Access®

EXCLUSIVE CHECKLIST REPORT

The Role Metadata Plays in the Information Lifecycle

KMWorld
KMWorld.com

Metadata: Making the Invisible Visible

By Marydee Ojala, Editor-in-Chief, *KMWorld*

The role metadata plays in finding information is generally hidden, yet it is of vital importance. Behind the scenes, metadata provides various attributes that define characteristics not immediately obvious but enormously helpful. It provides both the context and background that are necessary for findability.

Metadata is hardly a new concept. Most of us first encountered metadata when looking for a book in a library catalog. The idea of identifying authors, titles, and subjects dates back thousands of years, to the Library of Alexandria, which put tags on its scrolls so users could quickly find which scroll they wanted to unroll. Tagging documents has progressed significantly since then, as digitization opens up new vistas of what can be included. Simple tagging has evolved into the more sophisticated metadata we use today.

DATA ABOUT DATA

Often defined as "data about data," which may not be all that clear to those outside the records, information, and knowledge management areas, metadata serves as the bedrock for organizing, understanding, finding, and using the ever-expanding universe of enterprise information assets. Within organizations, the amount of data being created, the increasing complexity of that data, and the rising expectations about fast and accurate access to all types of information pose challenges that metadata can address. Without good metadata, vital information can hide from those who need it. As Randy Sanders, Director of Product Management, Access, points out, metadata is not necessarily static, but it is affected by the information lifecycle.

"Within organizations, metadata is critical to finding information in all its disparate formats and, it sometimes seems, disguises. It's vastly more encompassing than what you see on a boarding pass. It could include names of creators, format and length (pages of a text file, minutes of an audio or video file), date created, date disseminated, access restrictions, regulatory and legal compliance, and tags derived from a taxonomy."

Metadata is not just for knowledge management within organizations, however. This can be a useful starting point if you need to explain the concept of "data about data" to those in your organization who need to be convinced of its value. Metadata, you could explain, occurs in aspects of our everyday life as well. As an example, take the bar code on an airline boarding pass. It has information encoded on it that not only relates to the travel itself—flight number, originating airport, destination airport, time of takeoff and landing—but also personal data—the passenger's name, address, contact details, frequent flyer number, authorization for airline lounge access, passenger type (adult, child), passport number, meal preference, and possibly other data, depending on the airline.

Within organizations, metadata is critical to finding information in all its disparate formats and, it sometimes seems, disguises. It's vastly more encompassing than what you see on a boarding pass. It could include names of creators, format and length (pages of a text file, minutes of an audio or video file), date created, date disseminated, access restrictions, regulatory and legal compliance, and tags derived from a taxonomy. Metadata describes information in ways that span the entire lifecycle of information, from its creation and storage to its ultimate archiving or removal.

Some metadata rarely changes. The date created and the original date of dissemination are constants. In most cases, the creator's name won't change, but people do change their names, often because of marriage or divorce, although there can be other reasons. How important it is to keep track of this in the metadata will vary from one organization to another. In the publishing world, it's somewhat annoying when an author of a series of novels starts writing under one name and then switches midstream to another. This makes it difficult for potential readers to find all the titles in the series. One example is the mystery writer who goes by Alice Kimberly as well as Cleo Coyle. Plus, that writer isn't one person, it's a husband-and-wife team, just to add to the confusion.

ASSIGNING METADATA

Metadata is best assigned at the point of creation. That's when the topic is fresh in the minds of those involved. Some can be automatic, such as names, dates, and format. Creators can be reminded to add other types of metadata, particularly about retention (is this to be retained in perpetuity, destroyed after a set period of time, or called up for review in a year), governance (who is allowed to view this, what version is it, can it be copied), and storage (is it in a physical location and if so, where and with what identifying box number, is it online on the intranet, is it a scanned document sent to the cloud). As always, what metadata is used depends on the organization's infrastructure, policies, and regulatory position.

Other issues arise when assigning metadata. Just think of the different ways individuals describe topics, some of which may be more applicable in some situations, while others may be outdated. Quality circles, for example, was a term widely in use in management some years ago but rarely used today. Management hasn't lost its commitment to producing high-quality goods and services, it's simply that other words and phrases are not in use. In one sense, the KM use of Communities of Practice has expanded knowledge sharing to other important considerations, of which quality might be just one.

Terminology might have unique meanings within an organization that those outside it or new hires will not grasp. This could be a code name for a project or a name assigned by the company to an internal

"Tracking data lineage is another area where metadata can help immensely. Information does not, and should not, necessarily live forever. As information develops throughout its lifecycle, its importance to the organization and the world around it can decline, although it can also accelerate."

platform to customize it. Instead of saying "The Corporate Intranet," it might be called "Sparky" at an electronics firm, "AskJerome" for a company founded by a Jerome, or "ChemHereFirst" at a chemical company. In the very early days of ATMs, when they were still a novelty in the financial services industry, one enterprising bank called its ATM "Harvey Wallbanker."

Acronyms are particularly challenging. Encountering the acronym OAR as a standalone taxonomy term, it's unclear what it stands for. It could be Older Adult Retention, Other Activities Release from, Original Aspect Ratio, Office of Alumni Recognition, Open Air Research, Open Account Resolution, or Olympic Athletes from Russia. That's not even a complete list. Context is key to understanding the meaning. A financial firm would not confuse Open Account Resolution with Olympic Athletes from Russia. A university would not confuse Office of Alumni Recognition with Original Aspect Ratio. If rowboats are involved, an oar is not an acronym and does not appear in all capital letters.

DATA LINEAGE

Tracking data lineage is another area where metadata can help immensely. Information does not, and should not, necessarily live forever. As information develops throughout its lifecycle, its importance to the organization and the world around it can decline, although it can also accelerate. That boarding pass with all its metadata in the bar code? It has a very short lifespan, ending with the conclusion of the flight. A company history on its website? Its employees and customers hope for a very long lifespan.

Version control facilitates the process of developing a piece of information, ensuring that, as various individuals work collectively on it, the most recent version is not confused with earlier ones. How one version relates to other, similar information assets also relies on metadata. Tracking changes in terminology that links information using older taxonomy terms with newer ones is another important role for metadata. Department X that is now Department Y needs to be ensconced in the metadata so everyone is clear on the relationship.

Retention policies can have legal implications. Depending on whether an organization is in a regulated industry, bound by the regulatory agency to keep some information in perpetuity, or subject to other external constraints, knowing how often to dispose of stored data is critical. Metadata can be used to track changes during preservation processes as well, including data migration and refresh. It provides documentation on the disposition of information objects, whether they are archived, destroyed, or otherwise disposed of. It may also explain the reasons behind the disposition decision, leading to greater transparency and accountability.

METADATA FOR DISCOVERY

For employees and customers, the most valuable service provided by metadata is findability. Metadata is a memory-booster for some. Someone trying to find a document written by Jane Smith in the past 3 months, with the assistance of metadata, discovers the relevant information, but it was written by Janet Smithfield a year ago. Another searcher wants data about the company's product ABC123 but it is now called DEF456. Thanks to metadata, search results are automatically retrieved for both product names, allowing the searcher to determine which of the products fit the information need. Similarly, searchers may remember a video about the product but it turns out to be a podcast. Metadata can be a boon for discovering information objects in multiple formats.

People describe what they want to know in sometimes wildly different ways. Think of soft drinks, which could be known as sodas, pops, or soda pops, depending on where you are in the United States. They could also show up as brand names. Expand that audience to other countries and other languages, and the complexity in terminology increases. Good metadata means searchers have no need to think about all the possible variations or to worry about missing information because they didn't have the correct term.

The obverse of discovery is not finding information that is restricted. A publicly traded company may be in discussions to acquire another company or divest one of its divisions. This data should only be accessible to those with a need to know. Otherwise, the company is likely to be in trouble with the Securities and Exchange Commission (SEC) in the U.S. or its counterpart in another country.

VALUE OF METADATA

The value of metadata is immense. It may not be top of mind for creators, but showing how employees and customers benefit from discoverability, why retention schedules keep out-of-date information from appearing when it shouldn't, and when to comply with legal issues regarding information assets results in some powerful reasons to pay attention to having good metadata in place.

Metadata acts as a guide and a companion, regardless of where in the information lifecycle it appears. Although frequently invisible, hovering behind the scenes to describe content in ways that may not be evident in the material itself, enables the effective management, usability, and preservation of organizations' information assets. Without metadata, all the information within an organization can be visualized as rows of file drawers, all labeled "Miscellaneous"—except for the one drawer labeled "A to Z." It's a frightening thought. ▮

Marydee Ojala is editor-in-chief of *KMWorld* magazine and is conference program director for Information Today, Inc. She works on conferences such as Enterprise Search & Discovery, KMWorld, Data Summit, and AI & ML Summit. She is a frequent speaker at U.S. and international information professional events. In addition, she moderates the popular KMWorld webinar series.

The Role Metadata Plays in the Information Lifecycle

By Randy Sanders, Director of Product Management, Access

We've all been in the situation where you're desperately searching through a company shared drive and filing cabinets for a specific piece of information that you need to do your job. No matter what keywords you search for or which files you rifle through, it's nowhere to be found. Unfortunately, this issue is all too common.

The challenges organizations face can be overwhelming due to the sheer volume of information generated each day. Adding complexity, as data is created across multiple departments, systems, and formats, it can become fragmented, making it difficult to track, retrieve, and utilize effectively.

Without a comprehensive approach to information lifecycle management, organizations struggle with:

- ✓ **Inconsistent information governance—**Lack of standardized policies leads to compliance risks and inefficiencies.
- ✓ **Limited visibility and access—**Critical information stored in separate systems hinders productivity and decision-making.
- ✓ **Duplicate and redundant data—**Unstructured and siloed data results in wasted storage, higher costs, and inaccurate insights.
- ✓ **Security and compliance risks—**Inadequate control over sensitive data increases the risk of data breaches and regulatory penalties.

To address these challenges and strengthen your information management strategy, organizations should consider using metadata throughout the lifecycle.

With metadata, businesses can break down information silos, streamline workflows, and reduce risks tied to outdated or mismanaged data. From securing sensitive documents to automating destruction processes, metadata provides the clarity and control needed for effective information lifecycle management.

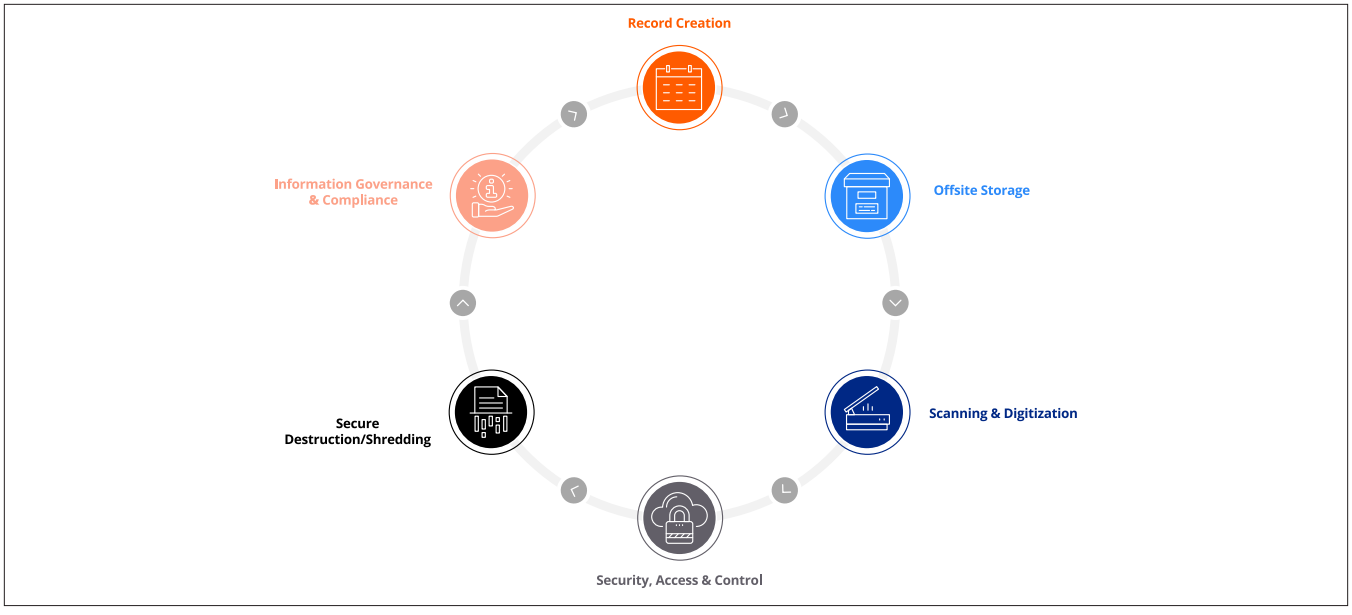
BUT FIRST, LET'S DEFINE THE INFORMATION LIFECYCLE

The information lifecycle describes how a record or document is handled throughout its existence, ensuring compliance, security, and efficiency, while reducing risks and costs.

In its most simplified form, the information lifecycle has three distinct stages: creation, use, and disposition. From there, you'll find that different organizations and professionals label and detail the cycle differently.

At Access, we depict the information lifecycle with six distinct stages, as follows:

- ✓ **Information Governance and Compliance—**Establishing policies and standards for the rest of the lifecycle stages to be built upon.
- ✓ **Record Creation—**Generating and capturing information in various formats.
- ✓ **Offsite Storage—**Securely storing physical records to optimize space and enhance security.
- ✓ **Scanning and Digitization—**Converting paper records into digital formats for improved accessibility.
- ✓ **Security, Access, and Control—**Managing permissions and ensuring data is protected while remaining accessible to those who need it.
- ✓ **Secure Destruction—**Properly disposing of records that have met their retention period to reduce risk.



This structured view of the information lifecycle provides a way for organizations to evaluate how they manage their information from both broad and narrow views. For example, organizations can narrow in on their information governance policies, then evaluate how they affect the management of information in other stages of the lifecycle.

METADATA IS THE KEY TO ENSURING INFORMATION IS MANAGED PROPERLY THROUGHOUT THE LIFECYCLE

Throughout the following sections, let's dive into the important role metadata plays in each stage of the information lifecycle.

Information Governance and Compliance

Metadata plays a critical role in information governance by providing the structure and context needed to manage records effectively. It helps organizations enforce policies, processes, and controls that ensure compliance with regulatory frameworks such as HIPAA, GDPR, CCPA, and more. By capturing key details, metadata supports auditability, security, and proper lifecycle management, thus reducing compliance risks.

Here are some best practices to implement at this stage of the information lifecycle:

- ✓ Leverage metadata to implement automated retention policies that align with regulatory requirements and business need.
- ✓ Define clear classification tags to streamline handling based on sensitivity.
- ✓ Maintain audit trails using metadata to ensure records are properly logged and can be retrieved for audits.

For example, when you're creating or editing reports, add metadata tags such as "Confidential," "Internal Use Only," or "Review Period: 6 months" to ensure proper handling and compliance.

Record Creation

Every document, email, or digital record generated in an organization begins with metadata. It provides context, making information easier to retrieve, categorize, and manage in the future. Applying metadata at record creation makes the future management of that record much easier. Here are some best practices to incorporate:

- ✓ Ensure metadata is automatically generated upon document creation, capturing key details like author, document type, and creation date.
- ✓ Establish standardized metadata tags to be used across all departments and systems of record, preventing inconsistencies and ensuring interoperability.
- ✓ Implement a metadata-driven system to track document versions, ensuring the latest version is always easily accessible.

To better control the creation of new records and convenience copies, consider this example: For frequently updated documents, like project plans, include version metadata (e.g., "Version 3") to ensure everyone is working with the most current iteration.

Offsite Storage

Most businesses (even the ones that are digital-first) retain some records in physical format. This could include legacy records, blueprints, or tape backups, for example. For physical records stored offsite, metadata plays a vital role in tracking, organizing, and efficient retrieval.

Here are some best practices to adopt at this stage of the information lifecycle:

- ✓ Utilize metadata to track the location of records stored offsite (e.g., storage box number, file number).
- ✓ Work with a storage vendor that implements barcode or RFID scanning to quickly retrieve physical records using metadata.
- ✓ Set alerts for when offsite records approach their retention period for timely disposition.

Here's a great example of how metadata can be applied to offsite records: A financial institution stores client tax records offsite in physical boxes, using metadata to tag each record with details such as "Tax Record, Box #34, Client ID: 12345." This metadata allows staff to quickly locate the specific box or record by searching for key identifiers. For records stored digitally in the cloud, metadata tags facilitate easy retrieval in a similar fashion.

Scanning and Digitization

Digitizing paper records is a common practice, but without proper metadata, scanned documents can become disorganized. Metadata provides essential context to digital files, such as document type, date of creation, author, and keywords. Without metadata, scanned records lack the necessary structure for efficient management, leading to time wasted searching for specific documents and increasing the risk of non-compliance or lost information. When scanning and digitizing important records, implement these metadata best practices:

- ✓ Use OCR technology to extract and assign metadata to scanned paper documents, especially when they have handwritten fields.
- ✓ Set up automated workflows that trigger document routing or approval based on extracted metadata.
- ✓ Ensure all digitized documents have well-defined metadata fields to make them searchable and retrievable quickly.

For example, consider a healthcare organization that digitizes patient medical records. Without metadata, scanning a paper chart could result in a digital file with only a generic file name like "Scan_001.pdf." This would make it nearly impossible for staff to quickly locate specific patient information. However, by applying metadata such as the patient's name, appointment date, and case number, each scanned record is tagged with valuable details that allow for fast, accurate searches.

Security, Access, and Control

Metadata plays a critical role in protecting sensitive information by controlling who can view, edit, or delete certain records. By embedding access permissions within the metadata, organizations can establish clear boundaries for sensitive data, ensuring that only authorized individuals can interact with specific documents. For better security, access, and control, incorporate these best practices:

- ✓ Use metadata to define and enforce access restrictions based on role or seniority.
- ✓ Leverage encryption and rights management metadata to prevent unauthorized copying, printing, or sharing.
- ✓ Set metadata-based rules that trigger alerts or actions when sensitive information is accessed or shared without proper authorization.

For example, a corporation handling quarterly financial reports can use metadata to enforce access restrictions, tagging it with metadata such as "Document Type: Financial Report," "Access Level: Executive Team Only," and "Confidential: Do Not Distribute."

When the report is shared, only senior executives and approved finance personnel can open it. If someone outside the group tries to access the file, they will receive a restricted access message.

Secure Destruction

Proper disposal of records is a critical step in managing an organization's information lifecycle, as it helps mitigate risks associated with outdated or sensitive information. Without a systematic process, organizations may inadvertently retain documents longer than necessary, increasing the potential for security breaches, legal complications, or non-compliance with regulations. This is where metadata plays an essential role in ensuring the timely and secure destruction of records.

Some best practices to adopt at this stage include:

- ✓ Use metadata to enforce the retention schedule, triggering alerts when documents are ready for destruction.
- ✓ Ensure audit trails are maintained in metadata to confirm that destruction processes follow compliance standards.
- ✓ Record who authorized and performed the destruction, ensuring a clear chain of custody for compliance.

Consider this example of how you can implement metadata in the secure destruction phase of the information lifecycle: A national retail chain uses metadata to manage the secure destruction of store transaction reports and vendor invoices. When due for disposal, records are tagged with "Destruction Status: Pending" and "Method: Certified Shredding." Then, an automated workflow flags them for approval. Once destroyed, a certificate of destruction provides the destruction date and other important information to retain for compliance purposes.

CONCLUSION

In conclusion, managing the overwhelming volume of data and ensuring proper handling throughout the information lifecycle is crucial for organizations of all sizes. By leveraging metadata, businesses can streamline processes, improve compliance, and mitigate risks associated with unorganized, siloed, or outdated information. Whether it's managing records at creation, securing sensitive data, or ensuring timely destruction, metadata provides the structure and context needed for more efficient, accessible, and compliant information management. By implementing the best practices detailed throughout this report, organizations can enhance operational efficiency, protect sensitive information, and ultimately achieve better control over their data.

With our expertise in records management, Access can assist in implementing metadata-driven processes that streamline operations, reduce risks, and maintain regulatory compliance at every stage of the information lifecycle. Contact us to see how Access can be your partner in managing and protecting your valuable business information. ▮

Contact Information
Access
4 First Ave,
Peabody, MA 01960
1.877.345.3546
www.accesscorp.com