

# Taxonomy and Functional Classification for RIM

Mile High Denver ARMA Chapter  
February 16, 2010

Bruce Schinkelwitz, CRM

## What this session is about

- Systems and Methods of organizing information and records
- Why is it important?
  - Organized information is easier to use and manage
  - Organizing information is the ONLY way to manage the 'Tsunami'
  - Records and Information Managers aren't trained on this
- We are going to:
  - Review and discuss what classification is
  - Experience how a classification schema goes together
  - Discuss specific challenges of RIM classification

## Core concepts

- Taxonomy:
  - *“a classification of things or concepts, as well as to the principles underlying such a classification”* (Wikipedia)
  - *“the rules or conventions of order or arrangement”* (Lambe)
- A taxonomy exists when: a set of information has a rule to its arrangement that is known or knowable (by its user group)
- Classification:
  - *A system of organizing things or concepts into groups or categories, or taxonomies, using established criteria*
- Taxonomy and Classification are all about WORDS and their MEANINGS in CONTEXT

## First Step: Let's Play Games

- Make a decision about what ONE 'type' of game it is.
- Some of you have Major game types (red Bold) on your card
  - Move to an open area in the meeting room
- 1. Each attendee: move to ONE group of 'like' game Types
- 2. Each group: Reflect on the suitability of the group makeup
  - Fill in index cards for alternate game 'types' within the group
- 3. How does this change the overall organization of games in the room?
- 4. Observations?

## Core concepts - 2

- Taxonomy and Classification
  - Are tools for understanding linkage of facts and ideas
  - Are Semantic
    - There is an inherent meaning to arrangement
  - Are developed out of necessity
    - Too much info to grasp and remember
  - Are inherently contextual
    - Meaning relates to the specific context
  - Are group-based
    - Same terms can be arranged differently by groups
    - But there is remarkable similarity between groups

## Core concepts - 3

- The 'Orders' of Order\*
  1. The thing itself - e.g., a book
  2. The information about the thing - e.g., a card catalog
  3. The electronic world - bits of data about the book; the book's metadata
- Orders 1 & 2: Single, 'physical' organization
- Order 3: Independent of the physical
  - Arrangement is presented based on the specific metadata

\*Weinberger: p. 18-19

## Records Classification Primer - 1

- A Classification is a system of organizing information using rules
- A Records Classification System (CS) has three objectives:
  - Provide a structure that guides how to organize records
  - Provide a means to easily find and use records
  - Provide a means to manage the records over time
- Characteristics of a Good CS:
  - Knowable
  - Consistent
  - Predictable
  - Flexible
  - As the CS grows, it incorporates a number of taxonomy tools to achieve goals

## Records Classification Primer - 2

- A CS uses terms (words) to represent concepts or groups (chunks) of information
- The terms used in a CS are critical
  - Their meaning has to be clear to all users
  - The relationship between terms also has meaning
- Every classification system has a structure
  - Most are a form of a hierarchy
    - Parent-child, broader-narrower, etc.
- The structure must reflect the relationship of terms that makes sense for business context

## Functional Classification

- Focuses classification on the business context to organize records
  - Theory: records are a natural result of business activity/process
  - ISO 15489, DIRKS
- Organization by Business Function, then Business Activity/Process
  - Hierarchy of
    - FUNCTION -> ACTIVITY -> (SUB-ACTIVITY) -> TRANSACTION (records)
- Why?
  - Stability of FUNCTIONS and Activities (unlike Org Chart)
  - Connects activity to information flows
  - Allows users to focus on their business process while supporting RIM goals
  - Provides a means to understand the bigger record context

## Functional Classification - 2

- **FUNCTION:** the top level of the CS
  - Equates to the highest level of business responsibility/mission that doesn't have a logical higher level
    - Examples: Administration, Financial Management, Human Resources, Legal, etc.
- **ACTIVITY:** the second level of the CS
  - The business process that is part of a higher FUNCTION
    - Examples: Accounting, Recruiting, Investigations, Program Management, Policies, Training
- **SUB-ACTIVITY:** the third level of the CS
  - The sub-process within an Activity
    - Examples: Accounts Receivable, Investigation Case File, etc.
- There can be lower levels (of sub-activity) - most go too deep

## Exercise 1: The 'fun' in Functional Classification

- Each participant has:
  - A card with a Function, Activity or Subject term in Black on it (*ignore the reverse side for now*)
    - FUNCTION terms are ALL CAPS
- 1. Those with FUNCTION cards move to a clear spot in the room
- 2. Those with Activity or Subject Terms,
  1. Determine (alone) which FUNCTION your term belongs to
  2. Stand with the appropriate FUNCTION group
- 3. What can we observe?

## Observations

- It is easy to identify relationships, BUT not when
  - Parts are missing
  - Relationship Context is not made clear
- It is easy to find alternative arrangements of a common term
  - We instinctively think that the term must fit in one place
- The cards don't impart much information
- The physical world imposes a strict reality:
  - A thing can be only in one place, and one relationship
- What about the electronic world?

## Observations - 2

- Electronic:
  - Classification is almost always Alpha-Numeric
  - Classification based NOT on inherent 'meaning', but by a METADATA VALUE (title=)
  - Classification typically capable of only a few pieces of data, just like cards (Name, Date Modified, Type, Size, etc.)
  - Technology typically blocks multiple use of a term (Name value)
- These are weaknesses in the TECHNOLOGY
  - Alpha-numeric is absolutely non-contextual
- Lessons:
  - The technology determines the level of capability of the Classification System
  - Electronic classification depends upon METADATA

## A Very Brief Primer on Metadata

- Metadata is the information we and computer systems use to understand and use data
  - E.g., Title, Date Modified, Unique ID
- In our exercises, we only have two metadata values:
  - Term Name = single value
  - Term Relationship = single-direction value (parent of...)
- BUT: a thing can have many different relationships
  - A book can be about many subjects (See [Map](#))
- Opportunity: use metadata to provide ways of arranging, viewing, enhancing a classification
- Metadata can also reduce number of classification levels

## Observations - 3

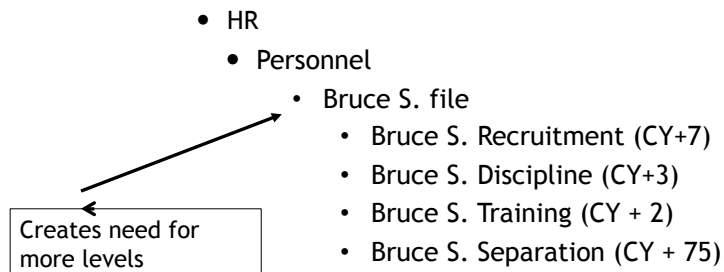
- A Simple Hierarchy (one tree branch) can't work for a large classification set
- The availability of advanced technology (ECM, relational DB) allows for a 'Poly-Hierarchy'
  - Terms can be repeated where they have a different context
- EXAMPLE

## The Challenge of Retention

- Review: Functional Classifications arrange relationships of functions and activities and records
  - The structure's RULE= Arrangement by Business Activity
- The RIM problem: We manage records rules, not activities
  - A retention is not an activity, but a RULE applied against the records involved in the activity
- Typical RIM approach: Classification = Retention series
  - Result:
    - RIM Classifications are hard to use, 'alien' to end users \*\*\*
    - Lose the records context in the arbitrary buckets
- Solution?: Incorporate retention as a piece of metadata associated with activity

## Classifying for Retention

- Retention is applied on folders, with one disposal rule per folder
  - Forces the creation of separate folders per retention rule:



## Classifying for Retention

- Functional classification allows reduction in levels and easier filing by users
- Retention rule becomes a metadata value, not a classification level
- Example:
  - HUMAN RESOURCES (FUNCTION)
    - Personnel Management (Activity)
      - Recruitment (Sub-Activity) - CY+7
        - Sub-folders by subject, case file, CY, etc.
      - Disciplinary Actions (Sub-Activity) -CY+3
      - Training (Sub-Activity) - CY+2
      - Separation (Sub-Activity)- CY+75

## Impact of Technology

- The technology you use to manage records governs what you are able to do with classification
- Reality: Content is managed in many technologies
- File Shares: Folder driven, little metadata
  - Structuring still possible
- ECM: Folder driven (many), metadata rich
  - Allow dynamic and flexible classification outside of limiting hierarchy
- Important: Manage classification system separate from the content repositories, using relational DB technology. WHY?
  - ECM = content focus; CS = structure & relationship focus
  - Supports 'Hub and Spoke' management

## Effective Classification Steps

- Start with a process analysis for a single business area
- Identify the end-user vocabulary for the process
- Build a simple hierarchical taxonomy based on the above
- Identify linkages between taxonomy and external processes, rules/laws, records requirements
- 'Grow the Onion' - Build outward from starting process
- Do not let The Perfect be the enemy of The Good

**Thank You!**

Bruce Schinkelwitz, CRM  
bruceschinkelwitz@me.com