

Enterprise Content Management with Microsoft SharePoint

*Overview of ECM Services and Features in Microsoft Office SharePoint Server
2007 and Windows SharePoint Services 3.0.*

A KnowledgeLake, Inc. White Paper

Published: March 2007



Introduction

The amount of content an organization must manage is growing at an incredible rate. The cost associated with managing content increases as the amount of content increases and regulatory compliance standards tighten. Corporations have ever increasing requirements to easily and quickly store, manage and retrieve enterprise content to solve business problems ranging from customer satisfaction to disaster recovery.

Microsoft made strides in the content management space with its Office 2003 suite. The Microsoft Office 2003 suite coupled with SharePoint Products and Technologies, Microsoft Content Management Server and Microsoft Exchange Server formed the general backbone of the Microsoft content management strategy. Even with all of these products working together, Microsoft had not stepped into the realm of true Enterprise Content Management. In general, these products introduced the Information Worker to Basic Content Services.

In February 2007, with the release of Microsoft Office SharePoint Server 2007 (MOSS), Windows SharePoint Services 3.0 (WSS 3.0) and the Microsoft Office 2007 suite, Microsoft officially entered the Enterprise Content Management (ECM) arena. Microsoft provides a robust, scalable ECM platform that rivals those of traditional ECM vendors. With the addition of mature SharePoint technology partners, Microsoft is a viable ECM contender today.

This paper reviews the capabilities of MOSS 2007 and WSS 3.0 for Enterprise Content Management, evaluating their ability to serve as an Enterprise Content Management solution today. The ability of Microsoft to meet the content management needs of customers relies on the solution requirements of each organization. Microsoft will continue to rely on its partners to enhance its ECM platform.

Enterprise Content Management Defined

Enterprise Content Management is defined by AIIM as *“the technologies used to capture, manage, store, preserve, and deliver content and documents related to organizational processes. ECM tools and strategies allow the management of an organization's unstructured information, wherever that information exists.”*

Enterprise Content Management (ECM) is the continuous business process of managing the complete life-cycle of unstructured or semi-structured content from capture or creation to archiving or destruction. Between these end points ECM includes the storage, use and reuse of content. The goal of ECM is to manage the ever-increasing amount of electronic content that continues to take on new forms. Content types range from simple text, emails, word processing and spreadsheet files to xml, html and even dynamic web page content. Images, voice mail and even streaming media are valid content requiring management.

ECM Components

Each year Gartner publishes a strategic white paper on the state of Enterprise Content Management market, titled “Magic Quadrant for Enterprise Content Management.” Gartner defines the following as components of an ECM suite:

- Document Management (DM)
- Records Management (RM)
- Web Content Management (WCM)
- Document-centric Collaboration
- Document Imaging (DI)
- Workflow (WF)

Additionally Gartner evaluates ECM suite vendors frequently providing additional components including:

- Integrated Document Archiving and Retrieval Systems (IDARs)
- Digital Asset Management (DAM)
- Electronic Forms
- Email Archiving and Management

These lists cannot be considered all inclusive and should only be viewed as suggested components. Conspicuously absent from the list is Search.

Short History of SharePoint and ECM

Microsoft entered the content management area in 2001 with Microsoft SharePoint Portal Server 2001 (SPS 2001). SPS 2001 provided basic content services to the information worker (IW). These services included basic document management with check-in/check-out, versioning, document profiles (metadata) and basic collaboration.

In 2003, Windows SharePoint Services 2.0, SharePoint Portal Server 2003 and Content Manager Server (CMS 2001/2002) were released. With these servers, Microsoft started providing document management, web content management and document-centric collaboration. Document imaging, workflow and records management components were provided by partners and vendors.

As the number of SharePoint licenses sold quickly grew to tens of millions, Microsoft moved towards making the SharePoint server products a backbone of its Office 2007 release and officially entered the Enterprise Content Management industry. With the release of Windows SharePoint Services 3.0 (WSS 3.0) and Microsoft Office Server 2007 (MOSS 2007), organizations now have a choice between Basic Content Services provided by WSS or Enterprise Content Management with MOSS. Many of the

components supplied by partners and vendors such as workflow and records management are included in WSS and MOSS.

Windows SharePoint Services 3.0 is a platform that provides core document management and collaboration features allowing organizations to provide basic document services such as document versioning, check-in/check-out and limited workflow. Applications including MOSS and Microsoft Project Server 2007 are built on top of WSS 3.0. Applications based on WSS 3.0 utilize and extend these core WSS 3.0 features.

Windows SharePoint Services is an additional feature that is included with Windows Server 2003. Organizations with a properly licensed Windows Server 2003 can download and run Windows SharePoint Services without additional cost. WSS 3.0 can provide basic content services to an organization such as:

- Document Management (DM)
- Document-centric Collaboration
- Basic Workflow
- Basic Search

Electronic forms are not included in WSS 3.0 but are available as a separate purchased component.

Organizations running MOSS 2007 can now benefit from enterprise content management features including:

- Document Management (DM)
- Records Management (RM)
- Web Content Management (WCM)
- Document-centric Collaboration
- Workflow (WF)
- Search
- Electronic Forms

Organizations can gain additional ECM functionality with document imaging and digital asset management from independent software vendors.

SharePoint and ECM

MOSS 2007 is built on top of WSS 3.0. With MOSS 2007, an organization will have all the features and functionalities of WSS 3.0 as well. (NOTE: "SharePoint" will refer to both WSS 3.0 and MOSS 2007 here forward. When applicable, "WSS 3.0" or "MOSS 2007" will be used to denote a specific product.)

Many of the features included in SharePoint including search, workflow and forms are used within the context of document management, records management, and web content management. This document will cover SharePoint and the three major areas of ECM: Document Management, Records Management and Web Content Management. Additionally this document will address Email Archiving and Management as it relates to ECM and SharePoint. Features including document-centric collaboration, workflow, search and forms are considered services used by the core ECM areas.

Document Management

Microsoft has included basic document management in one form or another since the very first version of SharePoint Portal Server. Document management is the management and organization of electronic documents which may include:

- Creation
- Storage
- Retrieval
- Editing
- Securing

SharePoint continues to improve upon the previous version's document management capabilities by supporting larger document libraries, improving search tools, and including tighter check-in/ check-out support.

Creation

Creating content involves both creating new content as well as capturing content from external sources such as paper documents.

SharePoint provides tight integration with Microsoft Office 2007 Suite applications for content creation and editing. Documents can be created from applications including Microsoft Word, Excel and PowerPoint and entered directly into document libraries within SharePoint. Tight integration allows Information Workers to create new content and easily access features such as versioning, metadata, and workflow information (see the Storage and Processing sections for further explanation).

The close integration between the Office clients and SharePoint is not limited to the Office clients only. Other applications, including in-house systems and partner/vendor solutions, can create the same close integration using the various application programming interfaces and web services.

Storage

Once a document has been created, it will need to be stored. SharePoint stores documents in special lists called document libraries. Document libraries provide the user interface elements such as views and security levels. Users can have various access permissions including none, view, edit and manage. New in SharePoint, document libraries give administrators the ability to apply permissions to folders and individual items.

Content types are new to SharePoint and provide a typing mechanism to a piece of content, and control various aspects associated with storage such as the required and optional metadata (columns or fields), workflows, and associated document library forms. List and library managers can select to manage a list of available content types and allow for multiple content types in a single document library.

Storing documents in SharePoint includes document versioning. In SharePoint, document libraries support major and minor versions. List managers can configure a document library to support versioning and the number of versions to retain.

Storage of documents in a document library may occur using the tightly integrated Microsoft Office applications that can create content and save it directly to a document library. Content can also be stored in document libraries using the document library upload pages, selecting New from the document library's menu bar or programmatically using the list-based objects or list web services.

Also new for 2007 is document translation. Documents destined for a document library can be processed prior to addition and translated into different formats. For example, a Microsoft Word 2007 docx file may be converted to PDF or XPS for storage. All documents can support linked copies of the document. These linked copies can be located in other document libraries but retain a link to the original document. Changes to the original document result in notification to users of the linked document.

The storage system in SharePoint has for the first time a user-initiated restoration of deleted documents. Users and administrators now have access to a two-phase Recycle Bin. The two phased Recycle Bin allows users to recover documents for a certain administrator-defined time period. After the time period has expired, administrators can still recover the documents for another defined period before recovery requires more expensive and time consuming backup restoration.

Retrieval

Documents stored in SharePoint document libraries will need to be found and retrieved for further editing, viewing or archiving. The search features in SharePoint are the primary mechanisms for retrieving documents. Browsing document libraries using views, filters and sorts can be considered another valid method, but this method becomes difficult with large numbers of documents in a library.

SharePoint provides scalable index and search services that allow property and full-text indexing and searching. WSS 3.0 search is a limited search compared to MOSS 2007 search. WSS exposes the search from each site as well as from an object model. MOSS 2007 exposes multiple search entry points including the Search Center which includes a collection of search related components, such as Content by Query Web Parts. MOSS 2007 search can be accessed using the object model as well as a web service. MOSS 2007 search features include:

- Full-text searching
- Basic property-based searching
- Search corpus scoping
- Term hit-highlighting

- Relevance ranking
- Suggested terms

Documents in SharePoint document libraries can be retrieved by accessing the library RSS feed as well as subscribing to alerts for changes in documents. Alerts can be sent to a user via email or viewed within the SharePoint user interface

Retrieval of the list results using search, RSS or alerts are provided in a format with associated metadata allowing the viewer to preview some document metadata without requiring the user to navigate to the library. Clicking the document's link generally will result in the document opening for editing or viewing depending on the document's type and user permissions.

Editing

Documents that have been created and stored may go through further editing. SharePoint provides more functionality to support the editing of documents.

Document management requires processes. Certain types of documents will need to have one or more processes applied to them during the editing phase. These processes may be simple content addition or verification, or may involve more complex processes such as workflow or automated steps. SharePoint workflows, which are an implementation of Windows Workflow Foundation, can be used for document-centric processing. WSS 3.0 does not install any default workflows and does not include the necessary object model and web services for Office client integration. MOSS 2007 does include basic workflows and the required object model and web services for interaction with Office clients such as Word, Excel and PowerPoint. To use workflows in SharePoint, developers or designers are required to create and associate workflows to content types or libraries.

MOSS 2007 supplies the following workflows during installation:

- Approval
- Collect Feedback
- Collect Signatures
- Disposition Approval

Workflows can provide automated routing and processing during the document's lifetime from approval routing to systems integration features. SharePoint is not limited to the few installed workflows listed above. Workflows can be developed in Visual Studio.Net 2005 as well as Microsoft SharePoint Designer 2007. The newly created workflows are integrated or added into the SharePoint environment.

Workflows are document or item-centric in SharePoint, meaning that workflows act with or upon documents or items and not arbitrary events. Workflows are generally associated with one or more lists or content types that are contained in a list. These workflows may be triggered by document additions or changes, but they can also be started manually. A workflow task list maintains the history and tasks for review. MOSS 2007, Microsoft Word 2007, Excel 2007 and PowerPoint 2007 natively display workflow information.

Electronic forms allow the users to gather document data as well as other supporting data. MOSS 2007 Enterprise version includes Microsoft Office Forms Server 2007 allowing users to gather data using a browser. Microsoft Office Forms Server 2007 can be purchased separately and used with WSS 3.0 or MOSS Standard Edition.

A document may have several editing steps involving many editors and changes. Document libraries provide key functionality including check-in/check-out and document versioning to help manage the changes to documents and avoid the loss of valuable content and editing time.

Document libraries support check-in/check-out functionality and may be configured to require check out before editing. A check out is an effective lock on the document preventing another editor from changing the content and avoiding a potential loss of content from concurrent editing. Checking in the document after editing will update the document for viewing and allow the editor to make metadata changes.

Versioning requires the system to maintain distinct copies of the documents during the editing process. SharePoint document libraries support both major and minor document versions but do not support a custom versioning scheme. Versioning can be configured to only allow a defined number of versions and pruning of versions.

Workflow, check-in/check-out and versioning are available in Microsoft Word 2007, Excel 2007 and PowerPoint 2007. Custom applications can achieve the same exposure using the MOSS 2007 and WSS 3.0 object models and web services allowing custom application integration and administration.

Securing

Vast improvements and enhancements have been made to the security of SharePoint with MOSS 2007. MOSS 2007 now supports “securable” objects, which is an object that can inherit security permissions from the parent object or define its own separate security.

Documents, folders and lists (including document libraries) are securable objects in MOSS 2007, allowing for more granular permissions required of ECM not seen in previous versions of SharePoint products. Security for these objects is generally set using the user interface but can also be set using the classes and objects in the API.

Pluggable authentication is another security feature that is new to SharePoint. While not as compelling as item level permissions, pluggable authentication allows the designers and administrators to pick from a varied set of supplied authentication methods including Basic, Windows Authentication, Passport and Forms. Developers can create custom authentication providers as needed.

Information Rights Management (IRM) is now integrated in MOSS 2007 document libraries. IRM allows an organization to control the permissions granted to a document. IRM is policy-based so that administrators can assign a policy to a document library to control how documents may be used. Examples of limiting or controlling a document using IRM may include disallowing printing or requiring the document to be downloaded after a certain period of time.

Records Management

At a certain point and time in the lifecycle of content, the content may need to become a record. A record is simply a piece of content that needs to be carefully controlled for legal or regulatory requirements. Records management is the setting of policies and standards, the assigning of responsibilities and authorities and the establishing of procedures and guidelines for records.

Records Center Site Template

MOSS 2007 provides many features and services to support the management of records. The Records Center is a site template this is central to records management and designed to collect, organize and control records within the organization. Using a site template dedicated to records management enables record and compliance managers to draw the line between content and records in order to easily control and administer required policies. Records management can also be applied to a site that is not based on the Records Center site template.

Information Management Policies

MOSS 2007 supports Information Management Policies, which are actions to be performed on a record and can be assigned on a content type, list or library. Information management policies limit the amount of required user interaction and assistance in achieving a higher level or compliance with its automatic application. Available policies in MOSS 2007 include:

- Audit

Setting an auditing policy to a list, library or content type allows record managers to collect the correct and required auditing events including events such as check-in/check-out, download, view item and view properties.

- Expire

Retention is controlled by Expiration policy, which allows records managers to define the retention period based on item metadata (i.e. created + 10 days) or programmatically. Records managers can establish the action that is to occur based on the expiration of an item, choosing from various supplied actions such as Delete or starting a known workflow.

- Label\Barcode

Labels and barcodes can be applied to items and can be based on known metadata. Barcoding is based on a provider model allowing developers to create custom barcodes as required. Barcodes can assist in associating physical entities to electronic records (i.e., paper document to the corresponding document imaging). When using Microsoft Office 2007 applications, the barcode and label can be injected into the document.

- User Communication Statements

User communication statements are notifications or reminders to the viewer. Statements might include “Do not disseminate” or other similar reminders.

MOSS 2007 provides extensibility mechanisms for developers to create custom policies to fit specific records management requirements.

Hold Orders

With Information Management Policies like Expiration, there are legal and regulatory compliance issues that require the hold of a document from the expiring process. MOSS 2007 supports the concept of holds which interrupts the expiration process.

Holds are defined in the Records Center. A hold itself does not affect any single record. Managers of the hold must select the items to be held. MOSS 2007 provides a search page with the ability to add search results to a hold. Records manager can use a MOSS 2007 search to return results and then select the documents to be placed on hold. Records managers can also view a hold report that is regenerated every 24 hours.

Document Routing

MOSS 2007 strives to limit the user interaction required to work with records. The simpler the interaction, the more likely users will comply with the organizations record managing policies. To achieve simple user interaction, MOSS provides a custom list or library action to send an item to the records center. Other applications can benefit from a web services or object model integration.

Once submitted to the records center, the document routing table determines the appropriate document library for submitting the record. This relieves the user from the responsibility of determining the appropriate library within the records center. Records are routed within the records center based on items in the Record Routing list. The content type of incoming records are compared to the Record Routing list items and then placed in the correct library. Metadata and audit history accompanies the document when placed in the records center.

Document Conversion

Documents are created and edited in formats that allow the editors to easily work with the content, such as Microsoft Word for text documents and Microsoft Excel for calculations. Records are generally static and do not require editing. Most records benefit from a format that cannot be edited easily and can be read and viewed by a generic audience that may not have the correct editing tools for the document format.

MOSS provides a document conversion service to convert documents from one format to another. MOSS 2007 includes the following conversions:

- InfoPath to Web Page
- Word Document to Web Page
- Word Document with Macros to Web Page
- XML to Web Page

The document conversion is a service supported by MOSS 2007 and developers can create and package custom document converters using the MOSS 2007 object model.

Email Retention

Email Archiving and Management is the responsibility of Microsoft Exchange 2007. The 2007 version supports records management using managed folders. Managed folders can be configured to send emails, metadata and attachments to the MOSS Records Center. When sent to the Records Center, the email is routed via the same records routing table as documents and can have Information Management Policies associated to the content.

Microsoft Outlook 2007 does provide assistance in Email Archiving and Management. Once managed folders are setup in Exchange 2007, which is out of the scope of this document. Email can be automatically or manually sent to the Records Center.

Web Content Management

The magnitude and complexity of web content is increasing in an organization. Web content management eases the creation, editing and publishing of web based content. Web Content Management prior to MOSS 2007 existed in the realm of Microsoft Content Management Services server. Web content management features and functionality were migrated to MOSS 2007.

MOSS 2007 provides portal templates for publishing sites. A publishing site is the top site plus any supporting sub sites. Designers and developers can modify or create new publishing sites. Each publishing site includes one or more content pages. Content pages are template pages that can be selected by an editor. Each page template can contain one or more field controls which are locations where content editors can add page specific content.

MOSS 2007 provides one single portal-level publishing template to create a publishing site. This template contains basic content pages. Designers and developers can create new templates, content pages and custom field controls. Publishing sites support master pages for consistent content branding. Developers can create custom navigation or extend some of the existing navigation controls providing a custom navigation experience.

Office Workflow provides the backbone of workflow, similar to how document management pages can be placed into a publishing or approval workflow allowing content.

Partners Fill Gaps in MOSS 2007

Gartner notes in its September 2006 white paper titled "Q&A: Microsoft's Content Management Software and Strategy" that Microsoft will continue to rely on partners to fill the gaps in its Enterprise Content Management platform. The key missing components include document imaging and digital asset management. Gartner further notes that Microsoft is "*beginning* to court developers and

independent software vendors to build on top of the new SharePoint infrastructure... and [that] developing Microsoft's partner channel to deliver certified, repeatable content management applications built on SharePoint products and technologies will be a challenge for Microsoft *for the next few years.*" Gartner fails to recognize that several independent software vendors have built scalable and extensible applications since the release of Office 2003. Through Microsoft development programs, such as the Technology Adoption Program for Office 2007, partners have had the opportunity to build and test their solutions for the newly released Office SharePoint Server 2007 platform as well. Microsoft ISV's will play a significant role in making Microsoft a viable ECM contender today. For more information on Microsoft partner solutions for MOSS 2007 and WSS 3.0, visit the Microsoft Office System Solutions Directory (<http://directory.partners.extranet.microsoft.com>) and search for partners with Information Worker Capabilities in "Enterprise Content Management and Forms."

Summary

All organizations from small businesses to large enterprises have some form of content that needs to be managed. While the benefits of Enterprise Content Management have long pushed large enterprises to make significant investments in ECM technology, the shift in the industry toward platform-based applications is having a significant impact on the traditional players in the industry.

SharePoint support for content management has increased dramatically from the original basic document management features found in Microsoft SharePoint Portal Server 2001. From basic document services to rich document management, records management and web content management with core services including collaboration, electronic forms and enterprise search, organizations can leverage MOSS 2007 to provide core portal services along with ECM.

MOSS 2007 is a horizontal ECM platform with a rich set of extensibility and integration points. Microsoft partners and vendors are quickly moving to leverage the MOSS core ECM services and components to integrate complementary ECM capabilities such as document imaging and digital asset management. While MOSS 2007 is a horizontal solution, the extensibility of the server provides the base ECM services and components for partners and vendors to create rich vertical solutions targeted to specific industries.

About the Author

Darrin Bishop is a product evangelist for KnowledgeLake, Inc. and has been recognized as a Windows SharePoint Services MVP for 2007. He is the author of *The Rational Guide to Building SharePoint Web Parts* published by Rational Press. Darrin has authored many magazine articles and blogs about SharePoint, ECM and other technologies. He has been working with SharePoint Technologies since the release of SharePoint Portal Server 2001 and has implemented global solutions using SharePoint Portal Server. He is a member of the Microsoft Developer Advisory Council for Office "14".

This White Paper is for informational purposes only. KNOWLEDGELAKE MAKES NO WARRANTIES, EXPRESS OR IMPLIED, AS TO THE INFORMATION IN THIS DOCUMENT.

© 2007 KnowledgeLake, Inc.. All rights reserved

The names of actual companies and products mentioned herein may be the trademarks of their respective owners.