Enterprise Content Management

COMP6205: Web Development

Dr A. Rezazadeh
December 17
Enterprise Content Management (ECM)

- ECM is a set of practices, processes, and methodology that make the technology **morph** into the most effective way to store, secure, and consume content.

  - **SharePoint** is a platform, a grab bag of features, and technology that can be moulded into an ECM solution.

  - ECM is also not moving what has been done in shared drives to a web-based modern platform.

  - ECM is the opposite of this; it requires up-front planning and practical application of information architecture.

  - It assumes a standardisation of **proven methods** for capturing, naming, and storing content.
Content Control

• A ECM system allows us to get control of our content by
  – Keeps track of content.
    • It “knows” where our content is,
    • what condition it’s in, who can access it,
    • and how it relates to other content.

• Furthermore, it seeks to prevent bad things from happening to our content.
Why use ECM?

• The assumption is that an ECM is a system where a knowledge worker’s effort to capture and categorise content is minimal but the amount of metadata capture is high, and the cost of finding and consuming the content is very low.

• These goals are often at odds with each other.
  – For example, expecting a user to enter content into SharePoint in the perfect way means that at time of capture they have to put in additional effort. Otherwise the findability of content suffers from long search times and getting, at best, duplicate content and, at worst, the wrong content.
Possible Scenarios for an ECM Adaptation

- Good ECM, bad adoption
- Ideal ECM
- Capture Effort
- Consumption Effort
- ECM failure
- Good adoption, bad ECM
Who does ECM target?

- ECM targets different people for a variety of reasons and use cases.
- ECM is a set of disciplines that are used to guide the development of Information Architecture and governance.
- A successful ECM platform supports the management of information for operational, transactional, and regulatory purposes.

<table>
<thead>
<tr>
<th>Users</th>
<th>Technical</th>
<th>Management</th>
<th>Governance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge Workers</td>
<td>IT Operations</td>
<td>Executive</td>
<td>Records</td>
</tr>
<tr>
<td></td>
<td>Developers</td>
<td>Departmental</td>
<td>Legal</td>
</tr>
<tr>
<td></td>
<td>Business Analyst</td>
<td>Project</td>
<td>Procurement</td>
</tr>
</tbody>
</table>
A Partial View of Enterprise Content in An Organization

Accounting Content
Human Resources Content
Engineering Content
Research and Development Content
Sales Content
IT Content
Legal Content
Publishing Content
Examples of different types of enterprise content

- A code of conduct policy manual
- An executive e-mail sent organization-wide announcing organizational change
- A product specification document
- A document with the meeting minutes from a project team’s status update meeting
- A user’s status update on the organisation’s blogging site
- Very diverse range of content, very different content characteristics, and content requirements
ECM stages

**Capture**
- Upload
- Office
- Scanning
- Native Documents
- Forms
- Streams

**Store**
- Information Architecture
- Versioning
- Formats
- Transformation

**Preserve**
- Compression
- Archive
- Remote Storage
A Phased Approach to ECM
Comparing ECM Costs and Value

The relation between cost and value in an ECM system over time
ECM life cycle

- **Capture**
  - File Upload
  - Microsoft Office
  - Native Documents
  - Electronic Forms
  - Document Scanning

- **Store**
  - Information Architecture
  - File Formats
  - Versioning
  - Transform
  - Compression

- **Process**
  - Workflow
  - Process Management
  - Business Intelligence
  - eDiscovery

- **Manage**
  - Access Levels
  - Change Control
  - Policy
  - Records Management

- **Deliver**
  - Editing & Viewing
  - Search
  - Collaboration
  - Publishing
  - Print

- **Preserve**
  - Reformat
  - Archive
Transitory Content Progressing to an Official Record
Content Control

• A ECMS provides core control functions, such as:

  – **Permissions**: Who can see this content? Who can change it? Who can delete it?

  – **State management and workflow**: Is it in draft stage? Is this content published? Has it been archived and removed from the public?

  – **Versioning**: How many times has this content changed? What did it look like three months ago? How does that version differ from the current version? Can I restore or republish an older version?
Content Control

- **Dependency management**: What content is being used by what other content? If I delete this content, how does that affect other content? What content is currently “orphaned” and unused?

- **Search and organisation**: How do I find a specific piece of content? How do I find all content that refers to X? How do I group and relate content so it’s easier to manage?

  Each of these aspects increases our level of control over our content and reduces risk. For example there is less chance that the shareholder report will be released early, or that the only copy of our procedures manual will be deleted accidentally.
ECM Improve Content Reuse

- Using content in more than one place and in more than one way increases its value.

- Some examples:
  - A news article appears on its own page, but also as a teaser on a category page and in multiple “Related Article” sidebars.
  - An author’s bio appears at the bottom of all articles written by that person.
  - A privacy statement appears at the bottom of every page on a website.

- The ability to reuse content is highly dependent on the structure of that content.
Content Automation and Aggregation

• We can allow users to consume content in other formats, such as PDF or other eBook formats.

• We can automatically create lists and navigation — more generally known as content aggregations.

• We can create multiple translations of content to ensure we deliver the language most appropriate to the current user.

• We can alter the content we publish in real time based on the specific behaviours and conditions exhibited by our visitors.

• A ECMS enables this by structuring, storing, examining, and providing query facilities around our content.
What a ECMS Doesn’t Do

• Create Content

• Create Marketing Plans
  – While some systems have marketing tools built into them, they still depend on human beings for direction.

• Effectively Format Content

• Provide Governance
  – “Governance” describes the access to and processes around your content: who has access to what, and what processes/steps they follow to make changes to it.
Feature Sets of An ECM System

The core feature sets of an ECM solution are as follows:

- Storage of documents
- Document viewing
- Document editing
- Document security
- Metadata model
- Versioning of documents
- Check-in/check-out of documents
The Content Lifecycle Model in SharePoint
Content Life Cycle - Capture

- Capture is the process of preparing, collecting, and indexing content before being stored in an ECM Platform.
- Capture into SharePoint can happen in the following six distinct ways, ordered by most common to least common:
  1. File upload
  2. Microsoft documents
  3. Document capture
  4. Natively created SharePoint documents
  5. Electronic forms
  6. Data streams
Content Life Cycle - Store

- Storage is not just the writing of a document’s content to a list or library.

- It also refers to all aspects of that document, including its security, its history, and its metadata.

- The following pieces of document storage are listed in the order that they should be implemented and used:

  1. Information Architecture
  2. Formats
  3. Versioning
  4. Transformation
Store - Information Architecture

• Ask yourself the following questions:
  – Does your company have a documented File Naming Convention policy?
  – Do the Shared Drives you work with follow a common Directory naming structure?
  – Can you easily navigate another department or workgroups Directory structure and find a file?
Information Architecture

• Q: How many ways can you organize unstructured information?

• A: Four — chronologically, alphabetically, numerically, and geographically.

• These four approaches are compounded by the many ways people use, format, and duplicate them.
Information Architecture in SharePoint

Repositories

Web Application(s)

Site Collection(s)

Site(s)

List & Libraries

Metadata Model

- Folders
- Document Sets
- Content types
- Managed Metadata

Visualization

- Views
- Columns

COMP6205 - ECM
SharePoint Building Blocks

• SharePoint 2016 includes many capabilities that can be combined to create enterprise business solutions.

• The core capabilities that these solutions are built on include the following:
  – **Sites** provide a structure for securing, storing, and organizing information and solutions.
  – **Lists** are containers for storing structured information.
  – Libraries are containers for storing and managing documents.
  – **Workflows** automate business processes.
  – **Records** management extends document management capabilities to manage documents through their full lifecycle.
SharePoint Building Blocks – Cont.

• **Alerts** can notify users when information has been added, changed, or removed.

• **Web Parts** and **Add-Ins** allow existing information to be organized and presented and provide additional business logic and functionality to be incorporated into the environment.

• **Search** provides the ability to locate information.

• **Personalization** and **social capabilities** allow users to create and manage their own information, communicate with others, and find others based on their skill and experience.

• **Data integration** capabilities allow information outside SharePoint to be incorporated into SharePoint solutions.
Content Life Cycle - Manage

- The Manage portion of Enterprise Content Management comprises all aspects around governance of the system.

- This includes informal and formal policies for users, the requirements for how content is captured, the requirements for how content is stored and secured.

- What is involved in records management, how and when content is deleted or consumed, and finally, but most ignored, how the system grows.

- Governance is defined as bringing together all the elements necessary to facilitate the long-term preservation, accessibility, and disposition of content.
Manage - Governance

• In conjunction with well-defined Information Architecture, a complete governance plan can help ensure that the user adoption is high and that the extensibility of the ECM solution is straightforward.

• Governance includes the following elements:
  – Records management
  – Security and access
  – Policies
  – Change control
Manage - Records Management

• Records management, like its broader parent, ECM, is a practice and methodology. However, records management requires a far more strict set of principles.

• A record is a **stamp in content, time, location, and metadata** for a document.

• When a document is declared as a record, its content will not change; the metadata, such as last modified data, will not change; and its logical and physical storage location will not change.

• Reassignment of security may also happen so that individuals who don’t have authority cannot access the records.
Content Life Cycle - Deliver

- The content delivery stage is the process of enhancing content with new information or consuming the content it already has.

- This includes editing of existing documents, changing of metadata, and sharing of the content with other users.

- The components of the Deliver stage are as follows:
  - Search
  - Editing and viewing
  - Publishing
Content Life Cycle - Process

- Content is not just viewed on an ad hoc basis, or edited, which is essentially another type of capture.
- Content consumption is incorporated into/conducted based on a decision-making process or line-of-business process.
- **Process** is taking what has been stored in content and incorporating it into another line-of-business activity.
- The elements of process are as follows:
  - Workflow
  - Business process management
  - Business intelligence
  - eDiscovery
Workflow

- This is the process of **routing a content item or transaction** through a series of **predefined steps** for approval between different layers of management.

  – This is very common in Human Resources, Finance, and Procurement.

  – The most common example of process is referred to as **approval workflows**.

- The purpose of workflows is to take metadata and content and turn them into action.

- A SharePoint workflow requires an existing business process to be well understood and defined.
Business Process Management (BPM)

• Business Process Management looks very similar to workflow, but it differs in that it allows for multidirectional processes, ability to version processes, and change control for processes.

• It is said that workflow is available out-of-the-box with SharePoint, but BPM usually comes via third-party solutions.

• The biggest technical difference is not just the workflows that can be created but also the management of those workflows.
  – The way we combine workflows to build BMPs.
Business Intelligence and BigData

• After content has been captured and processed, gaining *insights* from the content is a great way to take their value even further.

• Business intelligence (BI) is a broad category of technology that extracts greater value from content.

• BI really falls into the following three types:
  
  – **Dashboards and Key Performance Indicators** This is most often what people are referring to when they say BI. In SharePoint, this is to visualize large amounts of data in a graphical way.
  
  – **User-driven BI** This requires some data expertise and is usually performed using Excel and PowerPivot.
  
  – **Data mining** Data mining is the intelligent extraction of value, and in SharePoint, it relies on a third-party tool.
eDiscovery

- A very specific type of processing of content is called eDiscovery.
- Because eDiscovery is essentially search combined with records management.
- eDiscovery seems like some abstract term, but it really talks about an advanced form of search.
- eDiscovery is the identification, isolation, and locking of any content that pertains to a matter.
  - The most common instance of a “matter” is litigation. However, a matter could relate to the Freedom of Information Act, content audits, and so on.
Content Life Cycle - Preserve

- Content, just like everything else, has a shelf life. Most of the time, content that expires is deleted.

- This is beneficial to organisations from a compliance and legal perspective and as a part of ECM, but there is some content that is long lived.

- This type of content, when not consumed regularly, just takes up needless space in an ECM system.

- Content preservation is about taking that active content generated as part of ECM and moving it to a location in a format that can be accessed, although infrequently, in the foreseeable future.