

Outsell's Gilbane Group Beacon

Guidance on Content Strategies,  
Practices and Technologies

---

## Next Generation ECM for Mission Critical Applications

Open Source ECM Capabilities and Opportunities

---

By Dale Waldt, Senior Analyst

November 2010

Sponsored by

**nuxeo**  
Open Source ECM

**E**nterprise Content Management (ECM) generally refers to a system for managing the various types of content and documents that an organization relies upon to support its business objectives. ECM, the term, has been around for more than a decade and has evolved during that time. In early 2010, AIIM (the Association for Information and Image Management) updated its definition of ECM:

*"Enterprise Content Management (ECM) is the strategies, methods and tools used to capture, manage, store, preserve, and deliver content and documents related to organizational processes. ECM covers the management of information within the entire scope of an enterprise whether that information is in the form of a paper document, an electronic file, a database print stream, or even an email."<sup>1</sup>*

The three words that make up the term are descriptive of the role and scope of this definition. "Enterprise" addresses the broadest scope of an organization's activity, including internal users and external stakeholders and people throughout the information supply chain. "Content" has grown to include the broadest range of content types in common use today, including word processing and office documents, HTML and XML, graphics, video, records, and even databases. "Management" refers to a broad range of activities related to creating, updating, managing, extracting, and delivering the content.

Enterprise content management (ECM) systems have long proven their value as enablers of mission-critical business applications. If aligned properly with organizational business goals, they can provide significant benefits. Processes can be streamlined, even automated. Users can access content in ways that are optimized for completing their work. Customers can be given access to content in ways designed to enhance their experience and satisfaction. Supply chain and business partners can be integrated smoothly into the operations of the enterprise.

In the past, buyers of ECM were constrained by the limitations of the systems deployed, including poor interoperability, data formats that did not support new Web delivery capabilities or reuse of content, rigid licensing, and a big price tag. As the next generation of solutions takes hold across the ECM landscape, stakeholders are finding attractive options to expensive monolithic systems. The new ECM platforms are on par with the features, functions, and capabilities that are available in first generation ECM solutions with proprietary architectures, and also offer significant functionality, price and deployment advantages. In addition to competitive capabilities, these tools can leverage open source development and support community and achieve lower overall costs.

This paper targets business and technical decision makers who are evaluating ECM solutions and want to ensure that they are considering the full range of systems available today. It lays out the requirements for mission critical enterprise content solutions. It differentiates ECM as a class of technology application from other types of content management technologies. It describes the core functionality and capabilities that are essential for contemporary ECM practices, and shows how emerging modern ECM platforms overcome limitations that have compromised the full realization of these benefits in the past. A table comparing various ECM architectures and solutions illustrates that open source ECM platforms are solid candidate technologies for an ECM deployment project. Business and IT leaders can consider these platforms with confidence.

---

<sup>1</sup> AIIM: <http://www.aiim.org/What-is-ECM-Enterprise-Content-Management>



# The Promise of Enterprise Content Management Systems

Content management is a broad field. Several acronyms have emerged in recent years to describe and differentiate variations of content management systems, including Content Management Systems (CMS), Enterprise Content Management (ECM), Web Content Management Systems (WCMS), and even Component Content Management (CCM). The differences between these types of systems have mostly been related to the scope or type of the content being managed. For instance, initially CMS systems were designed to support the creation and management of content internally before it was forwarded to another part of the organization for delivery, but did not typically manage Web or mobile delivery requirements. WCMS systems manage and serve up content via the Web, and even mobile devices, but do not address content creation and management processes. CCM systems came about when organizations realized their content may consist of smaller fragments that might be reorganized and/or reused in a several publications or deliverables. ECM systems attempt to address the broadest scope of activities of all these different types of content management system in a single platform.

Advanced ECM platforms support a rich set of content creation, aggregation, classification, reorganization, transformation, and delivery features. This approach both unifies the management and access to diverse information and delivers it in a range of output formats. Changes in licensing models challenge the idea that ECM has to be an expensive and risky investment.

## Capabilities of ECM Today

This section describes capabilities for ECM systems where key functionality has advanced recently through standards, technology, and practices, and costs and risk have been reduced. It describes the current capabilities of the new generation of ECM platforms for consideration while evaluating ECM technologies.

### Multiple Document/Data Types

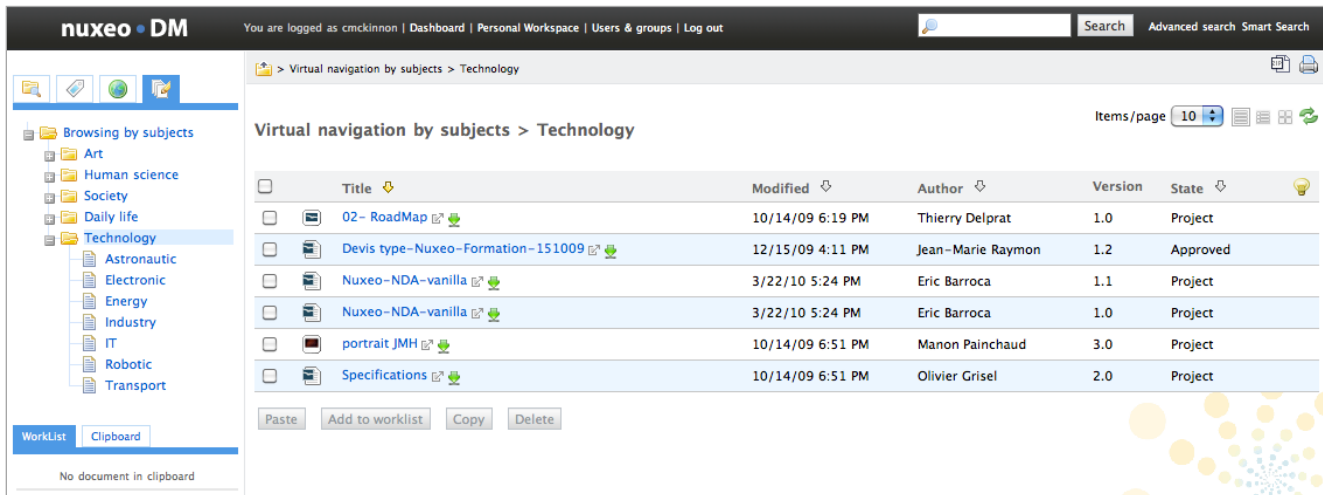
Content comes in many flavors. Organizations of all types may have word processing files, spread sheets, XML or HTML content, graphics, even videos and other binary formats. For an ECM to be effective, it must support a wide range of data formats and content types. Specific applications may be required to handle the wide range of possible formats and enable users to access or display the content.

### Multiple Content Views / Faceted Navigation

Content can be organized in many ways: topically, sequentially, by tasks, by authors, or by whatever structure makes sense for a particular set of data. How content is organized for storage in a repository can make a big difference in how it is accessed and viewed by users and internal consumers of information. Ideally, a repository should provide more than one method for accessing the information managed within it. Different users should have access and navigation methods that are optimized for their needs. Default views into the data should be automatically applied based on a user's roles or rights and be automatically configured to support an efficient and streamlined process. An ECM platform with powerful customization controls can ease configuration of different views and navigation methods, utilize the best of search and display capabilities, and make the ECM interface as intuitive as possible for each audience.

These custom views and navigation methods can be created a number of different ways, but generally rely on search techniques (described later) and descriptive metadata stored with the content. The more powerful and flexible tools for creating custom views allow dynamic updating, generation of persistent listings that can serve as navigation interfaces or "lists of contents," and sophisticated sorting and organizing capabilities.





*Example of a navigation interface generated automatically from topical metadata content.*

The interface into these custom views should also be highly configurable and allow descriptive elements of the content to be displayed with the content listings to aid users. Finally, the custom listings produced should be exported in open interchange standards such as ATOM, JSON, XLS, or even Comma Separated Values (CSV) files that provide access into the ECM platform from 3rd-party systems.

## Reliable/Secure Content Management

Enterprise applications tend to span many time zones and working environments. To support mission critical applications, ECMs must provide ways to ensure both availability and security. Management across heterogeneous yet compatible systems can be challenging, but may be essential to support timely, accurate, secure access to the information required for the task at hand. A robust, stable, high-performing data storage environment is needed to manage complex interrelated content. Methods and standards for integrating the systems themselves and the various content can provide a seamless user experience while providing management and accuracy efficiencies in a secure, distributed environment configured according to the security needs of the enterprise.

## Document Routing and Workflow Management

The most sophisticated ECM systems provide document routing and workflow management controls to support document-specific processes. Steps in these workflows can be automated or prompt a user to perform a task, such as editing, reviewing, approval, or other actions required by the business rules it is designed to enforce. These features can also help organizations manage related information together in case-centric processes, provide insight into progress, and provide guidance and reporting capabilities to help assure the work is completed correctly and on time. Workflow features should be highly customizable and extensible to allow efficient processing. Ideally, the ECM platform should support integration with third-party systems to eliminate the need to store duplicate content and support redundant application functionality.

## Metadata Management and Search Facets

Content stored in an ECM system with related metadata that can be filtered to select content meeting search criteria is often referred to as faceted search—the selection criteria being the means of identifying content that fits within a specific "facet." This function is very similar to manual searching, but instead of being



entered each time, the platform stores the search criteria in a persistent manner to allow that selection capability to be performed repeatedly. In this way, the ECM platform can provide persistent customized access and navigation tools that better meet the needs of more than one user audience.

## Content Analytics and Reporting

ECM systems that provide strong content analytics and reporting allow better administration of content. Content analytics may evaluate the content itself, related metadata, or actions of users creating and consuming the content. Keywords may be analyzed to provide semantic links between content. Or the content may be scanned to recognize and tag people, places, or significant elements for use as links or automatic classification and navigation support. Content analytic features increase the value of the ECM system, making it easier to access, manage, and even reuse content in support of mission critical business requirements.

## Distributed Users and Contributors

Mission critical applications are often used by audiences that are distributed geographically and need to be accessed across a wide area network or the internet. Some of these users may specialize in areas other than content management, such as legal departments, engineering, subject matter experts (SME) or users up and down the supply chain, or even the public at large. Access to the content in an ECM system can be improved through the use of social collaboration tools that provide access to the system to support communities of contributors and users. Collaboration features may include content editing and annotating capabilities, specific workflows and workspace features such as widgets, navigation tools or dashboards, or other applications. Some even include email and micro-messaging features (similar to Twitter). Distributed collaboration tools enable a community to work together across great distances and time zones.

## Flexibility and Extensibility

In the world of ECM, one size does not fit all. Enterprises can operate using highly customized procedures and infrastructure, and an ECM platform needs to support tools in different operating systems and frameworks. Robust configuration and customization tools for interfaces, searching, display, storage and workflow can allow the tools to be molded to the needs of the enterprise to create highly efficient federated work environments. Interfaces should be made intuitive to the specific audiences that will be expected to work with them. Workflow applications can guide users through a process or perform some steps automatically. Content results should be organized and integrated with related information. In short, the system should adapt to the processes and the people, not the other way around, and provide configuration and customization tools that allow content architects and developers to develop custom adaptations for business users.

## Open Standards for Improved Interoperability

The power of open interoperability and data formatting standards is evident with the state-of-the-art ECM platforms. An open standard is one that was developed by a standards community with a formal adoption and review process by a group of representatives from a cross section of users, developers, and implementers. There are several formally recognized standards development groups that have produced specifications for improving data formats and structure definition, methods for systems to interoperate, and other aspects of ECM. These include the World Wide Web Consortium (W3C)<sup>2</sup> and the Organization for the Advancement of

---

<sup>2</sup> W3C: <http://www.w3.org/>



Structured Information Standards (OASIS)<sup>3</sup>. Compared to proprietary formats and specifications, these open standards are more stable, consistent, widely understood, and address the broadest set of users. In short, open standards provide long-term reliability that enables stable platforms to be developed and deployed.

ECM platforms offer a wide range of capabilities that may be supported by one or more open standard. Data management is greatly enhanced using XML and other standard formats for content. Interoperability is improved using Web Services standards from the W3C and OASIS. OASIS also recently developed and ratified Content Management Interoperability Services (CMIS)<sup>4</sup>, a standard for communicating between content repositories in a reliable and efficient way, to share content across a network, and to access the data using powerful query functions against a wide variety of data formats. Adoption of CMIS and other open standards is growing rapidly and is essential to today's ECM platform. They are especially powerful and efficient when built into the design of the platform from inception instead of added later or provided through third-party middleware.

## Additional Considerations for Open Source Solutions

### Broad Support / Developer Community

Open source platforms can provide a diverse, experienced developer community to augment and reduce the costs of expensive consulting and support agreements. If coordinated well, this community can extend the functionality of the core platform, integrate third-party applications, and configure the platform for specific industry standards or processes. Others can act as support resources and help answer questions and provide examples. Significant efficiencies can be passed along to end user organizations by avoiding costly repetitive customization work and providing ability to leverage and build upon existing extensions or application enhancements. As the community's experience grows, so does the knowledge available to its members.

### Cost Effective Licensing

It was not unusual for a first-generation ECM system to cost a million dollars or more in the past. Now, through alternative delivery methods, such as open source licensing, upfront acquisition costs are greatly reduced, typically to a fraction of traditional costs. Companies who adopt next generation ECM platforms often favor open source licensing in order to focus their budget and project expenditure on the meaningful services and customization needed for specific use cases. The open source community has recognized that some enterprise activity will require a team of support developers. Many companies have made a thriving business supporting specific open source platforms. The result is that costs tend to be lower while support services remain high for many mature open source platforms.

## Gilbane Group Guidance

As the business ecosystems around ECM evolve, so must the tools that enable enterprise solutions. User expectations have become more sophisticated and demanding. Standards and systems can now provide seamless integration across heterogeneous environments. Content can be enriched and managed more effectively. And older approaches to architecting a system are becoming increasingly obsolete.

---

<sup>3</sup> <http://www.oasis-open.org/>

<sup>4</sup> CMIS: [http://www.oasis-open.org/committees/tc\\_home.php?wg\\_abbrev=cmis/](http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=cmis/)



At Gilbane, we suggest that any organization looking to deploy ECM technology today include open source platforms on the list of candidate solutions. They are indeed capable of meeting core requirements and delivering business value. In some cases, their approach and contemporary architecture can be very appealing to organizations that might have shied away from ECM in the past due to time to deployment, cost, and the effort to customize proprietary solutions. Open source platforms now make the business benefits of ECM available to new classes of buyers and users for whom they might have been out of reach in the past.

Regardless of which approach an organization ultimately takes, we encourage stakeholders to make sure to consider the following issues when planning ECM acquisition and adoption.

- **Define requirements for *enterprise activity*:** It is no longer adequate to simply consider the needs of the content creation department when capturing requirements and then try to figure out the needs of other stakeholders as an afterthought. Use an appropriately broad definition of what is in scope for the solution. Take into account the needs of content creation, enrichment, transformation and reorganization, and delivery. Different classes of users require methods for accessing information that are optimized to help them perform specific tasks. Stakeholders up and down the information supply chain may need specialized access and formatting of content to support their needs. Administrators will need content analytics to help curate the content.
- **Consciously evaluate pros and cons of platforms versus integrated application suites:** Established technology in the ECM space has evolved through mergers and acquisitions. These systems are typically offered as a suite of independent applications that have been made to work together. The new generation of ECM is based on platforms that are designed from the ground up to support interoperability, data access and sharing, and secure, role-based access across a distributed federated architecture. They provide improved interoperability, more efficient custom processes and tools, and a more reliable infrastructure for mission critical applications. Solutions designed to work in a federated, heterogeneous environment from the start remove constraints that may be imposed by legacy tools that were originally intended to work in a stand-alone manner but now must work together.
- **Exploit the potential of flexible licensing for cost management:** Today, even the most robust and well-supported open source solutions are usually much less expensive overall than proprietary systems. Open source software solutions across the entire IT spectrum have come a long way in the last decade. They have matured into proven, reliable platforms for complex, high volume, mission critical applications. Typically, up-front acquisition costs are much lower for open source tools than proprietary ones. For ECM buyers, the result is that ECM solutions can be deployed less expensively without compromising on features and reliability. Furthermore, the communities that develop around open source technology create a large body of experienced development resources to draw upon. Some open source tools are supported by a dedicated consulting and support organization, but even these support costs tend to be lower than for proprietary solutions.

The following table lists requirements for a robust, distributed, flexible ECM solution and compares a leading open source ECM platform to well-known proprietary systems. What is striking about this comparison is that the open source solution is clearly on par with the more established monolithic or proprietary ECM systems. If requirements are defined carefully and properly, organizations can find significant advantages in an open source platform without sacrificing reliability or functionality. And the resulting system can be flexible and extensible to accommodate new requirements over time, further protecting the investment in ECM technology.



# Summary of ECM Requirements for Mission Critical Applications

Below are major requirements for Enterprise Content Management (ECM) systems that meet the needs of today's demanding mission-critical applications. It is intended to assist information architects, application architects and senior technical decision makers in selecting an ECM solution. The table compares an open source ECM platform (Nuxeo EP) with suites offered by two leading ECM suite providers (EMC Documentum and Open Text ECM Suite). It is meant to encourage ECM stakeholders to consider the range of options available for long-term strategic application development and growth.

Legend: ● = Fully meets requirement, ○ = Partially meets requirement, Blank = Does not meet requirement.

Description	Priority	Nuxeo Enterprise Platform	EMC Documentum	Open Text ECM Suite
<b>Multiple Document Class Support:</b> Manage many document and content types including text, emails, office documents, PDFs, graphics and binaries, and related metadata, and accommodate new types easily.	High	●	●	●
<b>Multiple Access Methods / Views into Content:</b> Automatically produce multiple methods for accessing content (e.g., table of contents, faceted search, etc.) to support different users and processes.	High	●	●	●
<b>Security/Role-Based Access:</b> Storage of content in secure work areas with access controlled by user authentication enabled or restricted by user roles.	High	●	●	●
<b>Document Lifecycle/Workflow and Project Management:</b> Manage and control access and processing of content following to lifecycle states and processing steps, allow set up of multiple workflows by business users, enable grouping of documents into workgroups, projects, or other form to enable related content to be tracked and managed together.	High	●	●	●
<b>Content and Metadata Reporting/Searching:</b> Capture metadata and generate custom reports from content and metadata (e.g., by document creation date, processing date, etc.), tag and store content and metadata, manage values and taxonomies, process and export using metadata in tagged or other common formats, and to save searches to be run automatically or manually.	High	●	●	●
<b>Online/Offline Availability:</b> Available for use in both online and offline modes, preferably with offline mode synchronization and encryption.	Medium	●	○	○
<b>Distributed Collaboration:</b> Support distributed collaboration with intuitive end-user tools, Web-based access, and discussion and annotation features.	High	●	●	●





Description	Priority	Nuxeo Enterprise Platform	EMC Documentum	Open Text ECM Suite
<b>Imaging/Data Extraction:</b> Ability to scan and image documents, and to process and store them in common file formats, and to get textual information out of image files (e.g., names, dates, document ids, etc.).	Medium	○	●	●
<b>Open Standards for Extensibility/Interoperability:</b> Recognize and use open standards for content/metadata, interoperability, extending functionality, and integration with other systems or platforms (e.g., XHTML, Dublin Core, WebDAV, LDAP, CMIS, REST, OpenSocial, etc.).	High	●	●	●
<b>High-Volume Processing / Performance / Scalability:</b> Manage and process large amounts of content with frequent updates and transactions in a flexible and scalable system, and to work installed across multiple servers. Architecture includes platform / technical coherence, extensibility of data model and development tools, and differentiation from assembly of an acquired suite of components.	High	●	●	●
<b>Content Analytics and Reporting:</b> Ability to analyze, report, and manage content and metadata through querying and reporting features.	High	●	●	●
<b>Versioning/Auditing/Logging/Retention/Archiving:</b> Manage all versions of a document, revert to prior/subsequent versions, and track and report document creation, change, deletion, processing, etc. Retention and archiving rules for content per archiving and retention policies.	High	●	●	●
<b>Support/Documentation/Development Community:</b> Access to robust documentation, support tools and community resources for management and configuration, including an architecture / development platform that is widely adopted and well supported by skilled developers (vs. a proprietary or limited scripting language), with documented APIs and design tools to support backwards compatibility of customizations.	High	●	●	●
<b>Flexible/Protected Licensing/ Rapid Deployment:</b> Flexible licensing options that protect the user and encourage system expansion. Easy and rapid deployment ability to support enhancements and prototyping of features for a wide range of vertical or horizontal applications.	High	●	○	○



Outsell's Gilbane Group provides analyst and consulting services dedicated to content strategies, practices, and technologies and their application to high-value business solutions. We have delivered quality market education and insight to global enterprises, governments, non-profits, publishers, and information providers since 1987. Outsell's Gilbane Group is part of Outsell, Inc., the only research and advisory firm focused on the publishing and information industries. Our international team provides independent, fact-based analysis and actionable advice about competitors, markets, operational benchmarks, and best practices, so our clients thrive and grow in today's fast-changing digital and global environment.

Call +1 617.497.9443  
Fax +1 617.497.5256  
763 Massachusetts Avenue  
Cambridge, Massachusetts 02139

Call +1 650.342.6060  
Fax +1 650.342.7135  
**330 Primrose Road, Suite 510**  
**Burlingame, California 94010**

Call +44 (0)20 8090 6590  
Fax +44 (0)20 7031 8101  
25 Floral Street  
London WC2E 9DS

