Microsoft and Open Text: A New Age in Managing Unstructured Content

From Desktop Creation to Compliant Retention

A Microsoft and Open Text White Paper

Published: July 2006
## Contents

**Solving the Challenges of Enterprise Content Management**  
- The Microsoft and Open Text Alliance  
- Making Content Management Part of the User Experience  
- ECM Business Applications  
- ECM Repositories  

**Building an ECM Solution**  
- Architecture Built Using SharePoint Products and Technologies  
- Extending the ECM Baseline with Open Text  

**Extending the Use of ECM**  
- Metadata Management  
- Digital Rights Management  
- Federated Search  
- User Authorization and Authentication  

**Build for Today and Tomorrow**
Solving the Challenges of Enterprise Content Management

Chief Information Officers, Chief Compliance Officers, and other executives know that the effective use of information and greater ease in obtaining that content can drive revenue growth and profitability in their organizations. The global marketplace is being largely shaped and accelerated by the growing role of information and knowledge, and many organizations have invested in Enterprise Content Management (ECM) systems. ECM is unique from other systems because it manages the proliferation of unstructured content, like documents, Web pages, e-mail, spreadsheets, images, and structured data from enterprise systems, such as ERP, SCM, and CRM.

Many organizations that introduced ECM solutions for document management, Web content management, or collaboration have found that these systems can adequately manage the information contained within them. What remains a challenge, however, is how to access content that is relevant to business processes in disconnected e-mail servers, files servers, and stand-alone collaboration systems. Vendors of content management systems have been addressing this problem over the last 15 years with specialized tools and applications that either integrate content management functionality into existing applications or provide rich content management capabilities through additional clients. While those applications have now reached a high level of sophistication, the issue of an impractical user experience has remained.

Information workers need to have more business content available at their fingertips, quickly and easily, from within the desktop environment, which in most cases is Microsoft® Office. Meanwhile, the corporation seeks assurance that issues of compliance, control, and security can be addressed, especially as it relates to unstructured content. All levels of management must be certain that critical content is stored securely and for the long term, to support the transparency and integrity of corporate governance ethics and the risk management efforts of modern organizations. Additionally, all of this must be delivered in a cost-effective manner. The solution to all three scenarios lies in the forging of a new strategic alliance between the desktop powerhouse, Microsoft, and Open Text, the market leader in ECM.
The Microsoft and Open Text Alliance

Microsoft and Open Text are partnering to integrate all systems and tools involved in a compliant ECM system landscape. The goal of the alliance, formed in November 2005, is to provide ECM solutions that extend business processes from the desktop to back-end systems required to manage critical information in a compliant and long-term fashion.

The following diagram illustrates how key elements of a successful ECM deployment are addressed through the companies’ joint offering.

![Figure 1. The Microsoft and Open Text ECM Architecture](image_url)
Making Content Management Part of the User Experience

While ECM vendors like Open Text have always provided tools to integrate into applications, the capture of critical descriptive information about content (metadata) has rarely been utilized in the content authoring process. With the release of the 2007 Microsoft Office system, Microsoft is now extending this critical functionality into the most popular Microsoft Office applications.

Likewise, Open Text is extending the reach of enterprise content that is relevant to information workers through a series of vertical solutions and horizontal integrations that abstract where information is collected. One example of such integration is Open Text Enterprise Connect. This solution addresses the investment that organizations have made in enterprise applications such as ERP, CRM, and SCM by allowing content to surface from those applications back into the Microsoft Office system. Relevant information does not have to be gathered by opening the source applications; it can be used right in the familiar Microsoft Office desktop environment, which can result in improved efficiency and reduced costs.
Figure 3. Enterprise Connect enables relevant information to be accessed from commonly used desktop applications, regardless of where it is stored.

Dear Sir,

I am writing to you concerning the invoice Ref No: 90034/09 which we received from your organisation recently.

The invoice states that we have ordered 2 flat screen 17" monitors, and that as these have now been delivered you are requesting payment.

When cross checking this internally, we found out that we have originally ordered 3 flat screen monitors according to Order No. 8887.

Please could you confirm:

- Three monitors were originally ordered.
- That the delivery of the third monitor is still to take place.

Yours faithfully,

John Marquart

Financial Controller
COMPU Tech AG
ECM Business Applications

Business applications should be the core consideration of an ECM strategy, because these applications generate and drive the structure and related processes associated to the content. Likewise, they typically drive the adoption of ECM. These applications range from basic collaborative SharePoint sites to horizontally-oriented contract management applications or vertically-oriented regulated document applications for the Pharmaceutical and Life Sciences industry. Figure 4 illustrates an ECM stack covering all aspects, from the user interface to the content repositories. In this stack, Open Text extends the SharePoint Products and Technologies functionality to connect content across multiple leading systems such as ERP, e-mail, or digital asset management (DAM) systems, through the Open Text Functional Extensions. Together, Microsoft and Open Text can provide a platform to meet the needs of organizations.

Building on this functional foundation, content-centric business applications help organizations manage unstructured content, leading users through pre-defined processes and enforcing the capture of appropriate metadata as part the business logic of the application. Because these applications build on the Microsoft and Open Text framework, they can enable organizations to automatically store and classify content according to its long-term relevance, which can increase efficiency.
Figure 4. A Microsoft and Open Text ECM architecture

The integration of long-term storage and records management capabilities residing in the repository layer can enable compliance with strict content retention and e-discovery requirements.

The functional extensions enable Microsoft and Open Text to broaden the reach of the joint architecture across critical applications in the enterprise. All relevant content can be managed, including scanned paper content and documents in ERP systems.
Figure 5. Users can work with documents directly from SAP R/3

ECM Repositories

To provide a long-term structure and taxonomy while allowing flexibility in the way content-centric applications are architected, the joint ECM stack differentiates between repositories that contain long-term classified enterprise content and application specific repositories that contain transactional content. Transactional content is more than the structured database records containing the file and the associated metadata. The transactional repositories also store the business and user interface logic of the collaborative and business applications.

As the compliance considerations only relate to the content these applications generate and not the structure of the applications themselves, the content can subsequently be taken over by the long-term Information Lifecycle Management (ILM) repository, preventing the transactional repository from being compromised by long-term retention considerations. This provides the flexibility to change or retire the applications long before the retention schedule of the content expires, giving more flexibility when designing and maintaining business applications.

Many laws and regulations mandate that related content must be stored for long periods of time: between three and 15 years, or longer for content that is subject to specific industry regulations. The Microsoft and Open Text ECM architecture can manage the retention of structured and unstructured business content, provide...
litigation support, and allow retention hold capabilities through an enterprise-wide file plan and retention structure. The file plan and associated retention mechanisms are part of the ILM repository, which serves as the long-term repository behind the transactional repositories.

The ILM repository serves as the long-term metadata and records management repository. Storage services manage the longevity of the files through integrations into the appropriate storage hardware and the capacity to migrate content across different hardware over time. The content will inevitably outlive the lifetime of the storage media.

In contrast to non-integrated ECM architectures, Microsoft and Open Text offer a unique way of exposing records management and long-term archiving capability into the transactional repositories. Content identified by a business process or user as critical enterprise content worth retaining over long periods of time can be transparently transferred into the ILM repository without creating a duplicate of the content. This concept is generally referred to as “in-place records management” and meets both compliance as well as usability requirements.

The combination of the centralized records management functionality and the file archiving functionality in the ILM repository is referred to as the Enterprise Library in this white paper.

Building an ECM Solution

The ideal architecture of an ECM solution is highly dependent on:

- The scope of the solution
- The horizontal and vertical business problems
- The compliance requirements
- The size of the implementation

Architecture Built Using SharePoint Products and Technologies

Microsoft’s ECM design philosophy for Microsoft Office SharePoint® Products and Technologies 2007 has been to create an integrated ECM platform based on a common framework of components and technologies designed for scalability, interoperability, and stability. A unified architecture provides a common set of services, like Web part framework, integrated search, integrated user management and user rights, workflow, security model, and collaboration, along with a unified repository infrastructure. This ensures integration and consistency across the various ECM components like document management, records management, Web content management, and form management. It also supports individuals, teams, and business units, through intranet, extranet, and internet sites, without requiring the use of disparate systems for these different scenarios.
Thanks to this common architecture, customers can reuse applications, code, and site content, enabling a common development and deployment experience for developers and IT professionals respectively. It can also enable faster deployments, lower training costs, and better use of IT maintenance resources.

**Extending the ECM Baseline with Open Text**

Given the robust foundation of capabilities provided by Microsoft Office SharePoint Products and Technologies, organizations can increasingly begin to solve ECM-driven challenges with a desktop and enterprise environment characterized by the 2007 Microsoft Office system.

While the ECM baseline provided by Microsoft Office SharePoint Server 2007 serves many scenarios well, more complex and large scale implementations can derive a great benefit from an extended ECM infrastructure. In particular, large and complex infrastructures often need to store a holistic collection of enterprise content over a long period of time or according to other very strict requirements, incurring regulatory mandates, or legal obligations. In such deployments, the SharePoint repository can act as a transactional ECM repository and strategically leverage the Livelink lifecycle, metadata, and storage services provided by the Enterprise Library. The Enterprise Library manages and stores large volumes of content securely and efficiently, allowing the content to live outside of the collaborative environment without compromising functional aspects of the ECM application.

By combining the strengths of the repositories in this fashion, this tiered architecture offers the following benefits:

**Improved cost to scale** As an enterprise needs to scale long-term storage across a long period of time for a large amount of content across a variety of storage media, the cost to scale can become a hurdle for business applications and their repositories. Many business scenarios involve the secure management of a significant volume of documents. Consider a heavily litigated organization that uses the e-mail journaling functionality to capture a copy of every e-mail sent and received...
within the organization and retain it for multi-year periods. Alternatively, consider
heavily regulated industries, such as pharmaceutical/life sciences, that are required
to maintain content for very long periods of time. The number of documents to be
retained can quickly reach billions of files and terabytes of data, a costly, albeit
necessary, activity. In such cases, the cost and scalability of Livelink storage services
are an ideal complement to SharePoint Products and Technologies services, which
can be optimized for collaborative use scenarios.

**In-place records management** The SharePoint repository is exceptionally well-
designed to manage content, providing an interface that allows users to collaborate
and iteratively access, review, and revise unstructured content. The release of Office
SharePoint Server 2007 adds records management capabilities to successfully
manage documents that have reached a conclusive stage in their lifecycles and can
be classified as a record, creating a copy of the content as a record and providing the
retention and disposition capabilities necessary to manage it. In many cases,
organizations must manage content as a record beyond the ability to automate its
retention and disposition lifecycle, also providing the ability to secure the content from
deletion while the content continues to be referenced in business applications. With
Livelink lifecycle services, organizations can classify documents in SharePoint
Products and Technologies, Microsoft Exchange, or other content repositories and
enact all relevant records management functionality on the documents without
affecting the way that users work with them. In these cases, SharePoint site content
can be moved and physically stored in the Enterprise Library, managed by the
Livelink services. Although content lives within the Enterprise Library, it remains fully
accessible through the SharePoint site interface, and users can continue to work with
their content from SharePoint Products and Technologies as required.

**Compliant, long-term storage** When large volumes of content carry significant
risk, organizations should consider a centralized and long-term archiving strategy.
Capabilities such as single-instance archiving, automatic compression and ISO
image storage can enable more efficient content storage. Not only is content stored
more efficiently from a cost perspective, but it is done so in accordance with
compliance standards* that require to store certain documents in a write-once-read-many technology* that
require specific storage media or methodologies. The ability to lock files through
WORM (write-once-read-many) or WORM-enabled media helps enable compliance
with regulations that mandate additional storage security. For example, SEC-relevant
data can be automatically stored on WORM content storage for the required
mandated duration driven by the classification selected in the Enterprise Library.

**Comprehensive information management strategy** Much of an enterprise’s
mission-critical data is created and managed within other repositories such as ERP,
CRM, e-mail, DAM, file or legacy systems. Organizations need a comprehensive
approach to managing, retaining, retrieving, and destroying this content. Consider the
scope of a legal discovery order: inevitably this mandate will span various types of
content from many sources created across many applications. In the event of a
discovery order, it is imperative to secure any potentially responsive content from
deletion, and the only reliable and cost-effective way to do that is to have centralized
control over the information in the same way that it is provided for SharePoint site
content. An extension of an information management strategy is the provision of
access to this content when and where it is needed. As SharePoint Products and
Technologies increasingly becomes the interface for users working with corporate
content, it becomes important to be able to access all relevant information from this system. By consolidating these enterprise informational assets into a centralized information management framework, Open Text also enables users to access all relevant content directly from a SharePoint site, regardless of source.

**Comprehensive storage strategy** As most enterprises work with content originating from a variety of systems, they currently rely on a distributed and mixed storage strategy that includes numerous types of storage media, such as WORM media, enterprise disk, or capacity tape from different storage vendors, such as EMC, HDS, HP, IBM SUN, Netapp, and others. A long-term storage strategy requires the ability to virtualize storage, or allow content to be directed and migrated between storage media without impacting the way that users access that content. For example, consider how quickly hardware can become outdated, especially in cases where information must be stored for several decades.

### Extending the Use of ECM

Since announcing Open Text’s strategic partnership with Microsoft in November 2005, both companies have worked closely to plan future products, integrations, and solutions. Current and ongoing integrations focus on the following key areas.

#### Metadata Management

A limitation of file system-based document storage approaches is the technically-oriented metadata set associated with content objects, which constrains their use with records management policies and practices. Consider that the typical metadata associated with a document on a file system is restricted to the created date, modified date, size, and user, but does not contain any meaningful description of the content itself.

The 2007 Microsoft Office system enables records management-specific metadata to be embedded directly into documents and surfaced through a document template, making this information accessible and usable within applications like Microsoft Office Word. The metadata stored within the files can enable users to associate default document classifications driven by the user when creating the document, rather than being applied as the document is being checked into the repository.

#### Digital Rights Management

One of the fundamental challenges of records management for regulated or litigated organizations is to ensure that all instances of the document are destroyed in the disposition process of records and that those instances remain in the enterprise. Users often need to save and store files locally in order to do their jobs but in doing so, they remove these files from corporate control and create offline copies. Unintentional e-mail forwarding and laptop theft also pose great challenges. If a laptop is stolen or lost, stored messages can be accessed, passwords can be breached, and every bit of data in a local message store can be made publicly available.
Beyond the fundamental capabilities provided by rights management, like limiting whether a document can be printed or preventing an e-mail from being forwarded to other users, Microsoft Information Rights Management can also provide key functionality that supports records management requirements.

Information Rights Management can be facilitated by locking down documents completely and storing a “key” that unlocks the document on a centralized server. In addition to unlocking the document, the key also indicates what actions can be performed on the document while the RMS server records all instances where the document was accessed. Open Text and Microsoft tightly integrate rights management and records management to gain ultimate control over content and help drive the centralized destruction of all copies of a record by destroying the access key to it. Also, documents can be set to expire so that they can’t be used after being out of the repository for a determined amount of days. This, in effect, disables all local copies of documents, regardless of where they reside.

**Federated Search**

As a key benefit of a consolidated information management strategy, Microsoft and Open Text collaborate to provide federated search across the entire content platform, spanning all content through one consistent user interface. This again highlights a key to the partnership, hiding the complexities of ECM in the applications used every day. Federated search is a key aspect of a cost-effective and timely response to business requirements that demand immediate access to a distributed range of content. Consider again a regulatory audit or discovery order: responsive content may exist in numerous, traditionally disconnected systems. The ability to retrieve relevant information with a single search, no matter where it resides, can significantly streamline and simplify the process.

**User Authorization and Authentication**

User authorization and authentication is seamless throughout the entire platform, with single sign-on access to all appropriate content. Open Text and Microsoft take great care to enable a more seamless user experience with full-access protection to privileged information. Audit trails make it easier to track access to content throughout the platform, regardless of the user interfaces employed. This feature supports compliance requirements about who has done what and when.
Build for Today and Tomorrow

Together, Microsoft and Open Text are making ECM available to user desktops, driving the adoption of ECM even deeper into organizations. That means every information worker is involved in helping the organization become more compliant and more productive, enabling better content management. The investments Microsoft has made in the 2007 Office system platform enable a more seamless and improved user experience with the Open Text platform.

Open Text brings to the partnership an impressive breadth and depth of experience in ECM, garnered from 15 years of delivering ECM solutions to large enterprises around the world. Open Text provides applications on its respected content management infrastructure, extending the architecture to cover other content and providing a cost-efficient and compliant records management and storage architecture.

Microsoft and Open Text customers have long recognized the importance of a flexible and connected ECM infrastructure that fits strategically with enterprise investments. If deployed as another silo, enterprise content technologies can perpetuate the very problem they’re designed to solve: the ability to find content. But if deployed as part of a framework of a true ECM strategy, the right technologies can leverage existing IT investments and enable a richer user experience.

The strategic alliance of Open Text and Microsoft signals a revolutionary way to manage unstructured content. The solution partnership of these software powerhouses is ringing in an exciting new era in ECM usability.
About Open Text

Organizations today face an explosion of digital content and its impact on financial performance, process efficiency, and corporate risk. Open Text provides the bridge between any variety of enterprise systems and the desktop tools people use every day, enabling organizations to securely capture, organize, and manage all relevant content in the context of their business or industry, distilled into one concise view on the desktop. Our ECM solutions extend Microsoft to the enterprise and SAP to the desktop, helping companies control the digital explosion to maximize ROI, streamline processes, assure compliance and inspire innovation.

Open Text is the established leader in providing ECM solutions for all types of business and every major industry. With a history of success and innovation that has positioned us as a recognized industry leader and true global presence, Open Text is the world’s largest independent ECM provider, serving more than 13,000 customers in 114 countries. Learn more about us at www.opentext.com.

About Microsoft

Founded in 1975, Microsoft (Nasdaq "MSFT") is the worldwide leader in software, services and solutions that help people and businesses realize their full potential.