

Introduction

The value and importance of metadata is one of the most underrated aspects of information technology (IT) today. Often defined as “data about data,” it has been a lightning rod for misinterpretation and myth, which has undermined its significance and fueled misunderstandings about the impact it can have on enterprise content management (ECM) and other IT disciplines.

According to the March 2012 Gartner research report, *How Metadata Improves Business Opportunities and Threats*, “Metadata unlocks the value of data and, therefore, requires management attention.” It goes on to say, “Failure to manage metadata properly will hamper critical activities such as information

management, business process management and service-oriented architecture (SOA) initiatives.”

Experts say that the more valuable the information asset, the more metadata you’ll find associated with it, and therefore the more critical managing metadata becomes to your organization. Unfortunately for many companies, it continues to be an afterthought, particularly where ECM is concerned. As a result, many of them are missing out on the myriad business benefits that proper metadata management offers, from highly improved management of workflows, business processes and access permissions, to faster, more accurate information navigation and search, all of which directly lead to better overall business performance.

What is Metadata?

Wikipedia defines metadata as “structured information that describes, explains, locates, or otherwise makes it easier to retrieve, use or manage an information resource.” Examples of the most basic metadata associated with documents include informational properties such as *author*, *date created*, *date modified* and *file size*.

Those who have been in IT for more than a few years know that the concept of metadata is not new: it’s a term that emerged in the late 1960’s. It has been used since the advent of the Dewey Decimal System in libraries for classifying and categorizing materials based on title, author and subject so that they can more easily be located on shelves from among many thousands of others.

Libraries provide an early example of metadata applied in the physical form, but in our increasingly digital world, the primary purpose of metadata management remains the same: to more quickly navigate across data repositories in order to efficiently find, manage and track information. This, in fact, continues to be the essence of what makes metadata management so critical to today’s most widely used applications, for both small-to-midsize businesses (SMBs) and enterprises alike.

After all, an information asset is not of value unless it can be found and reused, and for businesses that need to operate at breakneck speeds, finding data fast can mean the difference between work getting done today versus tomorrow -- or not at all. Metadata offers a powerful and more effective way to access, organize and track vital business information and processes.

Also important is the value that metadata offers for creating associations and relationships between items and users across one or more repositories or related applications, such as an ERP or CRM system, as well as its benefits for instituting consistency in the way information is used, stored and shared. Metadata also provides clarity about data origins and data histories, and ensures workflows and business processes are properly followed and administered. For example, metadata may include information on the development and lifecycle of a document, including the users, processes and applications involved in its creation, revision and archival, retention and destruction, with granular details that drill down to the exact timestamp of changes and actions, such as reviews and approvals, as well as the access permissions involved in performing them.

In other words, metadata organizes and tracks the entire digital lifecycle of important business information, including the processes, procedures and users that affect it, providing a precise audit trail that can prove invaluable -- or mandatory in highly-regulated industries -- to your business at any point in time. Protecting and organizing this audit trail is yet another reason why metadata should be a cornerstone of your ECM strategy.

The value of metadata lies in its ability to more efficiently classify and organize information, as well as to yield deeper insight into the actions taking place across your business, providing more intelligence and higher quality information to fuel big data initiatives, automation, compliance, data sharing, collaboration and more. Yet, many ECM applications use metadata as an additional information layer and rely on traditional folder structures to organize information. Consequently, users consider metadata administration as extra work, which often leads to bad metadata quality. On the other hand, effective metadata-based ECM systems only ask users to describe the document or data object with tags and properties when saving it. This enables users to search the information in a manner that is most logical to them (i.e., project name, date, contract type, customer, etc...).

Managing Metadata Separately: A Best-Practice Approach

Metadata exists for all structured content in your organization, and most unstructured content (i.e., files) also contains basic metadata properties, such as name, size, type, date modified and attributes. Microsoft Office files (Word, Excel, PowerPoint, Outlook), for example, contain considerable metadata information, such as number of pages and words, and the name of the person who modified the document last.





But rather than looking at metadata simply from the perspective of the file attributes it can reveal (information on date created or modified, size, etc.), forward-thinking organizations view it from a holistic business perspective. When you consider that it contains specific properties that not only relate to critical elements of the organization, but that can also be proactively applied to drive processes and classify data intelligently – such as by project, customer, workflow, state and other factors – then you start to recognize the powerful role it can play in ECM.

This brings up another key point: what companies need to concern themselves with is not so much where metadata resides, or what attributes it contains, but *what* the information represents and *how* it should best be classified and managed. And metadata-based approaches to ECM that enable information to be categorized not by *where* it is, but by *what* it is, are gaining ground as the preferred choice over folder-based ECM systems with limited metadata management capability. That's because it makes for a more effective and faster search to describe *what* something is rather than guess *where* it's stored (which folder it's stored in, for example).

And to get to the “what” that's behind metadata, many are turning to a best practice approach of separate metadata management. This approach takes into account the entire scope of enterprise content, including addressing the idea of metadata associated with information

for which no file exists. For instance, an audit or a deviation is not a file, but an object for which metadata exists, so by definition, to support this, the system must manage metadata separately from the file itself.

And when metadata is not embedded in files, but managed separately, IT administrators gain more flexibility to:

-  Manage metadata structure using centralized tools.
-  Support adding metadata to all documents regardless of file format.
-  Add metadata to documents (or objects) that do not contain files (or that contain multiple files). This is useful when a document is actually a single paper copy that needs to be incorporated into the ECM system. Some ECM providers often refer to records management when discussing this capability, and others simply provide it as another way to manage a document.
-  Export files from the ECM platform without metadata tags.

Separate metadata management in ECM helps to ensure that all enterprise information is searchable, available and exportable – regardless of file type, format or object type -- underscoring again the idea that data is not valuable to an organization unless it can be found.



Getting Started with Metadata Management

To unlock the true value of metadata, organizations must first define how they will organize and use it. In simple terms, they need to decide on the types of information they need to manage and track, such as document types and business processes, along with the metadata that will enable them to accurately classify, organize and process such information.

According to the September 2011 Gartner research report, *Defining the Scope of Metadata Management for the Information Capabilities Framework (ICF)*, “a key objective of metadata management is to drive a consistent approach to the management of information assets, in contrast to the fragmented approaches adopted by many organizations.”

Semantics must be taken into account at this stage in the process. Groups within an organization may use terminology in their own unique way, and the same metadata item may be called different things by different departments. And while names and descriptors are deeply ingrained into a corporate culture, standardized vocabulary surrounding organizational metadata is essential to optimal classification and management of content across the enterprise.

Also important is the need to determine the degree to which – or how – metadata will be integrated across the enterprise. After all, metadata not only describes and classifies information, but also indicates how different information assets are related to each other. IT needs to carefully consider how

aspects of the business, including associated applications and platforms, will leverage and benefit from metadata in order to add the most value to information assets.

With all of the above in mind, developing a strategy for managing metadata should embrace the notion that things change over time, especially when it comes to ECM. The strategy should focus on creating a highly adaptable and configurable framework for the way in which an organization produces and manages information across its lifecycle. In this regard, the goal is to provide a simple yet effective structure for employees to follow that can be easily and flexibly modified as new types of content emerge and as the business evolves and grows over time.



Metadata-Driven Access Control Protects Confidential Information

One of the most important factors for any information management executive (or any executive for that matter) to consider is access control: who is authorized to access, edit and/or approve specific documents? These decisions impact content throughout its entire lifecycle – meaning that permissions can evolve and change over time, adding to the complexity.

While policies related to access permissions can vary depending upon your business, IT leaders today have found a measure of relief surrounding this challenging process with metadata, which can be leveraged to create a smarter, more successful permissions strategy for ECM and other enterprise systems. Metadata has moved to the forefront of permissions

control as it has become increasingly clear that traditional approaches to managing access to content are often too restrictive and inflexible. In fact, today's most advanced ECM systems offer new ways to derive access control settings from metadata, making the process of setting permissions for documents and other information both dynamic and automatic. Metadata-driven permissions and the associated audit trail and event log also help organizations prove that they actually follow the access control policies they have defined.

Information in traditional folder-based ECM systems typically inherits access permissions from the folder in which it resides within the ECM system's folder hierarchy. An additional approach supported

by traditional ECM systems is what is known as an Access Control List (ACL), which is basically a list of access permissions that can be assigned to a specific folder, which is then inherited by all the documents stored in that folder. These folder-based approaches are inherently inflexible and restrictive in that they rely on information residing in a single location, or folder, which presents a dilemma since information can only reside in one location, unless it is duplicated, and duplication creates new challenges associated with ensuring that all duplicates remain synchronized and up to date. This intractable problem simply can't be effectively solved with folder-based approaches. It's challenging enough when one is just classifying content, but the fundamental weakness of folder-based ECM systems becomes glaringly apparent with regard to access permissions.

In a metadata-driven ECM platform, access to content can be controlled by a combination of object-specific permissions and ACLs that are automatically determined by its metadata. The idea is that instead of inheriting access control settings from a containing folder, a document should have its final access permissions derived from its metadata, so a single document could be accessible to members of a project team, a certain group of managers, all of management and accounting, only to employees with a certain security clearance, or any combination of these. Further, permissions could automatically change based on the state the document is in a particular workflow, whether it is in draft form and

being reviewed, or approved and ready to be published. And simply assigning or changing the document's metadata could automatically adjust permissions as appropriate.

For example, an employment agreement document may have its permissions derived from several pieces of metadata: the specified document class ("employment agreement," in this case) may restrict access to the document to the HR department by default. Further, the employee field in the document's metadata may expand access to the document by offering the employee in question the ability to view, but not edit, the document. And more, the supervisor of the employee may automatically be granted appropriate access to the same document, based on the metadata of the employee object. All of this is fully dynamic in a truly metadata-driven ECM solution, meaning that changes to metadata can be instantly and automatically reflected in document permissions.

In a similar manner, you can control the project documentation through one single project data object. You can specify the members of a project with metadata properties and force the ECM to inherit permissions of the documents related to a project from the project object itself. In this way, you can dynamically change the permissions of all objects related to the project by adding or removing project members in the project object.

You can also implement role-based permissions in such a way that project managers have full access to all project content, and project engineers see just those document types that are relevant to their work.

Metadata the M-Files Way

At the forefront of thought leadership in the area of ECM metadata is M-Files, developer of the award-winning M-Files ECM solution. M-Files is built from the ground up with metadata in mind, taking the hassle out of metadata management for organizations in a variety of industries with a user-friendly approach that is both scalable and adaptable.

M-Files is based upon the notion that it is easier to describe "*what*" something is rather than guess "*where*" it is stored. This metadata-driven approach is comprehensive, from navigation (or browsing and search), to access permissions, workflow and replication.

Metadata-Driven Navigation and "Dynamic Views"

When compared to traditional folder-based approaches, metadata-driven "navigation" provides a vastly superior alternative to traditional folders by presenting content in "dynamic views" with familiar-looking "virtual folders," that are dynamically generated from metadata based on the context or need. In this way, a unique document can be found in different "locations," or "virtual folders," with no duplication of data. For example, a proposal can show up in a dynamic view displaying all documents, or specifically in a view displaying just proposals that is automatically organized hierarchically by date, by customer, by value, by sales team, by workflow state, etc. With this approach, "folders," or more correctly, "virtual folders," are simply a dynamic

product of metadata rather than a static location as in a traditional folder scheme.

Metadata-driven navigation as provided by M-Files solves an intractable shortcoming of traditional folder-based approaches that are limited to allowing a file to exist in only one location, or having copies of the file (or links to the file) reside in other folders. In M-Files, unique information shows up dynamically wherever and whenever it is needed without duplication, thereby eliminating a host of traditional ECM and workflow issues.

The definition of a view in M-Files consists of a filter that specifies what content is to be shown in the view, and optionally, one or more grouping levels that define how the content that meets the filtering criteria is to be categorized into virtual folders. One can think of

dynamic views as “queries” to the document repository, with the additional option of having the results presented in a folder-like hierarchy. Company-wide standard views can be set up and made accessible to everyone, or to particular departments or groups, or they can be created as personal views by an individual user. This allows all users to view the document repository based on their particular needs by creating their own personal views. Views are also always up-to-date and display virtual folders based on the content in the system. When a user tags the first document to a customer, a new customer folder is automatically added to all applicable views, so there is no need to update views manually.

Metadata-Driven Replication and Long-term Archiving

In M-Files, replication is also metadata-driven, enabling some data to be replicated to other remote document vaults based on its metadata. For instance, certain documents might be replicated to a cloud-based vault for publishing or archiving, or only certain data might be replicated from a main corporate vault to a subsidiary vault, while all data from the subsidiary vault is replicated to the main corporate vault.

Archiving content in M-Files is also based on metadata, so instead of simply archiving

certain folders to an external storage location, users can create dynamic archiving rules to control what documents and information is archived, such as archiving only documents that are classified as receipts that are more than three years old.

An almost limitless number of scenarios are possible that leverage metadata to control what information is actually being processed and how. For instance, the following are just some of the additional replication and archiving use cases that are enabled with the M-Files metadata-driven architecture.

- Cloud-based backup of an on-premise repository, or vice versa, (i.e., on-premise back-up of a cloud-based repository), which provides robust and instantaneous disaster recovery.
- Replication of data between two or more sites, such as a corporate entity and its subsidiaries, or a business with distributed sales offices or retail outlets.
- Publishing of select content from a corporate on-premise repository to a cloud-based project vault accessible by external partners, vendors or clients. In this scenario, a business can efficiently provide clients and partners with up-to-date price lists, product descriptions,

brochures and other material from the publishing vault at all times, while still maintaining a tightly controlled central document repository.

- Moving content and information from an active repository to an archive repository for long-term retention. This example addresses demanding records management requirements, as well as those associated with regulatory or quality compliance.
- Collecting information from multiple distributed repositories into one centralized enterprise repository. This use case is common in master data management (MDM) and environments with multiple contractors, vendors or partners working on separate but related systems or projects.

Leveraging Metadata for Secure, Automated Access Control

Access control is a vital component of information management, and M-Files incorporates a unique metadata-driven security architecture (patent-pending) for controlling content permissions. This one-of-a-kind architecture provides a revolutionary way for M-Files customers to manage access to confidential content.

Unlike traditional ECM systems, which are more restrictive and make use of antiquated security models based on folders, permissions control the M-Files way means that a document's final access control settings are derived from its metadata – and it is done so in a highly dynamic way, with changes to the metadata driving changes in document permissions – instantly and automatically.

M-Files is the only ECM solution on the market to provide this capability: alternative solutions offer only very limited aspects of it, which creates access control challenges for users and IT leaders alike. With M-Files, businesses get comprehensive metadata-driven permissions for documents and other objects, including the ability to specify faceted permissions. In other words, organizations can derive permission settings from multiple sources (i.e., from multiple metadata properties concurrently), as well as support for metadata-driven direct and indirect “pseudo-user” definitions such as “grant read access to [employee]” and “grant edit access to [employee's supervisor].”

The M-Files approach to permissions management takes the time and guesswork out of this demanding yet critical business imperative, leading to more effective and efficient access control, more successful security audits, as well as a

more fluid and dynamic ECM environment.

Maximizing Workflow and Business Process Efficiency with Metadata

Utilizing metadata not only provides a superior platform for searching, protecting, replicating and archiving vital information assets, but also for ensuring business processes and workflows are followed throughout the entire document and object lifecycle.

An organization's M-Files administrator can easily define workflows about task-related issues (i.e., reviews, approvals, signatures, etc...), and managers can monitor progress and approve completed tasks. Workflows within M-Files can be instituted for approval of contracts, circulation of purchase invoices, processing job applications and for myriad of other use cases that require the review, edit and/or approval by several entities.

With invoices, for example, workflow state options can be configured for “received,” “waiting for approval,” “approved,” “rejected” and “paid in full.” The administrator can define who can move a document into or out of a certain state. For the purpose of this example, only a manager may be able to approve invoices, or only an accountant can move an object into the “paid in full” state.

With the M-Files workflow feature, routine company tasks can be automated and task assignments can be given to the right persons at the right time - with no programming required. Once a task is assigned, the information and responsibility for completing a task are assigned to the correct user. In M-Files, assignments are “objects” that can be included in workflows, or can be independent of workflows. One object can have many assignments, and conversely, one assignment can be attached to several objects (i.e., several documents can be assigned for review with single assignment).

M-Files provides built-in assignment properties for “assigned to,” “assignment description,” “monitored by,” “deadline” and “task completed” - but can also be easily defined and configured to meet the specific needs of the company or task. M-Files users defined as participants within specific workflows are notified via email during the stages of the workflow that they are associated with.

Leveraging the metadata-driven workflow capabilities of M-Files to improve business process efficiencies eliminates bottlenecks, maintains consistency and quality in documentation and assures employees do not accidentally skip a step in important procedures.

Conclusion

Many organizations are just starting to recognize that metadata can form the cornerstone of a dramatically more effective ECM initiative. With a metadata-driven ECM solution in place, companies can leave the limitations and hassles of outdated folder-based approaches behind, while taking a quantum leap forward with a strategy that covers all the bases. M-Files is leading the charge in metadata-driven ECM with an easy-to-use and easy-to-implement solution that enables companies to unlock the power of their mission-critical information and processes.

The result is faster, more efficient search and navigation, as well as more effective access controls, replication and workflow management. And that translates into more time for your organization to innovate, sell and outperform the competition – in other words, to do what it does best.



For more information please contact:

M-Files Inc. | 5050 Quorum Drive | Suite 600 | Dallas, TX 75254
Phone: 972-516-4210 | Fax: 972-516-4211

M-Files Corporation | Hatanpään valtatie 26 | 33100 Tampere Finland
Phone: +358 3 3138 7500 | Fax: +358 3 3138 7550

sales@m-files.com | www.m-files.com