

Digital Content Management

Madhura K
*Assistant professor,
CSE Department,
Presidency University*

Abstract: Content management (CM) is the administration of digital and unstructured content throughout its lifecycle, from creation to permanent storage or deletion. The content involved may be images, video, audio and multimedia as well as text. An organization consists of 10's of terabytes of unstructured contents like documents images that are essential to its business and functioning.

A good CM tool ensures the digital content's integrity, availability and security. There are various CM tools available like Sharepoint, FileNet, Documentum, OpenText that allows the management of an organization's unstructured information, wherever that information exists. However most of these tools are costly in-terms of licencing fees(based most of the time on user count) and installation in Data Centre and maintenance. This is the reason why most of the public organizations, universities in India are still not purchasing such CM Products and instead rely on saving their organizations content via Paper trail or on computer drives, CD's, Google Drive, etc. This kind of storage mechanism doesn't ensure contents integrity, availability and security.

The main objective of the project is to create a Digital Content Management (DCM) product which is hosted in public cloud like MeghRaj(National Cloud by NIC),Amazon Web Services (AWS) or Microsoft Azure. The Digital content management product will be built using Lambda services, elastic search and The Content will be stored using cloud storage devices like S3/Glacier. The content will be encrypted to ensure integrity of data. The Product will have interfaces on mobile and web to allow users to push documents to the DCM tool and to retrieve it with accuracy at any point of time. Furthermore DCM will define an extensible taxonomy to define metadata (information about the content) so that users can define their own search criteria for searching the document.

1. Introduction

A Digital **content management system (DCM)** is a computer application that supports the creation and modification of digital content. It is often used to support multiple users working in a collaborative environment.

Digital Content management System (DCM) allows publishing, editing, and modifying content as well as its maintenance by combining rules, processes and/or workflows, from a central interface, in a collaborative environment.

DCM serve as a central repository for content, which could be, textual data, documents, movies, pictures, phone numbers, and/or scientific data. This can then be used by users of the enterprise/organization across different applications. The web published content can also be distributed to customers and business partners outside the organization. The core application of the DCM is to manage content during its entire lifecycle i.e. from creation through publishing. The content of the DCM can also be shared by e-commerce and customer relationship management systems (CRM). Web Publishing system enables you to establish a consistent look and feel throughout your site, but gives your non-technical content authors the power to publish and update their own content using simple, but powerful, browser-based tools. Some of the DCM systems integrate with content delivery applications to deliver the content via a web site.

Thus in this project following activities will be taken up-

- 1) Develop a dynamic taxonomy model that can be used by organizations to store their documents.
- 2) Design a security model to ensure document integrity and security.
- 3) Design and Develop DCM System in Cloud-MeghRaj(National Cloud by NIC), Amazon AWS or AZURE.

- 4) Design and Develop an interface that is compatible to content management interoperability standard (CMIS) for document ingestion and retrieval.
- 5) Built a publishing framework that can render and publish content to Web.
- 6) Demonstrate the capability of DCM by onboarding one of the departments of Acharya Institute.

It is envisaged that many educational institutions, public work departments, hospitals may be interested with proposed DCM model.

2. Origin of the Proposal

In a country like India, universities, public work departments and other organizations can't afford to buy expensive market based DCM Systems, thus limiting their capability in terms of document management which indirectly affect their efficiency and increase their operational costs. We propose to develop an online DCM System that is available to public/organization at very low cost. Furthermore the real world applicability of this system is tremendous in areas of education, public works, hospitals and any other field that documents (paper, fax, email, scan, and photo) records to be maintained and managed.

Even Government of India has been focusing on Digital India and has launched several initiatives towards the same. One of the example would be Digilocker initiative (<https://digilocker.gov.in/>) that has been used by many Organizations like CBSE, Ministry of Petroleum, Ministry of road transport etc. Digilocker is catered toward individual users whereas DCM system will cater the enterprise/organizations.

Furthermore Government of India has launched a public cloud service named MeghRaj an initiative by Ministry of Electronics and Information Technology. And has launched policy papers "GI Cloud Strategic Direction Paper" and "GI Cloud adaption and implementation roadmap" that showcase the ambition of Government of India towards Digital India. The Policy papers are available for download at <http://meity.gov.in/content/gi-cloud-meghraj>.

By Building and demonstrating low cost DCM, the project propose to assist organizations/enterprises to implement and adapt Digital philosophy and thus moving India towards Digital India.

3. Definition of the Problem

A good CM tool ensures the digital content's integrity, availability and security. There are various CM tools available like Sharepoint, FileNet, Documentum, OpenText, Alfresco, Liferay that allows the management of an organization's unstructured information, wherever that information exists. These tools can be deployed in traditional Data Center as well as on cloud. However most of these products are very expensive in-terms of-

- 1) licensing fees(based most of the time on user count)
- 2) Installation cost
- 3) Operational or maintenance Costs.
- 4) People effort/cost.

This is the reason why most of the public organizations, universities in India, Hospitals are still not purchasing such CM Products and instead rely on saving their organizations content via Paper trail or on computer drives, CD's, Google Drive, etc. This kind of storage mechanism doesn't ensure documents integrity, availability and security. Furthermore searching of document real time is very difficult and often one has to rely on user memory (where have they stored the document).

The Project attempts to solve this problem by building a Digital Content Management (DCM) System that is low-cost and easy to use out of the box.

4. Objectives

- 1) Built a low cost DCM- low on boarding and low operational cost.
- 2) Can be used by any organization/enterprise out of the box.
- 3) Hosted in public cloud like MeghRaj, Amazon Web Services (AWS) or Microsoft Azure.
- 4) Built using latest technologies like Lambda services, elastic search.
- 5) Encryption of document to ensure integrity of data.
- 6) Interfaces on mobile and web to allow users to push documents to the DCM tool and to retrieve it with accuracy at any point of time.

5. Existing System

There are various CM tools available like Sharepoint, FileNet, Documentum, OpenText, Alfresco, Liferay that allows the management of an organization's unstructured information, wherever that information exists. However most of these products are very expensive in-terms of licensing fees(based most of the time on user count), Installation cost, Operational or maintenance Costs, People effort/cost. Furthermore these products can't be used very easily out of the box. They have to be configured and customized to suit an enterprise/organization needs. This results in huge time delay.

6. PROPOSED SYTEM

The Proposed system will be a low cost on cloud DCM similar to what has been envisioned in Government of India Digital India Policy.

The Proposed system will be hosted in public cloud like AWS or Azure of MeghRaj(GOI initiative). The system will be built using Lambda services, elastic search and other new technologies available in market. The actual document will be stored using cloud storage devices like S3/Glacier.

Core Product functionality will include-

- 1) Taxonomy or Document model designer to define metadata (information about the content)
- 2) Search Designer so that users can define their own search criteria for searching the document.
- 3) An out of the box user interface on mobile and web to allow users to push documents to the DCM tool and to retrieve it with accuracy at any point of time.
- 4) API Interface to perform document operation remotely.
- 5) Interfaces on mobile and web to allow users to push documents to the DCM tool and to retrieve it with accuracy at any point of time.
- 6) Security model for authentication and authorization will be designed to ensure document security.
- 7) Document encryption to ensure integrity of data.
- 8) A viewer will also be included to allow document viewing capability.
- 9) Publishing capability to publish document to websites.
- 10) Administration user interface for management.

7. Importance of Proposed System

It is certain that a low cost easy to use digital content management can be a boon for Indian universities, public work departments and other organizations that can't afford to buy expensive market based DCM Systems The real world applicability of this system is tremendous in almost all areas like education, public works, hospitals and any other field that documents (paper, fax, email, scan, and photo) records to be maintained and managed. This will help the organizations achieve efficiency and transparency. And the decision making abilities will be improved with the help of fast access to documents in real time on need basis.

8. International Status

Lot of companies are moving towards digital content management. Companies like IBM, Microsoft, OpenText, Alfresco, Box are moving towards DCM area. These companies however have essentially commercial interest for moving towards digital content management.

1. IBM Content Foundation on Cloud enables businesses to leverage the power and scalability of IBM Content Foundation in a nimble, affordable, and consumable way.
2. Gartner: Box a leader in Content Collaboration Platforms for 2017 : Gartner valued 14 different vendors in the Content Collaboration Platforms market segment based on their completeness of vision and ability to execute on that vision.
3. The Association for Information and Image Management (AIIM) International, the worldwide association for enterprise-content management, defined the term in 2000. AIIM has refined the abbreviation *ECM* several times to reflect the expanding scope and importance of information management
4. **SharePoint** is a web-based, collaborative platform that integrates with Microsoft Office. Launched in 2001, SharePoint is primarily sold as a document management and storage system, but the product is highly configurable and usage varies substantially between organizations.

9. National Status

Even the Indian companies and government services are moving towards the digital content management techniques to maintain the confidentiality of enormous amount of data due to huge population of country.

1. **DigiLocker** is a "digital locker" service operated by the Government of India that enables Indian citizens to store certain official documents on the cloud. The service is aimed towards reducing the need to carry physical documents, and is part of the Narendra Modi-led government's Digital India initiative.
2. **GI Cloud** which has been named as 'MeghRaj'. This initiative is to implement various components including governance mechanism to ensure proliferation of Cloud in the government. The focus of this initiative is to accelerate delivery of e-services in the country while optimizing ICT spending of the Government. MeghRaj will ensure optimum utilization of the infrastructure and speed up the development and deployment of eGov applications.

10. Literature Survey

Many of the companies and government organizations are moving towards the services using digital content management.

1. **GI Cloud** which has been named as 'MeghRaj'. This initiative is to implement various components including governance mechanism to ensure proliferation of Cloud in the government.
2. **ECM(Enterprise Content Management)** System aims to combine the functionalities of the Web, content management and to systematically incorporate not only traditional publishing activities but e-mail, financial records, human resource, documents, etc., for an entire organization.
3. **IBM Content Foundation on Cloud** : IBM Content Foundation on Cloud enables businesses to leverage the power and scalability of IBM Content Foundation in a nimble, affordable, and consumable way.
4. **Alfresco Process Services (powered by Activiti)** is an enterprise Business Process Management (BPM) solution targeted at business people and developers. At its core is a high performance open-source business process engine based on Activiti with the flexibility and scalability to handle a wide variety of critical processes. Alfresco Process Services provides a powerful suite of end user tools and integrates with a range of enterprise systems, including Alfresco Content Services, Box and Google Drive.
5. **Gartner**: Box a leader in Content Collaboration Platforms for 2017 : Gartner valued 14 different vendors in the Content Collaboration Platforms market segment based on their completeness of vision and ability to execute on that vision.
6. **SharePoint** is a web-based, collaborative platform that integrates with Microsoft Office. Launched in 2001, SharePoint is primarily sold as a document management and storage system, but the product is highly configurable and usage varies substantially between organizations.
7. **DigiLocker** is a "digital locker" service operated by the Government of India that enables Indian citizens to store certain official documents on the cloud.
8. **GI Cloud** which has been named as 'MeghRaj'. This initiative is to implement various components including governance mechanism to ensure proliferation of Cloud in the government.

11. References

- [1]. "Research Paper on Content Management System" by Prof.Yogesh Vedpathak Assistant Professor University Of Pune, India,2014.
- [2]. International Journal of Advanced Research in Computer Science and Software Engineering – Enterprise Content Management On Cloud by Arulmozhi, P.G Scholar 2 nd yr CSE& Anna university, Volume 5, Issue 1, January 2015
- [3]. www.ibm.com/bs-en/marketplace/content-management
- [4]. www.alfresco.com/products/cloud/subscription
- [5]. www.box.com/home
- [6]. www.opentext.com/what-we-do/products/enterprise-content-management/ecm-in-the-cloud
- [7]. products.office.com/en-us/sharepoint/sharepoint-online-collaboration-software
- [8]. digilocker.gov.in/
- [9]. meity.gov.in/content/gi-cloud-meghraj