# the care and feeding of taxonomies

## taking an enterprise approach

Jane Stewart
jane@aboutness.ca
aboutness.ca

Aboutness

## Introduction

- Librarian
- Web information architect
  - First Polar Bear book – 1998
- Public servant
  - GOL era
  - Content management / information management
- Consultant / IA practitioner / taxonomist
  - Primarily web projects & web IM (and just plain M) challenges
  - Increasingly... enterprise IM and the bigger picture

2

## In the next 50 minutes...

- Fisher-Price history of knowledge organization
- What "taxonomy" means now
- Why taxonomy management matters
- Schools of thought (architecture/design)
- Taking an enterprise approach
- The future

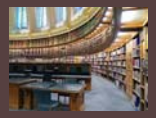
3

## Raise your hand

- Taxonomist
- Enterprise Taxonomy Specialist
- Encoding Scheme Manager
- Knowledge Organization System Specialist
- Ontologist
- Chief Semantic Officer

4

## 15-second history of classification



- First order
  - Arrangement of physical objects
- Second order
  - Experts organize knowledge
  - Catalogues, lists, folders...metadata!
- Third order – digital order
  - Infinite ways of describing things
  - Ability to link related things
  - For people
  - For computers
    - Advent of the Semantic Web

5

## Third order preoccupations

- Search / findability
  - Now
  - In the future
- Knowledge management
  - Knowing what we know
  - Synthesizing what we know
- Business process management
  - Efficiency
  - Reduction of duplication
  - Digital recordkeeping
- Integration & interoperability
  - Within the enterprise
  - Between enterprises

Taxonomies support all of the above

6

## What "taxonomy" means

- Then
  - A single, hierarchical term structure
    - Broader / narrower concepts
    - Lumping and splitting
  - Specific to a domain or discipline
  - Often concerned with "subject classification" (aboutness)
- Now
  - Looser: more about the purpose than the shape
  - Order and arrangement of knowledge
  - Not just subjects or topics
    - "Proper noun" entities (people, organizations, places)
    - Functions/activities
    - Document types
  - Describes what a resource "is" as well as what it's "about"

7

## How are taxonomies used?

- Information retrieval
- Information management
- Search systems
- Navigation
- Picklists / autocomplete suggestions
- Recommendation systems
- Data integration
- Multilingual systems

8

## Why taxonomy management matters

- Mitigates against the fragmentation and chaos caused by defining concepts one application / function at a time
- Contributes to "sense-making" in the enterprise; builds coherence
- Unlocks meaning and value in information assets
- Supports the goal of delivering the right information, in context, for the people who need it

9

"The solution
to the overabundance of information is
more information."

David Weinberger, *Everything is Miscellaneous*

10

## Designing taxonomies

- In the enterprise, taxonomies are a component of information architecture (IA)
- IA schools of thought
  - "User experience" (UX)
    - Concerned with people & behaviour
    - Support processes and tasks ("taskonomy")
    - Specific contexts of use
  - Enterprise IA
    - Holistic and multi-layered
      - Conceptual, logical, physical
    - Concerned with assets and their nature
      - Structured & unstructured
    - Support integration and interoperability

11

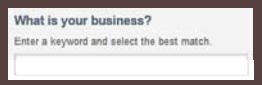
## UX taxonomy design considerations

- Mental models
  - How people think things work
- Tasks
  - What people are trying to do (business processes)
- Attention & focus
- Memory
- Making choices
  - Differentiation / ambiguity of language
    - Homonyms: Mercury = planet / element / chemical / automobile
    - Synonyms: automobile = car = motor vehicle
  - User language
    - "Boat" vs. "pleasure craft"

12

## Example: BizPaL Sector Selection

- Q: What is your business?
- Before
  - Pick list of 700+ NAICS codes
  - Formal, not everyday language
    - Example: "Mobile Food Vending"
- After
  - Single search box
  - Type-ahead matching
  - Informal, everyday user language mapped to formal codes: "chip wagon"




13

## Enterprise IA & taxonomy

- Enterprise IA *"links an enterprise's information assets to the business processes that need them and the IT systems that use and management them."*
- Includes technical & application architecture
- Metadata and taxonomy support both underlying structure and context-specific uses

Gene Leganza, Forrester Research, *Topic Overview: Information Architecture*

14

## We ❤ facets

- Enterprise taxonomies are typically "faceted"
- Facets have emerged as an effective third order solution, based on a second order system designed in 1933
- Not "where do we want to put it?" but "how might we want to find it and use it?"
- Resource-centric

Dr. S. R. Ranganathan

15

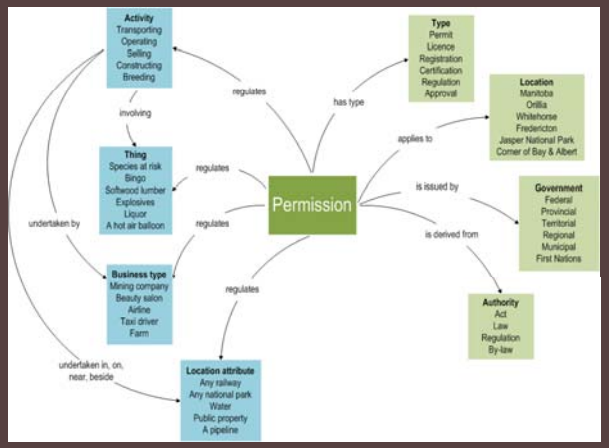
## Facets are powerful

"Four facets of 10 nodes each have the same discriminatory power as one taxonomy of 10,000 nodes."

Joseph Busch's *Golden Law of Facets*

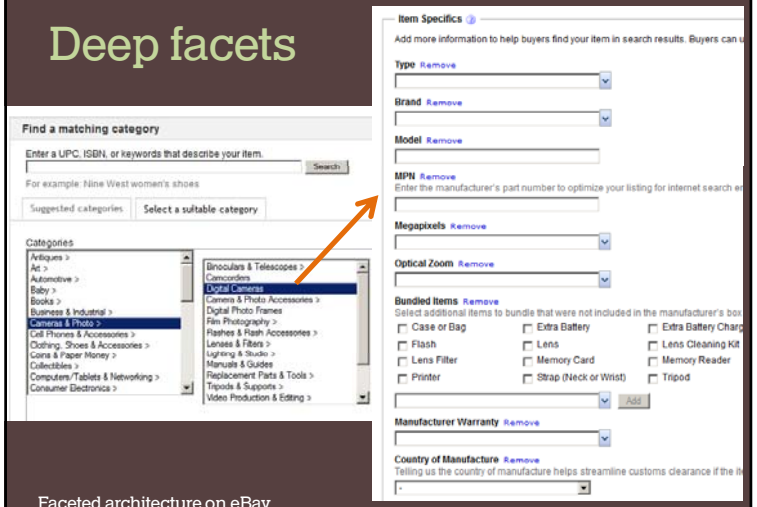
16

# Facet analysis






# Deep facets




Faceted architecture on eBay



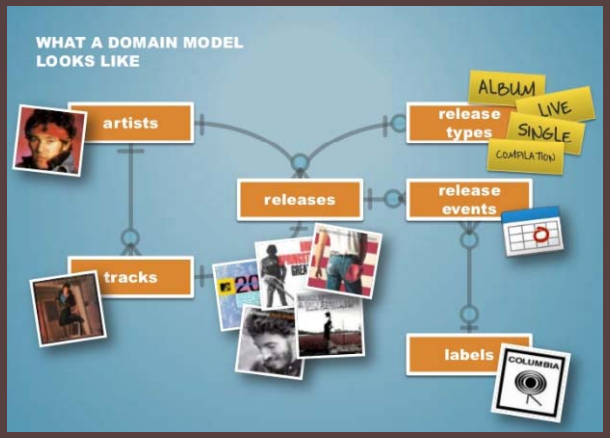
# Faceted enterprise taxonomy


- Each facet is supported by a controlled vocabulary
  - Lists
  - Hierarchical taxonomies
  - Thesauri



# The future of IA

- New kinds of models
  - Domain models
  - Ontologies
- Define relationships *between* things
  - Facets = classes
  - Many types of relationships beyond the traditional controlled vocabulary types:
    - Hierarchical
    - Associative
    - Equivalent

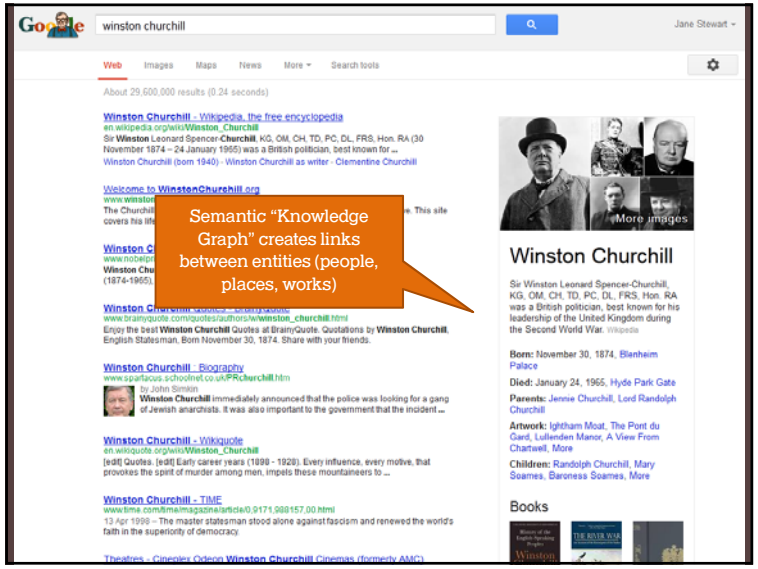



Mike Atherton, *Beyond the Polar Bear (IA Summit 2011 presentation)*

21

# Domain model for government


22

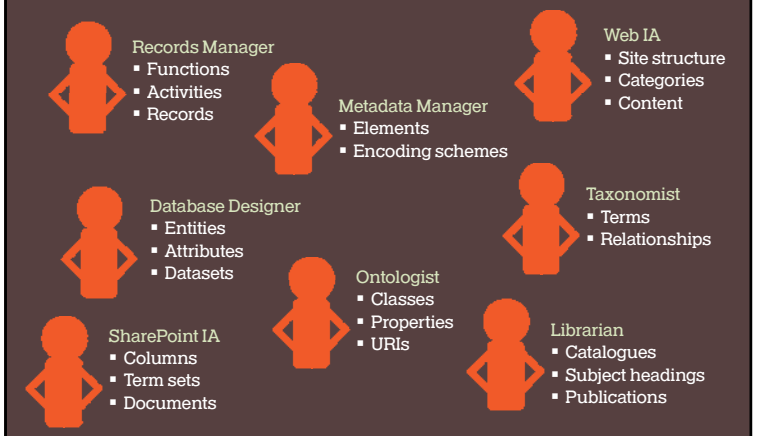



# Taking an enterprise approach

- A responsibility centre for managing taxonomies and controlled vocabularies on behalf of the enterprise
- Challenges
  - Building shared meaning
  - Scope of effort
  - Services
  - Governance
  - Resources
    - Skills
    - Tools

24

## Whole lotta semantics going on




25

## Scope

- What's in and what's out?
- Core vs. non-core
  - Enterprise-wide vs. function-specific
- Semantic archaeological dig
  - What's actually in use now?
  - Embedded in what systems?
    - EDRMS, datasets, Web CMS, other applications?
  - Are there overlaps?
  - What shape are they in?
    - Maintained or not? Effective or not?

26

## Where do they come from?

- Pre-existing / external
  - GC, standards bodies, other orgs
  - Accessed programmatically, or embedded?
  - Used as is, or adapted?
- Internal
  - Design method used?
  - Ownership, maintenance processes?
  - Consistency, quality of application?

27

## Governance

- Defining how taxonomy management fits with:
  - Enterprise information architecture
  - Data management
  - Metadata management
  - Recordkeeping
  - Libraries
  - Web publishing

28

## Governance cont'd

- Core Taxonomy Management
  - Enterprise-wide vocabularies are centrally designed, managed and controlled
  - Taxonomy management function performs maintenance, monitoring
  - Supports application and use
- Non-core Taxonomy Management
  - Central function is *aware* of all vocabs in use, but does not manage
  - Ability (and authority) to say things like
    - There is a pre-existing vocabulary for this concept
    - This is our preferred vocabulary for this concept
  - Use "stewardship" model for taxonomy management
  - Provide guidance and support

29

## Taxonomy management services

- Design and develop
- Monitor and maintain
- Publish and integrate
- Advise and support

## Resources - Skills

- Communication, consultation and negotiation
- Taxonomy and classification theory and standards
- Enterprise terminology and/or subject expertise
- UX IA methods and techniques
- EIA principles and practices
- Taxonomy management software and capabilities
- Enterprise information management systems

31

## Resources - Tools

- Excel

Dedicated taxonomy management tools

- Low-end tools can help manage lists, taxonomies and thesauri
  - Standard relationships (Z39.19)
    - Broader/narrower (BT/NT)
    - Equivalent (USE/UF)
    - Related (RT)
  - Flexible output formats

32

## High-end tools

- Support domain modelling and ontology development
  - Define and manage custom relationships and properties
- Support governance of semantics
  - Workflow, versioning
- Apply semantics to resources
  - Auto-classification of large document sets
- Support integration
  - Programmatic access and integration with enterprise systems

33

## Embrace iteration

"It is better to do something and tweak it for the rest of your life than to get 30 people into a room to figure out everything you're ever going to need."

Tim Falconer

## Conclusion

- Look for the big picture
- Semantics will become more important over time
- Treat your taxonomies and controlled vocabularies as enterprise assets, not system-specific assets
- The third order is still very young: we are all pioneers

35

## Thank you!

Jane Stewart
jane@aboutness.ca

Aboutness

36