Controlled Vocabularies Need Context

Taxonomies to Ontology

By Lynda W. Moulton

LWM Technology Services

Context for Case Studies

- Definitions

- How taxonomy relates to knowledge management

- Types of taxonomy and other categorization systems

- Relationship of taxonomy and search

Definitions:

- Categorization
- Classification
- Content
- Controlled Vocabulary
- Cross -reference
- Index (Search)
- Metadata Elements
- Ontology
- Taxonomy
- Thesaurus
- Validation List

Categorization

Process of grouping materials into one or more classes or topical areas

e.g. Products, Professions, Specialties

Classification

A structured and reasoned system of organizing materials according to their <u>single</u> strongest attribute

e. g. Dewey Decimal system, Yellow Pages

Content

The informational matter in a collection of materials

e.g. Training materials, Product documentation, Technical reports

Controlled Vocabulary

Authorized and standardized list of terminology used to define the content of one attribute of materials in a collection. A collection may be defined by more than one controlled vocabulary

e.g. Subject, Corporate names

Cross -reference

In a controlled vocabulary list, a directive from one term in the list to use another

e.g. Ibuprofen **USE** NSAID, H_2O **USE** Water

Index (Search)

A finding device. A set of information that directs the user to the object of the listing. In a term index, each entry points to one or more entries by a virtual or physical location

e.g. Back of the book index, Author listing, Product catalog

Metadata Elements

Structured categories defined to contain information about one aspect or attribute (e.g. publisher, subjects) of an information resource (e.g. book, document, image)

e.g. Fields in a content management system

Ontology

Structural specification for expressing complex relationships among concepts

e.g. Bill of materials, Process flow diagram, Periodic table
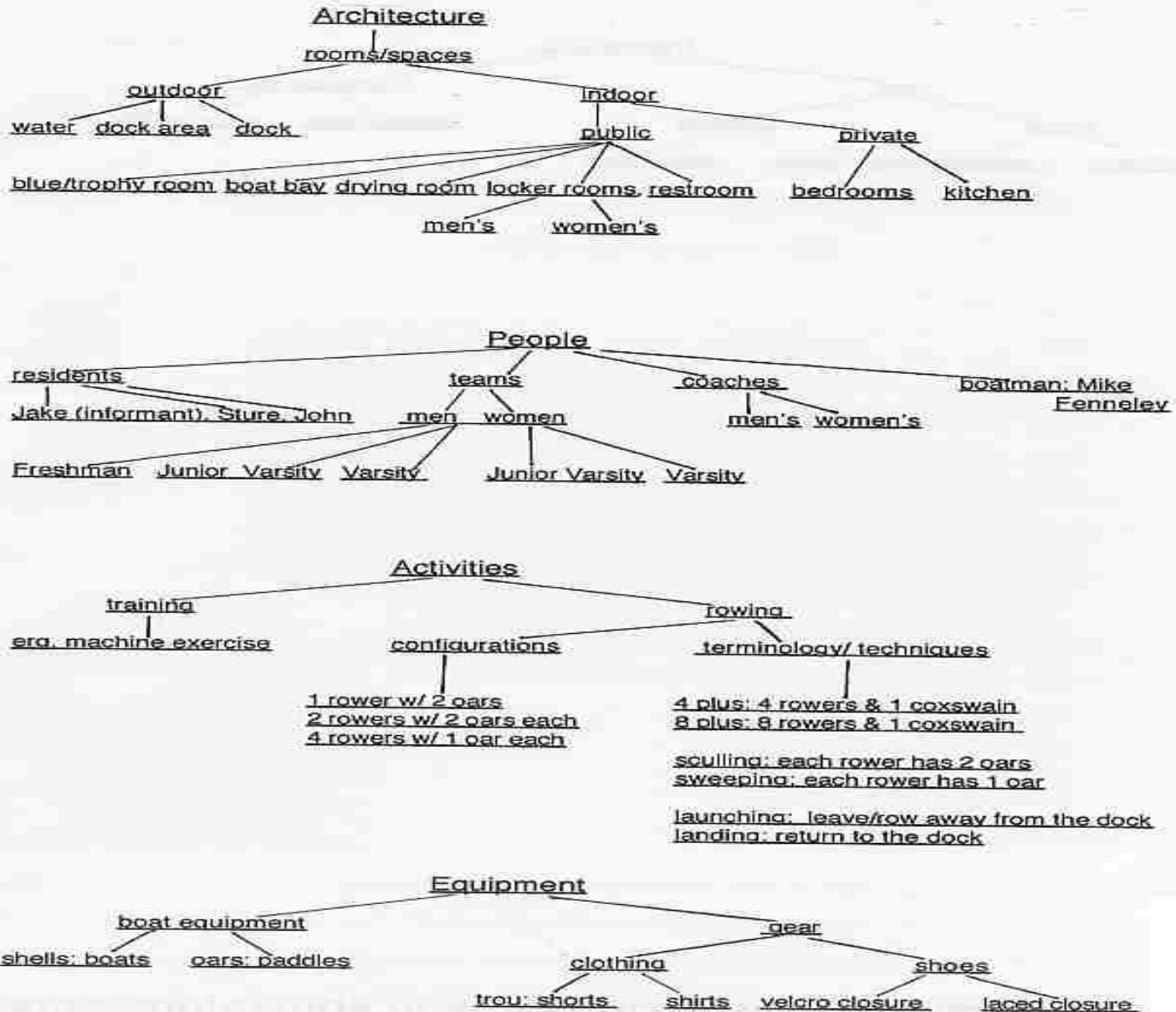
Taxonomy

Hierarchy of terms for classifying one attribute of information resources (e.g. subjects, names). A type of controlled vocabulary with a schematic layout. May reflect facets

e.g. Topics, People, Places

Taxonomy of Jake Wetzel's cultural knowledge of the behavior setting:

Architecture

rooms/spaces

- outdoor
  - water
  - dock area
  - dock
- indoor
  - public
    - blue/trophy room
    - boat bay
    - drying room
    - locker rooms
      - men's
      - women's
    - restroom
  - private
    - bedrooms
    - kitchen

People

- residents
  - Jake (informant), Sture, John
- teams
  - men
    - Freshman
    - Junior Varsity
    - Varsity
  - women
    - Junior Varsity
    - Varsity
- coaches
  - men's
  - women's
- boatman: Mike Fenneley

Activities

- training
  - erg. machine exercise
  - configurations
    - 1 rower w/ 2 oars
    - 2 rowers w/ 2 oars each
    - 4 rowers w/ 1 oar each
- rowing
  - terminology/ techniques
    - 4 plus: 4 rowers & 1 coxswain
    - 8 plus: 8 rowers & 1 coxswain
    - sculling: each rower has 2 oars
    - sweeping: each rower has 1 oar
    - launching: leave/row away from the dock
    - landing: return to the dock

Equipment

- boat equipment
  - shells: boats
  - oars: paddles
- gear
  - clothing
    - trou: shorts
    - shirts
  - shoes
    - velcro closure
    - laced closure

13

Operational Regimes

Operational Units

Responsive Deployable
Agile and Versatile
Lethal Survivable
Sustainable

System of Systems Platforms

Ground	**Air**	**Sea**	**Warrior**	**Other**
Manned & Unmanned Mobile & Non-Mobile	Manned & Unmanned Aircraft	Ships, Submarines	OFW, Land Warrior Marines	UGS, Munitions

Platform Applications

Hybrid Electric Propulsion	Environment Management	Dynamic Armor	EM & ETC Guns	Active Protection	C4 ISR	Signature Management	Directed Energy Weapons

Technologies

Switches : Capacitors : Batteries : Power
Converters : Fuel Cells : Fuel Reformation
Thermal Management : Power Control:
Power Generation

LWM Technology Services   © 11/2004

14

Thesaurus

Hierarchically structured controlled vocabulary of terms that are used to describe information resources. A more comprehensive form of taxonomy with deeper relationships and cross-references. Does not assume a graphic layout

e.g. MeSH (Medical Subject Headings, PsychInfo)

Subject: Creativity

USED FOR: Innovativeness, Originality

To search for the subject, or one or more term
Click on a term to see more information.

Subject

☐ Creativity

Less Specific (broader) terms

☐ Personality Traits

Related terms

☐ Ability (+)

☐ Artistic Ability (+)

☐ Divergent Thinking

☐ Gifted

Subject: Cognitive Style

USED FOR: Learning Style

To search for the subject, or one or mo
Click on a term to see more informatior

Subject

☐ Cognitive Style

More Specific (narrower) terms

☐ Cognitive Complexity

☐ Conceptual Tempo

☐ Field Dependence

☐ Impulsiveness

☐ Reflectiveness

Thesaurus Rex - ☺

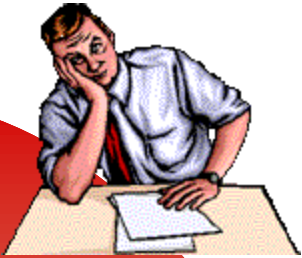**http://msowww.anu.edu.au/library/thesaurus/trexlarge.gif**

Validation List

Any controlled vocabulary that a computer algorithm uses to verify acceptability in a database field

e.g. Product names, Project codes, State codes

Still wondering how taxonomy relates to knowledge management?

- Textual <u>maps</u> to information content

- Content (information) represents an aspect of knowledge - <u>*knowledge tokens*</u>

- Helps us <u>visualize</u> topical relationships when we search for those knowledge tokens

Types of taxonomy and other categorization systems

- Simple lists of terms in a hierarchical format

- Thesaurus

- Ontological relationships

- Statistical organization - chart

- Types (media) of resources

- Location

- Genealogical

- etc.

Relationship of taxonomy and search

- *Taxonomy* is a type of controlled vocabulary
- When searchable *content is categorized* a type of taxonomy (term list, thesaurus or ontology) is employed as the organizing entity
- *Searching by categories* uses a map of a searchable index often showing relationships among terms
- The goal is to *orient and simplify search* for those who don't know where to begin
- It is not rocket science but reading is required

Thoughts for Developers and Designers of Knowledge Bases

- There are numerous ways to organize content - **context for audience is imperative**

- Technology tools can help automate the indexing and search process but **human beings need to design** or validate the systems of organizing information

- Taxonomies are organic and **must evolve** within the context of use

- Automated **tools for building and maintaining complex relationships** are in their infancy

Thank you for Listening

Questions?

lwmtech@lwmtechnology.com