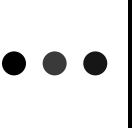


Turning Data into Knowledge: Creating and Implementing a Metadata Strategy

Anne Marie Smith
<http://www.annemariesmith.com>
<http://www.alabamayankeesystems.com>

ALABAMA YANKEE SYSTEMS, LLC



Agenda

- Introduction
- Metadata Overview
- Principles of Metadata Management
- Metadata Strategy
- Key Roles in Metadata Management
- Metadata Quality
- Important Aspects of Metadata
- Conclusion

1/12/2006 (C) 2005 - 2006 Alabama Yankee Systems, LLC 2

● ● ●

Anne Marie Smith Background

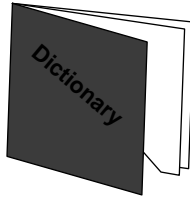
- BA and MBA Management Information Systems (La Salle University)
- PhD, Management Information Systems (NorthCentral University)
- Consultant in data management, metadata, data warehousing, requirements analysis, systems re-engineering
- Certified Project Management Professional (PMP)
- Certified Data Management Professional (www.iccp.org)
- Author of over 40 publications and papers
- Instructor on varied information systems management topics

1/12/2006
(C) 2005 - 2006 Alabama Yankee Systems, LLC
3

● ● ●

Definition of "Metadata"

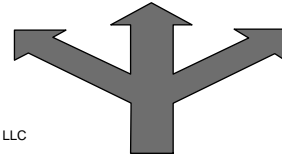
- Metadata: "data about data"
- Collection of information about the data stored in an application or database
- Examples:
 - Data element definition
 - Data element business name
 - Data element systems abbreviation
 - Data type and data size
 - Data element source location



1/12/2006
(C) 2005 - 2006 Alabama Yankee Systems, LLC
4

● ● ● | Data Orientation Movement

- Traditional systems development focused on processes, not data as foundation
- Data orientation is a major cultural shift
 - change in methods used to design and deliver systems
 - change in perception, access and asset management of data and information



1/12/2006

(C) 2005 - 2006 Alabama Yankee Systems, LLC

● ● ● | Enterprise View of Data

- Data is a sharable resource
- Effective use and reuse of data requires an enterprise view of essential concepts, entities and attributes for cross-application and cross-departmental functionality
- Enterprise view **MUST** be generated at the top of an organization, with active commitment and support for efforts



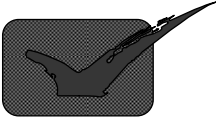
1/12/2006

(C) 2005 - 2006 Alabama Yankee Systems, LLC

6

● ● ● | **Goals of Metadata Strategy**

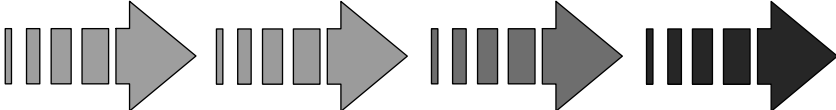
- Provide business orientation as foundation for sharing data assets across entire organization
- Offer recognition of value of data and its components
- Develop map for managing expanding information requirements
- Highlight importance of central data administration
- Address data quality, data integrity, data reuse
- Measurement of value of information



1/12/2006 (C) 2005 - 2006 Alabama Yankee Systems, LLC 7

● ● ● | **Strategic Plan for Information**

- Evaluate and choose a systems methodology
- Customize methodology for data orientation and business focus
- Develop and implement metadata strategy
- Integrate new methodology in systems development and maintenance



1/12/2006 (C) 2005 - 2006 Alabama Yankee Systems, LLC 8

• • • | Data Administration Objectives

- Build a framework for accurately capturing, sharing, distributing, securing and leveraging a company's data resources, including creating and refining data models and establishing and maintaining an enterprise data repository
- Strategically plan for a company's future information requirements
- Support data self-reliance and efficient use of data in business practices



1/12/2006

(C) 2005 - 2006 Alabama Yankee Systems, LLC

• • • | Principles of Metadata Management

- Establish a metadata strategy / policy
- Create or adopt a metadata standards set
- Establish and maintain data / metadata stewardship role and function
- Establish and maintain collection of relevant metadata
- Publish relevant metadata
- Maintain metadata standards set
- Maintain metadata content in all systems

1/12/2006

(C) 2005 - 2006 Alabama Yankee Systems, LLC

10



Keys to Metadata Management

- Develop metadata strategy before embarking on evaluating, purchasing and installing complex management products
- Products / tools are necessary for effective management of metadata
- Senior management commitment to metadata management is required - cost and time issues
- Explanation of WHY metadata is needed and the purpose of each type of metadata
- Policies to ensure acronyms and abbreviations are interoperable
- Procedures to ensure policies are implemented correctly

1/12/2006

(C) 2005 - 2006 Alabama Yankee Systems, LLC

11



Enterprise Metadata Strategy

- Offer recognition of value of data and its metadata (corporate assets)
- Establish information requirements and instill data management / metadata management practices into organization
- Address data quality, data reuse, data integrity issues for legacy applications and new development
- Promote enterprise use of data management tools and techniques
- Highlight importance of central data administration
- Measurement of value of information

1/12/2006

(C) 2005 - 2006 Alabama Yankee Systems, LLC

12

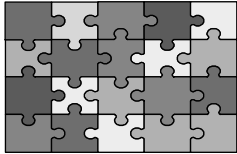
● ● ● | **Metadata Challenge**

- Broad: some metadata for every production application and structure
- Deep: detailed metadata for critical production applications and structures
- Decision points:
 - What groups have “pain”? Why does “pain” exist?
 - What processes can be leveraged for maintenance?
 - What metadata is readily available and will be useful?
- Who owns metadata in an organization?

1/12/2006 (C) 2005 - 2006 Alabama Yankee Systems, LLC 13

● ● ● | **Metadata Strategy Components**

- Metadata Responsibility and Usage
- Metadata Elements
- Metadata Sources
- Metadata Quality
- Metadata Storage and Products
- Metadata Standards and Procedures
- Metadata Training
- Metadata Measurement



1/12/2006 (C) 2005 - 2006 Alabama Yankee Systems, LLC 14

● ● ● | Kinds of Metadata

- Direct:
 - Byproduct of original source – information systems that create, maintain and dispense data
- Indirect:
 - Created after direct metadata – derived from system artifacts, uses of system, analysis and decision-making

1/12/2006

(C) 2005 - 2006 Alabama Yankee Systems, LLC

15

● ● ● | Metadata Committee

- Data Administration
 - Data Analysts
 - Repository Administrator
- Database Administration
- Data Stewards
- Rotating Application Development programmers



1/12/2006

(C) 2005 - 2006 Alabama Yankee Systems, LLC

16

● ● ● | Metadata Responsibilities

- Data Stewardship: responsible for consistency and integrity of data objects under their control; usually from functional business units
- Data Custodians: responsible for syntactical value of data; usually from Data Administration and/or Database Administration

Jointly establish standards and procedures for proper use, quality control and integration (“stewardship team”)

1/12/2006

(C) 2005 - 2006 Alabama Yankee Systems, LLC

17

● ● ● | Key Metadata Management Tasks

- Establish data owners, custodians and users
- Formalize data stewardship committee
- Define corporate data and process standards
- Define edit philosophy and user education philosophy
- Centralize information object identification and control processes and impact assessment

1/12/2006

(C) 2005 - 2006 Alabama Yankee Systems, LLC

18

● ● ● | Key Metadata Management Tasks

- Define basic enterprise information through an enterprise data model
- Develop application data and process models based on enterprise model - define application information
- Oversee application development and implement change control procedures
- Define data and metadata access processes and choose appropriate tools

1/12/2006

(C) 2005 - 2006 Alabama Yankee Systems, LLC

19

● ● ● | Metadata Responsibility and Use

- Data Ownership – syntactical value of data
- Data Stewardship – consistency and integrity of metadata
- Establish standards and procedures
- Proper use, quality control and measurement
- Integration with methodology and business

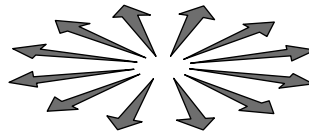
1/12/2006

(C) 2005 - 2006 Alabama Yankee Systems, LLC

20

● ● ● | Metadata Elements and Sources

- Business definitions and business names
- Systems definitions and systems names
- CASE models and element relationships
- Data element sources and targets (programs, files, databases, etc...)
- Methods for consolidating Metadata from multiple sources



1/12/2006

(C) 2005 - 2006 Alabama Yankee Systems, LLC

21

● ● ● | Quality: Imperative!

- Low-quality metadata creates:
 - Replicated dictionaries / repositories / metadata storage
 - Inconsistent metadata
 - Competing sources of metadata “truth”
- High quality metadata creates:
 - Confident, cross-organizational development
 - Consistent understanding of the values of the data resources
 - Metadata “knowledge” across the organization

1/12/2006

(C) 2005 - 2006 Alabama Yankee Systems, LLC

22

● ● ● | Metadata Quality - 1

- Data Accuracy: how well data values represent the actual business requirements
- Data Completeness: how well available data meets current and future business information demands
- Data Currency: how timely are data values
- Data Consistency
- Data Integrity: conformation of source data values to those allowed by business rules (data characteristics, default values, referential integrity constraints, derived data values)
- Data Source and Target

1/12/2006

(C) 2005 - 2006 Alabama Yankee Systems, LLC

23

● ● ● | Metadata Quality - 2

- Data Version: which data values best represent the actual state at a point in time
- Data Translation: how data values were converted between variations of the same data characteristics (for format and value variations)
- Data Integration: steps to reduce data redundancy and data variability (e.g., 10 versus 7 positions; data code translation)
- Data Summarization
- Data Recasting: primary key changes, content and meaning changes, missing periods of data

1/12/2006

(C) 2005 - 2006 Alabama Yankee Systems, LLC

24

● ● ● | Metadata Standards

- Established by Metadata Committee
- Naming standards
 - Abbreviations (systems and business)
 - Class words
 - Code values
 - Business names and definitions
- CASE modeling standards
- Entity and element creation/modification procedures
- Storage and accessibility of standards and procedures



1/12/2006

(C) 2005 - 2006 Alabama Yankee Systems, LLC

25

● ● ● | Choosing or Creating a Standard

- What schema should an organization use for metadata (i.e. how do we structure metadata)?
- Use a standard schema (good for interoperability) or develop one (may provide better support for local needs)
- Choice may require significant compromise – combination approach may be acceptable

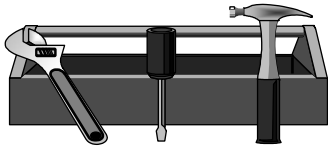
1/12/2006

(C) 2005 - 2006 Alabama Yankee Systems, LLC

26

● ● ● | **Metadata Storage and Products**


- Central versus distributed storage
- Active versus passive storage
- Physical requirements of storage facility
- Metadata products, capabilities and integration
 - Repository
 - CASE dictionary
 - Warehouse manager
 - Query tool data dictionary



1/12/2006 (C) 2005 - 2006 Alabama Yankee Systems, LLC 27

● ● ● | **Metadata Training**

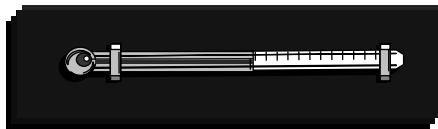
- Systems
 - Data Administration
 - Database Administration
 - Application Development and maintenance programming
 - Executive
- Business
 - Primary users (e.g., data stewards)
 - Secondary users
 - Executive



1/12/2006 (C) 2005 - 2006 Alabama Yankee Systems, LLC 28

● ● ● | Metadata Measurement

- Measure use and effectiveness of metadata
- Measure success of quality control program
- Measure acceptance and success of stewardship
- Integration of metadata strategy into systems methodology and business practices



1/12/2006

(C) 2005 - 2006 Alabama Yankee Systems, LLC

29

● ● ● | Cultural Change Required

- Data orientation and a focus on metadata require a conscious cultural change from individually developed applications to a unified acceptance and usage of data as the foundation of information and knowledge
- Change is unsettling to all creatures - expectations and reactions must be anticipated, addressed and resolved
- To be effective, cultural change must be managed in a spirit of co-operation and mutual accountability

1/12/2006

(C) 2005 - 2006 Alabama Yankee Systems, LLC

30

● ● ● | **Conclusions**

- Metadata Strategy can assist in achievement of data and information / intelligence goals
- Metadata Strategy should be part of a comprehensive customized systems methodology and developed in conjunction with the business community
- Establishment and implementation of metrics and a permanent committee for implementation are fundamental to the success of a Metadata Strategy



1/12/2006

(C) 2005 - 2006 Alabama Yankee Systems, LLC

31

● ● ● | **Questions and Answers**



1/12/2006

(C) 2005 - 2006 Alabama Yankee Systems, LLC

32

● ● ● | Metadata Reference Sources

- <http://dublincore.org/>
- <http://metadata-standards.org>
- <http://www.ukoln.ac.uk/metadata/>
- <http://www.ulb.ac.be/ceese/meta/meta.html>