ARMA/Liberty Bell
Presentation – September 27, 2018

TAXONOMIES 101

GEORGE D. DARNELL, CRM

HUMANS NEED TO NAME THINGS

- Genesis 2:20 King James Version (KJV)

20 And Adam gave names to all cattle, and to the fowl of the air, and to every beast of the field;

AND THEN THEY NEED TO ORGANIZE THEM

- Animal
- Vegetable
- Mineral

GUESS WHAT?

- That's a Taxonomy.
- But what does that have to do with Records and Information Management?
- Let's do a quick review of filing.

CLASSIFICATION OR CATEGORIZATION AS A FORM OF TAXONOMY

- Process of putting like things, e.g., records of similar subject or category together—dates to Greek Classical Period.
- Two common schemes for filing:
  - Case or project files organized alphabetically or numerically, e.g., Claims or Contracts
  - Subject file organized alphabetically or numerically, e.g., Travel or Training

CASE AND SUBJECT FILES

- Documents in a Case File are kept in chronological order and the file is closed and held for reference for a period of time until some event occurs, e.g., claim is settled or project ends.
- Documents in a Subject File are kept in chronological order and the file is closed at the end of the calendar or fiscal year and disposed of after a period of time.

SO, WHAT IS A TAXONOMY?

- Taxonomy:
  - Simply the practice and science of classification of things or concepts, including the principles that underlie such classification.
- Origin:
  - Borrowed from science (biological).
    - See Entrez Taxonomy
    - Or Bloom's Taxonomy.

TYPES OF TAXONOMIES

- Flat – single level
- Hierarchical – multiple levels (Most Common)
- Faceted – linked to other taxonomies
- Networked – complex relationships among taxonomies

OTHER CLASSIFICATION TERMS

- Controlled Vocabulary
  - A collection of preferred terms that indicates which terms are preferred and which are variants of the preferred terms.
  - Often used with taxonomies to find correct term.
  - Most common example is yellow pages.

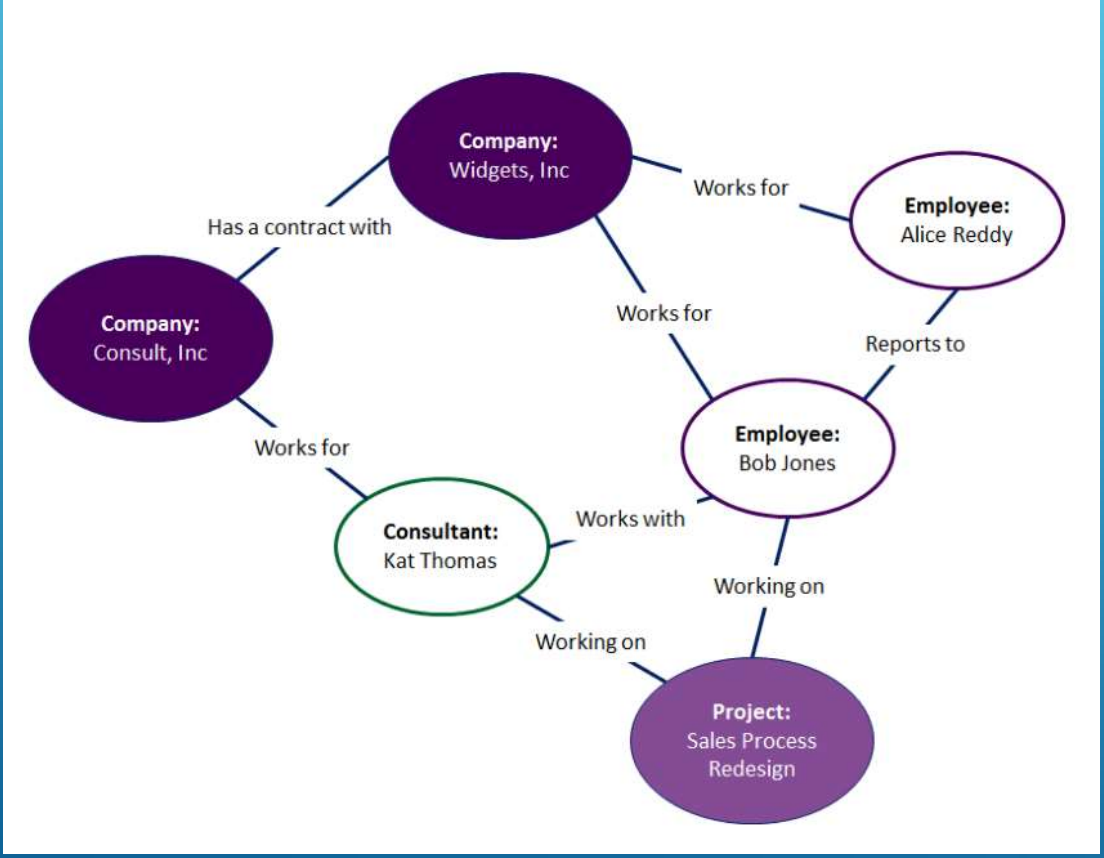
OTHER CLASSIFICATION TERMS(CONTD.)

- Thesaurus
  - A type of controlled vocabulary that shows the hierarchical (parent-child), associative (related terms) and equivalent (synonymous) relationships among terms.
  - Useful in building taxonomies to ensure relationship among terms at different levels is correct.
  - Example: Mobile Phone, Cell Phone, Cellular Phone, Personal Phone

OTHER CLASSIFICATION TERMS (CONTD.)

- Ontology
  - A hierarchical classification that is more complex and subtle than a taxonomy. It explains relationships between objects by mapping relationships, such as "part of" or "located in". Also called knowledge mapping.
  - Used by data miners to track relationships among entities.

ONTOLOGY EXAMPLE



EXAMPLES OF TAXONOMIES

- in Education
  - http://www.bloomstaxonomy.org/
- In Science
  - http://www.ucmp.berkeley.edu/help/taxaform.html
- On the web
  - USA.gov
  - Yahoo.com
- The Yellow Pages
  - Yellow Pages

FILE PLANS AS TAXONOMIES

- Before Taxonomies there were File Plans.
- File Plan Objectives:
  - User information retrieval
  - Record set integrity (Provenance)
  - Uniform classification system

FILE PLAN ARCHITECTURE

- Major Headings (7 to 10)
  - Use major business functions, e.g.,
    - Administration
    - Engineering
    - Marketing
    - Operations
    - Policy and Plans
    - Research and Development

FILE PLAN ARCHITECTURE (CONTD.)

- Divide Major Headings into Primaries, e.g.,
  - Administration
    - Budget & Finance
    - Human Resources
    - Logistics
    - Procurement
    - Security

FILE PLAN ARCHITECTURE (CONTD.)

- Primaries into Secondary & Tertiary, e.g.,
  - Administration
    - Budget & Finance
      - Accounts Payable
        - (Files arranged by vendor or date range)
      - Accounts Receivable
        - (Files arranged by customer or date range)
      - Budget, Corporate
        - (Files arranged by fiscal year)
      - Budget, Departmental
        - (Files arranged by department)
      - General Ledger
        - (Files arranged account or year)
      - Payroll, etc.

DEWEY DECIMAL EXAMPLE
700 THE ARTS

730 Sculpture and plastic arts

- 731 Processes of sculpture
  - 731.4 Methods and Techniques of sculpting
    - 731.44 Casting
      - 731.441 Plaster and cement casting
      - 731.442 Castings and synthetic plastic

BUILDING A TAXONOMY

1. Start with list of standard terms – "controlled vocabulary"
2. Add relations among terms "thesaurus"
   - Xref non-standard terms (mobile phone example)
   - Point to related terms (e.g., indexing →taxonomy)

BUILDING A TAXONOMY

3. Connect files to taxonomy terms.
   - Indexers build taxonomies to terms and subjects in a publication (alphabetical index).
   - Catalogers build taxonomies to publications in a library (subject index in card catalog).
   - Records Managers build taxonomies to records (file series for retention schedules).

BUILDING A TAXONOMY

4. Train users to put files in appropriate place in taxonomy.
   - May not be intuitive.
   - Some round pegs will fit in square holes.
5. Continue maintenance of taxonomy structure and repository files.
6. Provide for legacy ingestion.

ORGANIZATIONAL VS. FUNCTIONAL TAXONOMIES

- Organizational
  - Tied to structure of organization chart
  - Not suited to enterprises that re-organize frequently
- Functional
  - Tied to business functions
  - More insulated from reorganizations

TAXONOMY DO'S AND DON'TS.

- DO
  - Inventory files before organizing them, i.e., know your content
  - Use terms your users understand, i.e., know your users
  - Distribute files evenly across taxonomy
  - Use uniform depth (4 is about max)
  - Have a place for everything (and usually only one place)
  - Make taxonomy intuitive and easy to remember

TAXONOMY DO'S AND DON'TS (CONTD.)

- DO
  - Have content at every level
  - Allow for expansion
  - Balance breadth and depth
    - 3 wide & 5 deep (animal, vegetable, mineral) is harder to use than
    - 5 wide & 3 deep
  - Review periodically
  - Use configuration management/change control

TAXONOMY DO'S AND DON'TS (CONTD.)

- DON'T
  - Have empty categories
  - Use "Miscellaneous" as a category
  - Use more than four levels
  - Use the same term at different levels, e.g., Administrative
  - Permit uncontrolled changes

AMBIGUITY EXAMPLE

- Home budget using Quicken or Money
  - Create a category for new auto loan
    - Under Loans (new sub category) or
    - Under Auto (new sub category)?

AUTO LOAN EXAMPLE

- Auto
  - Fuel
  - Insurance
  - Service
  - Registration
- ’’’
- Loans
  - Mortgage
  - Personal
  - Education

RETRIEVING RECORDS: TAXONOMY VS. FULL TEXT

- Taxonomy
  - Limited set of search terms
  - Browsing enabled
  - Discovery of like records
  - "Authority" based categorization using standard terms
- Full Text
  - Unlimited set of search terms
  - Only terms used in document found
  - Words with multiple meanings cause false hits, e.g., "stock"
  - Depends on author to use correct terms
  - NIST Study – only 20% found

THE ROLE OF METADATA IN ORGANIZING FILES

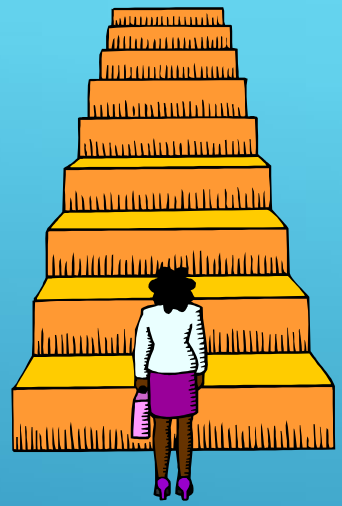
- Metadata = data about data
  - Provides information about information (content) objects, e.g., title, author, publication date, recording media, subject matter.
  - Two familiar examples are
    - Dublin Core Standard Metadata (15 items)
    - Microsoft File Properties (not fully used by most of us)
  - 7-10 Values is optimum

THE ROLE OF METADATA IN ORGANIZING FILES

- Including Metadata in information objects or in a separate index enables users to limit searches to
  - Files within a date range
  - Files with a specific title
  - Files by a specific author
  - Etc.

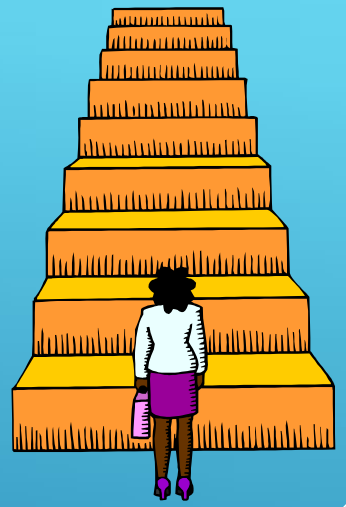
10 STEPS TO TAXONOMY PROJECT SUCCESS

1. Define scope and business objective
2. Perform information audit (inventory)
   - What exists? – Where is it? – How used?
3. Draft high-level taxonomy architecture
   - Organize around major functions
   - Research external sources
   - Reconcile terms into a controlled vocabulary
4. Develop proof-of-concept taxonomy
   - Involve stakeholder in testing and feedback



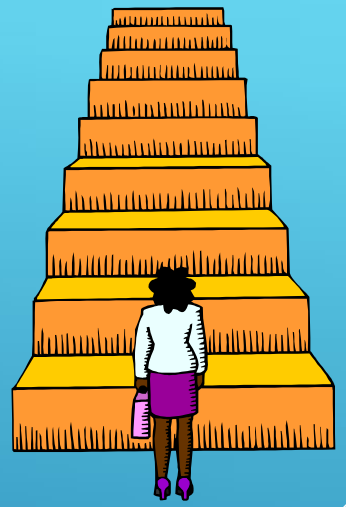
10 STEPS TO TAXONOMY PROJECT SUCCESS

5. Involve stakeholders – agree on success factors
6. Keep project history and use configuration management
7. Add detail to taxonomy
   - Establish rules for applying terminology
   - Establish metadata standard
8. Know when to quit



10 STEPS TO TAXONOMY PROJECT SUCCESS

9. Establish maintenance and governance process
   - Include stakeholders to make taxonomy as dynamic as necessary
10. Evaluate technology and assess vendors
   - Know what you need before you select
   - Continue to manage the project – don't abdicate to vendor



RESOURCES

- https://ontopia.net/topicmaps/materials/tao.html
- http://www.loc.gov
- http://argus-acia.com/index.html
- WWW Virtual Library
- http://www.kmworld.com/Articles/Editorial/What-Is/Taxonomy-101-The-Basics-and-Getting-Started-with-Taxonomies-98787.aspx

QUESTIONS?