Mapping FRSAD Model and Other Abstract Models

[FRSAR = Functional Requirements for Subject Authority Records]

Marcia Zeng, Kent State University, USA

Maja Žumer, University of Ljubljana, Slovenia

Based on the work of the IFLA FRSAR Working Group

ISKO UK Conference 2009, 22-23 June 2009, London

Outline

- 1. Introducing the FRSAD model
- 2. Mapping to other models (BS 8723, SKOS, OWL, DCMI-AM)

- (Note: only selected slides will be talked in the 30 minutes presentation.)

1. Introducing the FRSAD model

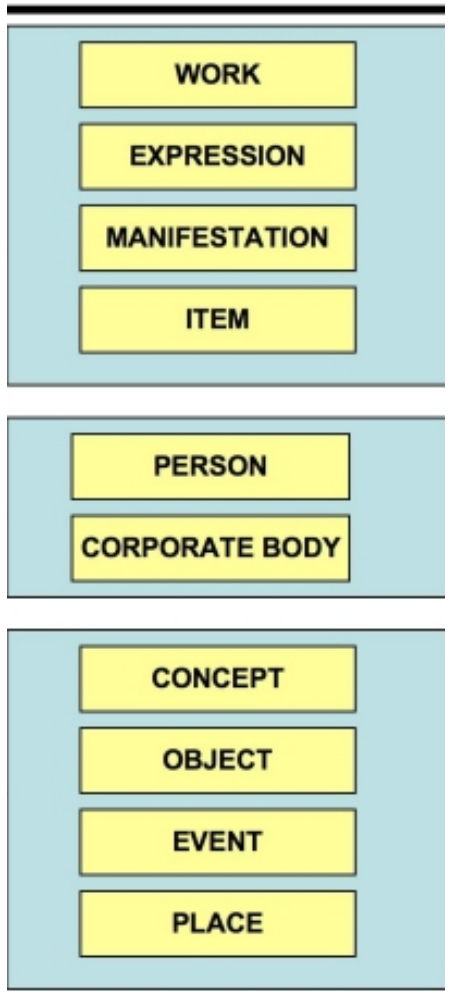
FRSAD = Functional Requirements for Subject Authority **Data**

FRBR

- Functional Requirements for Bibliographic Records (FRBR)
  - Approved by IFLA in 1997
  - Published in 1998
  - Conceptual model of the 'bibliographic universe'

IFLA. (1998). *Functional Requirements for Bibliographic Records*: *Final Report*. IFLA Study Group on the Functional Requirements for Bibliographic Records. München: KG Saur. http://www.ifla.org/publications/functional-requirements-for-bibliographic-records

FRBR' three groups of entities



The "FRBR family"

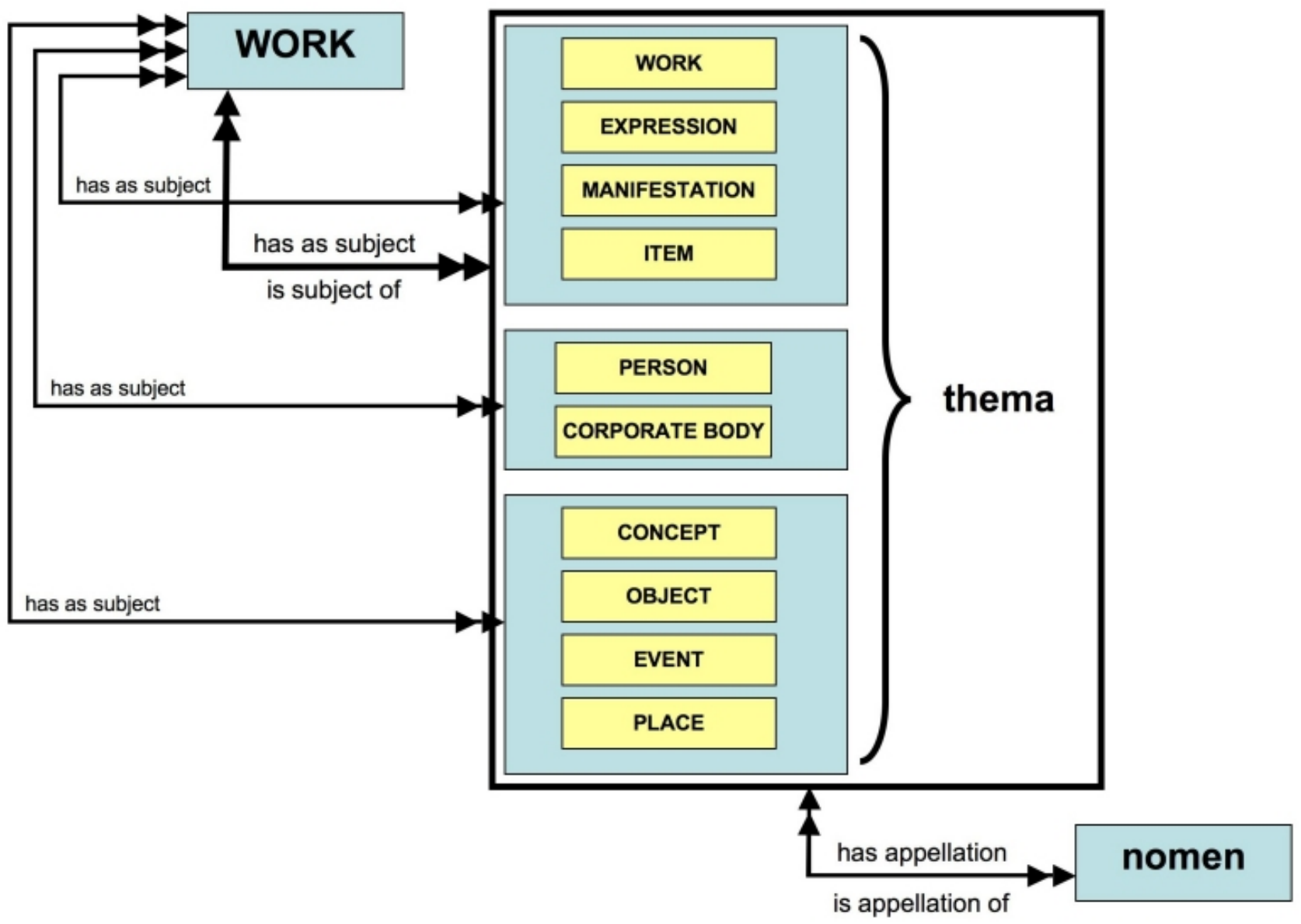
- FRBR: the original framework
  - All entities, focusing on Group 1 entities
- FRAR (FRAD): Functional Requirements for Authority Records/Data
  - Focusing on Group 2 entities
  - Just published yesterday (June 22, 2009)
- **FRSAR: Functional Requirements for Subject Authority Records/Data**
  - Focusing on Group3 entities
  - Established April 2005
  - Draft Report is released for commend today (June 23, 2009)
  - http://nkos.slis.kent.edu/FRSAR/

FRSAR Working Group

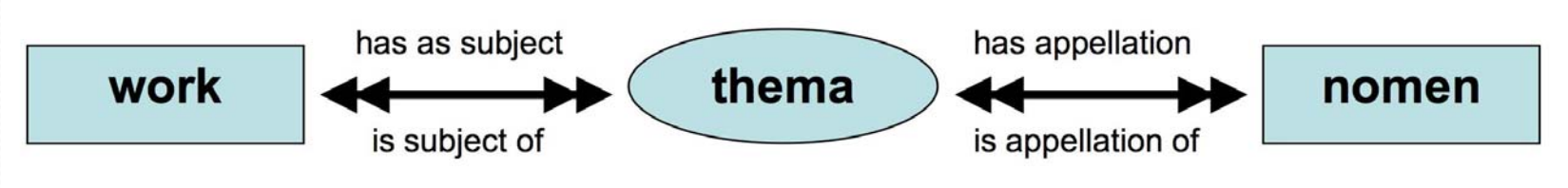
FRSAR = Functional Requirements for Subject Authority Records

- **Terms of Reference**
  - to build a conceptual model of Group 3 entities within the FRBR framework as they relate to the ***aboutness*** of works,
  - to provide a clearly defined, structured frame of reference for relating the data that are recorded in subject authority records to the needs of the users of those records, and
  - to assist in an assessment of the potential for international sharing and use of subject authority data both within the library sector and beyond.

FRSAD's relation to FRBR

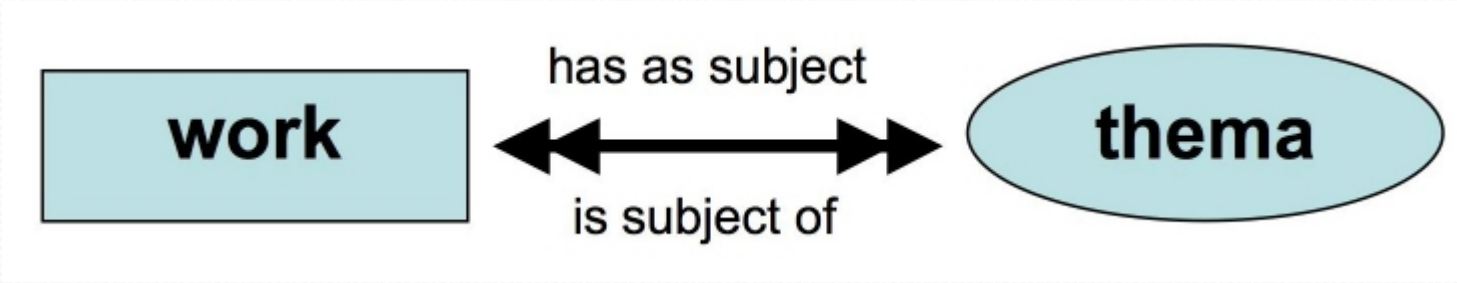


FRSAD Conceptual Model



FRSAD Part 1:

***WORK* has as subject *THEMA* / *THEMA* is subject of *WORK*.**

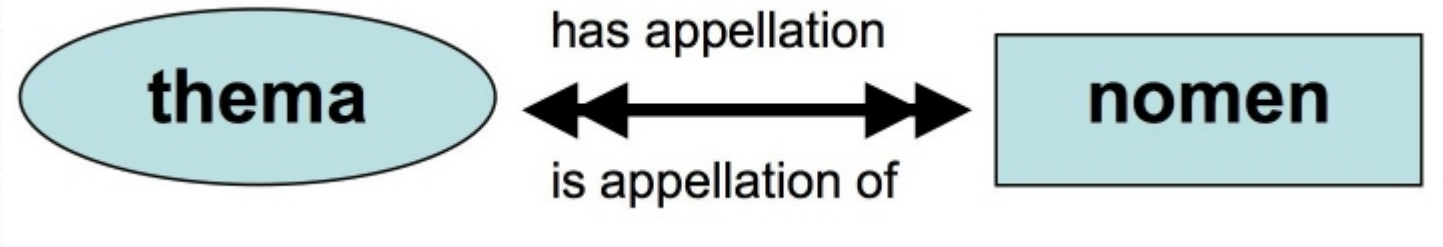


This model confirms one of the basic relationships defined in FRBR:
WORK* has as subject *THEMA* / *THEMA* is subject of *WORK.

- *Thema* = "any FRBR entity as used as a subject of a *work*".
- *Thema* includes any of the FRBR entities:
  Group 1 and Group 2 entities and,
  in addition, all other subjects of *works*.

FRSAD Part 2:

***THEMA* has appellation *NOMEN* / *NOMEN* is appellation of *THEMA*.**

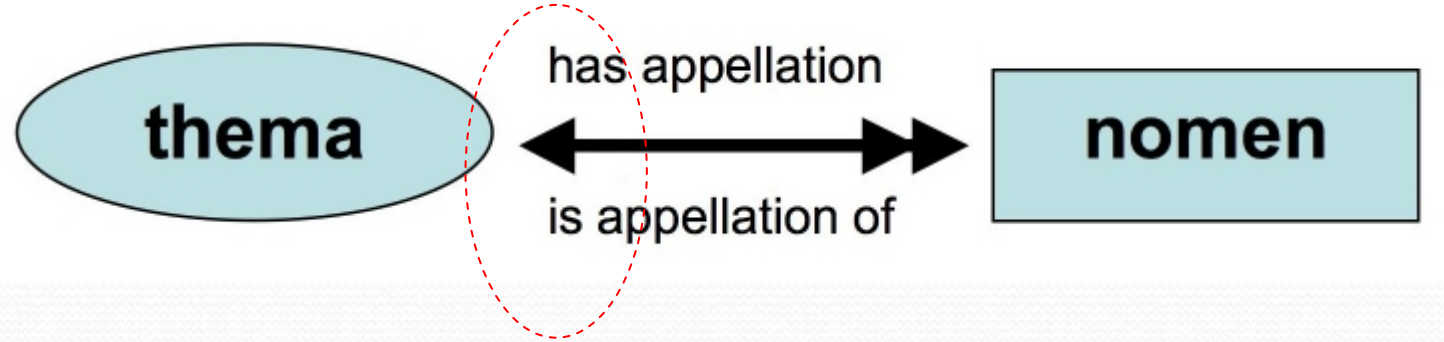


This model also proposes a new relationship: ***THEMA* has appellation *NOMEN* / *NOMEN* is appellation of *THEMA*.**

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

Part 2b

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



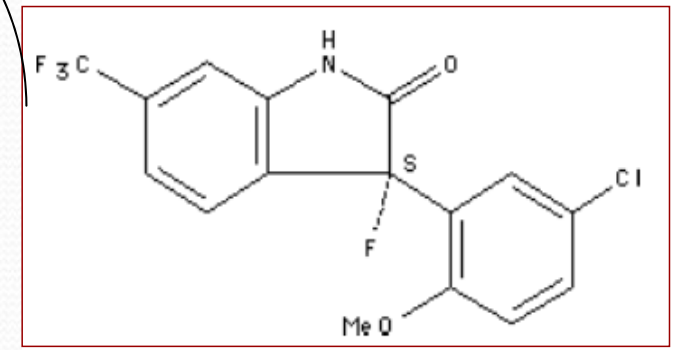
***NOMEN* = any sign or sequence of signs (alphanumeric characters, symbols, sound, etc.) by which a *thema* is known, referred to or addressed.**

Example: various nomens for the same thema

```
Accession Number    (AN): 2005:3738  USAN
Publication Year    (PY): 2002
Generic Name        (CN): Flindokalner
OTHER NAMES:
Chemical Name       (CN): 2H-Indol-2-one, 3-(5-chloro-2-methoxyphenyl)-3-fluoro-
                          1,3-dihydro-6-(trifluoromethyl)-, (3S)-
Chemical Name       (CN): (3S)-3-(5-Chloro-2-methoxyphenyl)-3-fluoro-6-
                          (trifluoromethyl)-1,3-dihydro-2H-indol-2-one
Trade Name          (CN): MaxiPost (Bristol-Myers Squibb)
Code Designation    (CN): BMS-204352
CAS Registry No.    (RN): 187523-35-9
Molecular Formula   (MF): C16 H10 Cl F4 N O2
Lin. Str. Formula (LSF): C16 H10 Cl F4 N O2
Molecular Weight    (MW): 359.71

Absolute stereochemistry.  Rotation (+).
```

different types of nomen



nomen representation="graphic"

Source: STN Database Summary Sheet: USAN (The USP Dictionary of U.S. Adopted Names and International Drug Names)

Choice of terms (thema, nomen)

- Different and overlapping meaning of ‘subject’, ‘topic’, ‘concept’, ‘class’, etc.
- Different views on granularity
- ‘Name’ understood as ‘proper name’

Therefore:

- Terms from Latin that do not have to be translated and are not loaded with other meanings

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

- Hierarchical
  - Partitive
  - Generic
  - Instance
- Associative

Other thema-to-thema relationships are implementation-dependent

Example: An online display record of the AAT concept “Mercury”

ID: 300011026 **Record Type:** concept

mercury (<mercury and amalgam>, nonferrous metal, ... Materials)

Note: Pure metallic element having symbol Hg and atomic num
liquid at ordinary temperatures. Use also for this metal as proc
combination with other substances, to make various objects a

Terms:

mercury (**preferred**,C,D,U,LC,English-P)
Hg (C,UF,U,A,English)
quicksilver (C,UF,U,English)
argento vivo (C,D,U,Italian-P)

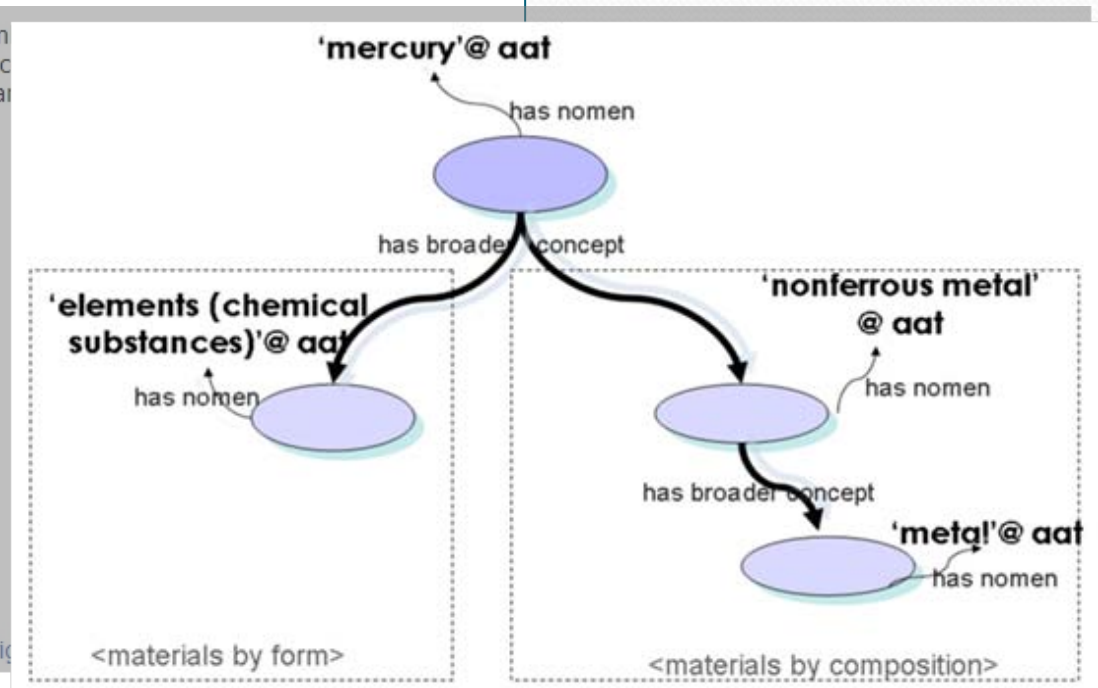
Facet/Hierarchy Code: M.MT

Hierarchical Position:

Materials Facet
.... Materials
..... materials
..... <materials by composition>
..... inorganic material
..... <metal and metal products>
..... metal
..... <metal by composition or or
..... nonferrous metal
..... <mercury and amalgam>
..... mercury

Additional Parents:

Materials Facet
.... Materials
..... materials
..... <materials by form>
..... <materials by chemical form>
..... elements (chemical substances)
..... mercury

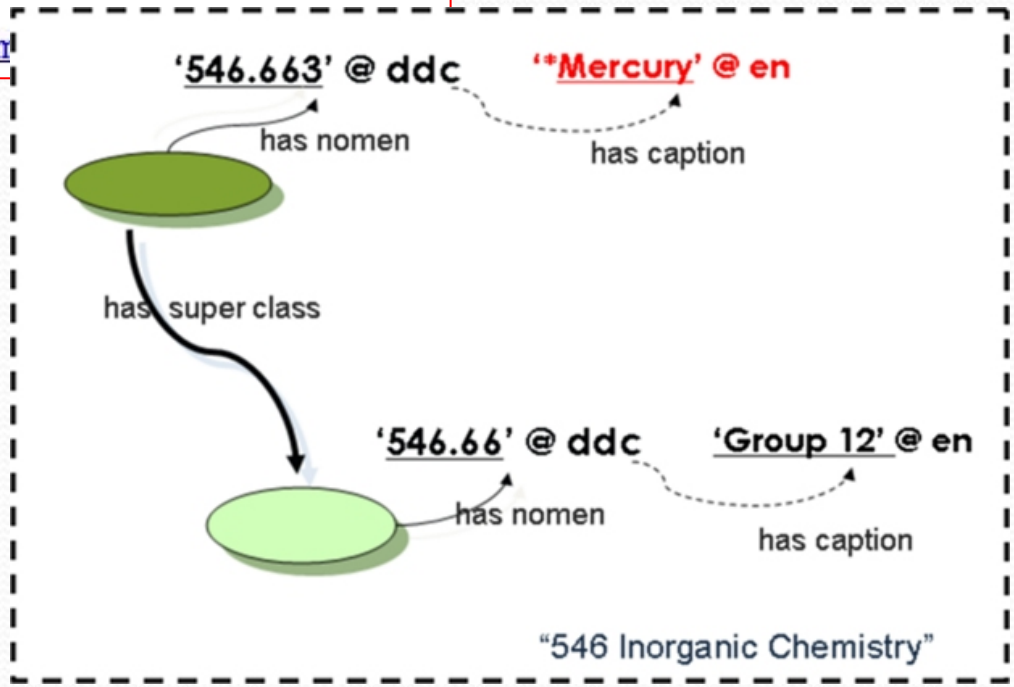


Source: Art and Architecture Thesaurus Online

Example: From WebDewey, a record for classes related to "Mercury“ (element)

Class Number: 546.663
Segmented Number: 546/.663
Caption: *Mercury

Main Classes
500 Science
540 Chemistry
541-547 Chemistry
546 Inorganic chemistry
546.6 Groups 8, 9, 10, 11, 12, 13, 14
546.66 Group 12
546.663 ***Mercury**
546.6635 Mercury (Elem



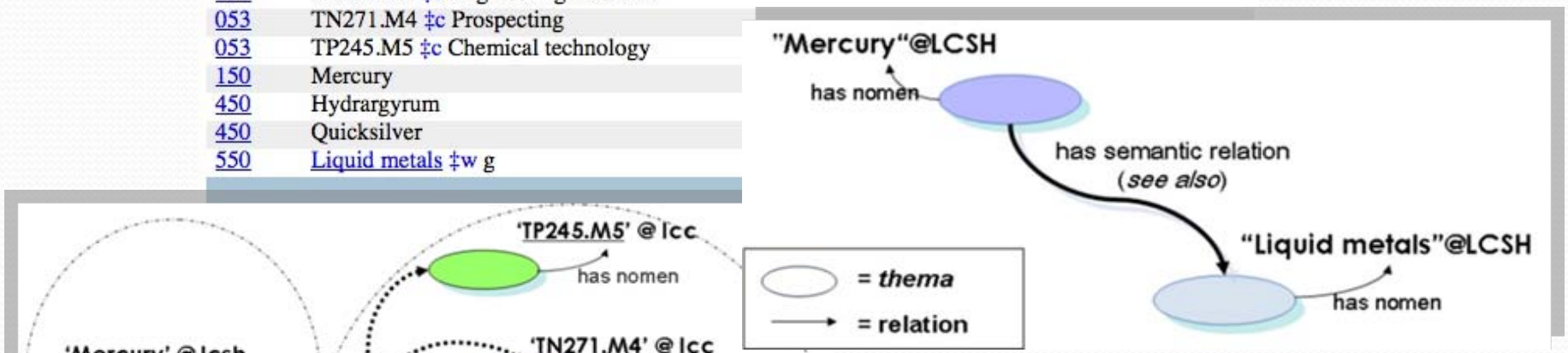
Example: An online display record of the LCSH Concept “Mercury” and associated LC Classification classes

Authority File: Mercury MARC Help

ARN 2015071

Rec stat	c	**Entered**	19860211	**Replaced**	19900221112154.6		
Type	z	**Upd status**	a	**Enc lvl**	n	**Source**	
Roman	|	**Ref status**	a	**Mod rec**		**Name use**	b
Govt agn	|	**Auth status**	a	**Subj**	a	**Subj use**	a
Series	n	**Auth/ref**	a	**Geo subd**	i	**Ser use**	b
Ser num	n	**Name**	n	**Subdiv tp**	n	**Rules**	n

010	sh 85083794
040	DLC ‡b eng ‡c DLC ‡d DLC
053	QD181.H6 ‡c Chemistry
053	TA480.M4 ‡c Engineering materials
053	TN271.M4 ‡c Prospecting
053	TP245.M5 ‡c Chemical technology
150	Mercury
450	Hydrargyrum
450	Quicksilver
550	Liquid metals ‡w g



Nomen-to-nomen relationships (include but not limited to)

- Partitive
- Equivalence

Equivalence can be specified further, e.g.:

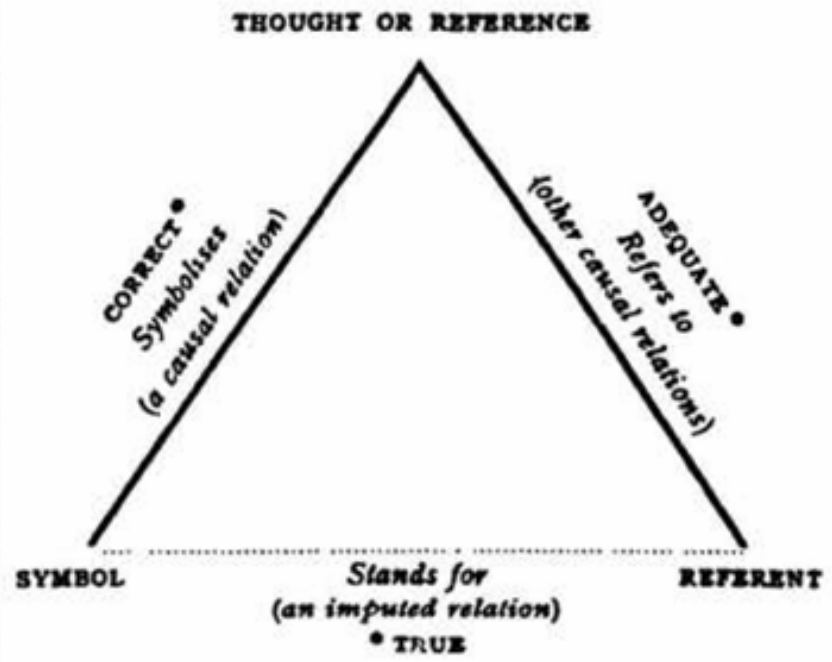
- replaces/is replaced by
- has variant form/is variant form
- has derivation/is derived from
  - has acronym/is acronym for
  - has abbreviation/is abbreviation of
  - has transliterated form/is transliteration of

The importance of the *THEMA-NOMEN* model to the subject authority data

- to separate what are usually called *concepts* (or *topics*, *subjects, concepts, classes [of concepts]*) from what they are known by, referred to, or addressed.

2. Mapping to Other Models

2.1 Ogden & Richard's (1923) triangle of meaning



•the referent of an expression (a word or another sign or symbol) is relative to different language users.

• multiple terms may refer to the same object or idea,

• a single term may refer ambiguously to more than one object or idea,

• terms may be confusing because they are out of date

2.2 British standard *BS8723-5*: *Structured vocabularies for information retrieval – Guide. Part 5: Exchange formats and protocols for interoperability* (DD 8723-5:2008).

- It includes what is needed for modeling:
  - (1) a whole thesaurus,
  - (2) arrays of thesaurus concepts, and
  - (3) records that document a thesaurus entry.
- In the model, each *concept* in a structured vocabulary (especially in a thesaurus) is represented by one preferred term per language, and by any number of nonpreferred terms. The notation, scope note and broader/narrower/related term relationships ***apply to the concept*** as a whole, rather than to its preferred term.

FRSAD and BS8723-5 Model

- Both represent these relationships:
  - (1) *thema*-and-*nomen* (a record documenting a concept and its *nomen(s)*,
  - (2) *thema*-and-*thema* (hierarchical (broader, narrower, and top concepts)) and associative (related concepts), and
  - (3) *nomen*-and-*nomen* (preferred and non-preferred, variant lexical forms, and in various languages).

2.3 SKOS (Simple Knowledge Organization System)

-- **provides a model for expressing the basic structure and content of concept schemes** such as

- Thesauri
- Classification Schemes
- Taxonomies
- Subject Heading lists
- Folksonomies, and
- other similar types of controlled vocabulary.

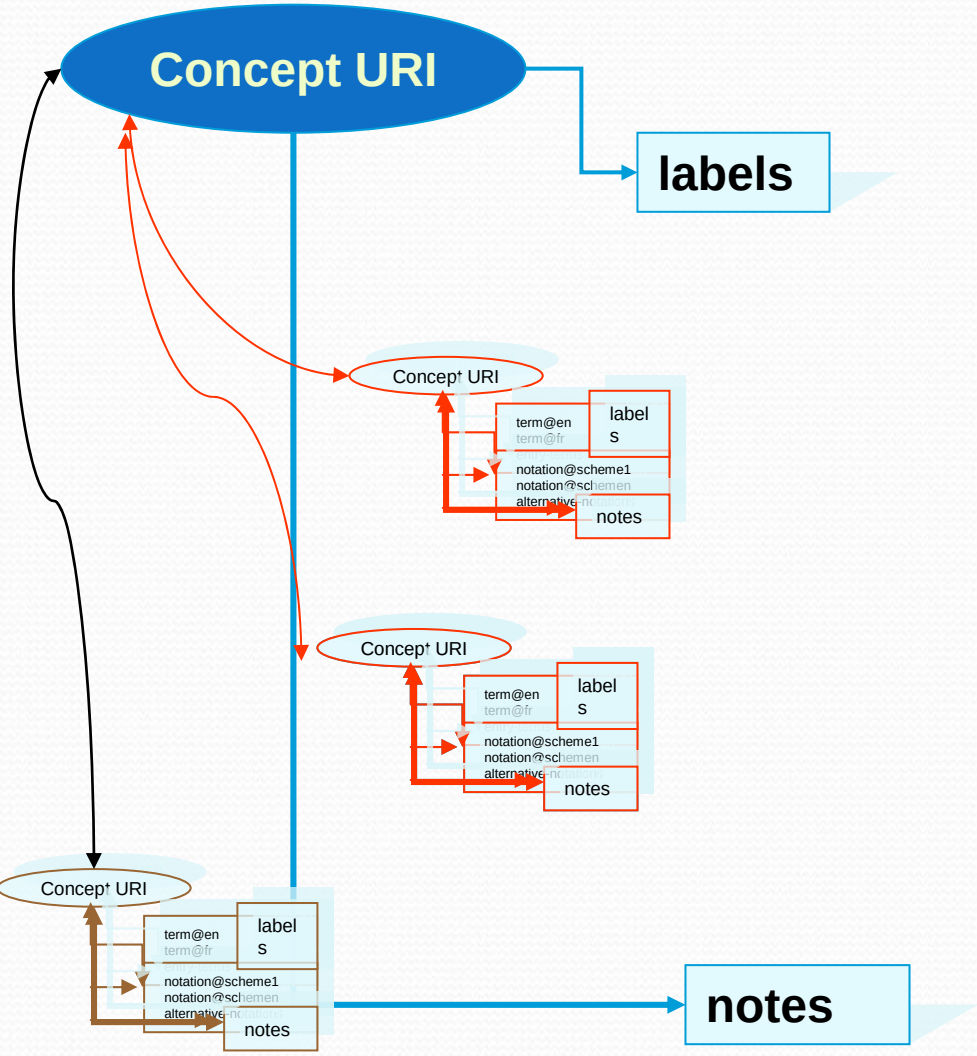
-- is an **application of RDF**

-- allows concepts to be

- **composed**
- **published** on the Web
- **linked** with data on the Web and
- **integrated** into other concept schemes.

--SKOS Reference, http://www.w3.org/TR/skos-reference/

SKOS for a thesaurus entry



[concept]

Preferred label

Alternative label

Broader concept

Narrower concept

Related concept

Scope note on concept

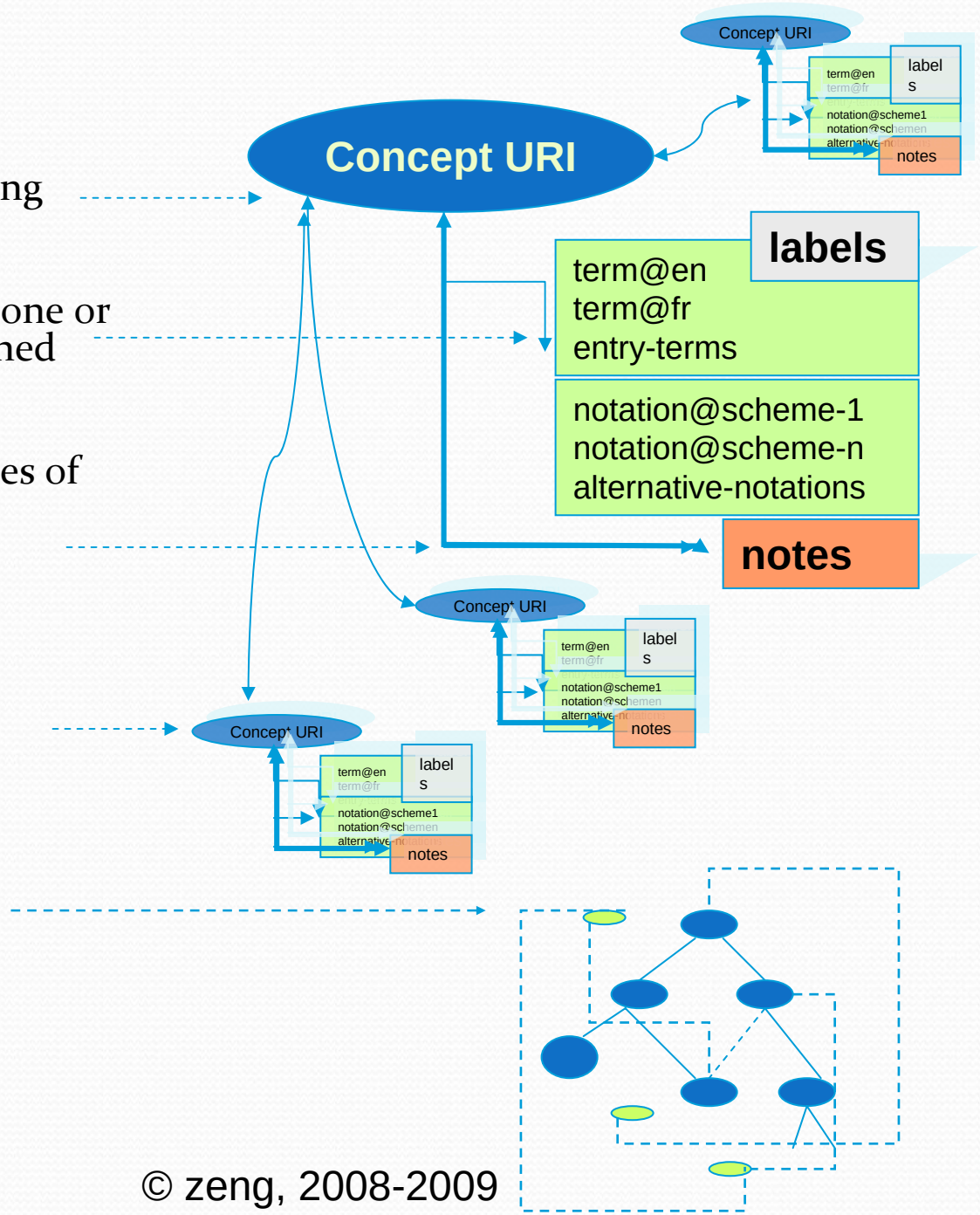
SKOS Synopsis (1)

Using SKOS,

- **concepts** can be identified using URIs
- **labeled** with lexical strings in one or more natural languages, assigned **notations** (lexical codes)
- **documented** with various types of note
- **linked** to other concepts

and

organized into
informal hierarchies and
association networks
(to be continued →)



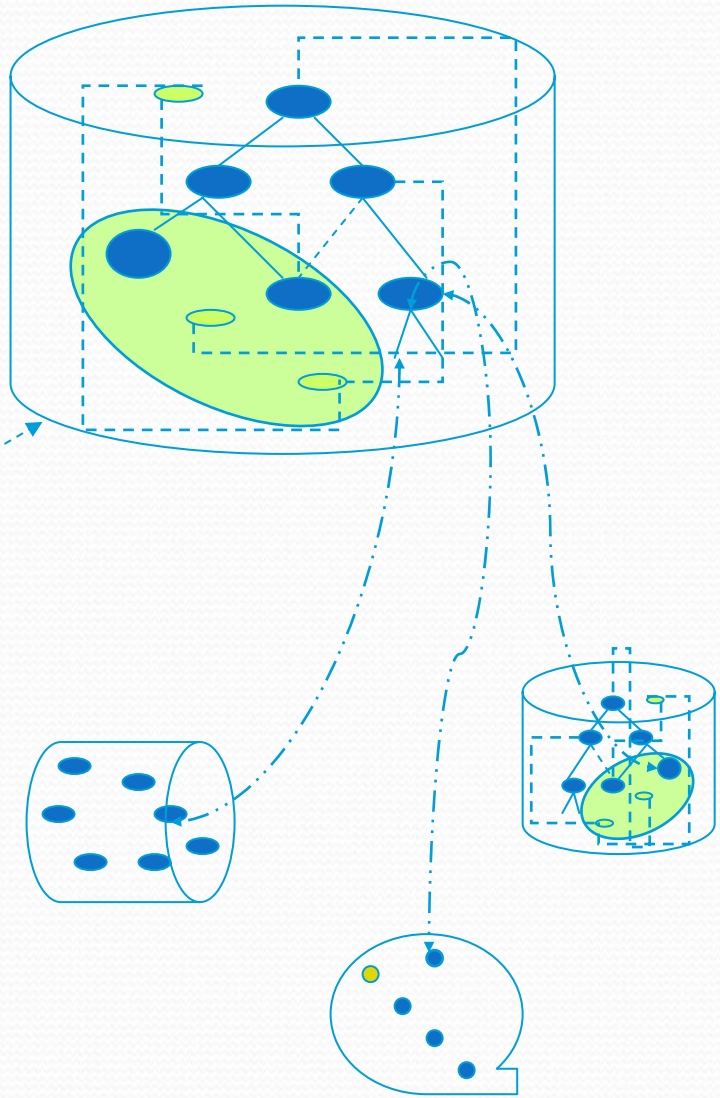
SKOS Synopsis (2)

(continued from previous slide,
-- concepts are organized into informal hierarchies and association networks)

•aggregated into **concept schemes**,

- [grouped into labeled and/or ordered **collections**,]

• and **mapped** to concepts in other schemes.



SKOS and FRSAD models

- SKOS model is based on a concept-centric view of vocabulary, where primitive objects are not labels; rather, they are concepts represented by labels.
- These can be matched to what have been defined in the FRSAD model, in terms of *thema*, *nomen* and their attributes.
- SKOS also has specific properties to represent all the semantic relationships, which matches the ones defined by FRSAD as well.

2.4 OWL Web Ontology Language

- standard ontology languages
- endorsed by the W3C
  to promote the *Semantic Web* vision.

At least two different user groups

- OWL used as data exchange language
  (define interfaces of services and agents)

- OWL used for terminologies or knowledge models

OWL Classes

OWL is an ontology language that is primarily designed to describe and define classes. Classes are therefore the basic building blocks of an OWL ontology.

OWL provides axioms (statements that say what is true in the domain) that allow relationships to be established between class expressions, including:

- SubClassOf,
- EquivalentClasses,
- DisjointClasses, and
- DisjointUnion.

- In OWL, classes and property expressions are used to construct class expressions, (sometimes also called descriptions, and, in the description logic literature, complex concepts).
  - ObjectIntersectionOf, 
  - ObjectUnionOf, and 
  - ObjectComplementOf 

  - ObjectOneOf -- contains exactly the specified individuals

Class Relationships

- **Inheritance**
- **Disjoint**
- **Equivalent**

OWL and FRSAD

- For the issues of the complexity and granularity of *themas* and comprehensive semantic relationships between and among *themas* that FRSAD attempted to cover, OWL has great matches.

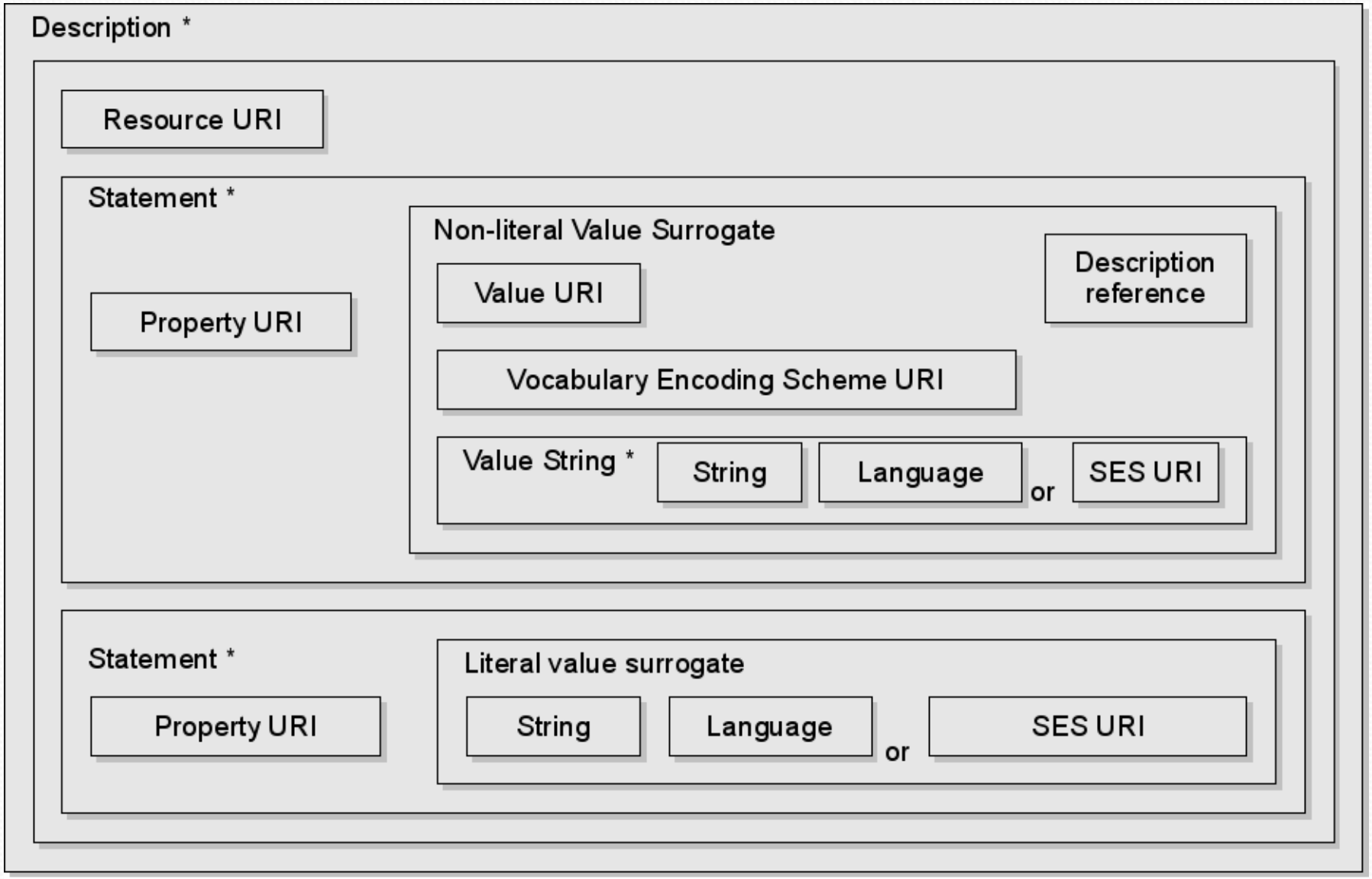
- **Described resource**
- **Property** = type of relationship
- **Value** = other resource

2.5 DCMI Abstract Model

- Formal modeling basis for Dublin Core metadata
- Like a “grammar” for Dublin Core
- Strong link with parallel development of RDF (Resource Description Framework)

The constructs of a record

- **Described resource**
- **Property** = type of relationship
- **Value** = other resource



SES = Syntax Encoding Scheme

* = repeatable

Source: Nilsson, 2007: slide 12

DCMI-AM and FRSAD

- The FRSAD model corresponds to the DCMI Abstract Model by allowing any *thema* to be independent of any nomen, including any syntax that a *nomen* may use.
- Thus this conceptual model will facilitate the sharing and reuse of subject authority data amongst not only the subject vocabularies themselves, but also metadata resources.

Conclusion

- The FRSAD model is developed with the goal to assist in an assessment of the potential for international sharing and use of subject authority data both within the library sector and beyond.
- The FRSAD model will:
  - enable the consideration of the functions of subject authority data and concept schemes at a higher level that is independent of any implementation, system, or specific context, and
  - allow us to focus on the semantics, structures, and interoperability of subject authority data.

Draft Report available at:

- **FRSAR: Functional Requirements for Subject Authority Data (FRSAD)**
  - http://nkos.slis.kent.edu/FRSAR/

Acknowledgement

- This paper is based on the work of the FRSAR Working Group and suggestions from the Advisory Group, established by the IFLA Division IV. Bibliographic Control and especially the Section of Classification and Indexing.
- FRSAR Working Group members: Leda Bultrini, Lois Mai Chan, Jonathan Furner, Edward O'Neill, Gerhard Riesthuis, Athena Salaba, Diane Vizine-Goetz, Ekaterina Zaytseva, Marcia Lei Zeng, and Maja Zumer.
- FRSAR Advisory Group members: Victoria Francu, Hemalata Iyer, Dorothy McGarry, David Miller, Päivi Pekkarinen, and Barbara Tillett.
- The communication and sharing with the FRANAR Working Group chaired by Glenn Patton.
- IFLA, OCLC, and Kent State University have provided funding, facilities, and tremendous support.

- **Described resource**
- **Property** = type of relationship
- **Value** = other resource

References

- Baker, Thomas. 2005. "Diverse vocabularies in a common model: Dublin Core at 10 years." Presentation at *DC-2005: Vocabularies in Practice*. Available at http://dc2005.uc3m.es/program/presentations/2005-09-12.plenary.baker-keynote.ppt
- Nilsson, Mikael. 2007. DCMI Basic Syntaxes Tutorial. *DC-2007: International Conference on Dublin Core and Metadata Applications: Application Profiles and their Application in Practice*. 27-31 August 2007, Singapore. Available at http://www.dc2007.sg/T2-BasicSyntaxes.pdf
- Dublin Core Metadata Initiatives (DCMI) http://dublincore.org/
- *OWL 2 Web Ontology Language Structural Specification and Functional-Style Syntax.* (2009). Motik, B, Patel-Schneider, P.F. and Parsia, B. eds. W3C Working Draft 21 April 2009. http://www.w3.org/TR/owl2-syntax/
- BS8723 Official Development Website. (2008). http://schemas.bs8723.org/Home.aspx
- *SKOS Simple Knowledge Organization System Reference.* (2009). W3C Candidate Recommendation 17 March 2009; http://www.w3.org/TR/2009/CR-skos-reference-20090317/
- *SKOS Reference* (2009). W3C Candidate Recommendation 17 March 2009. Available at:
- Miles, Alistair. (2008). The Web and SKOS, ISKO London, July 2008. Available at: *www.iskouk.org/presentations/**miles**_web_and_skos_200807.pdf*
- Miles, Alistair. (2005) SKOS Core Tutorial, DC-2005, Madrid. Available at: http://www.dublincore.org/resources/training/dc-2005/tutorial4_eng.pdf