

# APPLICATION PROFILES

[based on FRSAD model]

# FOR SUBJECT DOMAINS

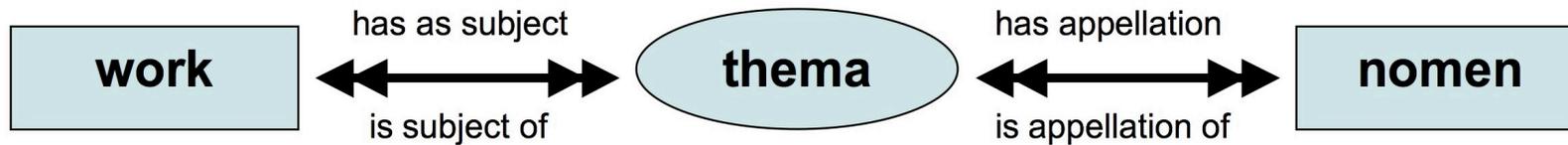
Marcia Zeng, Gordon Dunsire, Maja Zumer

# Questions to be discussed



1. Why APs are needed for subject authority data?
2. How formally (or informally) can this style of “application profile” be defined?
3. In what ways are application profiles for **subject domains** different from APs for **descriptive metadata**?

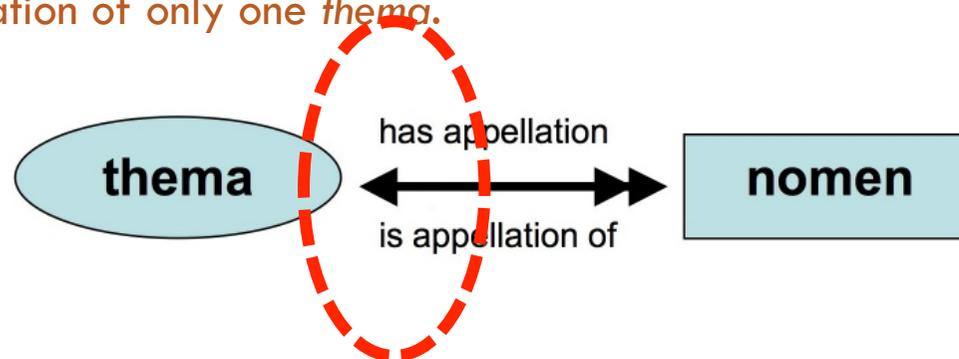
# FRSAD Conceptual Model



*Thema* = "any entity used as a subject of a *work*".

*NOMEN* = any sign or sequence of signs (alphanumeric characters, symbols, sound, etc.) that a *thema* is known by, referred to or addressed as.

Note: in a given controlled vocabulary and within a domain, a *nomen* should be an appellation of only one *thema*.

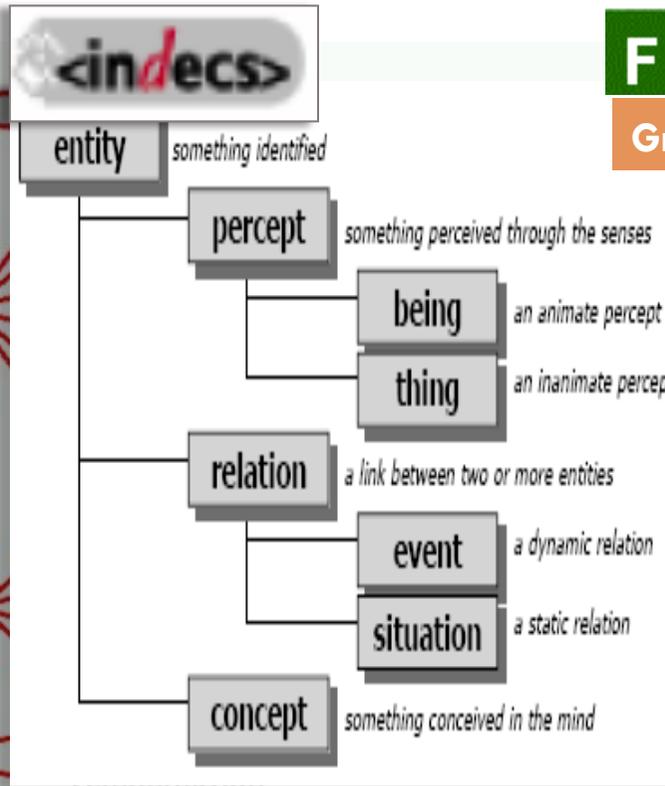


# 1. Why APs are needed?

4

## a. thema *types*

Depending on the implementation, *themas* can be categorised in various ways, even in the same discipline/subject domain



**F R B R**

Group 3 entities

object

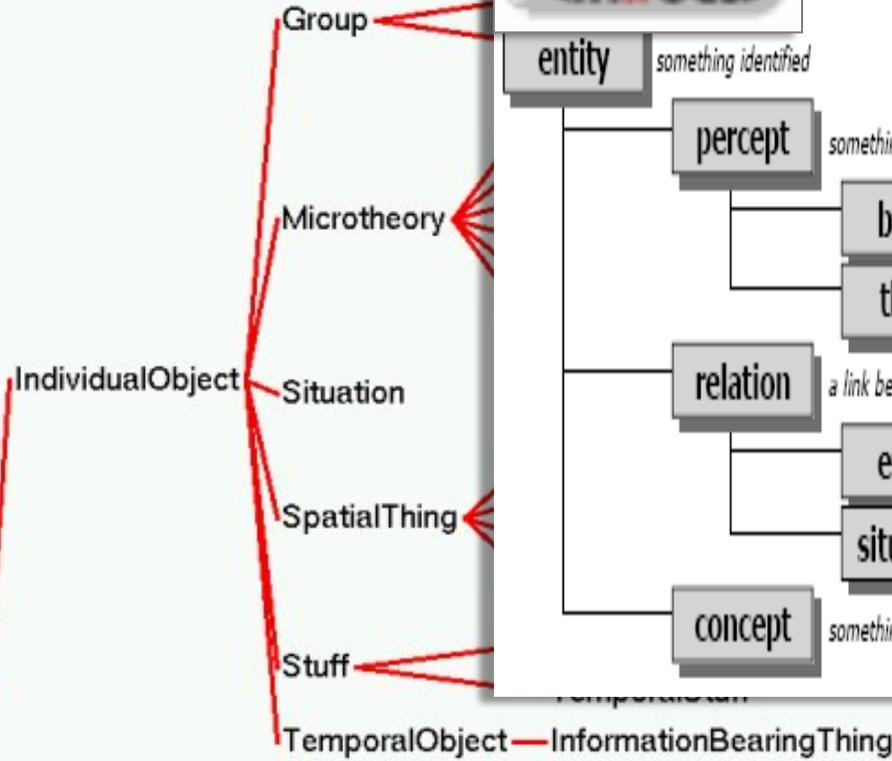
event

concept

place

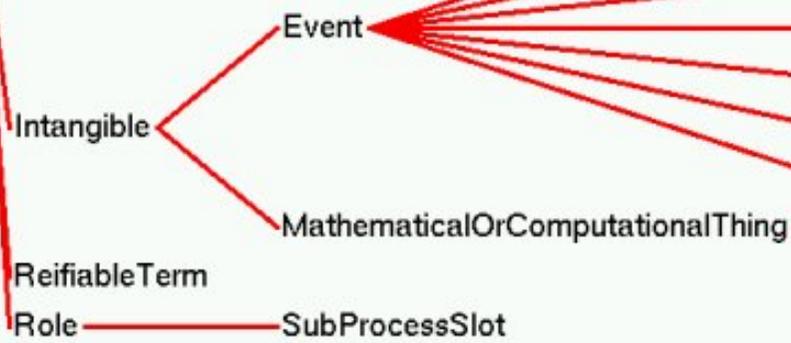
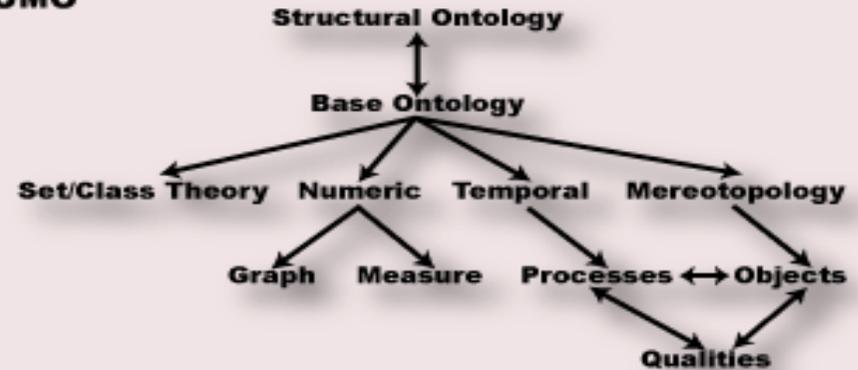
“thing”

Thing



Suggested Upper Merged Ontology

**SUMO**



# UMLS

## □ Entities

- ▣ Physical Object
  - Organism
  - Anatomical Structure
  - Manufactured Object
  - Substance
- ▣ Conceptual Entity
  - Idea or Concept
  - Finding
  - Organism Attribute
  - Intellectual Product
  - Language
  - Occupation or Discipline
  - Organization
  - Group Attribute
  - Group

## □ Events

- ▣ Activity
- ▣ Phenomenon or Process

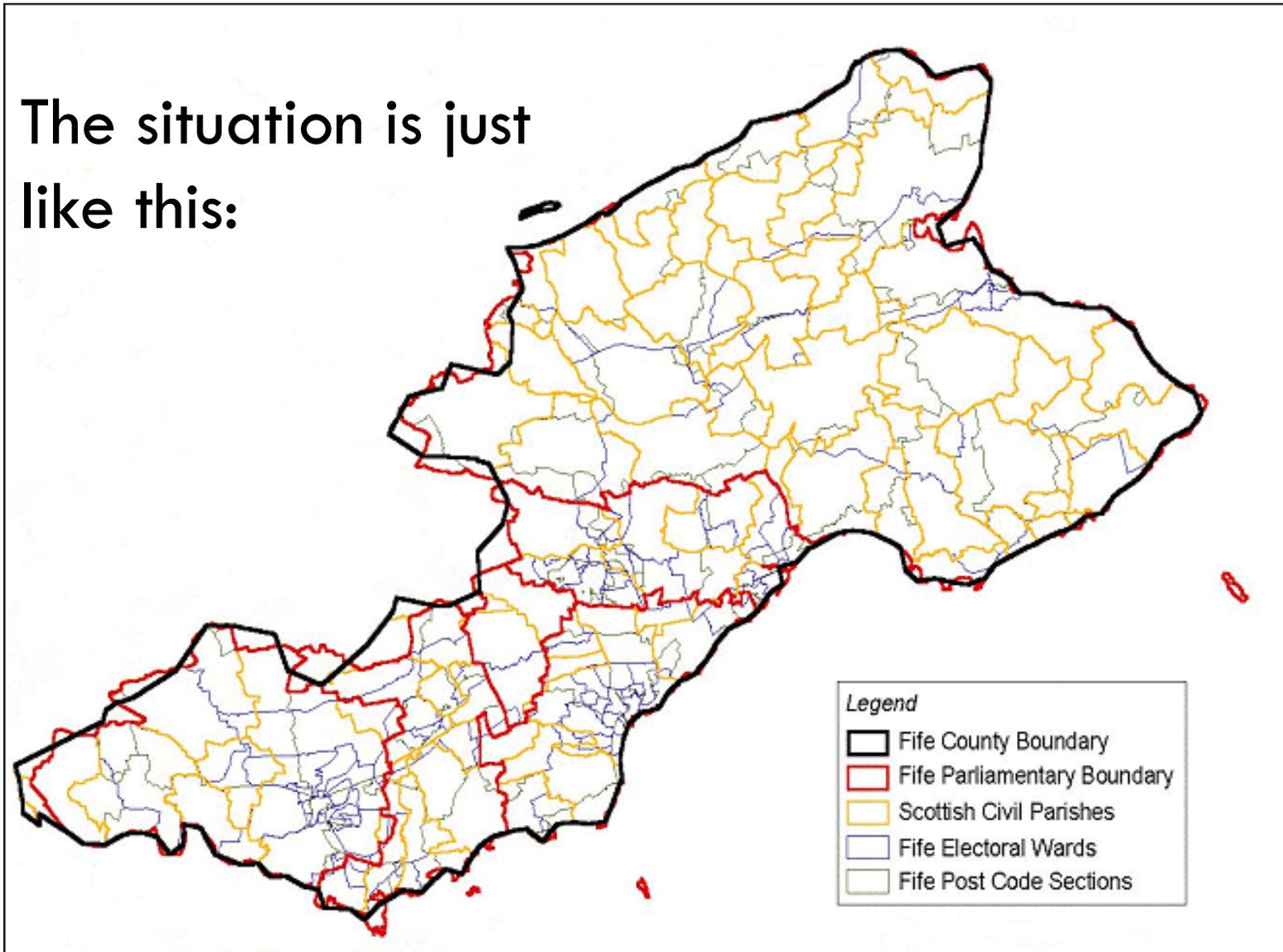
## The Foundational Model of Anatomy (FMA)

- **Anatomical Entity**
  - Non-physical anatomical entity
  - Physical anatomical entity
- **Attribute Entity**
  - Cell morphology
  - Cell shape type
  - Cell surface feature
  - Concept name
  - Miscellaneous term
  - Organ part phenotype
  - Physical attribute relationship
  - Physical state
  - Structural relationship value
- **Dimensional Entity**
  - Line
  - Point
  - Surface
  - Volume

## International Classification of Diseases (ICD)

- DISEASES AND INJURIES
- PROCEDURES
- + EXTERNAL CAUSES OF INJURY AND POISONING
- + FACTORS INFLUENCING HEALTH STATUS AND CONTACT WITH HEALTH SERVICES

The situation is just  
like this:



Andy Corbett, James Reid, David Medyckyj-Scott, Cressida Chappell (Universities of Edinburgh and Essex): **Geo-Crosswalk: A gazetteer service and server for the UK.** JCDL2002 NKOS Workshop July 18, 2002, Portland, Oregon.

<http://nkos.slis.kent.edu/DL02workshop.htm>

# (cont.) 1. Why APs are needed?



## b) thema- to - thema relationships

General relationships between *themas* (applicable to all types)

- Hierarchical
  - ▣ Partitive
  - ▣ Generic
  - ▣ Instance
  - ▣ Polyhierarchical
- Associative (=other)

Other thema-to-thema relationships are implementation-dependent

# FAO Country Profiles -- *The Geopolitical Ontology*

<http://www.fao.org/countryprofiles/geoinfo.asp>



## Area types:

- Groups
- Territories

## Data associated to areas:

- Names (multilingual)
- International codes
- Coordinates
- DBPedia ID
- Currency names and codes
- Adjectives of nationality
- Basic statistical data

## Relations:

- Groups membership
- Land borders
- Historic changes: predecessor, successor, valid since, valid until

## ADL Digital Gazetteer

### Relationships between entries

#### Inherently spatial

- ▣ Containment
- ▣ Overlap
- ▣ Proximity
- ▣ Directional

#### Explicitly stated

- ▣ **PartOf**
  - AdministrativePartOf
  - AdministrativePartitionMemberOf
  - AdministrativeSeatOf
  - ConventionallyQualifiedBy
  - SubfeatureOf
  - GeophysicalPartitionMemberOf
- ▣ **PhysicallyConnectedTo**
  - FlowsInto

# Nomens in different types of KOS

themas

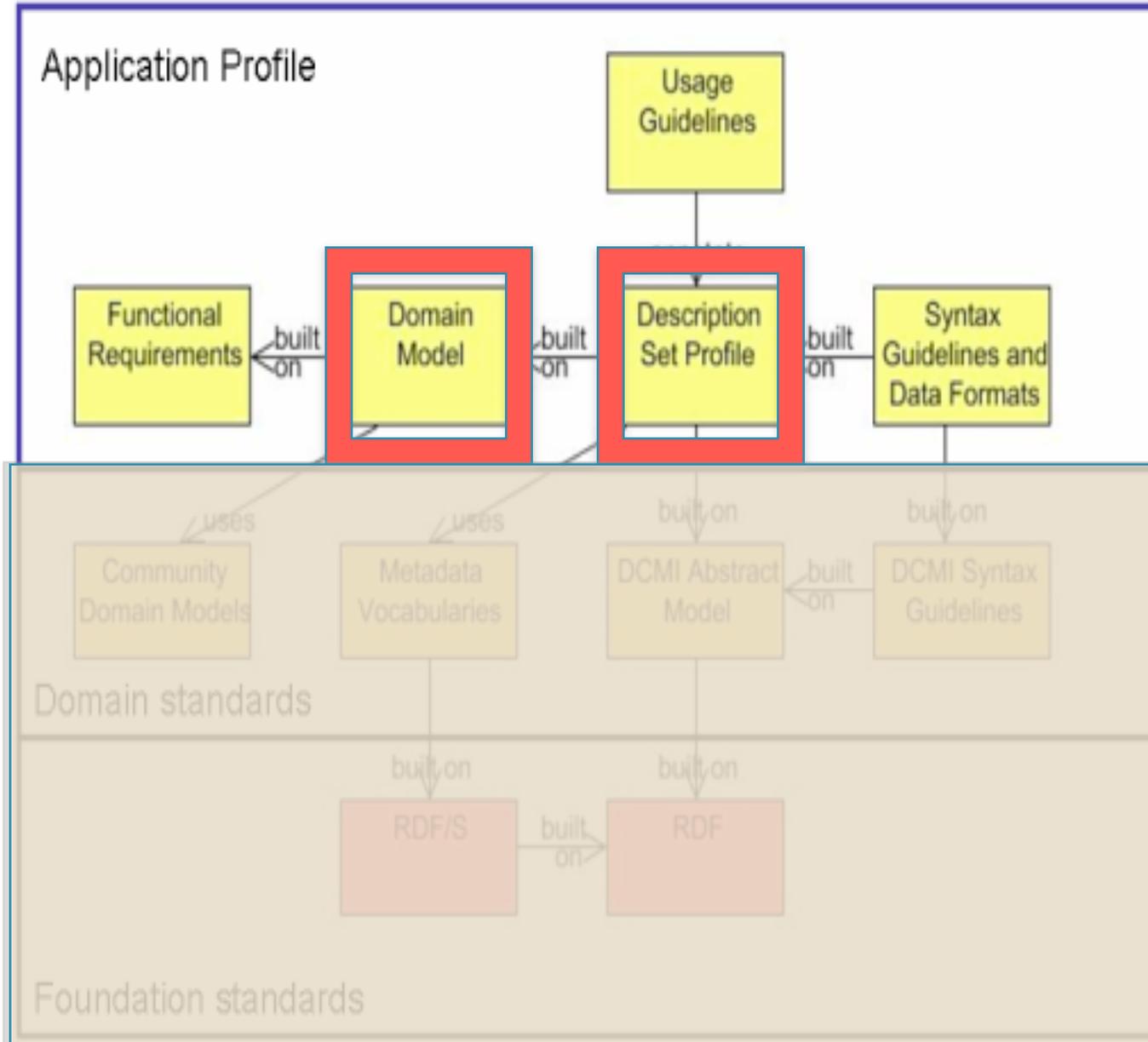
represented by:

- thesauri:
- classification schemes:
- subject heading systems:
- taxonomies:
- ontologies:
- picklists:
- ... ..

terms (preferred & non-preferred)  
notations  
terms of pre-coordinated strings  
category labels (w or w/t notations)  
terms or identifiers  
terms  
... ..



## 2. How formally (or informally) can this style of application profile be defined?



**Functional Requirements\***

**Domain Model\***

**Description Set Profile\***

**Usage Guidelines**

**Encoding syntax guidelines**

*\*mandatory*

## (cont.) 2. How formally (or informally) can this style of application profile be defined?

### DCAP

- **Functional Requirements\***

(describes what a community wants to accomplish with its application)

### FRSAD-AP

- vocab control for retrieval, organizing/categorizing, navigation, reasoning, provenance ...

## (cont.) 2. How formally (or informally) can this style of application profile be defined?

### DCAP

- **Domain Model\***  
(characterizes the types of things described by the ~~metadata~~ properties and [characterizes] their relationships)

### FRSAD-AP

- FRSAD is a general model;
- Need more specific ones for
  - ▣ different types  
(e.g. classification vs. thesaurus vs. subject headings)
  - ▣ different subject domains  
(e.g., medical vs. consumer health)
  - ▣ different structures  
(e.g., enumerated vs. synthetic vs. semi-synthetic)

E.g., What are the basic entities in a classification system?

*thema* : class  
*nomen*: notation

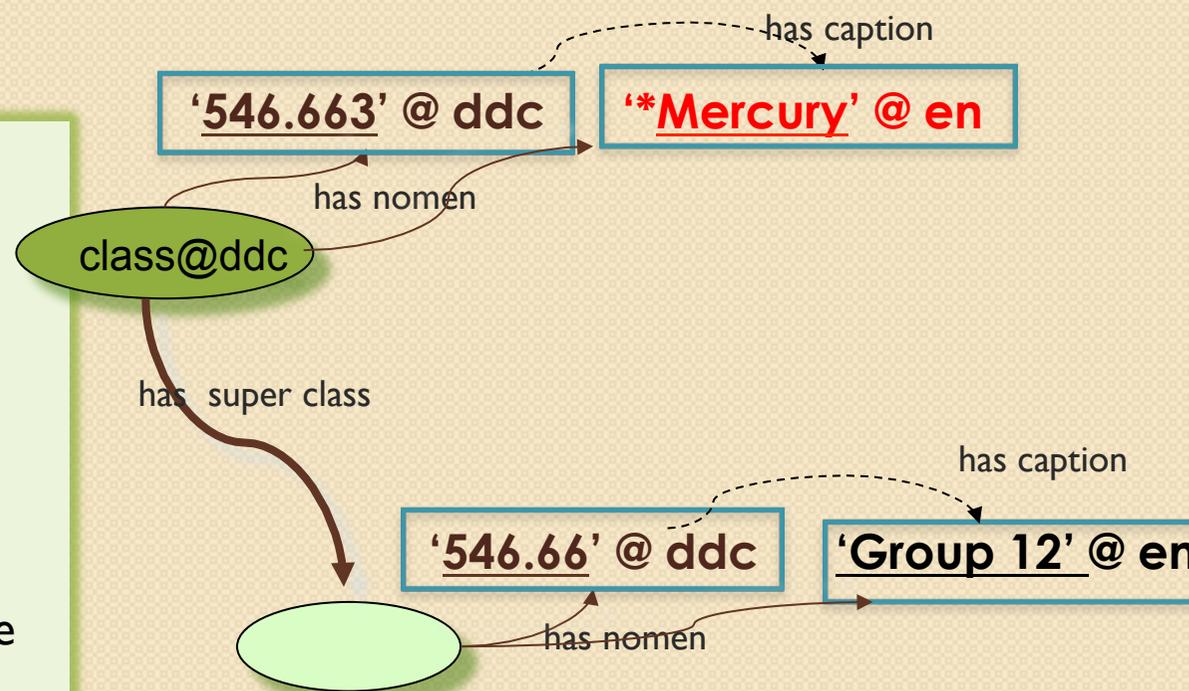
or

*themas* :  
 class  
 . including built classes[1]  
 . memberInClass[2]  
 ...  
*nomens*:  
 notation  
 caption  
 nameOfMember-inScopeNote  
 index term  
 ... ..

|                          |          |
|--------------------------|----------|
| <b>Class Number:</b>     | 546.663  |
| <b>Segmented Number:</b> | 546/.663 |
| <b>Caption:</b>          | *Mercury |

|          |  |
|----------|--|
|          | <u>Main Classes</u>                          |
| 500      | <u>Science</u>                               |
| 540      | <u>Chemistry</u>                             |
| 541-547  | <u>Chemistry</u>                             |
| 546      | <u>Inorganic chemistry</u>                   |
| 546.6    | <u>Groups 8, 9, 10, 11, 12, 13, 14</u>       |
| 546.66   | <u>Group 12</u>                              |
| 546.663  | <u>*Mercury</u>                              |
| 546.6635 | <u>Mercury (Element)--physical chemistry</u> |



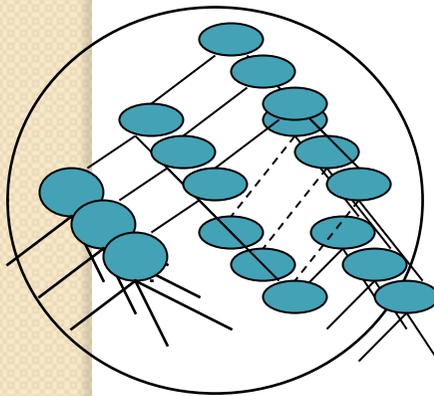
A notation has its *semantic value* and an *ordinal value*

## e.g., How to describe the orders/sequences of coordinate classes (not just hierarchical relationships)

Semantically meaningful orders in a classification system

### Classes are arranged according to

- stages in a process (e.g., brewing processes, packaging of product processes);
- time or evolutionary sequence (e.g., ancient Greek sculptures, paleontology, stars);
- degree of complexity (e.g., geometric figures),
- size (e.g., town, cities, metropolis, and other administrative units)
- According to Literary Warrant principle (e.g., arrange literature according to publication amount)
- According to User Warrant principle (e.g., arrange services and products according to popularity)



## (cont.) 2. How formally (or informally) can this style of application profile be defined?

### DCAP (cont.)

- Description Set Profile\*
- (enumerates the metadata terms to be used)

### FRSAD-AP (cont.)

- **Properties** of entities
  - APs may need specific attributes and/or values, e.g., for notation & caption
  - [other questions]

## *Nomen* general attributes (include but not limited to)

- ▣ Type of nomen (identifier, controlled name, ...)\*
- ▣ Scheme (LCSH, DDC, UDC, ULAN, ISO 8601 ...)
- ▣ Reference Source of nomen (Encyclopedia Britannica...)
- ▣ Representation of nomen (alphanumeric, sound, visual,...)
- ▣ Language of nomen (English, Japanese, Slovenian,...)
- ▣ Script of nomen (Cyrillic, Thai, Chinese-simplified,...)
- ▣ Script conversion (Pinyin, ISO 3601, Romanisation of Japanese...)
- ▣ Form of nomen (full name, abbreviation, formula...)
- ▣ Time of validity of nomen (until xxxx, after xxxx, from... to ...)
- ▣ Audience (English-speaking users, scientists, children ...)
- ▣ Status of nomen (provisional, accepted, official,...)

\*note: examples of attribute values in parenthesis

# Example: Notations -- Rules

- Classification numbers may be built according to rules
- Example from DDC:

## **821.008 Collections of English poetry**

is built with

**82** (following the instruction at **820.1-828 Subdivisions of English literature**)

plus **100** (following the instruction at **T3B--1001-T3B--1009 Standard subdivisions; collections; history, description, critical appraisal**)

plus **8 Collections of literary texts** from the add table at **T3B--1-T3B--8 Specific forms.**

821 English poetry

821.008 English poetry--collections

821.00803543 Love--poetry--English literature--collections, . . .

821.0080355 English poetry--social themes--collections, . . .

821.008036 English poetry--nature--collections, . . .

821.0080382 English poetry--religious themes--collections, . . .

821.00**9** English poetry--history and criticism

821.04 English poetry--lyric poetry, . . .

821.0708 Humorous poetry--English literature--collections, . . .

Source: [One Zero or Two?](#) Dewey Blog. September 28, 2006

[http://ddc.typepad.com/025431/ddc\\_tip\\_of\\_the\\_week/](http://ddc.typepad.com/025431/ddc_tip_of_the_week/)

# General Nomen relationships

19

- Partitive
- Equivalence

Equivalence can be specified further, e.g.:

- Replaces/is replaced by
- Has variant form/is variant form
- Has derivation/is derived from
  - ▣ Has acronym/is acronym
  - ▣ Has abbreviation/is abbreviation
  - ▣ Has transliterated form/is transliteration

**APs may need more specific relationships, e.g., for notation & caption**

## (cont.) 2. How formally (or informally) can this style of application profile be defined?

### DCAP (cont.)

- Usage Guidelines
- Encoding syntax guidelines

### FRSAD-AP (cont.)

- Usage Guidelines
- Recommendation:  
e.g., SKOS & extensions;  
MADS, BS8723-5,  
ISO25964, ...

### 3. In what ways are application profiles for **subject domains** different from APs for **descriptive metadata**?

Descriptive metadata

Subject domain vocabularies



### 3. In what ways are application profiles for **subject domains** different from APs for **descriptive metadata**?

Descriptive metadata



Subject domain vocabularies



Describing a thema

-- what a concept is about

-- where it belongs to

Serious sameAs issue

-- senior@schemaA =? senior@schemaB

-- sunflower@mesh =? sunflower@aat

Integrity rely on the domain model and properties around a thema and a nomen

# Questions to be discussed



1. Why APs are needed for subject authority data?
2. How formally (or informally) can this style of “application profile” be defined?
3. In what ways are application profiles for **subject domains** different from APs for **descriptive metadata**?

# FRSAD's relation to FRBR

