METADATA MATTERS

Introduction

Metadata is a *strategic* imperative for any organization looking to manage and exploit its knowledge more effectively. Optimity Advisors understands this imperative and has both the strategic offerings and the resources required to successfully implement metadata for your business operations.

Indeed, these are good times for metadata. It was General Petraeus' alleged affair a few years ago that brought metadata into the public spotlight with Internet protocol addresses, thanks to the metadata, pointed to network locations that gave the affair away. No longer was it an obscure term reserved for librarians and data scientists modeling content for discovery. It became a notable topic of discussion at the water cooler. And, then there was Edward Snowden, whose notoriety from the disclosures regarding the NSA program on collecting metadata on telephone calls, is responsible for bringing the term metadata to true public discourse, more than Melvil Dewey or any other public figure has been able to achieve. However, while it made its way into the modern lexicon, it is as misunderstood as it is controversial. Senator Dianne Feinstein argued that, "This is just metadata – there is no content involved," responding to the NSA's blanket surveillance of Americans' phone records and Internet activity. But, metadata is more than that and it does matter; there is always content involved and with that comes value, purpose, and meaning.

The opportunity for content owners, marketing technologists, and all those managing content is in understanding the value metadata provides their assets, and how it can empower their digital operations from creation, to discovery, through distribution.

Metadata is …

Metadata, simply stated, is information that describes other data; data about data. It is the descriptive, administrative, and structural data that define your assets.

- **Descriptive** metadata describes a resource for purposes such as discovery and identification (i.e., information you would use in a search). It can include elements such as title, creator, author, and keywords
- **Structural** metadata indicates how compound objects are put together, for example, how a digital image is configured as provided in EXIF data, or how pages are ordered to form chapters (e.g. file format, file dimension, and file length)
- **Administrative** metadata provides information that helps manage an asset. Two common subsets of administrative data are **rights management** metadata (which deals with intellectual property rights) and **preservation** metadata (which contains information needed to archive and preserve a resource)

And yet, metadata is an asset unto itself—and an important one, at that. It provides the foundation and structure needed to make your assets more discoverable, accessible and, therefore, more valuable. In other words: Metadata makes them "smart assets." Simply digitizing video, audio, graphic files, and more only scratches the surface of their value as digital assets. Their full potential is realized only by their use, and they can only be used if they can be found. The robustness and relevance of the metadata associated with an asset is what makes it findable, and therefore usable.

Related Metadata Concepts

Taxonomy is the classification of information into groups or classes that share similar characteristics. It is a way to organize information to best solve a business problem based on user needs by exposing relationships between subjects. A well-designed taxonomy brings business processes into alignment allowing users to intuitively navigate to the "right" content. The best reason for creating and implementing a single, standard taxonomy across the enterprise is that it provides good business value. But more than that, it enhances and improves enterprise search and enables quick information discovery. Taxonomy provides the consistent and controlled vocabulary that can power the single source of truth as expressed in a DAM or CMS. It is a **key enabler** for organizing any large body of content. It is required for meaningful information management and critical to effective "findability."

A *Controlled vocabulary* is for drop-downs/pick lists, and the use of "preferred terms" is a good way to provide authority and consistency to your digital assets. Each tag could point to a different topic, but, fundamentally, it's the same principal element of the subject under review that is relevant. If the topic is "Country", and you only have eight countries with which you work, then those eight countries are your controlled list. **Control**, and stronger yet, **authority**, is needed to describe your assets. **You need to know what it is you are describing and how it may best be described.**

Structured data refers to information with a good level of control and organization, for example, a "date" value in an "Expiration Date" field. Structured data is usually found in a controlled data environment with inherent meaning and purpose. *Unstructured data* lacks that control and meaning offering a confused sense of purpose and requires analysis or interpretation to restore meaning. Using the example above, if a "date" is discovered with no "field" in which to provide that control and structure, what does that tell you? The interest is in wrangling all that data to create a more structured sense of purpose for the content in your organization; it makes information more relevant, palpable, understandable, and useable.

Master Data Management (MDM) is the technical discipline/method of enabling an enterprise (both the Business and IT) to link all of its critical data to one file, most often referred to a "Master" file for centralized governance.

- **Master Data** is business critical data that is governed and shared across multiple systems, applications, or departments within an organization. Master Data can be identifiers, attributes, relationships, reference data, and yes, metadata!
- **Master Data Management (MDM)** is the set of processes, tools, and governance standards/policies that consistently define, manage, and distribute Master Data. This could be expanded to discuss the various implementation details. (Note: No reference to a "master file" or anything as typical implementations these days are much more complex).

You can't have MDM without good metadata. Everything starts with data modeling, and data modeling is inherently tied to metadata (ISO-IEC 11179).

Regarding *Standards*, never miss the opportunity to learn more about some of the available metadata standards specific to your industry or application. A **standard** is an agreed level of quality, created by industry leaders and used as a measure or norm in comparative evaluations. Standards needs to be evaluated on a regular basis through the ongoing work of the Metadata Manager/Taxonomist. Use an **industry standard (or standards)** that applies and then extend it as needed. Consider participating in the standards community by publishing your updates with the governing body.

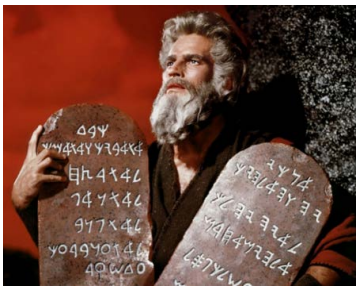
Language Is Always Changing

Research shows that workers waste more than 40 percent of their time searching for existing assets and recreating them when they are not found[1]. This lost productivity and redundancy from the non-discovery of assets is expensive. The key to avoiding these unnecessary costs is good metadata to aid in search and retrieval and workflow. It is estimated that every year, 800 neologisms (new words and phrases) are added to the English language showing how little of our modern language is written in stone; instead, it is always evolving. It is critical to understand that metadata is a snapshot in time representing the business processes and goals at a particular time. In an ever-changing business environment, metadata must be adaptable and must evolve over time to stay relevant to the digital assets that they support. If maintained and governed well, then metadata can be a very real contribution to your business goals.

The best way to plan for future change is to apply an effective layer of **governance** to your metadata. There is more to maintaining the metadata than just maintaining the taxonomy and metadata specifications – **you must manage the change.** Vocabularies must change over time to stay relevant, and processes must be created to manage this change. This is also true for new terminology being added to assets as well as synonyms and/or slang terms and more.

[1] http://www.kmworld.com/Articles/Editorial/Features/The-high-cost-of-not-finding-information-9534.com

As an example, the following images are all "tablets" and so similar by name, but all so very different in meaning depending upon the subject and time at hand.



They are all by definition "tablet(s)," but each of them unique in meaning and value to the user. You need to know what it is you are describing and how it may best be described. If more than 40 percent of time is wasted searching for existing assets and recreating them when they aren't found, then it behooves you to invest the time to get to know your assets, your access, and your users and create an environment that provides meaningful results. The key is good metadata! Your data wants to be found! Develop an incremental, extensible process that identifies and enables users, and engages stakeholders with feedback loops, user testing and evaluations for your metadata. Remember that metadata is a "snapshot" in time. **Take the time to manage your language and control the change.**

Key Findings across Industries and Clients

Optimity Advisors has been working with clients across many industries from media and entertainment, health, insurance, consumer packaged goods (CPG), and pharmaceuticals to help them with digital strategy and information management. Over time, we have discovered common themes among metadata and digital problems:

- Little or no metadata planning for new process or systems
- No clear ownership over digital assets and their management
- No current documentation on metadata/controlled vocabulary
- Lack of documentation and control over assets
- Poor labeling of folders and assets
- Agencies, such as external advertising or digital agencies, are often the default single source of truth

Metadata matters, and in some situations, it is at the wrong end of the business requirements gathering process. In other situations, it is often an oversight; not necessarily a negligible action, but worthy of concern for any practice looking to manage digital assets effectively within an organization.

A Necessary Defense

Metadata is the best way to protect yourself and defend your digital assets from information anxiety and mismanagement. If a good offense is your best defense with metadata as with anything else, then it is worthy to invest the time, energy, and resources to identify, define, and organize your assets for discovery. Metadata serves asset discovery by:

- Allowing assets to be found by relevant criteria
- Identifying assets
- Bringing similar assets together
- Distinguishing dissimilar assets
- Giving location information

The struggle is managing within a "**big-data**" landscape where the data is as complex as the digital workflows it supports. This landscape may not only include your internal one but the wider geography of your partners and third-party entities that crawl for your data on a very public Internet. Add to data complexity, the increasing rate at which it is produced and the diversity of the formats being used. You know that your assets are critical to your business operations, and you want them to be discovered at all points within a digital lifecycle from creation, to discovery and distribution. To accomplish this, you need a discernable sense of trust and certainty that your data is accurate and usable. Metadata matters and is your best chance for a return on investment on the assets you have created and also a line of defense against lost opportunities. Think about the digital experience for your users and ensuring they identify, discover, and experience your brand the way in which it was intended. It is a necessary defense.

Metadata Design: Where to Start?

The path to good metadata design begins with the realization that your digital assets need to be identified, organized and made available for discovery. The following questions serve as the beginning of that design.

- **What problems do you need to solve?** Ensure you know the business goals of your organization and how metadata may contribute to those goals. The goal is to be "cohesive" and not "disjointed."
- **Who is going to use the metadata and for what?** Discover who the audience for the metadata is and consider how much metadata you need; the best strategy is accurate intelligence.
- **What kinds of metadata are important for those purposes?** Metadata may well be expensive; make your model extensible and allow for its natural growth and evolution over time.

Metadata is the foundation for your digital strategy. It is needed to deliver an optimized and fully engaging consumer experience. There are other critical steps to take as well, including building the right team, making the correct business case, and performing effective requirements gathering – but nothing can replace an effective metadata foundation for your digital strategy. As previously stated, you want your assets to be discovered; they want to be found. Content may still be king, but the user is also worthy because if you have great content and no one can find it, the value of the content is as good as it not existing. Metadata will help ensure that you are building the right system for the right users.

Conclusion

Metadata matters and it is neither a trend, nor a buzzword. Metadata is the most real application of asset and data management that enables creation, discovery, and ultimately distribution and consumption. Metadata demands attention in effective business solutions. Our metadata and taxonomy experts at Optimity Advisors work with you to develop and sustain your content strategy, from the beginning of the process with the information audit to the data analysis and ultimately to the organization of your content for discovery. We stay engaged with you and your content stewards to ensure that metadata remains current, meaningful, and actionable within the ever-changing digital landscape.

John Horodyski is a Partner with Optimity Advisors with over 13 years of management strategy experience in Digital Asset Management (DAM), metadata and taxonomy design, digital and social media marketing and brand strategy. John has provided strategic direction and consulting on DAM implementations including metadata and taxonomy design for a variety of Fortune 500 clients from consumer packaging goods, to the pharmaceutical industry and media and entertainment. John is also part of the Adjunct Faculty at San Jose State University where he teaches a graduate course in Digital Asset Management.

optimity ADVISORS

Washington, DC
1600 K Street, Suite 202
Washington, DC 20006
Phone: 202.540.9222
Fax: 202.540.9223
Email: info@optimityadvisors.com

New York, NY
183 Madison Avenue, Suite 1205
New York, NY 10016
Phone: 212.239.3371
Email: newyork@optimityadvisors.com