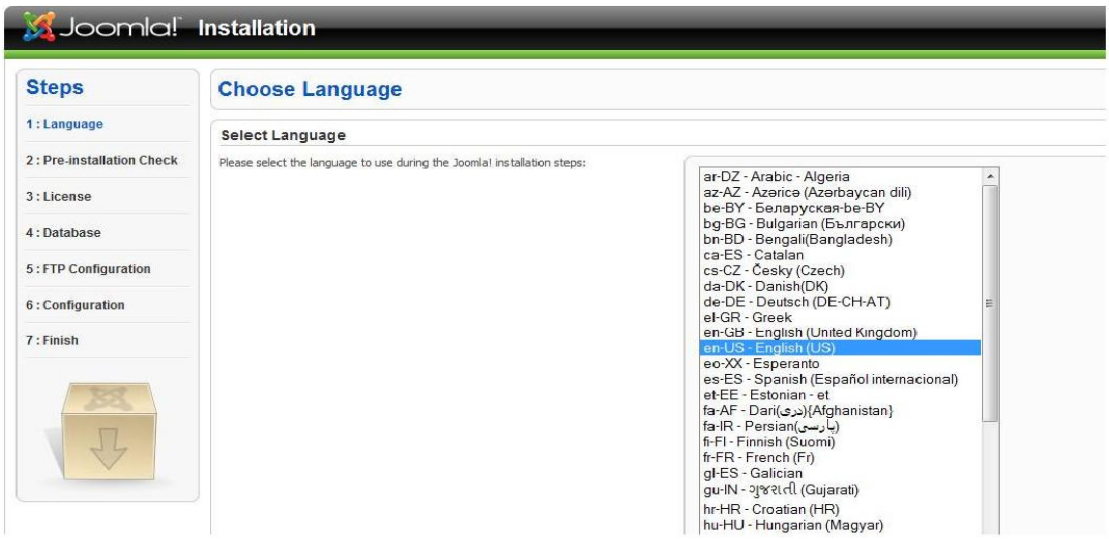

Figure 4.3 Joomla Installation: Choose Language

After selecting Language for installation the installer will do the pre-installation check to make sure all required components are available.

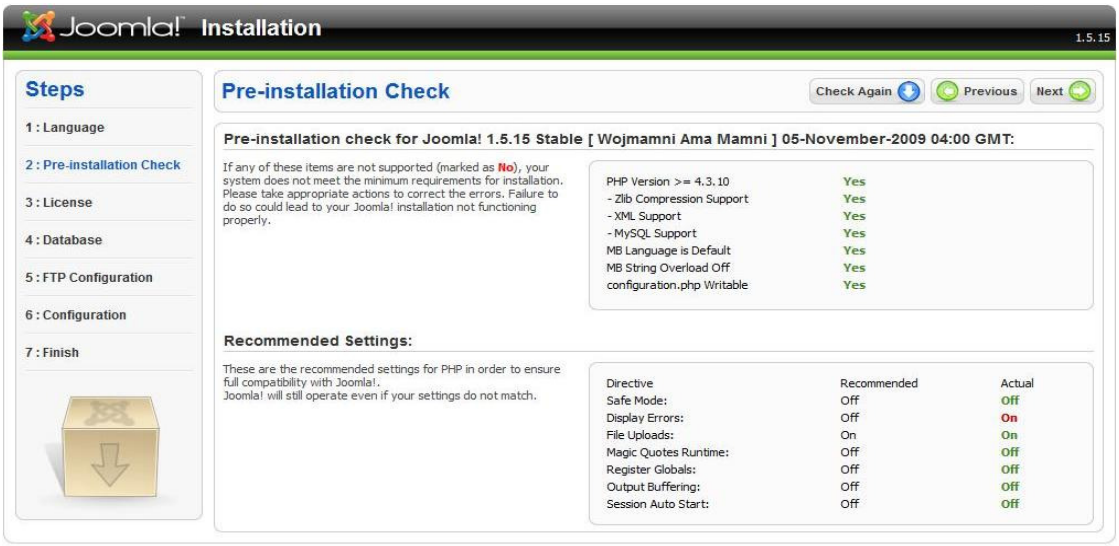

Figure 4.4 Joomla Installation: Pre-installation Check

Once pre-installation check is done with no warning installer will ask for Database details.

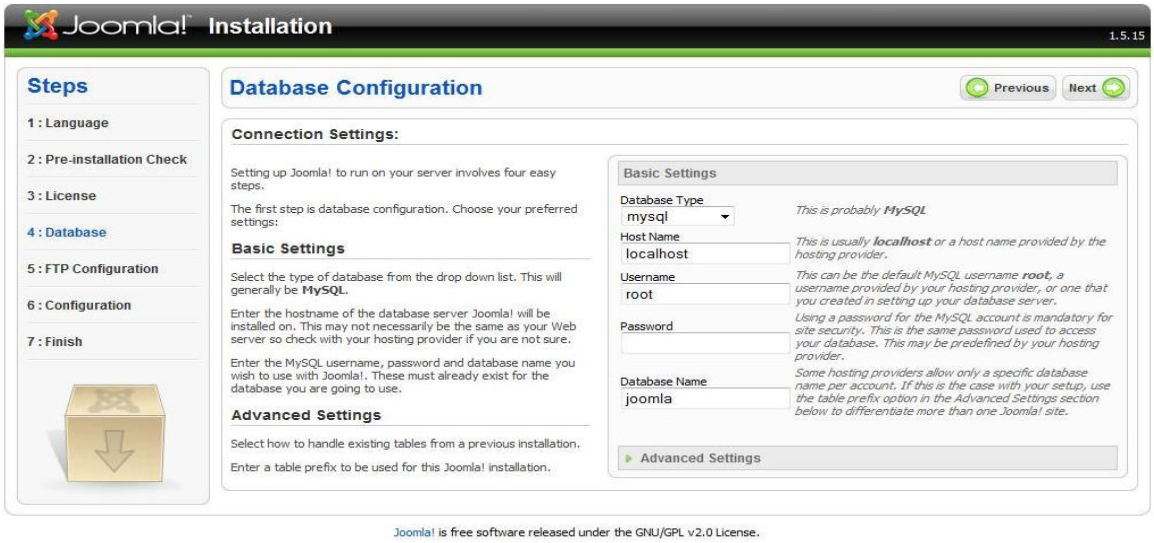

Figure 4.5 Joomla Installation: Database Configuration

In the next step the installer will ask to create a FTP user. This is not required when working on a local server.

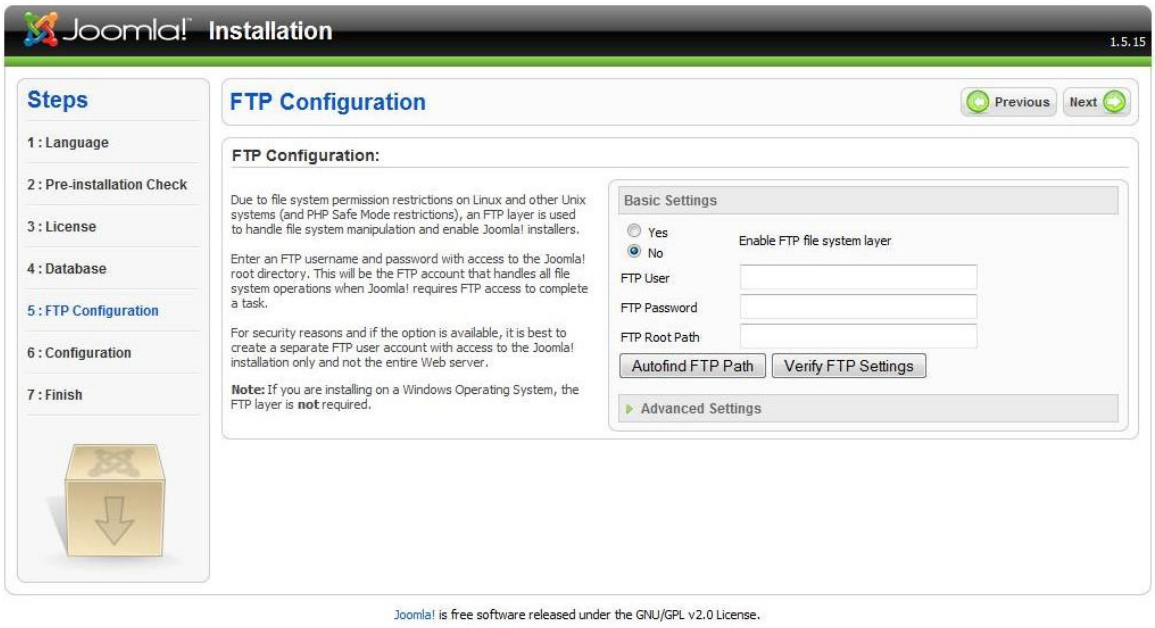

Figure 4.6 Joomla Installation: FTP Configuration

For final step installer will ask for Site Name and details for admin account.

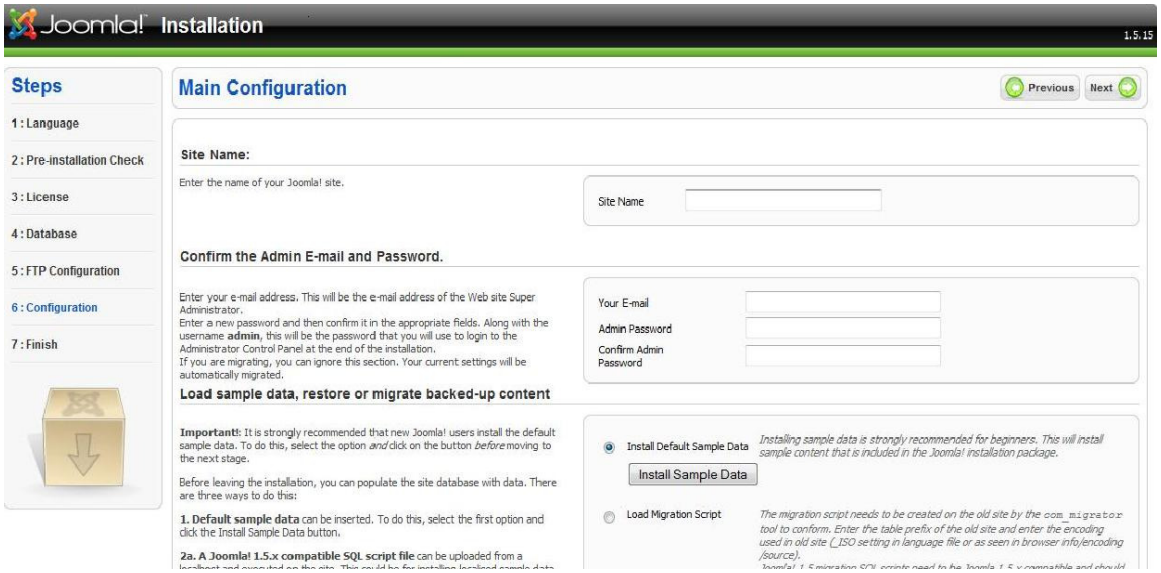

Figure 4.7 Joomla Installation: Main Configuration

Chapter 5

WordPress CMS

WordPress is an Open Source CMS powered by PHP and MySQL and licensed under GPL [23]. It is one of the most popular open source applications for blogging purposes. Its core functionality can be extended by integrating the contributed themes and plug-ins. It has a very robust community of users and developers who so far have contributed 8000+ plug-ins and around 1200 free themes. The documentation about using and configuring WordPress is very well written. There are numerous tutorials available to help the non-technical user to get started in a shorter amount of time. WordPress won best Open Source CMS Award for the year 2009 [24].

WordPress templating system is very flexible and allows user to rearrange the widgets without changing the PHP or HTML code and canbe easily switched between different themes. Unlike Drupal and Joomla, WordPress allows user to edit the PHP and HTML files from administrative area for advanced customization. WordPress has good support for Web 2.0 community features including an integrated search engine with a friendly, clean permalink structure. It also allows users to tag, post, and categorize articles. WordPress Multi User (WMPU) allows multiple blogs under one installation and their moderation them from a single dashboard.

WordPress installation is very easy and can be done either manually or by using various user-friendly installers. It can be installed on Microsoft IIS server by using Microsoft

Web Platform Installer which detects any missing dependencies before installing WordPress. Many hosting companies provide automated installation through their control panel. WordPress has applications for Blackberry, Iphone/Ipod and Android through which users can have access to some of the features of Admin Panel.

5.1 WordPress File Structure

It is very essential to understand the basics of how the directories and files are structured on WordPress site. When WordPress is installed, there will be a default file structure either on the local machine or on the server. Below is an example showing how each folder has all the important documentation structured and organized.

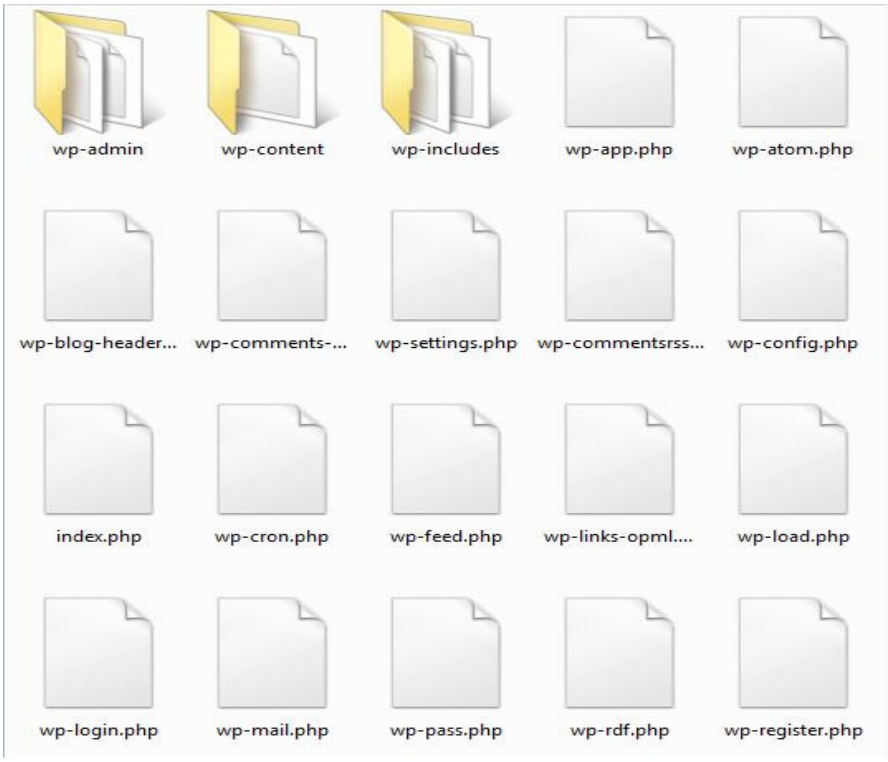

Figure 5.1 WordPress File Structure

The files are described below.

wp-admin: This directory contains the files responsible for generating the admin features.

wp-content: This directory contains the contributed plug-ins and themes.

index.php: Front to the WordPress application. These files do not do anything, but loads wp-blog-header.php which tells WordPress to load the theme.

wp-atom.php: Redirects to Atom feeds. This file is deprecated and only exists for backwards compatibility.

wp-blog-header.php: Loads the template and WordPress Environment

wp-comments-post.php: Handles Comments post to WordPress and prevents duplicate comment posting

wp-load.php: Bootstrap file for setting the ABSPATH constant and loading the wp-conFigure.php file. The wp-conFigure.php file will then load the wp-settings.php file, which will then setup the WordPress Environment.

wp-login.php: This files handles user authentication, registering, resetting passwords, forgot password and other user handling.

wp-register.php: As the name the suggests, this file is responsible for displaying the registration page.

wp-rss.php: This file redirects the user to the RSS feed source. This file is deprecated and only exists for backwards compatibility.

wp-settings.php: Used to set up and fix common variables and include the WordPress procedural and class library.

wp-signup.php: It contains 16 functions that together enables the user signup functionality.

wp-trackback.php: This file is responsible for handling the trackbacks and Pingbacks sent to WordPress

5.2WordPress Key Features

Search Engine Optimization (SEO): WordPress is search engine friendly by default. It manages integrated links and provides clean permalink structure.

Automatic filters: These filters are included in core package which facilitates accurate styling and formatting of content in articles.

Administration Features: WordPress has excellent administration panel if the basic purpose of the website is only to publish general information**.**

5.3WordPress Users

Many big companies and organizations are using WordPress to power their blog. A few of them are eBay, yahoo, Digg Blog, Wall Street Journal, Ford, People Magazine, Samsung, NYTimes Blog, Playstation, and cPanel.

5.4 Analysis of WordPress:

1. Installation: The installation process of WordPress is very easy and anyone who has experience in installing other systems he can install WordPress in around 15 minutes. There is no technical expertise needed for installation, the user should just know how to connect the File Transfer Protocol (FTP) and create databases.

2. Maintenance and Update: It is advisable to backup the site by downloading just one directory from web server and then store it in the database. There have been two major upgrades to date which broke some themes and modules. Like Drupal, WordPress also gives notice to its user regarding the upgrade, and it provides maintenance for older versions by providing security updates. WordPress has automatic one-click upgrades.

3. Community Strength and Contribution: WordPress is supported by a very large community of developers and users [23]. There are discussions and forums where people help each other by posting questions and answers. This community also welcomes newcomers. This community is governed by 'Automatic,' a for-profit company [24].

4. Usability: The administration area is very user-friendly; even a novice user can figure it out. It's very easy to find the page or article which administrator wants to edit within the visitor's view of the website. It supports Microsoft Word format so the in-built

function allows the copy and paste of text from a Word document. Adding a new page or section in navigation bar is very easy.

5. Scalability: This CMS can handle large amount of traffic. Its internal cache stores infrequent changing data.

6. Web 2.0 Features: WordPress is excellent blogging platform but does not provide robust support for advance features. It supports single and multi-author blogs. In WordPressa content administrator can give permission to allow visitor to post comments. The system includes spam filter and comment moderation support. Outgoing RSS feeds are supported by the core package and add-on plug-ins can be used to display others' RSS feed. With add-on modules, a social networking site can be built. Website visitors can not submit content through the front end.

7. Security: The security updates are frequent but not published on WordPress.org.

8. **User Roles and Workflow:** WordPress supports only three user roles for administration. First, roles of this user can draft and publish their own as well as others' content. Second, users can draft and publish their own content but not others'. Third, users can only draft the content but not publish. Apart from these, no other administrative role can be created and no plug-in can provide more roles. WordPress doesn't support any notification system to flag content. It allows users the ability to roll back to the prior version of website.

9. WordPress Usage Statistics

There are over 22 million WordPress publishers as of February 2010: 10.6 million blogs hosted on WordPress.com plus 11.4 million active installations of the WordPress [13].

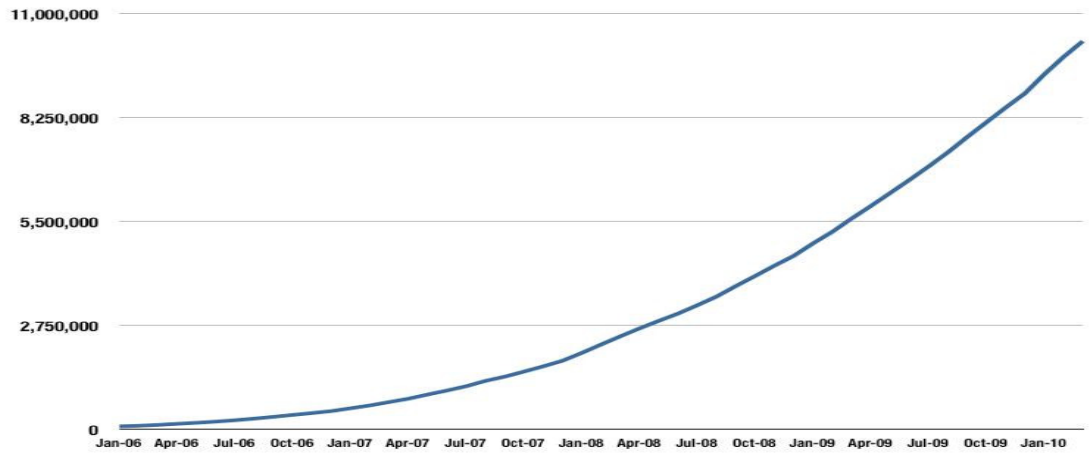

Figure 5.2 WordPress Usage Statistics[26]

5.5WordPress Installation

Installing and configuring the WordPress is much easier and faster as compared to Drupal and Joomla. Below are the three easy steps to install and setup the WordPress on local server running on Windows Operating System

Step 1: Create a local server.

To setup a local server user can download the XAMPP executable file for the Windows Operating System from http://www.apachefriends.org/en/xampp-windows.html. After installation this executable will create a directory call 'htdocs' along with others under XAMPP directory.

Step2: Create a Database

After installing the XAMPP a user can easily manage the database through phpMyadmin. phpMyadmin is also an Open Source GUI for managing MySQL databases. phpMyadmin client can be access by visiting http://localhost/phpmyadmin . Through phpMyadmin

client a user can easily create a new database and assign a user with all necessary privileges to access the database. User must keep note of username, password and database name as this will be required during installation.

Step3: Download WordPress Package.

A user can download the most stable version of WordPress (2.9.2 as of writing this thesis) from WordPress.org and extract all the files in the htdocs directory.

To begin the installation process user can navigate to http://localhost and installer will ask for the database details which include the database name, username, password, database host and table prefix for multi site installation.

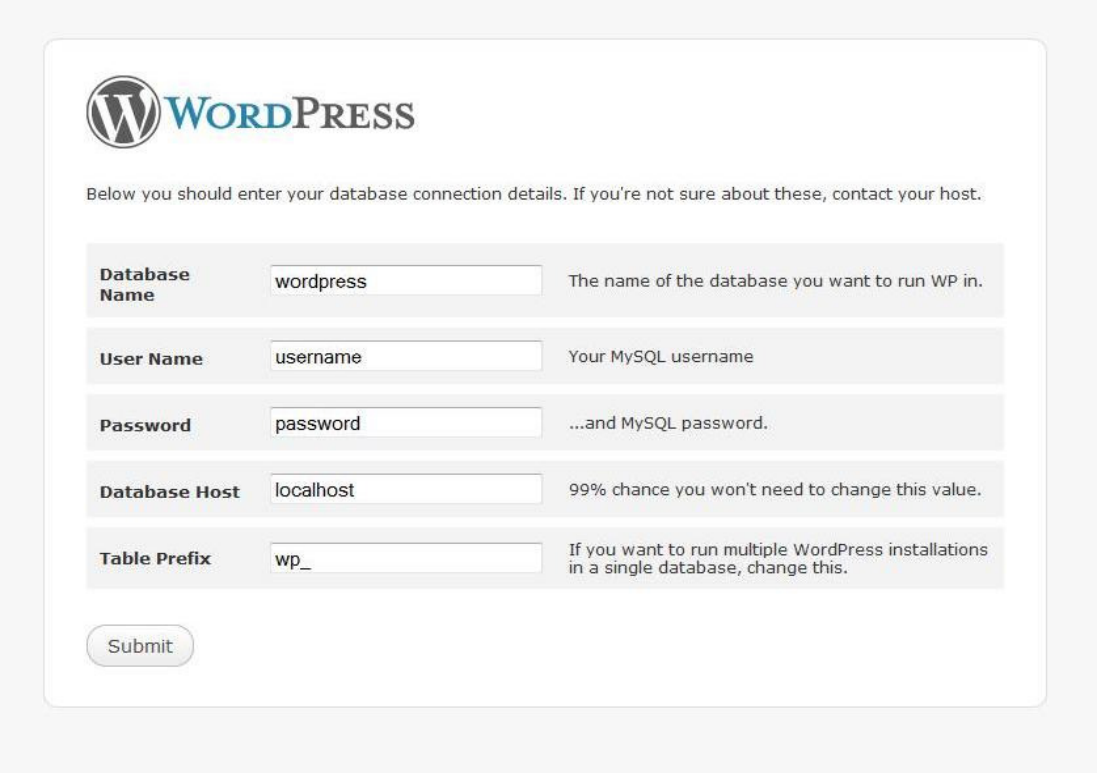

Figure 5.3: WordPress Installation: Database Configuration

It will update the wp-conFigure.php file with the form data and sets it appropriate read and write permissions.

Once the database connection is setup it will ask for Blog Title and email address. These settings can be changed later on.

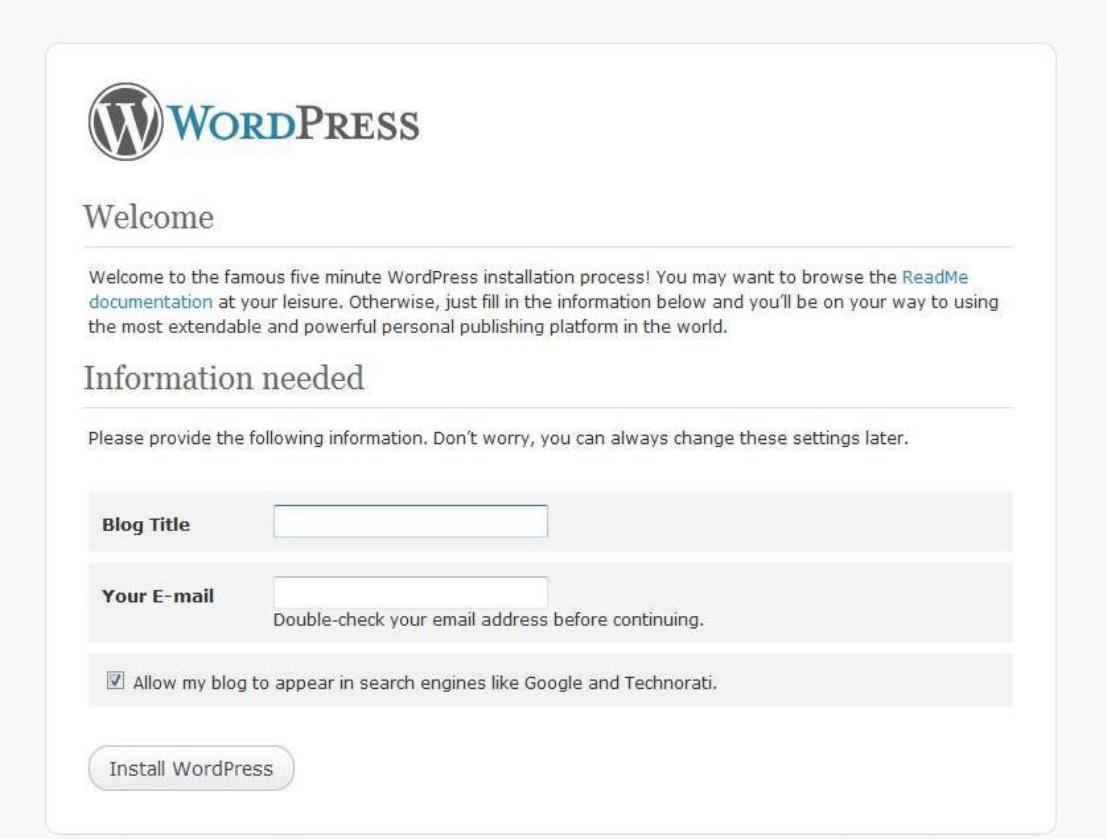

Figure5.4: WordPress Installation: Site Information

On filling up the site information the installer will generate a username and a password for admin. The password can be change after log-in as admin.

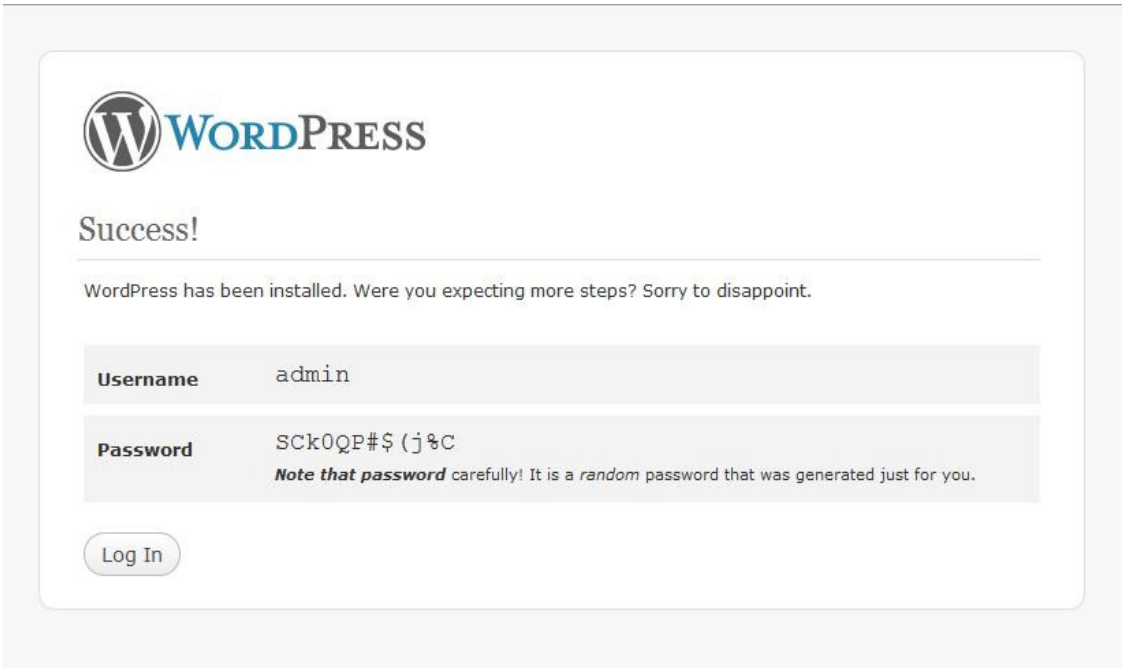

Figure 5.5: WordPress Installation: Installation Successful

Chapter 6

Comparison between CMSs

Below is the table of comparison of three Open Source CMSs based on the areas defined in earlier chapters and a survey we conducted on small set of corporate. We interviewed six participants to investigate the primary factors that are important in selecting a CMS to build a website. Three of these interviewers were from education industry who implements more than one of the three systems covered; two were from media industry who evaluated the CMSs and chosen to implement a particular system; and one was from non-profit organization who built a member management system using one of the CMSs described in this thesis.

No	CMS	Drupal	Joomla	WordPress
1	**Installation**	Excellent	Excellent	Excellent
2	**Maintenance and Update**	Good	Good	Excellent
3	**Community Strength and Contribution**	Excellent	Excellent	Excellent
4	**Usability**	Excellent	Good	Good
5	**Scalability**	Excellent	Fair	Good
6	**Web 2.0 features**	Excellent	Good	Good
7	**Security**	Good	Good	Good
8	**User roles and Workflow**	Excellent	Fair	Fair

Table 6.0 Drupal, Joomla and WordPress Key Features Comparison

The table also summarizes our opinion after developing two websites at Rutgers using Drupal CMS. We chose particularly Drupal CMS for our projects because Rutgers University chose Drupal for its primary website.

Chapter 7

BIO-1

Bio-1 is an organization whose goal is to make Central New Jersey (CNJ) the next “hot spot” for the global bioscience industry, by creating high-quality, high-paying jobs and a skilled workforce. The five-county BIO-1 partnership is named for the Route 1 corridor from Rutgers to Princeton, around which most of CNJ’s biotech firms are clustered. The CNJ region, comprising Hunterdon, Mercer, Middlesex, Monmouth, and Somerset counties, has received $5 million, available under the Workforce Innovation in Regional Economic Development (WIRED) program from the United States Department of Labor (USDOL). The WIRED grant is used to transform the rich array of existing bioscience education and training and economic development initiatives into a into a world class bioscience talent development system [3].

Bio-1 maintains a job portal dedicated to Bio Industries and students. Employers from industry, academia and government who have opportunities to offer – full or part-time jobs, fellowships, internships, research opportunities, and more - can match their needs with talented candidates from New Jersey. The free site is a valuable tool for job seekers and employers – applicants can post resumes and search for jobs either by selection criteria or by any keywords, and employers can post jobs or log in to search for candidates by any keywords they require.

A unique feature of the website that may appeal to students and employers with internship opportunities is the bioscience skills inventory. Students may indicate their skill set by completing a series of check off boxes to enhance their resume and facilitate employer searches. Employers may do keyword searches for quicker matching on specific skills. The search results will list candidates who have identified keywords on their resume, skills inventory or both [4]. As part of this thesis, the internship matching portal of Bio-1 was designed and developed using Drupal CMS.

7.1 Project Requirements

The main purpose of this application is to have a system that will provide a common platform to Bio related companies and students to interact. The website needs to have a registrations system through which students and employer can create their profile on the site. The registration process is a two step process. In first step, a user creates an account either as student or employer. On submitting the registration form the user gets confirmation email of their application. The user will be able to access the site only after the admin approves the application and assign the appropriate role. The employers should be able to post a job and search resume database using keywords. The application should allow the employers to see all their jobs and edit them once they login. The students should be able to create a profile and upload the resume either in doc or pdf format. Registered or anonymous user should be able to search for available jobs using key words or filtering options from the drop down menu. The home page should display the five most recent jobs with a link at bottom to all available jobs. The users should be able give

their feedback about the website using the feedback form. The submission of feedback form should be automatically routed to admin email address. We summarize the requirements below:

- Ability to index the content of uploaded files (resume) in pdf and doc format
- Easy to use admin screen
- Show five latest jobs on homepage
- Employer role to allow user post job and search resume database
- Student role to allow user upload the resume and create a profile.
- Scheduling a cron.php script to run every 20 minute and index the new content on the website.

7.2 Solution Implementation

We studied the requirements in detail and configured the required modules to meet the needs. We created two roles: employer and student with appropriate permissions. As mentioned in the requirements, users with the employer role will be able to access the resume database and create a job post on the website. On the other hand, the users with student role will be able to create a profile and search jobs based on keywords and drop down filters. We created two custom content types call ‘Employer Profile’ and ‘Student Profile’ with appropriate Content Construction Kit (CCK) fields for registration purpose. To index the student’s profile, we used the Drupal’s built in search module. And to index the resumes (uploaded either in pdf or doc format), we implemented File Search module

along with Apache Solar module. These two modules converts the attached file content into text format which than can be indexed by the Drupal's default search engine. The search results for employers contain the results from the students profile and resumes. For the job search page, we created a View with appropriate filter options and fields to display the available job. The user can sort the result by clicking on the table header or narrow down by using the filters. Section 7.4 explains more about the implementation of Views and custom content types.

7.3 User Interface Screens:

The layout of homepage is two columns with the header and primary navigation at top. The right-side bar contains the information about the Bio-1 and navigation for employer and job seekers.

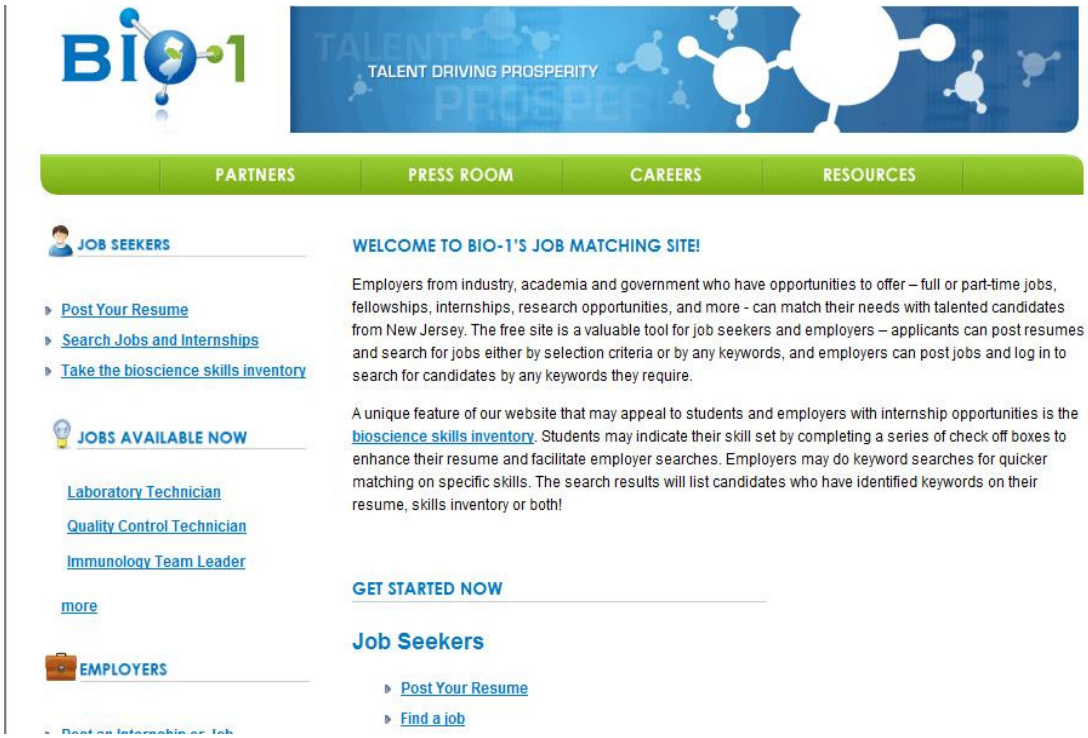

Fig 7.1 Bio-1 Home Page

The login screen allows existing users to login or request a new password. The new user can create an account by clicking on 'Create new account' tab.

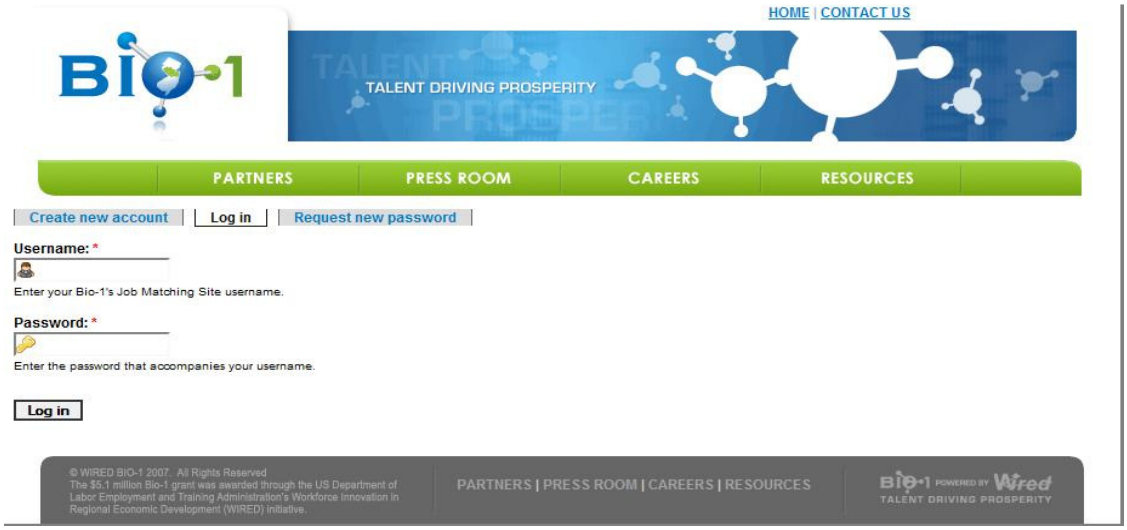

Fig 7.2 Bio-1Login Screen

Registration Screen allows new user to create an account either as student for employer. On submitting the form, user will get the confirmation email and will have to wait for admin approval.

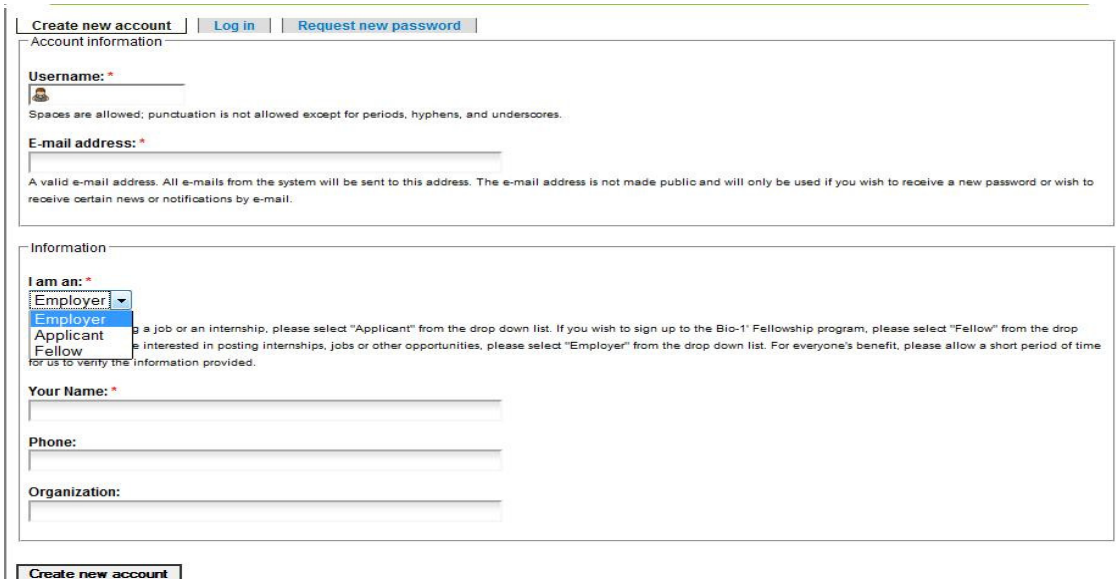

Fig 7.3 Registration Screen

The job search result page is built using View and CCK modules. It allows user to search the available jobs either through keywords or drop down options. User can also sort the result by clicking the table header title.

Job Title	Organization	Location	Posted	Type
Research & Development	AstraZeneca		07/13/09	Research Opportunity
Protein scientist	Neurotez Inc	North Brunswick	12/01/09	Research Opportunity
Senior Manager; Corporate Communications (PT – Job Share)	AstraZeneca	Delaware City, Delaware	03/09/10	Part time job
Experienced Pharmaceutical Operations [Contractor only]	NSD Bio Group, LLC		12/12/09	Other
Assistant Chief Medical Officer (Interim)	Ashton Tweed	Pennsylvania	12/17/09	Other
Contractor Opportunity, Clinical Project Management	Amicus Therapeutics	Cranbury, NJ	02/24/10	Other
Kelly Scientific Resources Future Scientist Program	Kelly Scientific Resources		04/07/09	Internship
Post-Graduate Intern	AstraZeneca	Delaware City, Delaware	11/12/09	Internship
2010 MBA Marketing Summer Intern, NJ	Johnson & Johnson	Skillman & Morris Plains, NJ	11/13/09	Internship
Intern - Targeting and		Delaware City		

Fig 7.3 Job Search Page.

7.4 Content Types Implemented:

Informational pages and blogs are each a different content type. Each item of content, called a 'node' belongs to a single content type and defines various default settings.

User Profile: This content is used by student to create their profile and upload the resume. It consists of the following CCK fields:

- **Name:** This field contains the First and Last Name of Student

- **Email**: Text field to display student's email address
- **School:** Display the student's current School Name
- **Education Level:** Select list to display the education level.
- **Skills**: Multi choice filed to select matching skills
- **Resume**: File upload field to upload the resume

Job: This content type is use by the employer to create a job and publish it. It consists of following CCK fields:

- **EmployerName**: Text field to display the name of HR
- **CompanyName**: Text field to display the company Name
- **JobDescription**: Text area to explain Job role.
- **Location**: Text field to display job location
- **EducationLevel**: Select list to specify required education for particular job
- **JobType**: Select list to choose the category of job i:e Internship, Full-time, Part-time, Contract, Fellow-ship or Research Opportunities

7.5 Views Implemented

This module enables to control the presentation of lists and tables through appropriate settings of filter and fields. For this project two views were implemented – one for displaying latest jobs and another for displaying all open jobs with filter options.

1) **Latest Jobs** – This view was created to display five most recently published jobs on home page as a block on left sidebar. To achieve this view as block following options were set for filter and fields

Filter:

- Node Type – A Job content type was selected for this option
- Node Published – This option was set to true.

Fields:

- Node Title – This was selected to display Job Title
- Content Field – This was selected to display Organization name

2) Job Board with keywords search and filter options

This view enables student to see all available jobs, sort them with keyword search, or use filter to narrow down the results.

Filter:

- Node Type – A Job content type was selected for this option
- Node Published – This option was set to true.
- Exposed filter for Job Type CCK field.

In addition to above filters search functionality was also implemented to help students narrow down results based on a keyword.

Fields:

- Node Title – This was selected to display Job Title
- Company-Name CCK Field – This was selected to display Organization name
- Job-Type CCK Field – Type of Job i:e full-time, internship, fellowship etc
- Location CCK Field – Location of the Job
- Date CCK Field – Job posting date

7.6 Indexing Resume

The most challenging and interesting task was to implement a search functionality which will not just index the pages but also the attached files (resume), uploaded in pdf and doc format. Different methods were tested and studied for this purpose.

1) **Google Mini**

This is a product by Google which can index 200+ types of files and supports document-level security so that users only see the content they're authorized to view. The drawback of this method was it was not free and secondly the indexing results were stored on Google server instead of the local server.

2) **Apache Solar + File Search Module**

These two modules are contributed by the Drupal community and are free to use. The file search module along with the Apache Solar [5] search allows indexing of the attached files and the ability to display them as a search result.

We implemented Apache Solar module for following reasons.

- Apache Solar module was free to use.
- No extra hardware requirement.
- This module stores the indexing the result on local server.
- Developer can easily customize the look of search result page.

7.7 Search Module

The search module indexes the content of the website and attached files. The location of this file is public_html/modules/search [19]. The source code is explained in Appendix A

Chapter 8

VIZLAB

VIZLAB (VIZLAB.rutgers.edu) is a website developed for Visualization Lab at Rutgers University which gives information about the lab, its research, publications, professor and students. This website demonstrates the use of Drupal CMS for simple informatory website. This website allows students to search the publication based on a particular area. The accordion script was implemented along with Views to display the publication with different filter options.

8.1 Project Requirements

The aim of this project is to create a website for Visualization lab at Rutgers University with easy to use admin screen. We summarize the requirements below:

- Allow user to upload the publication
- Categorize the publications
- Ability to have revision of content
- Upload the research related code
- Video and Photo gallery
- Information about ongoing research
- Details of lab personnel

- Search publications based on keyword

8.2 Solution Implementation

To create a publication page we used default content type called ‘Page’ and integrated CCK file field in it. The file filed enables the user to upload the publications either in doc or pdf format. We set the upload file limit to 10MB. To create the video and photo gallery we implemented the Views with appropriate filter options and used Lightbox2 module for displaying the images. We also enabled the Drupal’s search functionality for searching publication and content of the website based on the keyword. Section 8.4 discusses more about the implementation of Views.

8.3 User Interface Screen

The home page of VIZLAB website project is a two column layout. The ride-side bar contains the information about the VIZLAB and left-side bar displays a block for quick links and ongoing researches.

Fig 8.1 VIZLAB Home Page

Registration Screen allows new user to create an account as student. On submitting the form, user will get the confirmation email and will have to wait for admin approval.

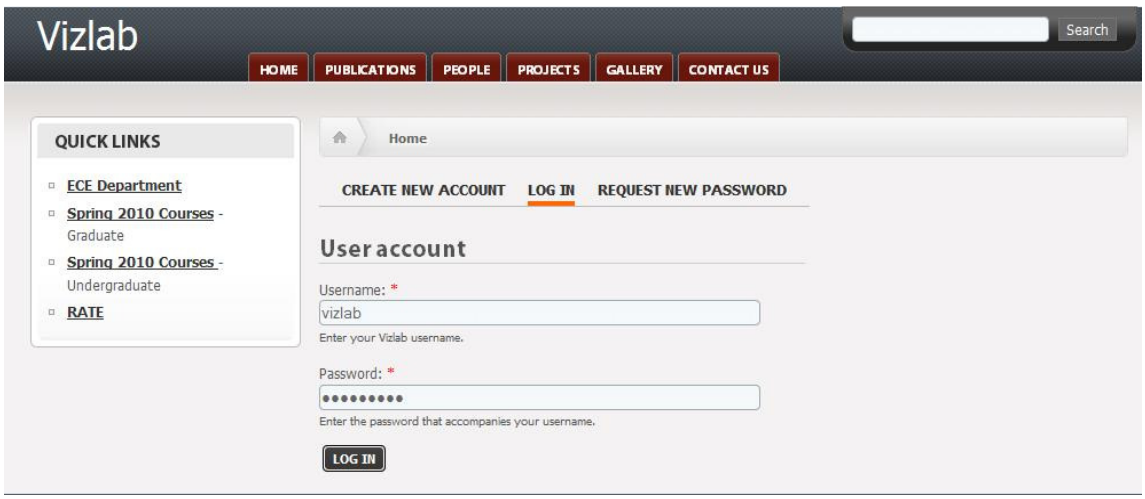

Fig 8.2 VIZLAB Login Screen

The publication page list all the published papers and thesis wrote under the direction of Prof. Deborah Silver. The user can read the paper by clicking read more link next to the

title of the publication.

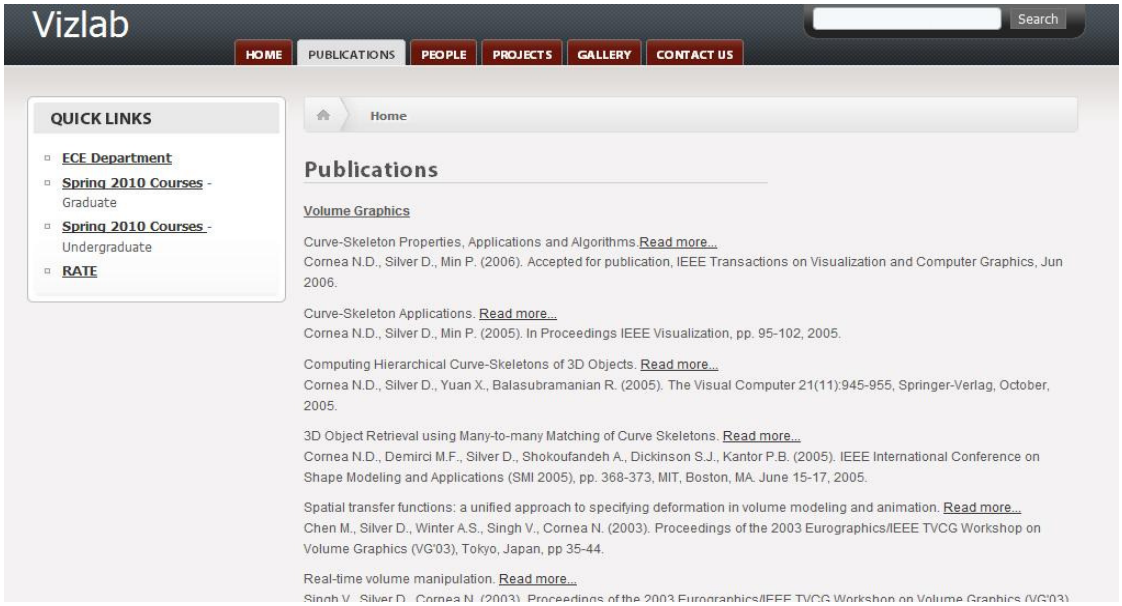

Fig 8.4 VIZLAB Publication Page

8.4 Modules Implemented

Following contributed modules were implemented for this project in addition to usage of default modules.

- **File Upload**: This allows the administrator to upload the publication easily on the server
- **Views and CCK** : To display the publication list in table format
- **Search:** Helps user to narrow down search based on key word.

To create any publication a content type was created with the following CCK fields

- Title – To display the title of the publication
- Abstract - A text area to write an abstract of publication
- Appearance – A text field to display where publication appeared

- File Upload – To upload the publication in pdf, .docx or doc format.

8.5 File Upload Module

This module provides with a CCK field which can be integrated in any content type. The admin can set the limit for file size and also the formats [5]. For this project the file size is set to 10MB and only doc and pdf formats are allowed. The code is explained in Appendix B

Chapter 9

Conclusion

This thesis explained the detailed study of CMSs and their importance in building and maintaining websites. The study of widely used CMSs and the analysis of the features of an individual system can help an individual or organization to choose an appropriate CMS for their specific web application. A specific CMS may provide a better fit for one application while another CMS may be more suitable for a different purpose. The implementation of Drupal CMS for two projects helped to compare and analyze the features and functionality of the different distributions.

Appendix Table of Contents

APPENDIX A

Search Module Code for Bio-1 Website

The search module lets users search for specific content on the website. One can search both for users and for particular words. When user is on the "content" tab of Search, user will be able to search for words appearing in the default rendering of node content on the website, which would include the default rendering of any CCK fields, Location fields, Taxonomy, etc., as well as comments. When one is on the "users" tab of search, the search result will display the name of registered user and their email addresses [19].

```
/**
 * Implementation of hook_menu().
 */
function search_menu() {
  $items['search'] = array(
    'title' => 'Search',
    'page callback' => 'search_view',
    'access arguments' => array('search content'),
    'type' => MENU_SUGGESTED_ITEM,
    'file' => 'search.pages.inc',
  );
  $items['admin/settings/search'] = array(
    'title' => 'Search settings',
    'description' => 'Configure relevance settings for search and other indexing options',
```

```
  'page callback' => 'drupal_get_form',
  'page arguments' => array('search_admin_settings'),
  'access arguments' => array('administer search'),
  'type' => MENU_NORMAL_ITEM,
  'file' => 'search.admin.inc',
);
$items['admin/settings/search/wipe'] = array(
  'title' => 'Clear index',
  'page callback' => 'drupal_get_form',
  'page arguments' => array('search_wipe_confirm'),
  'access arguments' => array('administer search'),
  'type' => MENU_CALLBACK,
  'file' => 'search.admin.inc',
);
$items['admin/reports/search'] = array(
  'title' => 'Top search phrases',
  'description' => 'View most popular search phrases.',
  'page callback' => 'dblog_top',
  'page arguments' => array('search'),
  'access arguments' => array('access site reports'),
  'file' => 'dblog.admin.inc',
  'file path' => drupal_get_path('module', 'dblog'),
);

foreach (module_implements('search') as $name) {
  $items['search/'. $name .'/%menu_tail'] = array(
    'title callback' => 'module_invoke',
```

```
      'title arguments' => array($name, 'search', 'name', TRUE),
      'page callback' => 'search_view',
      'page arguments' => array($name),
      'access callback' => '_search_menu',
      'access arguments' => array($name),
      'type' => MENU_LOCAL_TASK,
      'parent' => 'search',
      'file' => 'search.pages.inc',
    );
  }
  return $items;
}

function _search_menu($name) {
  return user_access('search content') && module_invoke($name, 'search', 'name');
}

/**
 * Wipes a part of or the entire search index.
 *
 * @param $sid
 *  (optional) The SID of the item to wipe. If specified, $type must be passed
 *  too.
 * @param $type
 *  (optional) The type of item to wipe.
 */
function search_wipe($sid = NULL, $type = NULL, $reindex = FALSE) {
```

```
  if ($type == NULL && $sid == NULL) {
    module_invoke_all('search', 'reset');
  }
  else {
    db_query("DELETE FROM {search_dataset} WHERE sid = %d AND type = '%s'",
$sid, $type);
    db_query("DELETE FROM {search_index} WHERE sid = %d AND type = '%s'",
$sid, $type);
    // Don't remove links if re-indexing.
    if (!$reindex) {
      db_query("DELETE FROM {search_node_links} WHERE sid = %d AND type =
'%s'", $sid, $type);
    }
  }
}

/**
 * Marks a word as dirty (or retrieves the list of dirty words). This is used
 * during indexing (cron). Words which are dirty have outdated total counts in
 * the search_total table, and need to be recounted.
 */
function search_dirty($word = NULL) {
  static $dirty = array();
  if ($word !== NULL) {
    $dirty[$word] = TRUE;
  }
  else {
    return $dirty;
```

```
  }
}

/**
 * Implementation of hook_cron().
 *
 * Fires hook_update_index() in all modules and cleans up dirty words (see
 * search_dirty).
 */
function search_cron() {
  // We register a shutdown function to ensure that search_total is always up
  // to date.
  register_shutdown_function('search_update_totals');

  // Update word index
  foreach (module_list() as $module) {
    module_invoke($module, 'update_index');
  }
}

/**
 * This function is called on shutdown to ensure that search_total is always
 * up to date (even if cron times out or otherwise fails).
 */
function search_update_totals() {
  // Update word IDF (Inverse Document Frequency) counts for new/changed words
  foreach (search_dirty() as $word => $dummy) {
```

```
    // Get total count
    $total = db_result(db_query("SELECT SUM(score) FROM {search_index} WHERE
word = '%s'", $word));
    // Apply Zipf's law to equalize the probability distribution
    $total = log10(1 + 1/(max(1, $total)));
    db_query("UPDATE {search_total} SET count = %f WHERE word = '%s'", $total,
$word);
    if (!db_affected_rows()) {
      db_query("INSERT INTO {search_total} (word, count) VALUES ('%s', %f)", $word,
$total);
    }
  }
  // Find words that were deleted from search_index, but are still in
  // search_total. We use a LEFT JOIN between the two tables and keep only the
  // rows which fail to join.
  $result = db_query("SELECT t.word AS realword, i.word FROM {search_total} t
LEFT JOIN {search_index} i ON t.word = i.word WHERE i.word IS NULL");
  while ($word = db_fetch_object($result)) {
    db_query("DELETE FROM {search_total} WHERE word = '%s'", $word->realword);
  }
}

/**
 * Simplifies a string according to indexing rules.
 */
function search_simplify($text) {
  // Decode entities to UTF-8
  $text = decode_entities($text);
```

```
  // Lowercase
  $text = drupal_strtolower($text);

  // Call an external processor for word handling.
  search_invoke_preprocess($text);

  // Simple CJK handling
  if (variable_get('overlap_cjk', TRUE)) {
    $text = preg_replace_callback('/['. PREG_CLASS_CJK .']+/u', 'search_expand_cjk',
$text);
  }

  // To improve searching for numerical data such as dates, IP addresses
  // or version numbers, we consider a group of numerical characters
  // separated only by punctuation characters to be one piece.
  // This also means that searching for e.g. '20/03/1984' also returns
  // results with '20-03-1984' in them.
  // Readable regexp: ([number]+)[punctuation]+(?=[number])
  $text = preg_replace('/(['. PREG_CLASS_NUMBERS .']+)['.
PREG_CLASS_PUNCTUATION .']+(?=['. PREG_CLASS_NUMBERS .'])/u', '\1',
$text);

  // The dot, underscore and dash are simply removed. This allows meaningful
  // search behavior with acronyms and URLs.
  $text = preg_replace('/[._-]+/', '', $text);

  // With the exception of the rules above, we consider all punctuation,
  // marks, spacers, etc, to be a word boundary.
```

```
  $text = preg_replace('/[' . PREG_CLASS_SEARCH_EXCLUDE .']+/u', ' ', $text);

  return $text;
}[19]
```

APPENDIX B

File Upload Module for Vizlab Website

FileField provides a universal file upload field for CCK. It is a robust alternative to core's Upload module and an absolute must for users uploading a large number of files [22].

Features [22]

- Configurable upload paths allow you to save files into per-field or per-user directories
- Per-field and per-node file size limits
- Extensive API for extending field widgets and managing files
- Full revision/translation file management
- Views support
- Ajax Uploads

Code:

```
<?php
// $Id: filefield.module,v 1.209 2009/10/20 17:46:22 quicksketch Exp $

/**
 * @file
 * FileField: Defines a CCK file field type.
 *
 * Uses content.module to store the fid and field specific metadata,
```

```
 * and Drupal's {files} table to store the actual file data.
 */

// FileField API hooks should always be available.
include_once dirname(__FILE__) . '/field_file.inc';
include_once dirname(__FILE__) . '/filefield_widget.inc';

/**
 * Implementation of hook_init().
 */
function filefield_init() {
  // File hooks and callbacks may be used by any module.
  drupal_add_css(drupal_get_path('module', 'filefield') .'/filefield.css');

  // Conditional module support.
  if (module_exists('token')) {
    module_load_include('inc', 'filefield', 'filefield.token');
  }
}

/**
 * Implementation of hook_menu().
 */
function filefield_menu() {
  $items = array();

  $items['filefield/ahah/%/%/%'] = array(
```

```
    'page callback' => 'filefield_js',
    'page arguments' => array(2, 3, 4),
    'access callback' => 'filefield_edit_access',
    'access arguments' => array(3),
    'type' => MENU_CALLBACK,
  );
  $items['filefield/progress'] = array(
    'page callback' => 'filefield_progress',
    'access arguments' => array('access content'),
    'type' => MENU_CALLBACK,
  );

  return $items;
}

/**
 * Implementation of hook_elements().
 */
function filefield_elements() {
  $elements = array();
  $elements['filefield_widget'] = array(
    '#input' => TRUE,
    '#columns' => array('fid', 'list', 'data'),
    '#process' => array('filefield_widget_process'),
    '#value_callback' => 'filefield_widget_value',
    '#element_validate' => array('filefield_widget_validate'),
  );
```

```
  return $elements;
}

/**
 * Implementation of hook_theme().
 * @todo: autogenerate theme registry entrys for widgets.
 */
function filefield_theme() {
  return array(
    'filefield_file' => array(
      'arguments' => array('file' => NULL),
      'file' => 'filefield_formatter.inc',
    ),
    'filefield_icon' => array(
      'arguments' => array('file' => NULL),
      'file' => 'filefield.theme.inc',
    ),
    'filefield_widget' => array(
      'arguments' => array('element' => NULL),
      'file' => 'filefield_widget.inc',
    ),
    'filefield_widget_item' => array(
      'arguments' => array('element' => NULL),
      'file' => 'filefield_widget.inc',
    ),
    'filefield_widget_preview' => array(
      'arguments' => array('element' => NULL),
```

```
  'file' => 'filefield_widget.inc',
),
'filefield_widget_file' => array(
  'arguments' => array('element' => NULL),
  'file' => 'filefield_widget.inc',
),


'filefield_formatter_default' => array(
  'arguments' => array('element' => NULL),
  'file' => 'filefield_formatter.inc',
),
'filefield_formatter_url_plain' => array(
  'arguments' => array('element' => NULL),
  'file' => 'filefield_formatter.inc',
),
'filefield_formatter_path_plain' => array(
  'arguments' => array('element' => NULL),
  'file' => 'filefield_formatter.inc',
),
'filefield_item' => array(
  'arguments' => array('file' => NULL, 'field' => NULL),
  'file' => 'filefield_formatter.inc',
),
'filefield_file' => array(
  'arguments' => array('file' => NULL),
  'file' => 'filefield_formatter.inc',
```

```
    ),

);
} [22]
```

References

[1] An article comparing content management systems, www.idealware.org

[2] Angela Byron, Addison Berry, Jeff Eaton, Nate Haug, James Walker, and Jeff Robbin. Using Drupal. O'Reilly, December 2008

[3] Bio-1, http://www.bio-1stop.org

[4] Bio-1 Job matching site, http://www.bio-1stop.org/jobs/

[5] Drupal, www.drupal.org

[6] Drupal, http://en.wikipedia.org/wiki/Drupal

[7] Drupal: Getting Started, http://drupal.org/getting-started

[8] Drupal Handbooks 2009, http://drupal.org/handbooks

[9] Drupal Tutorial, http://drupalsn.com/learn-drupal/drupal-tutorials/drupal-vs-world-why-choose-drupal

[10] Drupal Usage Statistics, http://trends.builtwith.com/cms/drupal

[11] Gobala Krishnan. When CMS met SEO: Using the power of Content Management Systems, http://www.addme.com/newsletters/issue340.htm

[12] Hodge Silver, Hasin Hayder. WordPress 2.7 Complete

[13] Joomla, www.joomla.org

[14] Joomla CMS, http://www.wilsonmar.com/joomla.htm

[15] Joomla installation guide, http://www.siteground.com/tutorials/joomla/joomla_manual_installation.htm

[16] Joomla Usage Statistics, http://trends.builtwith.com/compare/Joomla!

[17] Model-view-controller, http://en.wikipedia.org/wiki/Model–view–controller

[18] Robert T. Douglass, Mike Little, and Jared W. Smith. "Building Online Communities with Drupal, phpBB, and WordPress", Apress, 2006

[19] Search Module Code, http://drupal.org/handbook/modules/search

[20] Step Two Designs, http://www.steptwo.com.au/papers/cmb_usability/index.html

[21] The art of Joomla, http://www.theartofjoomla.com/

[22] Upload File Module, http://drupal.org/project/filefield

[23] Webmaster Tips, http://tips.webdesign10.com

[24] Wiki dot, http://programmingexamples.wikidot.com/

[25] WordPress, www.wordpress.org

[26] WordPress, http://en.wikipedia.org/wiki/Wordpress

[27] WordPress 2.7 User Handbook, http://wordpress.org/docs/en/handbook/2.7/#preface

[28] WordPress Usage Statistics, http://en.wordpress.com/stats/

[29] Worldwide telnet CMS, http://www.wwtelenet.com/articles/web-site-design-and-development/the-benefits-of-a-solid-content-management-system.html