OUTSOURCED PHARMA

Guest Column | March 23, 2018

Where EDMS Fails: Data Integrity Pitfalls To Avoid In Metadata For Life Science Products

By Kip Wolf, Tunnell Consulting

When creating and managing electronic documents, document metadata deserves as much attention as document content. Firms that do so can improve compliance and even gain potential competitive advantage by realizing electronic documents (with appropriate metadata) as real business assets.

As we all know, life sciences firms generate enormous amounts of data. A colleague of mine once quipped that GMP actually stands for "generate more paper." In modern times, few firms rely exclusively on paper documents, with most making some use of an electronic document management system (EDMS). Even in electronic form, "all regulated data is subject to GxP requirements for data integrity and good documentation practices."[1] With the electronic document comes metadata, which is "data about data that provide the contextual information required to understand those data."[2] Examples of metadata in EDMS include date/time stamp, author, title, and the like. Metadata can also include contextual data to classify or categorize the document content, such as therapeutic area, indication, or registered country, to name a few.

We have found that many firms attempt to use the implementation of a new EDMS to establish an authoritative source of content for critical documents (i.e., a single source of truth). This idea is rooted in the GMP concept of primary records, meaning "the record which takes primacy in cases where data that are collected and retained concurrently by more than one method fail to concur,"[3]

particularly where the metadata is concerned. Before we get into the details of primary data as it influences metadata decisions, let us set a baseline understanding of how EDMS works and the difference between content and metadata.

EDMS 101: Understanding Content And Metadata

For the purposes of this discussion, we are not going to delve deeply into the nuanced differences between, for example, document management, content management, and metadata management systems. Rather, we are going to provide an oversimplified concept to illustrate the importance of metadata to take advantage of some opportunities presented by EDMS and avoid some common pitfalls.

Every EDMS is based on a fundamental principle of content (e.g., what made up the old paper document) and metadata (i.e., the context for that document). Lose either one of these components, or the link between them, and you have lost the value of the object altogether.
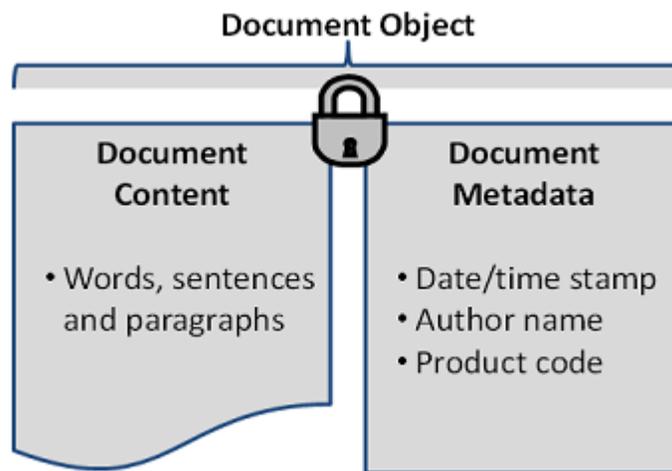

Figure 1: Simplified example of the relationship of document content to document metadata, together making up the document object

Losing the document content has an obvious impact. The critical document (e.g., regulatory submission information) may be lost and unavailable for inspection, annual reporting, and the like. This presents obvious compliance and operational risks. Losing the document metadata renders the document content without context. The critical document may exist, but without the metadata necessary to classify or categorize it by author or related product. The individual components of document content and document

metadata make up the document object. When managed effectively, efficiently, and in a compliant manner, this document object can be realized as a true business asset. We have found that firms that do so realize greater ease in document location and retrieval, inspection readiness, and agile business operations, to name a few benefits.

Metadata Opportunities And Mistakes Affecting Data Integrity

EDMS integrators are not to blame for mistakes in metadata effectiveness. A new EDMS project presents great opportunity to the sponsoring firm, and EDMS integrators are (from what we've seen) generally adept at the technical implementation of their solutions. The risks to metadata implementation come from the lack of the sponsor's understanding of: (1) the relationship of document metadata to document content; (2) the relationship to remaining requirements for compliance with standard good documentation practices/CGMP regulations; and (3) the great opportunity for realizing maturity of the concept of primary records as it relates to metadata. The first risk restates the concepts shown in Figure 1: Simplified example of the relationship of document content to document metadata, together making up the document object. The second risk states the obvious relationship to predicate rules and good documentation practices standards. This article attempts to unpack the details around the third risk related to the concept of primary records as it relates to metadata.

For the sake of simplification, we will continue to use an example of document content related to regulatory submission, as it is a robust example of document content that extends usefulness and utility beyond a single functional area. Regulatory submission documents are (for example) referenced for licensure for commercial marketing, related to the manufacture of commercial products, and relied on for ongoing maintenance of regulatory approval (e.g., annual reports). This type of document transcends the frequently siloed structure of business divisions and functional units and presents a fine example for the purposes of this discussion.

When implementing a new EDMS, or when creating or migrating document content to an EDMS, the question of metadata is always first and foremost. Mapping attributes or document metadata becomes a common exercise, and questions are asked of the document owner (e.g., who is the author, what is the appropriate product code to attribute to the document content, etc.). The opportunities and mistakes in data integrity occur here at the attribution of the metadata.

Let's extend our example by considering the product code metadata attribute. The attribute may be entered by free text (no control) or by selection from a predetermined list or drop-down. Free text is the bane of data integrity objectives, as it provides little to no control of the metadata being entered, and therefore provides risks to the value of the metadata (think of how many versions of your name

exist on your credit cards or bank accounts and trying to group them by consistent name). Using predetermined lists or drop-downs is a far more preferred method, but sourcing the data for the predetermined list or drop-down becomes the principle opportunity and frequent source of mistakes when it comes to metadata and data integrity.

Table 1: Example of Data Mapping Model to Inform Metadata Integration

Attribute (general term)	System field name(s) with authoritative source in RED.			
	EDMS	ERP	LIMS	ProdReg db
Product Name	Product_Name	prod_name	name.prod	Prod-Name
Product Family	Product_Family	prod_family	fam.prod	Prod-Fam
Product Code	Product_Code	prod_code	code.prod	Prod-Code
Trade Name	Trade_Name	trade_name	name.trade	Trade-Name
Generic Name	Generic_Name	generic_name	name.gen	Gen-Name

If we create the list of product codes from a source that is not authoritative (e.g., type the list for the drop-down from an old hard copy list of product codes), there remains a strong probability of error and a difficult operation to maintain the list. However, if the list is created dynamically by integrating with an authoritative source like ERP (see Table 1), then the list will always be updated with the most current and accurate representation of product codes.

An Ounce Of Prevention

By spending as much or more time defining the requirements and controls around the document metadata as is spent around the document content, and keeping some of these principles in mind, companies can create significant value in the form of realizing document objects as true business assets. This approach not only ensures greater compliance with predicate rules, CGMP, good documentation practices, and data integrity requirements, but it also has the potential to provide a real competitive advantage by more accurately, efficiently, and effectively organizing document content, simplifying document control, and enabling specific document list queries.

References

1. *ISPE GAMP® Guide: Records and Data Integrity*, International Society for Pharmaceutical Engineering (ISPE), 2017, www.ispe.org.

2. WHO Technical Report Series No. 996, 2016, Annex 5, https://www.gmp-compliance.org/guidelines/gmp-guideline/who-guidance-on-good-data-and-record-management-practices.

3. MHRA GMP Data Integrity Definitions and Guidance for Industry, Revision 1.1, March 2015, www.gov.uk/government/publications/good-manufacturing-practice-data-integrity-definitions.

This article is part of a series that aims to discuss specific data integrity improvement opportunities within individual quality systems (e.g., document control, records management, materials management) with specific operational examples (e.g., master batch record design, efficient review of executed batch records, improved schedule adherence).

About The Author:

Kip Wolf is a senior managing consultant at Tunnell Consulting, where he leads the data integrity practice. Wolf has more than 25 years of experience in quality assurance and regulatory affairs, GMP and IT compliance, technical operations, and product supply. His areas of expertise include business transformation, new business development, organizational change leadership, and program/project management. He has held various management positions at some of the world's top life sciences companies. Wolf can be reached at Kip.Wolf@tunnellconsulting.com.