

Content Management System : Comparative Case Study

Manish Nath

Department of Computer Science and Information Technology, Jaypee Institute of Information Technology, Uttar Pradesh, INDIA
manish.nath87@gmail.com

Anuja Arora

Department of Computer Science and Information Technology, Jaypee Institute of Information Technology, Uttar Pradesh, INDIA
anuja.arora@jiit.ac.in

Abstract— In this paper, the various features essential for a content management system are explored. This includes a comparison of various open source CMS , their features and a conclusion on improvements that will be made in order to alleviate problems . Our aim was to find out the best available open source CMS and implement it for our website. For this we conducted a study on 13 java based open source CMS available taking into account 29 features like the technology used for creation of the CMS and features like anti plagiarism , Digital Asset Management(DAM) , etc. for user convenience .

Keywords- Content Management System, digital asset management, anti plagiarism ,features ,Content.

I. INTRODUCTION

This paper outlines the purpose of development of the CMS and its future prospects for the website. A study of 13 different CMS and their features are covered in the paper . These features will be used to conclude the best content management system from a user and developer's point of view. Also creation criteria for emancipation of current CMS is proposed.

II. THE PROCESS OF SELECTING A CMS

[4]A study on existing open source CMS through a testing phase. Then 29 features were enlisted as selection and creation criteria. The availability of these features enlightened us on the best open source CMS available

A. Open Source Survey

For our project we compared 13 open source CMS developed in java which used basic java and advanced java both. We recorded totally 29 features for a CMS and compared those features in these 13 CMS and generated a small report which feature can be taken from which CMS.

[8]1. Atleap, [9]2.Hoarder, [10]3.HonoCMS, [11]4.Hono Generate, [12]5.Magnolia, [13]6.Co-coBlog, [14]7.Roller Weblogger, [15]8.SimpleWeb, [16]9.Cofax, [17]10.OpenCms, [18]11.NuxeoECM, [19]12.Alfresco, [20]13.dotCMS. The Following Table Represents availability of various features in the 13 CMS.

FEATURES	1	2	3	4	5	6	7
automatic UI generation	no	yes	no	no	yes	no	no

Dynamic control	yes						
Automatic User signup	no						
forgotten password notification	no						
email confirmation	no						
Backup	yes	no	yes	yes	no	no	no
Navigator	no	no	yes	no	no	no	no
Admin Control	no	no	yes	no	no	no	no
Content management	yes						
Digital Asset	no	no	yes	no	no	no	no
User manager	no	no	yes	no	no	no	yes
Template	no	no	yes	no	no	no	yes
Platform Independence	yes	no	yes	yes	yes	no	yes
Multiple Database repository	yes	no	yes	yes	no	no	no
Code auto generation	no	no	no	yes	no	no	no
Multiple user	yes	no	yes	no	yes	no	yes
File uploads	yes	no	yes	no	no	no	no
File download	yes	no	yes	no	no	no	no
document management	yes	no	no	no	no	no	no
comments	no	no	no	no	no	no	yes
web content management	yes	no	no	no	no	no	no
Spring	yes	no	no	no	no	no	no
Hibernate(ORM)	yes	no	no	no	no	no	no
Search (LUCENE)	yes	no	no	no	no	no	no
JSF	yes	no	no	no	no	no	no
Anti Plagiarism support	no						
contact us	yes	no	no	no	no	no	no
Forums	yes	no	no	no	no	no	no
site map with auto generation	yes	no	no	no	no	no	no

FEATURES	8	9	10	11	12	13
automatic UI generation	no	yes	yes	no	no	no
Dynamic control	yes	yes	yes	yes	yes	yes
Automatic User signup	yes	no	no	no	no	no
forgotten password notification	yes	no	no	no	no	no
email confirmation	yes	no	no	no	yes	no
Backup	no	no	yes	no	yes	no
Navigator	no	no	yes	no	no	no
Admin Control	yes	yes	yes	no	yes	no
Content management	yes	yes	yes	yes	yes	no
Digital Asset	no	no	no	no	no	no
User manager	no	yes	yes	no	yes	no
Template	no	yes	no	no	no	no
Platform Independence	no	yes	yes	no	yes	yes
Multiple Database repository	no	no	no	no	yes	yes
Code auto generation	no	no	no	no	yes	no
Multiple user	no	yes	yes	yes	yes	no
File uploads	yes	yes	yes	yes	yes	no
File download	no	yes	yes	yes	yes	no
document management	no	yes	yes	yes	yes	no
comments	no	no	yes	no	no	no
web content management	no	yes	yes	yes	yes	yes
Spring	no	no	no	no	yes	no
Hibernate(ORM)	no	no	no	no	yes	no
Search (LUCENE)	yes	no	no	yes	yes	yes
JSF	no	no	no	yes	yes	no
Anti Plagiarism support	no	no	no	no	no	no
contact us	no	no	no	no	no	no
forums	no	no	no	no	no	no
site map with auto generation	no	no	no	no	no	no

B. Selection Criteria and availability status

- Automatic UI generation: User interface generated automatically for a modification in code. The best suited is hoarder from the one's surveyed.
Availability: 4
- Dynamic control: Dynamic control and handling of data provided. This is present in mostly all products surveyed.
Availability: 13

- Automatic user signup: Present only in one product . Allows the user to sign in automatically at the time of opening of the site.
Availability: 1
- Forgotten password notification: This feature supports users who have forgotten there password. This is a notification sender of a new random password. This is present in only one product.
Availability: 1
- Email confirmation: Present in only one product this feature confirms a new user id to the user.
Availability: 1
- Backup: Data backup facility for a user to save his data. The cms manages the data and saves it through orm. The best available product for this feature is atleap
Availability: 4
- Navigator: Site navigator provides a map to the user showing his current location on the site map. The best product for this feature is honocms
Availability: 2
- Admin Control: Feature of high importance. Admin control is a must for a secured cms . The best product for this feature is atleap.
Availability: 5
- Content management: A necessary feature and present in all CMS.
Availability: 12
- Digital Asset: Manages mpeg and avi files from the user. Provided only in one product (honocms).
Availability: 1
- User manager: Manages the user account information and authentication .The best product for this feature is honocms.
Availability: 5
- Template: Manages previously stored templates and provides it to the user for dynamic interfacing. The best product for this feature is honocms.
Availability: 3
- Platform Independence: Platform independent code. All products are platform independent written in java.
Availability: 9
- Multiple Database repositories: multiple storage of database insures backup and safety from crashes. Honocms is the best product available for this feature.
Availability: 5
- Code auto generation: auto generation of code on drag and drop facility for the interface. The management of code is present in cms. Present in only one product honocms.
Availability: 2
- Multiple users: multiple user access management through locking facility. The best available product is atleap for this feature.
Availability: 8

- File uploads: management of files uploaded by the user. The best product available is atleap for this feature.
Availability: 7
- File download: management of file downloads by a user. The best product is atleap for this feature.
Availability: 6
- Document management: management of private documents of the user for his ongoing work. The best product is atleap for this feature.
Availability: 5
- Comments: facility to provide comments. Feature present only in roller web logger.
Availability: 2
- [5]Web content management: management of web content like web pages, images, etc. the best product is atleap for this feature.
Availability: 6
- Spring: framework to impart easy coding functionality provided through IOC and AOP. Atleap is the best available product for this feature.
Availability: 2
- Hibernate (ORM): object relational mapping tool to take care of database burdens. at leap is the best available product for this feature.
Availability: 2
- Search (LUCENE): search of web content, files, classes, services etc. within the system. The best product is atleap for this feature.
Availability: 5
- JSF: java server faces is a framework to simplify development of ui . The best product is atleap for this feature.
Availability: 3
- Contact us: facility to contact the site maintenance in case of any query. Present in atleap only.
Availability: 1
- Forums: forums for discussion by the user. Present in atleap only.
Availability: 1
- Site map: site map to guide the user through the site. Best product is atleap for this feature.
Availability: 1

III. PROPOSED SYSTEM

- Content Format Checking: This feature can be used where content can be delivered in two or more formats like any document can be delivered in .pdf, .doc, .tex etc. or media files can be delivered in .avi , .dat, etc. Automatic identification of submitted format and conversion to desired format can be beneficial for the user.
Availability: 0
- Anti Plagiarism support: [7] Plagiarism that is plagiarizing or composition retrieval, opinion, etcetera from other people and makes it is likely composition and opinion him-self. This feature

enables a conference managing website to check for plagiarized content submitted on their system.
Availability: 0

From the open source survey conducted it was clear that alfresco was clearly ahead in terms of technologies it was built on and the features that it supported . [2]Alfresco is a platform for Enterprise Content Management (ECM). ECM describes many content based technologies like:

- Document Management (DM): Capturing, organizing, and sharing binary files. These files are typically produced from office productivity software, but the scope of the files being managed is unlimited.
- Web Content Management (WCM): Managing files and content specifically intended to be delivered to the Web. The main purpose of WCM is to enable non technical users to be able to easily publish their own web content.
- Digital Asset Management (DAM): Managing graphics, video, and audio. Functionality specific to the needs of working with rich media such as thumbnailing, transcoding, and editing can be easily handled.
- Records Management (RM): Managing content as a legal record. Like DAM, RM starts with DM and adds functionality specific to the world of RM such as retention policies, records plans, and audit trails.
- Imaging: This includes capturing, tagging, and routing images of documents from scanners.

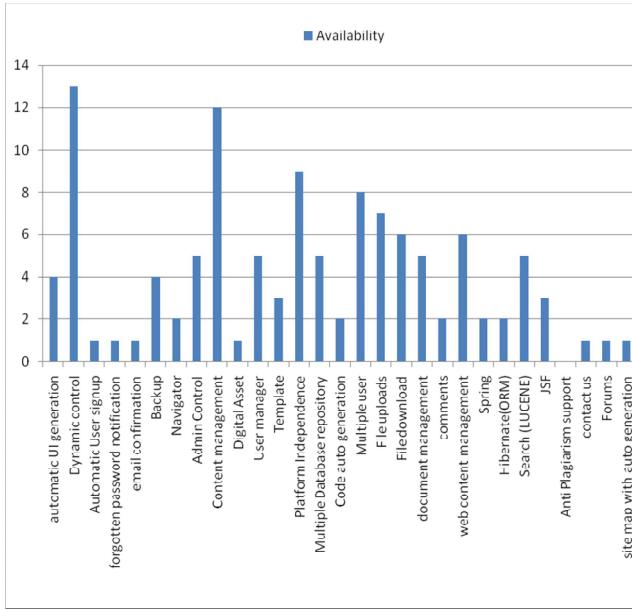
Alfresco is built on state-of-the-art open source components such as Spring, Hibernate, Lucene, JSF – often the components of choice website developers today. It offers one repository for the whole team. This repository is a modern platform for Web Content Management 2.0.

This new open source WCM platform delivers significant benefits in the areas of:

Time – Weeks to hours for change sets to be deployed
Cost – One Virtualization Server
Quality – One repository and One virtualization preview
Compliance – Common Governance Framework.
One virtualization server for past, present and future
Service Oriented Auditing, Migration – Simple file based migration

IV. RESULTS

This graph denotes the availability vs features . from the availability of various features we can clearly identify an optimum Content management system . These features have been identified by studying and implementing all 13 CMS listed in our case study.



V. CONCLUSION

It can be clearly seen that technologies like Hibernate , Spring and features like DAM, Anti plagiarism were absent in most CMS surveyed. We propose Alfresco as the best available open source CMS. We have successfully implemented Alfresco WCM for our website.

- We have defined the workflow which keeps schedule of the work in progress.
- We have defined users and roles for users to add delete update or modify content within the website.
- We have defined different sandboxes i.e. user sandbox and staging sandbox where in the user copy and published copy is different.
- Each user can modify data according to his access rights.
- Admin can lock any part of the content where he doesn't want any modification. That content becomes read only for the rest users.
- Admin has full control over each user and his activity.
- Only Admin assigns roles to users.
- While modifying a particular file the file is downloaded to the user system via user sandbox and the user can check in the file again after modification to the staging area.
- Any file can be uploaded to user sandbox by a user depending on his access rights as defined by the admin. These can be submitted to the workflow
- Lucene Search index is generated for the content and users to search within the system.

VI. FUTURE WORK

Our next aim is to implement Document management for conferences. This will contain automatic format checking for the Research papers submitted for the conference. Any discrepancy in the format shall lead to automatic email generation for the user about the error.

[7]Plagiarism is plagiarizing or composition retrieval, opinion, etcetera from other people and makes it is likely composition and opinion him-self. For this algorithms like Karp Rabin can be used to determine the percentage of copied content . Also one can use paid or free/paid services available online like Turnitin etc. Another way is to use the google API to match for a research paper for web based copying. Future work shall include implementation of [6]anti-plagiarism as web service for the website which shall identify any sort of copying within the data submitted by the users and reject it if copy percentage found above defined limits. This anti plagiarism application is expected to be applicable at the education world. The aim of this application is to look for percentage of similarity two the files.

REFERENCES

- [1] Greg DeitzTexas A&M University MS: Identification Developing a Content Management System for SIGUCCS,3142 Texas A&M University College Station, TX 77843 979.845.3893
- [2] Jeff Potts : Alfresco Developer Guide: Packt Publishing
- [3] Bob Boiko: Content Management Bible, 2nd Edition.Wiley Publishing . INC
- [4] Ann Amsler, Ron Nichols : Blazing a Web CMS Trail at the University of Delaware
- [5] Jurriaan Souer, Lutzen Luinenburg , Johan Versendaal, Inge van de Weerd, Sjaak Brinkkemper iiWAS 2008 Proceedings of iiWAS2008
- [6] Turnitin <http://turnitin.com/static/index.html>
- [7] Benny Mutiara , Sinta Agustina : Anti Plagiarism Application with algorithm Karp-Rabin at thesis in Gunadarma University.
- [8] Atleap (<http://java-source.net/open-source/content-managment-systems/atleap>)
- [9] Hoarder (<http://hoarder.sourceforge.net/>)
- [10] HonoCMS (www.honoweb.net/)
- [11] HonoGenerate (<http://www.honoweb.net/index.php?menuid=504&level=3>)
- [12] Magnolia (<http://www.magnolia.info/>)
- [13] CocoBlog (<http://www.beblogging.com/blog/docs>)
- [14] RollerWeblogger (<http://www.rollerweblogger.org/page/project>)
- [15] SimpleWeb (<http://www.russellbeattie.com/notebook/20020813>)
- [16] Cofax (<http://www.cofax.org/>)
- [17] OpenCms(<http://www.opencms.org/opencms/opencms/>)
- [18] NuxeoECM(<http://java-source.net/open-source/content-managment-systems/nuxeo-ecm>)
- [19] Alfresco(<http://java-source.net/open-source/content-managment-systems/alfresco>)
- [20] dotCMS(<http://java-source.net/open-source/content-managment-systems/dotcms>)