Content Management Systems: Drupal Vs Jahia

Mrudula Talloju
Department of Computing and Information Sciences
Kansas State University
Manhattan, KS – 66502.
mrudula@ksu.edu

Abstract

Content Management Systems (CMS) is the category of software that helps to manage all unstructured information. Various CMS are developed in various programming languages which support different databases. Here, two open source CMS, Drupal and Jahia, are compared based on their architecture, features supported, technologies used and ease of use. Finally, an example (Department of Computer Science website) is developed in Jahia.

1. Introduction:

CMS is the category of software that helps to manage all unstructured information or content. This information exists in many digital forms: text documents, images, audio and video files, and many other file types and formats. CMS supports the creation, management, distribution and publishing of the information. Web CMS typically focuses on the online content targeted at either corporate website or intranet.

Various CMS software are developed in programming languages like PHP, Perl, ASP.NET, VB.NET, Java, etc and support databases like MySQL, Oracle, PostgreSQL, etc. Also there are free-and-open source and commercial CMS software.

Jahia is a 100% Java based and open source CMS software. It supports HyperSonic SQL, MySQL, PostgreSQL, Oracle and Microsoft SQL server databases. Commercial editions of Jahia are also available at low cost, which supports additional features. The architecture of Jahia, technologies used and features supported, benefits and drawbacks are discussed in Section 2.

Drupal is also open source CMS software licensed under the GPL. It is developed in PHP and supports MySQL and PostgreSQL databases. It is maintained and developed by a community of thousands of users and developers. The architecture of Drupal, its terminology, benefits and drawbacks are discussed in Section 3.

Section 4 presents a comparison of the two CMS software. Section 5 presents an example CMS developed in Jahia.

1

2. Jahia:

Jahia is open source CMS software that integrates various servers and technologies which make it powerful. The technologies used in Jahia and the way they are mapped into its architecture are discussed below.

2.1 Jahia Technologies and Features Supported:

- *Enterprise Content Management Server:* This server manages the unstructured information of an organization. Jahia relies on its own content repository which manages all the content objects, their interactions (create, update, delete, import and so on) and their properties (lock, languages, permissions and so on). It uses technologies like WYSIWYG editor (based on the FCK editor), JTidy integration, AJAX powered GUI and sitemaps, SSO support and JSP templates with taglibs. These technologies empower Jahia to provide better features like:
    - i. Complete web based interface.
    - ii. Multilanguage Support.
    - iii. WYSIWYG editing.
    - iv. Automatic sitemaps management.
    - v. Staging: Authors may create or modify content in a staging environment. The author has three possible views of his current web page: live view, draft view and edit view.
    - vi. Default set of templates: Jahia is provided with a default set of templates which can be customized to fit the needs.
    - vii. Dynamic URL checking to avoid 403 and 404 errors.
    - viii. Fields lock in/ lock out: This is used to avoid the two users modifying the same content and also to lock the fields under validation in order to let the managers review the content.
    - ix. Time based publishing: Editor can setup a start and end date on every object of content.
    - x. Versioning capabilities.
    - xi. XML import/export: The content of a site or a page can be dynamically imported or exported in an XML format.

- *Document Management Server:* Jahia includes a document management server which supports WebDAV protocol, a protocol that provides functionality to create, change and move documents on a server. In Jahia, the uploading and managing of files can be handled directly from the windows web folders. It uses technologies like PDFBox for indexing pdf files. The features Jahia supports are:

  i. Index and search for files: Jahia automatically indexes all office and pdf files and allows making advanced searches in the repository of files.

  ii. Automatic thumbnail creation.

  iii. Drag and drop the files.

- *Search and Indexation Server:* Jahia provides multilanguage support. Indexing the structured content which dynamically changes on the same page according to the language, mode or user is difficult. This is done by Jahia Indexation Server. It also allows us to search the content based on hidden metadata like author, validation date and so on. Search and indexation server uses Apache Lucene based kernel for implementing the required search and indexation functions. The features it provides are:

  i. Automatically index all Jahia content.

  ii. Give distinct search results based on the user, language or navigation mode.

  iii. Saved searches.

- *Portal Server:* Jahia supports portal technology. It integrates Portlet Builder which allows developing new portlets with less or no technological skills. It supports SQL portlets and web portlets.

- *Collaborative Suite:* Collaborative Suite extends the features of Jahia portal server to deliver ready to use collaboration and productivity based features. It helps to create a discussion forum, calendar server, chat room and so on.

- *Business Process Management (BPM) Server:* Jahia BPM server helps to create and manage complex business process or validation workflows. These workflows can be attached to any page or content object in Jahia. The features supported include:

  i. Manage complex workflows.

  ii. Monitor the business processes.

- *Cache Proxy Server:* Jahia includes a front-end cache proxy server to reduce response time and support massive user loads. This server automatically caches and assembles HTML fragments without regenerating them from the underlying Jahia application server and database. The cache proxy server is not available in the community (open source) edition. It is available in the commercial editions of Jahia.

2.2 Jahia Architecture:

Jahia can be divided into seven categories. Details of each category are discussed below.

1. _Third Party Systems:_ Jahia relies on several third party systems like: a database, access to File System (since Jahia stores all the large texts or the configuration files on the file system), a front-end web server (for dynamic URL rewriting) and a LDAP server to support single sign-on (SSO).

Figure 1: Jahia Architecture
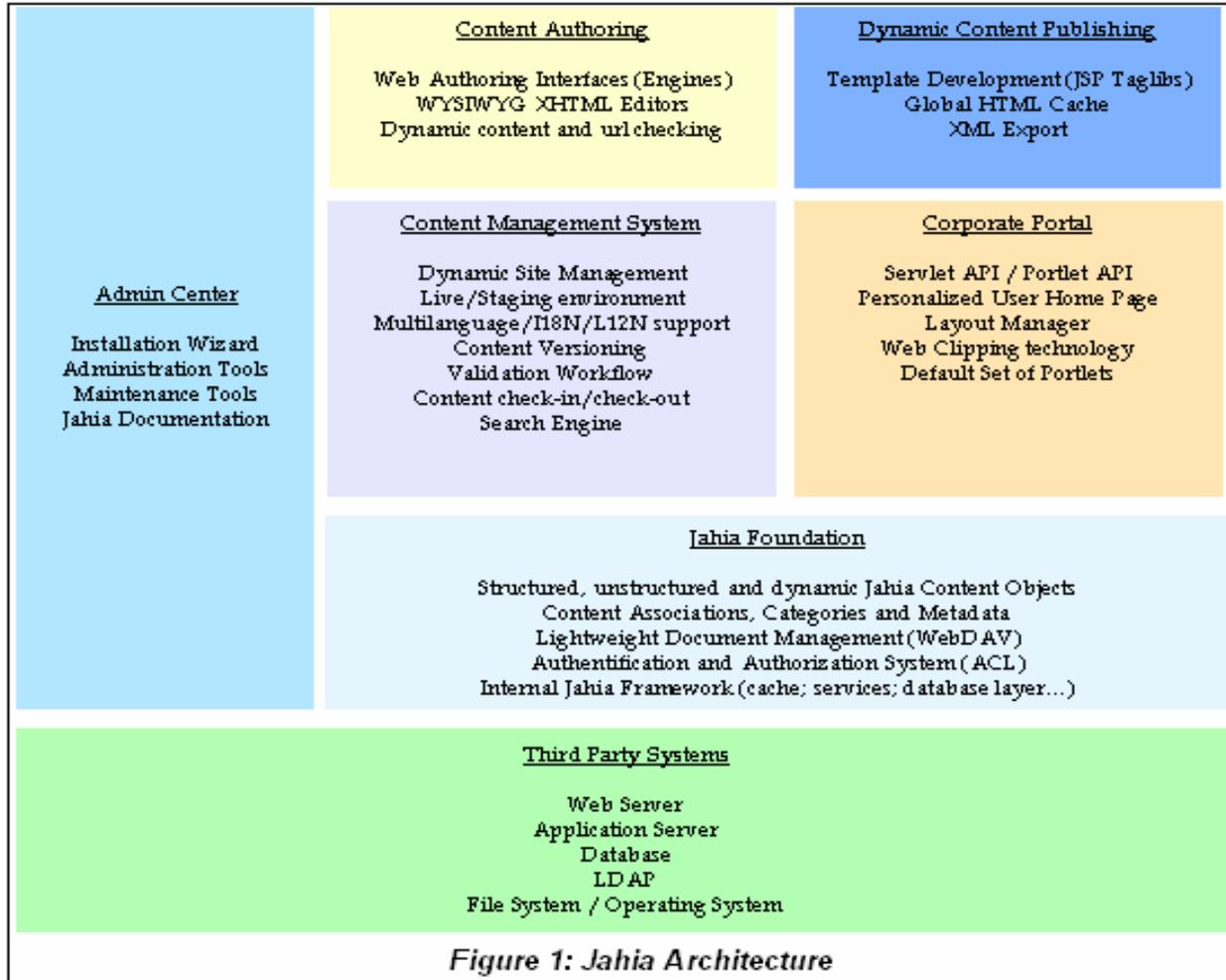
2. _Jahia Foundation Server:_ The foundation server provides all the generic services to access and manage the various Jahia content objects and metadata. The document management layer manages all the binary files. Authentication and authorization system provides services to authenticate and manage user permissions.

3. _Content Management System:_ CMS supports Multilanguage (internationalization I18N and localization L12N); the Live/Edit/Preview modes and the built-in workflow engine to move from one mode to another and content versioning. Further, CMS has an integrated search engine.

4. _Corporate Portal Server:_ Jahia offers personalized home or group page features to create portals. A complete layout manager interface lets users to personalize their own web pages and move/resize their portlets. Jahia comes with pre-built web applications like discussion forum, FAQ, Time Sheet and so on. Since the applications are open source, it helps developers to see how they are implemented and create their own applications.

5. _Content Authoring:_ Jahia provides several GUIs to help users input and manage content. These interfaces are called "engines". Each content object will have its own custom engine i.e., a date field will launch a date picker interface while a text field will launch WYSIWYG HTML editor. Before storing into the file system, the inputted HTML text is parsed by JTidy to remove some tags (like <font>). The hard coded links are directly managed by Jahia.

6. _Dynamic Content Publishing:_ Jahia dynamically aggregates all the content objects including static content and dynamic portlets and renders them through JSP template. XML export feature allows to output a page or a content object in XML which can be reused else where.

7. _Admin Center:_ Admin center provides administration and maintenance tools. It provides central control of multiple virtual web sites from single interface; site language settings, administration of users and groups and other low-level database and cache management.

2.3 Benefits:

Jahia includes various servers and technologies which provide various features and hence making it a powerful CMS. All the features listed in section 2.1 can be considered as benefits. Further, Jahia is well documented and also provides online guided tours.

2.4 Drawbacks:

Even though Jahia is a powerful tool for building and maintaining a CMS, it has few drawbacks.

- Some of the features like saved searches, advanced search form, HTML proxy/cache server are not available in the open source edition.
- An application developed in Jahia cannot run in CGI mode or other low-end systems.
- Jahia does not allow users to upload internal content and files via FTP.
- Jahia does not have any applications to support creating arbitrary data entry applications and database reports.
- Jahia does not support the security feature, Email Verification i.e., it does not send an activation key to users to make sure that they have entered a valid email address.

3. Drupal:

Drupal is a modular content management system and blogging engine. It is developed in PHP and is maintained and developed by a community of thousands of users. Drupal uses a terminology of "node", "block", "role" and "theme" to represent content, box of content, privileges of a user and page layout respectively. Drupal terminology and its modular architecture are discussed below.

3.1 Drupal Terminology:
- _Node:_ Content in Drupal is represented as "nodes". Different node types available in Drupal are blog entry, book page, forum (which in turn is a set of nodes and their comments), page, poll, story and so on.
- _Block:_ Blocks are the boxes of content which can be rendered into certain regions of the page i.e., header or side bars or content column. For example, a block can represent links for navigation in the side bar. Blocks can be enabled or disabled. Only enabled blocks are shown. Certain blocks can be configured to automatically disable themselves temporarily during high server loads. This feature is called auto-throttle.
- _Role:_ Roles represent the privileges of a user. Depending on the role, few blocks may not display.
- _Theme:_ Theme describes the look and feel of the website. Drupal comes with a set of default themes which can be used to change the layout of the website. These themes are written in the PHPTemplate engine or the XTemplate engine.
- _Modules:_ Modules are plug-ins for Drupal to extend its core functionality. Modules can be enabled or disabled depending on the application. Many modules developed by the users and developers of Drupal are available on the website, which can be downloaded and used to add additional features.

3.2 Drupal Architecture:
Drupal's architecture is modular. It has core modules which have functions for content and user management. Other modules can be plugged in for specific features. All these modules can be enabled or disabled depending on the requirement.
- _Drupal Core:_ Drupal core contains modules for content and user management, taxonomy, localization, templating and syndication.
  - ➤ _Taxonomy_ module allows user to create categories to organize content by type. For example, music files can be categorized based on genre. Drupal further allows defining vocabulary, a set of categories, to organize content. Drupal's flexible taxonomy allows creation of unlimited separate classification schemes. Each of these schemes can be arranged as a list or tree. Administrators can then use these schemes to associate with each content type and users can then browse this tagged content by its taxonomy terms.

➢ *Localization* module allows the translation of website to make it accessible in one or more languages.

➢ *Syndication:* Drupal provides "news aggregator", a powerful RSS news reader that can gather fresh content from news sites or blogs and make it available from the Drupal website.

Figure 2: Drupal Architecture
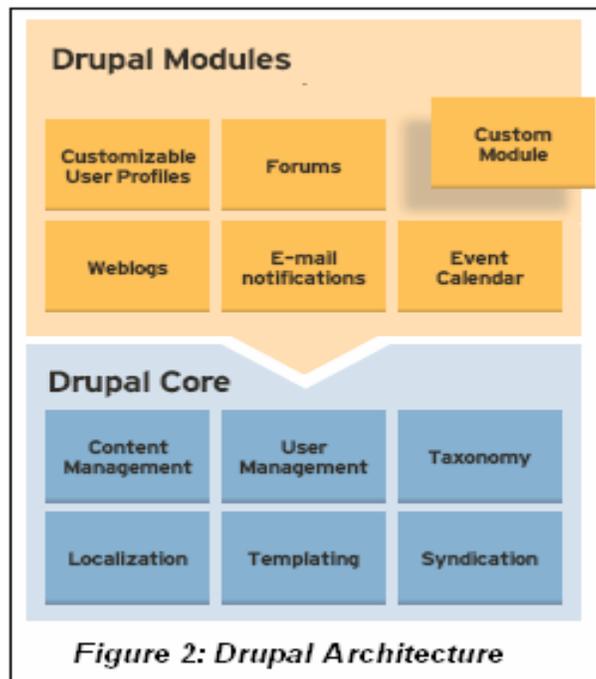
- *Drupal Modules:* Drupal modules focus on the specific features. The Drupal core binds all the modules together and hence modules integrate seamlessly into Drupal. There are many modules, contributed by the users and developers, available on the website. These can be downloaded and used as it is or they can be customized.

3.3 Benefits:

The benefits of Drupal are:

- Drupal's architecture is truly modular. This allows any module to be easily plugged in and used.
- Web-based administration.
- Multilanguage support.
- It provides applications to support creating arbitrary data entry applications.
- It allows user to upload content and files via FTP.
- It has a built-in search engine.
- Many modules are available on the website which can be downloaded and customized. Few examples of such modules include photo manager, advertisement, amazon search and other e-commerce modules.

- Many themes are available on the website which can be downloaded and used to change the look-n-feel of the website.

- Full discussion forum features are built into Drupal to create community sites.

- WYSIWYG editors are available as free plug-ins. This allows users to create and publish formatted content without knowing HTML, CSS or XML.

- Drupal supports content versioning.

- Drupal supports creating user and search-engine friendly URLs.

- Drupal has an integrated search engine that indexes the content and allows users to search the indexed content.

- It supports the security feature: Email Verification i.e., it sends an activation key to the users to make sure that they have entered a valid e-mail address.

- Drupal can print browser-based reports with information about the referrals, content popularity and how visitors navigate the site.

- Unlike Jahia, Drupal provides all the features it claims to support, in the open source edition.

3.4 Drawbacks:

Few of the drawbacks of Drupal are listed below.

- Configuration of Drupal was difficult in the earlier editions of Drupal. This is resolved in Drupal 5. However, it is still complex when compared to Jahia.

- Browser compatibility is an issue in Drupal. Few of the modules work well in Internet Explorer but not in other browsers like Mozilla Firefox.

- Drupal does not support zip archives. Zip archives are used for mass uploading the static content of the application/website.

- To develop new themes, the developer needs to have a good knowledge of PHP and PHPTemplate engine.

- The architecture of Drupal, its working and the technologies used are not well documented. This makes the new user difficult to understand the working of Drupal.

4. Jahia Vs Drupal:

      The following table compares Jahia and Drupal based on system requirements, applications, ease of use, interoperability and performance.
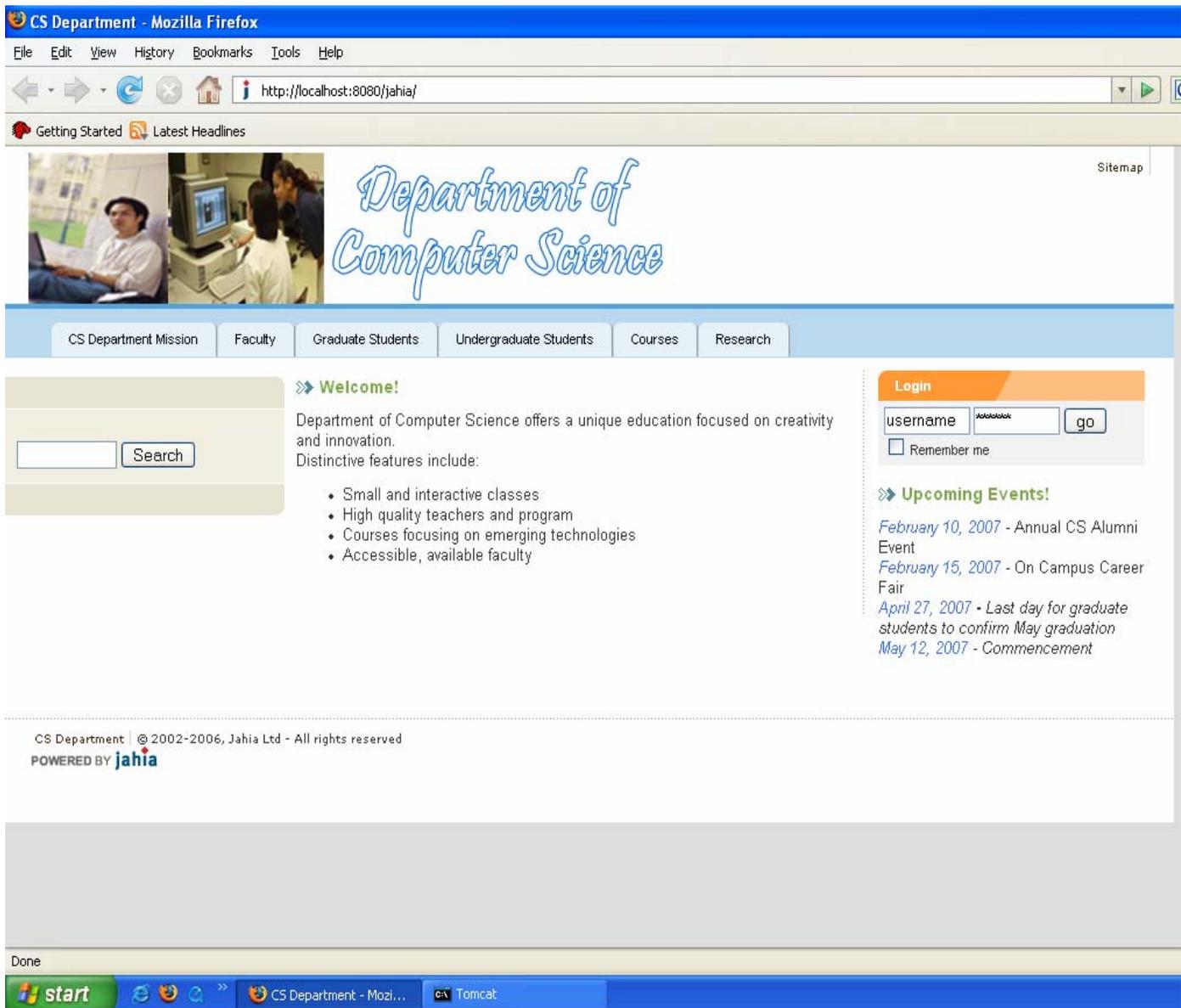
	Jahia	Drupal
System Requirements		
Application Server	J2EE	PHP 4.3.3 +
Approximate Cost	Free, Commercial editions also available from 0 - 40k	Free
Database	HyperSonic SQL, MySQL, PostgreSQL, Oracle, Microsoft SQL Server	MySQL, PostgreSQL
Operating System	Platform Independent	Platform Independent
Programming Language	Java	PHP
WebServer	Apache, IIS or any J2EE server	Apache, IIS
Built-in Applications		
Blog	Yes	Yes
Chat	Free Plug-in	Yes
Content Management	Free Plug-in	Yes
Data Entry	Free Plug-in	No
Document Management	Limited	Yes
Events Management	Free Plug-in	No
Graphs and Charts	No	No
HTTP Proxy	No	Yes
Photo Gallery	Free Plug-in	Yes
Polls	Yes	Yes
Search Engine	Yes	Yes
Site Map	Free Plug-in	Yes
Syndicated Content (RSS)	Yes	Yes
Ease Of Use		
Drag-n-Drop Content	No	No
Friendly URLs	Yes	Yes
Image Resizing	Free Plug-in	Yes
Spell Checker	Free Plug-in	Yes
Undo	Limited	Yes
WYSIWYG Editor	Free Plug-in	Yes
Zip Archives (for mass upload)	No	Yes
Interoperability		
Content Syndication (RSS)	Yes	Yes
FTP Support	Limited	No
Performance		
Advanced Caching	Yes	Yes
Load Balancing	Yes	Yes
Static Content Export	No	No

Table 1: Jahia Vs Drupal

5. Example Application in Jahia (*Department of Computer Science website*):

Department of Computer Science website provides information required for the graduate and undergraduate students. It further includes the information about the courses offered, research and faculty. "siteAdmin" is the administrator of the website. Each professor has a web space for their web pages to include details of the courses they teach and other information. "prof1" and "prof2" are two professors (users) who have all the privileges to update their respective pages.

The home page of the website is shown below:

Figure 3: Department Home Page

Faculty page gives the list of professors and links to their homepage. This page can be updated by "siteAdmin" only. When logged in as administrator, the page has edit mode.
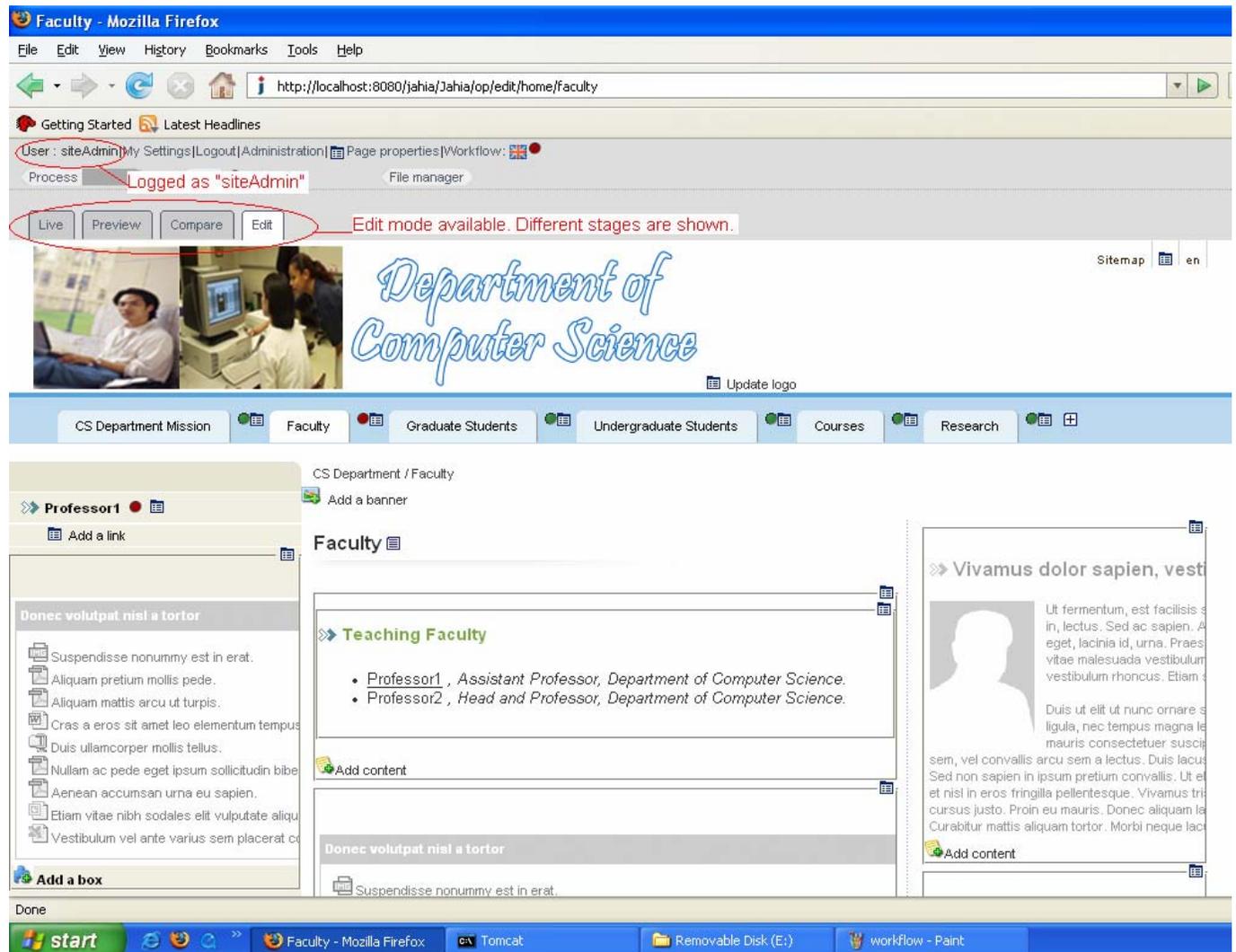

Figure 4: Faculty page when logged in as administrator

When logged in as "prof1" or "prof2" or any other user who does not have privileges to update the page, the edit mode is not displayed. This is shown in Figure 5.

This shows how the website developed in Jahia can be administered from anywhere through browser unlike the traditional websites.
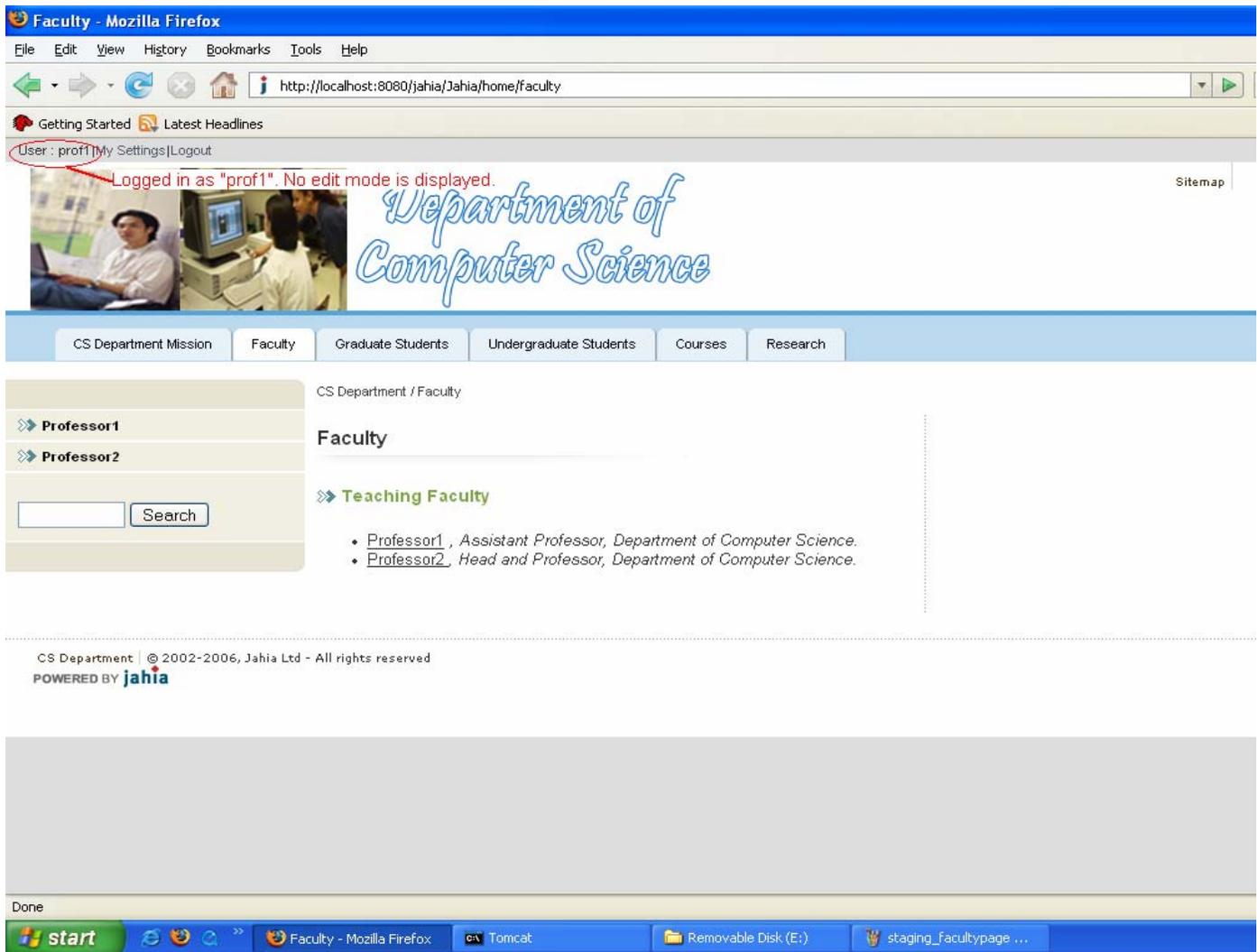
11

Figure 5: Faculty page when logged in as "prof1"

6. Conclusion:

CMS soft wares provide features that enable users with little or no technical knowledge to administer, develop and maintain complex applications. While Drupal is widely used to develop personal websites, blogs, discussion sites and social networking sites, Jahia is used to develop intranet and internet applications in various sectors like public, educational, manufacturing, banking and financial services and so on. However, neither of these two supports a very useful feature: drag-n-drop content. Including this feature and other useful features (like graphs and charts) can be the next step in the CMS software development.

References:

1. Wikipedia - Content Management, http://en.wikipedia.org/wiki/Content_management.
2. Wikipedia - Content Management Systems, http://en.wikipedia.org/wiki/Content_management_system.
3. The Open Unified Enterprise Content Management and Portal Server - JAHIA, http://www.jahia.org/jahia/jsp/index.jsp.
4. An open source content management platform - Drupal, http://drupal.org/.