A GUIDE TO ENTERPRISE CONTENT MANAGEMENT



 \mathbf{EMC}^2_{\circ} EMC Proven Professional Knowledge Sharing 2011

Mouli Ponnupandy Senior Program Manager Enterprise Content Management NTT Data Americas mouli.ponnupandy@nttdata.com



Table of Contents

Abstract	3
What is Enterprise Content Management?	
Consume	4
Process	5
Deliver	6
Constituents of ECM	7
Document Capture	7
Document Management	7
Digital Asset Management	8
Web Content Management	8
Web Experience Management	8
Enterprise Collaboration	9
Records Management	9
Document Output Management	9
Business Process Management	10
Enterprise Search	10

List of Exhibits

Exhibit 1: What is Enterprise Content Management?	4
Exhibit 2: Content Sources	5
Exhibit 3: Processing Content	6
Exhibit 4: Content Delivery	7

Disclaimer: The views, processes or methodologies published in this article are those of the author. They do not necessarily reflect EMC Corporation's views, processes or methodologies.

Abstract

Art is beautiful only when it reaches its audience. Likewise, technology is useful only when it achieves intended business benefits. A customer making technology investments often relies on industry reports, competitive analysis, vendor capability, and product data sheets before arriving at an informed choice of product or technology. These data sources are full of buzz words and industry jargon which a customer needs to cut through to really understand what he wants to buy and if the investment choice is good.

This article intends to present Enterprise Content Management (ECM) from an end user perspective and aims to educate on constituents of content management. It offers insight into document capture, document management, digital asset management, web content management, web experience management, enterprise collaboration, records management, document output management, business process management, and enterprise search. It explores how features of different functional components overlap and interconnect. This helps customers identify which functional components are essential to meet the content management needs of their organization. As business and technology experts, we need to understand enterprise content management from a buyer's perspective to effectively service customer needs.

As a consultant, I have found the first question my customers need answered is, "what are data, metadata, and content?" Computers are designed to process elemental data. Humans are inefficient at processing data but adept in processing content. To make it simple, computers are *analytical*—breaking things into parts to process; humans are *contextual*—putting things together to understand the big picture. We both speak different languages. So for us to communicate effectively we need a common language, metadata, which in basic terms is data about data, or let's say, context to the content. Structured metadata is stored systematically in a database table which interacts with unstructured content to us in a usable format.

What is Enterprise Content Management?

Content Management aims to manage content through its lifecycle—from creation to delivery and disposition—by associating it with appropriate metadata. The primary purpose of an Enterprise Content Management System is to **Consume**, **Process**, and **Deliver** content in a usable form and to manage any integration with existing business applications.



Exhibit 1: What is Enterprise Content Management?

Consume – There are three specific ways in which a user can bring electronic content in to the Content Management System¹ (CMS). **Create** – The most direct and probably the easiest way is to create the electronic content from within the CMS². **Capture** – There are certain business processes which require handling of paper-based content objects. The paper-based documents are scanned and introduced into the CMS as electronic content. **Import** – A good amount of content files reside in various legacy and business applications as silo. The electronic content in these disparate systems are imported or migrated into the CMS.

¹ Content Management System – A collection of people, process, and programs that helps manage content.

² CMS – Content Management System



Exhibit 2: Sources of Content

Process – Content by itself does not add any value to end users. It needs to be processed in some form for it to be fit for use. **Classify** – The preliminary processing is to add some method to madness by way of classification. Content consumed is organized in the form that best suits business needs. **Collaborate** – People from across boundaries need to share business content to get the job done. Collaboration facilitates processing of content per desired business outcome. **Manage** – The system should manage both content and people who access content. Authentication takes care of data security and authorization takes care of people security.



Exhibit 3: Processing of Content

Deliver – The content is processed in a neutral environment so that multiple output formats and channels are supported by single source content. The idea is to have single source content. **Personalization** – The content is then delivered in intermediate format which allows end users to personalize presentation and delivery channel of choice based on conditional processing (of metadata). The idea is to reuse, recycle, and repurpose content as there are multiple users and delivery channels for enterprise content. **Compliance** – Content (most, not all) becomes a record when it reaches end of active lifecycle and needs to be preserved either due to internal policies or external regulations. An ideal compliance initiative should have provisions for both retention and archival. Retention aims at holding on to a record for a fixed period of time and then ensuring that it is purged from the system, effectively bringing record life to an end. Archival aims to move non-essential records to low cost storage devices either automatically based on business rules or manually driven by administrative needs.



Exhibit 4: Delivery of Content

Constituents of ECM

The definition for ECM has evolved over time to include various functional components that make up today's content management needs of an enterprise. What was considered not to be a part of ECM yesterday happens to be part of ECM today. The rationale to bring these standalone components into the fold of ECM is due to the increasing need for integration and interoperability between various functional components. Let us see what comes under preview of ECM today.

Document Capture – Content existed in paper form before we moved on to electronic content and will continue to exist for the foreseeable future. Capture (Scanning/Imaging) technologies enables us to bring in legacy content to a readily useable electronic form. They also perform pre-processing, validation, and sanity checks on the data being captured from the images before ingesting them into CMS. Once these images are inside CMS they are treated as just another electronic content subjecting them to all native CMS operations.

Document Management – Document management serves as the gateway through which most electronic content enter the CMS framework. Common office documents, spreadsheets, presentations, text files, drawings, forms, and XML files constitute the bulk of files handled by

document management. They offer services to create and manage: object model, security, version, virtual document, relationship, rendition, annotation, workflow, lifecycle, and audit trails. Document management is the basic building block around which other functional components of ECM are stacked. Though there is significant difference between ECM and document management, they are often used interchangeably as the former started its evolution from document management.

Digital Asset Management – Digital assets refer to non-textual media files such as images, photos, graphics, artwork, audio, and video content. Document management does not encompass digital asset management due to the unique nature of multimedia files such as very large file size, need for file format conversion, and digital rights management (DRM). Hence there is a need for a more sophisticated application which can handle rich media assets. Digital Asset Management (DAM) focuses more on discovery, reuse, repurpose, assembly, and distribution of digital assets in addition to offering traditional document management functionality.

Web Content Management – Web Content Management (WCM)³ deals with how to create, manage, and publish web content to public-facing websites, intranets, and extranets. Web content management systems evolved with the aim to segregate creative content authoring from technical publishing while still giving both technical and business users a similar set of tools and technologies to perform their task. In essence it is electronic publishing for dummies! WCM systems facilitate content creation through templates, automates publishing web content to a predefined navigation structure, ensures consistency if there are multilingual content, and allows for syndication.

Web Experience Management – Some time back it looked like web experience management was indeed web content management morphed into a new buzz word but perceptions are changing to reflect the fact that Web Experience Management (WEM)⁴ is a superset of WCM. WCM has been focused on the content creation whereas WEM is more towards content consumption. To keep it simple, WCM caters to Creators and WEM caters to Consumers. Some of the key differentiation features of WEM are content personalization based

³ WCM – Web Content Management

⁴ WEM – Web Experience Management

either on user identification or content navigation patterns, user feedback, managing user generated web content such as wikis, blogs, comments, community tags, opinion polls, and web analytics, and performance improvements for delivery of high volume rich media content.

Enterprise Collaboration – Collaboration existed even before we understood what was content. Collaboration is a creative process in which content is exchanged with one or many participants where each participant adds value to original content there by contributing to final content which is of more value than individual contribution. Collaboration should be free flowing, natural, and be able to reach users with relevant information whenever they need it and wherever they need it. Collaboration can either be content-centric, which largely happens to be based on a predefined repetitive business process, or people-centric, which is spontaneous and creative. Common forms of collaboration include but are not limited to emails, threaded discussions, calendars, online meetings, instant messages, and white boarding. In simple terms, collaboration aims to connect people with content.

Records Management – Any content has potential to become a record when it reaches the end of its active lifecycle and is retained so long as it has business value. The decision to declare content as a record is either due to self-imposed policies or as a result of legal obligations set by mandatory regulations. A record serves as the collective memory of an organization at a specific point in time. Records management systems should have the ability to manage both paper-based and electronic records. Records management provides facilities for creating a system-wide classification schema such as file plans, policies for naming, retention, security, containment, storage and disposition, ability to manage and administer physical records, reporting, auditing, and notification.

Document Output Management – What started in paper ends in paper! This is not an accurate definition of Document Output Management (DOM)⁵ though it is often widely understood and interpreted this way. It does much more than just print documents. DOM provides a set of tools to extract, compose, format, publish, transform, and deliver compound documents that are personalized for individual customers. DOM is categorized into three functional groups based on the nature of the output generated: structured – structured output is high volume, recurring, scheduled batch process which produces personalized documents

⁵ DOM – Document Output Management

mostly for existing relationships; interactive – interactive output is template-driven and often workflow-based to enhance or approve the preformatted content which is being sent to a customer; on-demand – as the name signifies, on-demand output is based on real time external input which could either be a customer action or as a result of a system-specific business rule.

Business Process Management – What is a business process? A specific user- or systemgenerated input triggers a finite set of value adding, cross functional, structured activities performed according to predefined business logic to generate a meaningful product or services for an end user. Business Process Management (BPM)⁶ and workflow management both aim to automate business process. BPM is a superset of workflow management which provides sophisticated functionalities to analyze, build, simulate, and monitor business processes that runs on a process engine which provides for integration to external applications whereas workflow management provides limited functionalities to design and execute workflows. BPM and workflow are not synonymous though the terminologies are often used interchangeably. Workflow focuses on task-oriented automation whereas BPM goes beyond automation to aim for process optimization and governance.

Enterprise Search – Content is classified and organized as per the business rules at the time of ingestion into CMS. The sheer volume of the content makes it difficult to rely only on classification as a means to find relevant content. An alternate way to find content is by way of searching CMS. Search can either be based on metadata that is stored in the database to describe the content files or a full text index created out of actual content files. Search is designed to seamlessly integrate into core CMS in which content and metadata for search is provided by CMS. The logistics required for creating an index, parsing search query, performing analysis on an index, categorizing relevant search results, and delivering content searched for based on the underlying permissions in CMS is taken care of by search function.

I hope this brings clarity on what essentially is Enterprise Content Management for a naïve user!

EMC believes the information in this publication is accurate as of its publication date. The information is subject to change without notice.

⁶ BPM – Business Process Management

THE INFORMATION IN THIS PUBLICATION IS PROVIDED "AS IS." EMC CORPORATION MAKES NO RESPRESENTATIONS OR WARRANTIES OF ANY KIND WITH RESPECT TO THE INFORMATION IN THIS PUBLICATION, AND SPECIFICALLY DISCLAIMS IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Use, copying, and distribution of any EMC software described in this publication requires an applicable software license.