Modelling Knowledge Organization Systems and Structures

-A discussion in the context of conceptual and data models

Maja Žumer
University of Ljubljana
Slovenia

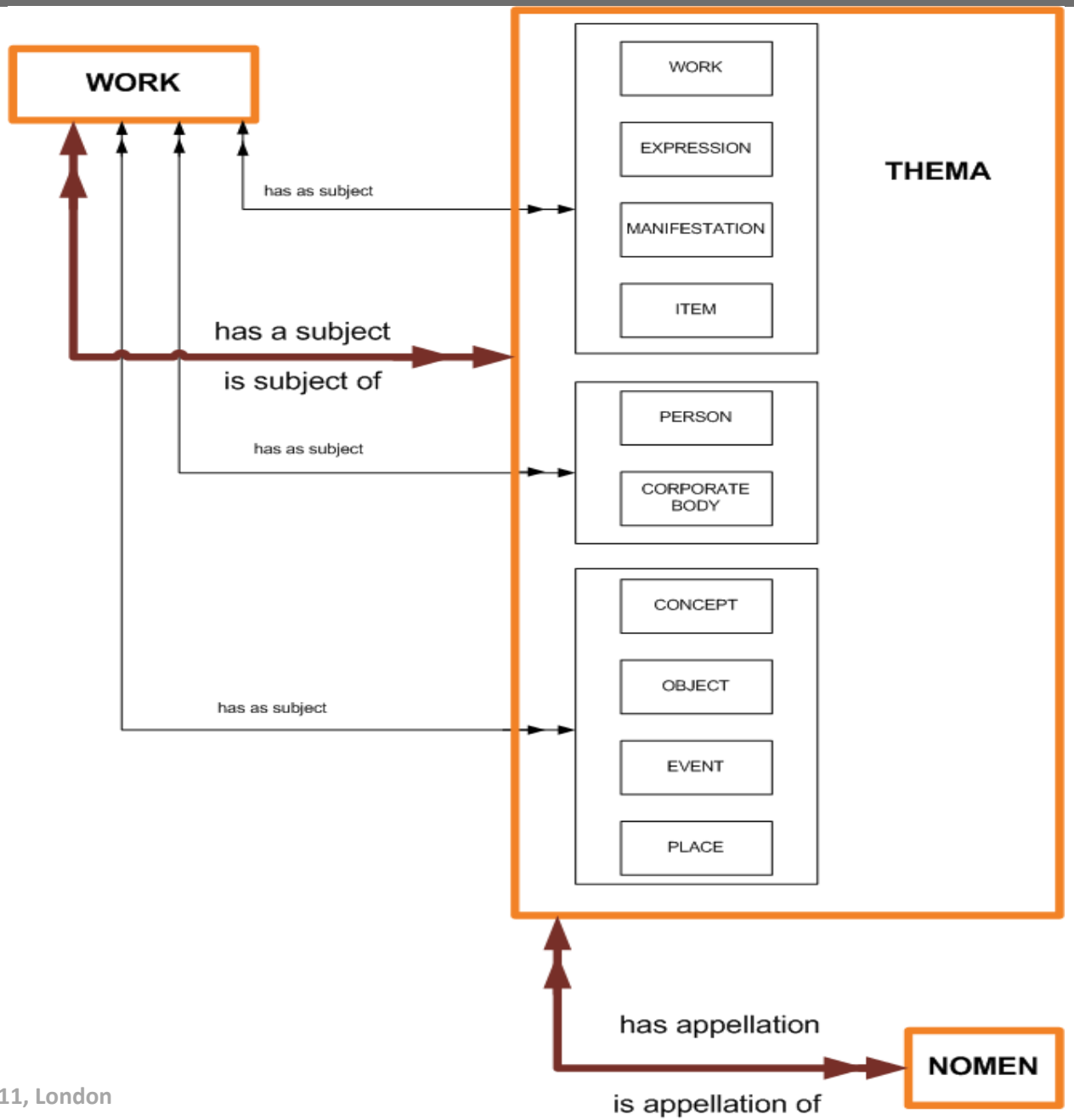
Marcia Lei Zeng
Kent State University
USA

ISKO-UK Biennial Conference, July 4-5, 2011, London

CC BY NC ND

I. Background

FRSAD– generalisation of FRBR



The core of the FRSAD conceptual model

FRSAD Part 1:

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

FRSAD Part 2:

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

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*.

The 'has appellation' relationship between *thema* and *nomen* in a controlled vocabulary:

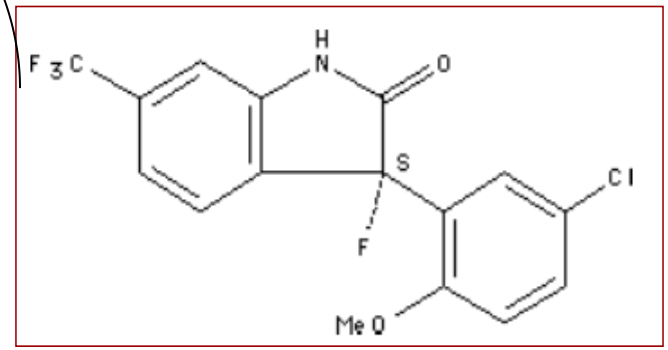
***NOMEN* = any sign or sequence of signs (alphanumeric characters, symbols, sound, etc.) that a *thema* is known by, referred to or addressed as.**

Example:

Nomen 1-8

```
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 (+).
```



Nomen 9

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

Basic elements in any KOS structure --analyzed from the FRSAD perspective

- Entities
  - Thema
  - Nomen
- Relationships:
  - Thema – to – nomen
  - Thema – to – thema
  - Nomen – to – nomen
- Attributes
  - Attributes of thema
  - Attributes of nomen

II. FRSAD conceptual model → implementation in data models for subject authority data

SKOS and the thema-nomen relationship model

1. SKOS and the thema-nomen relationship model

Simple Knowledge Organization System Reference, 2005, 2009

-- concept-centric

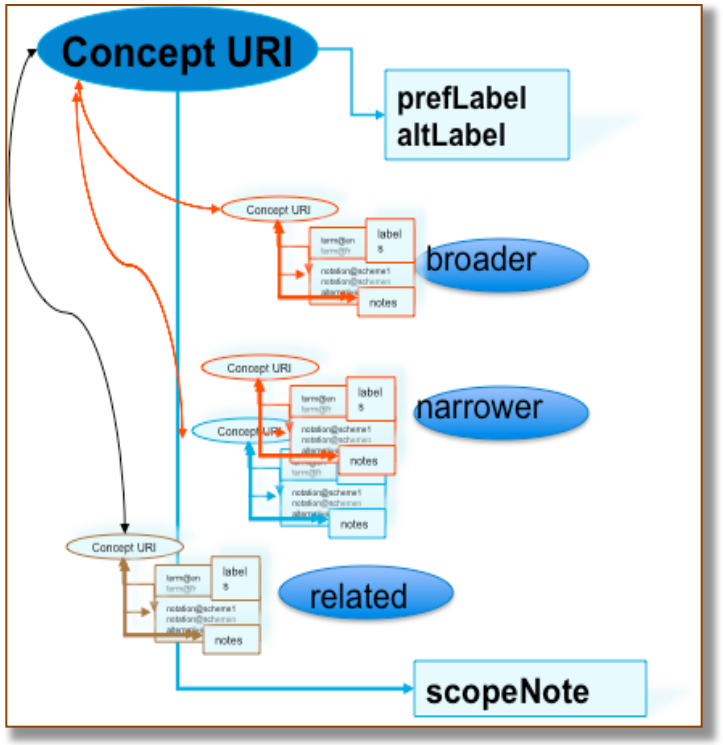
- Primitive objects are not labels; rather, they are concepts represented by labels.
- Rooted from thesaurus standards, where such an emphasis was not clearly stated or modeled due to the mix of relationships of concepts
  - E.g., thesauri seemed to mix of relationships of concepts (Broader Term (BT), Narrower Term (NT), and Related Term (RT)) and between the concept and its labels (Use and Used For (UF)).
- SKOS: for semantic relationships:
  - skos:broader and skos:narrower for hierarchical relationships
  - skos:related for associative (non-hierarchical) relationships

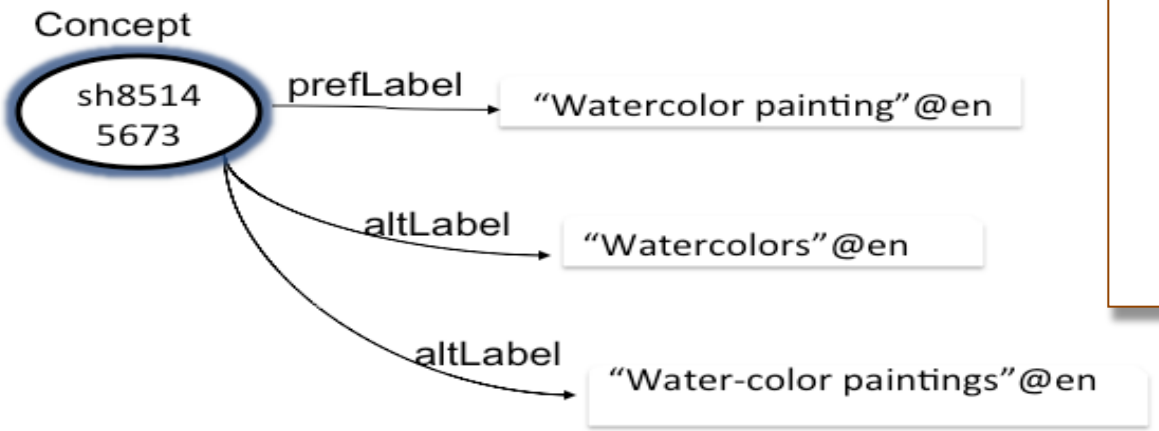
These convey the same relationships between themas defined in the FRSAD model.

SKOS Core (or no extension for labels)

Simple Knowledge Organization System Reference, 2005, 2009

- *Concept centered*
- *Labels are attributes of concepts*
- *Labels are literals; they do not have their own attributes.*





SKOS Core (or no extension for labels)

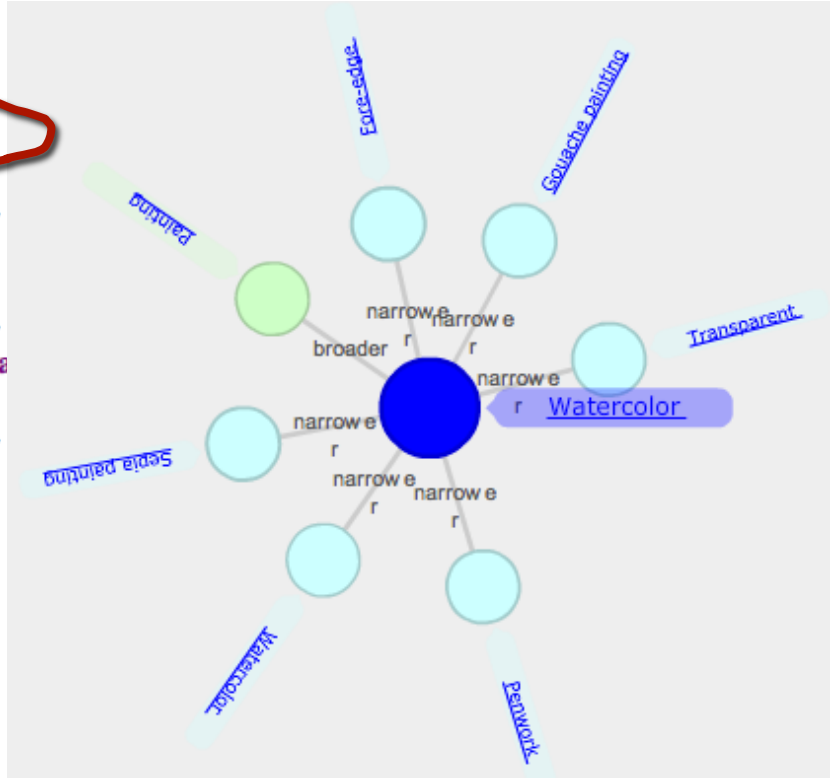
- KOS vocabularies that have implemented this SKOS core (or no extension for labels) model
  - Library of Congress Subject Headings (LCSH)

  at http://id.loc.gov/authorities/

A. Example from LCSH. http://id.loc.gov/authorities/sh85145673#concept

1. The first part of the record shows all related concepts. Each concept has a URI, and has a skos:preLabel. The cml:lang is a property of it.

```
– <rdf:RDF>
  – <rdf:Description rdf:about="http://id.loc.gov/authorities/sh85056013#concept">
      <skos:prefLabel xml:lang="en">Gouache painting</skos:prefLabel>
    </rdf:Description>
  – <rdf:Description rdf:about="http://id.loc.gov/authorities/sh85120108#concept">
      <skos:prefLabel xml:lang="en">Sepia painting</skos:prefLabel>
    </rdf:Description>
  – <rdf:Description rdf:about="http://id.loc.gov/authorities/sh2008006343#concept">
      <skos:prefLabel xml:lang="en">Penwork (Decoration)</skos:prefLabel>
    </rdf:Description>
  – <rdf:Description rdf:about="http://id.loc.gov/authorities/sh85096661#concept">
      <skos:prefLabel xml:lang="en">Painting</skos:prefLabel>
    </rdf:Description>
  – <rdf:Description rdf:about="http://id.loc.gov/authorities/sh85050481#concept">
      <skos:prefLabel xml:lang="en">Fore-edge painting</skos:prefLabel>
    </rdf:Description>
  – <rdf:Description rdf:about="http://id.loc.gov/authorities/sh89001408#concept">
      <skos:prefLabel xml:lang="en">Transparent watercolor painting</skos:prefLa
    </rdf:Description>
  – <rdf:Description rdf:about="http://id.loc.gov/authorities/sh99013408#concept">
      <skos:prefLabel xml:lang="en">Watercolor quilting</skos:prefLabel>
    </rdf:Description>
```



2. The second part of the XML/RDF record shows the labels and relationships around the concept (1/3)

RDF/XML record:

```
- <rdf:Description rdf:about="http://id.loc.gov/authorities/sh85145673#concept">
    <skos:narrower rdf:resource="http://id.loc.gov/authorities/sh85056013#concept"/>
    <skos:narrower rdf:resource="http://id.loc.gov/authorities/sh85120108#concept"/>
    <skos:narrower rdf:resource="http://id.loc.gov/authorities/sh2008006343#concept"/>
    <skos:narrower rdf:resource="http://id.loc.gov/authorities/sh85050481#concept"/>
    <skos:narrower rdf:resource="http://id.loc.gov/authorities/sh89001408#concept"/>
    <skos:narrower rdf:resource="http://id.loc.gov/authorities/sh99013408#concept"/>
    <skos:altLabel xml:lang="en">Watercolor paintings</skos:altLabel>
    <skos:altLabel xml:lang="en">Water-color paintings</skos:altLabel>
    <skos:altLabel xml:lang="en">Watercolors</skos:altLabel>
    <skos:altLabel xml:lang="en">Water-colors</skos:altLabel>
    <skos:altLabel xml:lang="en">Water-color painting</skos:altLabel>
    <skos:broader rdf:resource="http://id.loc.gov/authorities/sh85096661#concept"/>
    <skos:prefLabel xml:lang="en">Watercolor painting</skos:prefLabel>
    <dcterms:modified rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">1998-05-0(
    <owl:sameAs rdf:resource="info:lc/authorities/sh85145673"/>
    <skos:inScheme rdf:resource="http://id.loc.gov/authorities#topicalTerms"/>
    <skos:inScheme rdf:resource="http://id.loc.gov/authorities#conceptScheme"/>
    <rdf:type rdf:resource="http://www.w3.org/2004/02/skos/core#Concept"/>
    <skos:closeMatch rdf:resource="http://stitch.cs.vu.nl/vocabularies/rameau/ark:/12148/cb1193092
    <dcterms:created rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">1986-02-11
  </rdf:Description>
</rdf:RDF>
```

HTML display for human readers:

Watercolor painting

URI
<http://id.loc.gov/authorities/sh85145673#concept>

Type
Topical Term

Alternate Labels
- Water-color painting
- Water-color paintings
- Water-colors
- Watercolor paintings
- Watercolors

Broader Terms
- Painting

Narrower Terms
- Fore-edge painting
- Gouache painting
- Penwork (Decoration)
- Sepia painting
- Transparent watercolor painting
- Watercolor quilting

LC Classification
ND1700

Created
1986-02-11

Modified
1998-05-06 14:03:47

(cont.) 2. The second part of the XML/RDF record shows the labels and relationships around the concept (2/3)

RDF/XML record:

```
- <rdf:Description rdf:about="http://id.loc.gov/authorities/sh85145673#concept">
    <skos:narrower rdf:resource="http://id.loc.gov/authorities/sh85056013#concept"/>
    <skos:narrower rdf:resource="http://id.loc.gov/authorities/sh85120108#concept"/>
    <skos:narrower rdf:resource="http://id.loc.gov/authorities/sh2008006343#concept"/>
    <skos:narrower rdf:resource="http://id.loc.gov/authorities/sh85050481#concept"/>
    <skos:narrower rdf:resource="http://id.loc.gov/authorities/sh89001408#concept"/>
    <skos:narrower rdf:resource="http://id.loc.gov/authorities/sh99013408#concept"/>
    <skos:altLabel xml:lang="en">Watercolor paintings</skos:altLabel>
    <skos:altLabel xml:lang="en">Water-color paintings</skos:altLabel>
    <skos:altLabel xml:lang="en">Watercolors</skos:altLabel>
    <skos:altLabel xml:lang="en">Water-colors</skos:altLabel>
    <skos:altLabel xml:lang="en">Water-color painting</skos:altLabel>
    <skos:broader rdf:resource="http://id.loc.gov/authorities/sh85096661#concept"/>
    <skos:prefLabel xml:lang="en">Watercolor painting</skos:prefLabel>
    <dcterms:modified rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">1998-05-0(
    <owl:sameAs rdf:resource="info:lc/authorities/sh85145673"/>
    <skos:inScheme rdf:resource="http://id.loc.gov/authorities#topicalTerms"/>
    <skos:inScheme rdf:resource="http://id.loc.gov/authorities#conceptScheme"/>
    <rdf:type rdf:resource="http://www.w3.org/2004/02/skos/core#Concept"/>
    <skos:closeMatch rdf:resource="http://stitch.cs.vu.nl/vocabularies/rameau/ark:/12148/cb1193092
    <dcterms:created rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">1986-02-11T
  </rdf:Description>
</rdf:RDF>
```

HTML display for human readers:

Watercolor painting

URI
<http://id.loc.gov/authorities/sh85145673#concept>

Type
Topical Term

Alternate Labels
- Water-color painting
- Water-color paintings
- Water-colors
- Watercolor paintings
- Watercolors

Broader Terms
- Painting

Narrower Terms
- Fore-edge painting
- Gouache painting
- Penwork (Decoration)
- Sepia painting
- Transparent watercolor painting
- Watercolor quilting

LC Classification
ND1700

Created
1986-02-11

Modified
1998-05-06 14:03:47

(cont.) 2. The second part of the XML/RDF record shows the labels and relationships around the concept (3/3)

RDF/XML record:

```
- <rdf:Description rdf:about="http://id.loc.gov/authorities/sh85145673#concept">
    <skos:narrower rdf:resource="http://id.loc.gov/authorities/sh85056013#concept"/>
    <skos:narrower rdf:resource="http://id.loc.gov/authorities/sh85120108#concept"/>
    <skos:narrower rdf:resource="http://id.loc.gov/authorities/sh2008006343#concept"/>
    <skos:narrower rdf:resource="http://id.loc.gov/authorities/sh85050481#concept"/>
    <skos:narrower rdf:resource="http://id.loc.gov/authorities/sh89001408#concept"/>
    <skos:narrower rdf:resource="http://id.loc.gov/authorities/sh99013408#concept"/>
    <skos:altLabel xml:lang="en">Watercolor paintings</skos:altLabel>
    <skos:altLabel xml:lang="en">Water-color paintings</skos:altLabel>
    <skos:altLabel xml:lang="en">Watercolors</skos:altLabel>
    <skos:altLabel xml:lang="en">Water-colors</skos:altLabel>
    <skos:altLabel xml:lang="en">Water-color painting</skos:altLabel>
    <skos:broader rdf:resource="http://id.loc.gov/authorities/sh85096661#concept"/>
    <skos:prefLabel xml:lang="en">Watercolor painting</skos:prefLabel>
    <dcterms:modified rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">1998-05-0(
    <owl:sameAs rdf:resource="info:lc/authorities/sh85145673"/>
    <skos:inScheme rdf:resource="http://id.loc.gov/authorities#topicalTerms"/>
    <skos:inScheme rdf:resource="http://id.loc.gov/authorities#conceptScheme"/>
    <rdf:type rdf:resource="http://www.w3.org/2004/02/skos/core#Concept"/>
    <skos:closeMatch rdf:resource="http://stitch.cs.vu.nl/vocabularies/rameau/ark:/12148/cb1193092
    <dcterms:created rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">1986-02-11T
  </rdf:Description>
</rdf:RDF>
```

HTML display for human readers:

Watercolor painting

URI
<http://id.loc.gov/authorities/sh85145673#concept>

Type
Topical Term

Alternate Labels
- Water-color painting
- Water-color paintings
- Water-colors
- Watercolor paintings
- Watercolors

Broader Terms
- Painting

Narrower Terms
- Fore-edge painting
- Gouache painting
- Penwork (Decoration)
- Sepia painting
- Transparent watercolor painting
- Watercolor quilting

LC Classification
ND1700

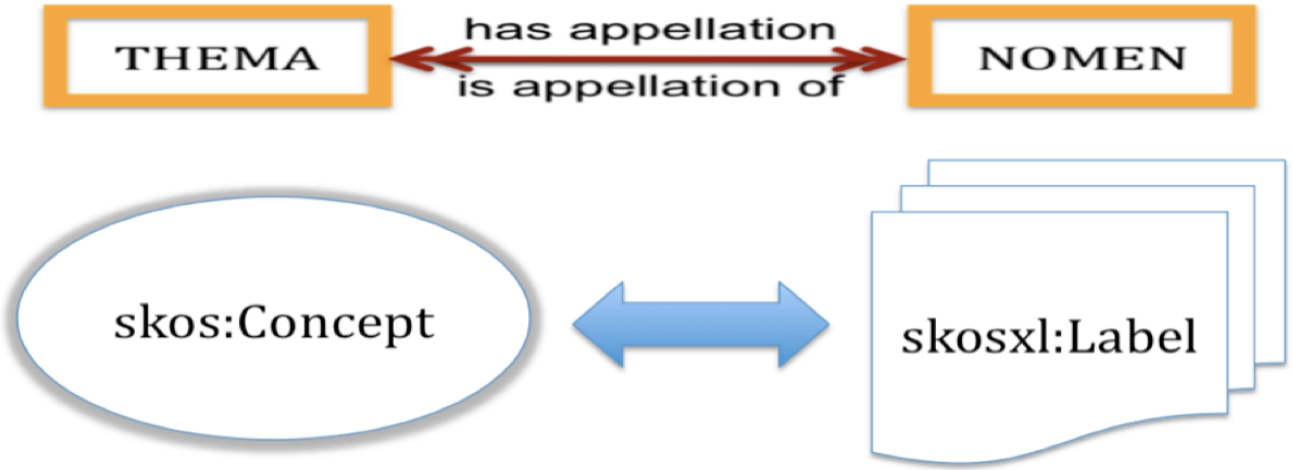
Created
1986-02-11

Modified
1998-05-06 14:03:47

(cont.) SKOS and the thema-nomen relationship model

2. SKOS eXtension for Labels (SKOS-XL) and relationships of nomens

-- *thema* corresponds to skos:Concept class
-- *nomen* corresponds to skosxl:Label



Ref: eXtension for Labels (SKOS-XL) , in: *SKOS Simple Knowledge Organization System Reference*, 2009, Appendix B

nomen as an entity

- The FRSAD model defines *nomen* as an entity.
- *Nomens* have attributes as well as relations between or among themselves while representing the same *thema*.
  - e.g., a preferred label of a concept may have various literal forms, synonyms, status of release, and administrative data
- FRSAD has provided a few common possible relationships and is flexible: implementation- or domain-specific relationships can be added.
- With the availability of SKOS-XL, such situations can be handled appropriately.

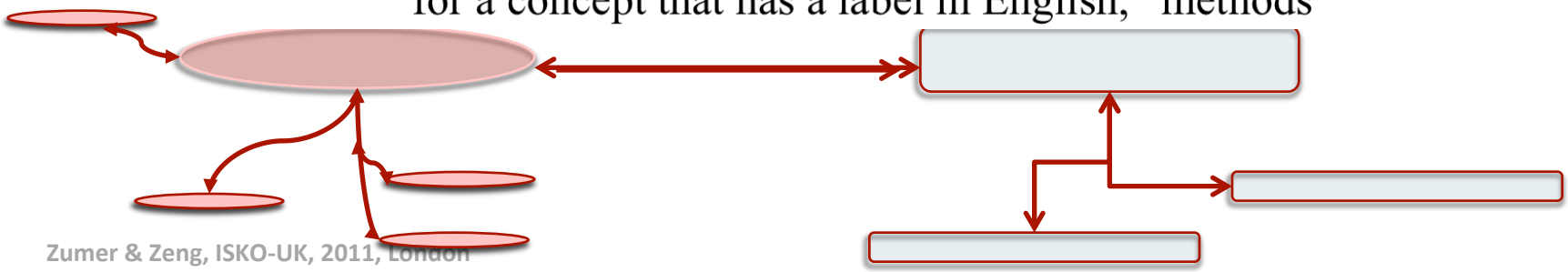
A Simplified Example

thema

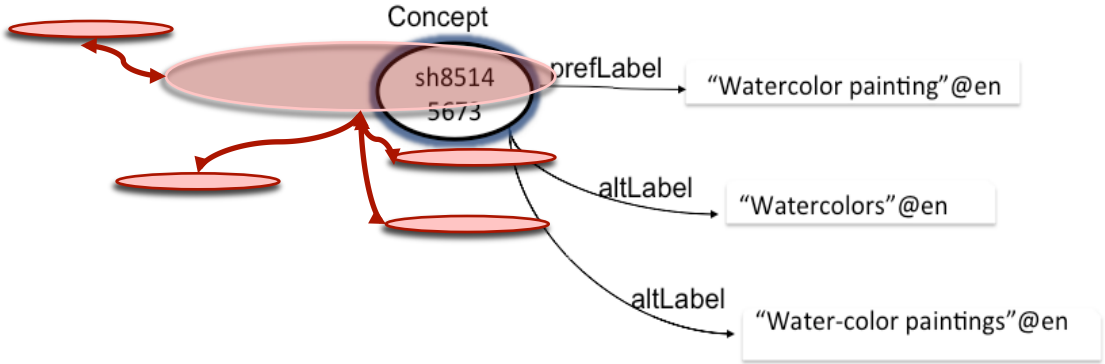
nomen

<xyz/sch/c_4788> <skos-xl#prefLabel> <xyz/sch/xl_en_123> ;
<xyz/sch/xl_en_123> <skos-xl#literalForm> "methods"@en ;
<xyz/sch/xl_en_123> <rdf-syntax-ns#type> <skos-xl#Label> ;
<xyz/sch/xl_en_123> <xyz/voc/hasSynonym> <xyz/xl_en_1285319878064> ;
<xyz/sch/xl_en_123> <xyz/voc/hasStatus>
"Published"^^<http://www.w3.org/2001/XMLSchema#string> ;
<xyz/sch/xl_en_123> <xyz/voc/hasDateCreated>
"1981-01-09 00:00:00"^^<http://www.w3.org/2001/XMLSchema#dateTime> ;
<xyz/sch/xl_en_123> <xyz/voc/hasCodeScheme>
"4788"^^<http://www.w3.org/2001/XMLSchema#int> .

Figure 7. Simplified statements, based on AGROVOC example, for a concept that has a label in English, "methods"

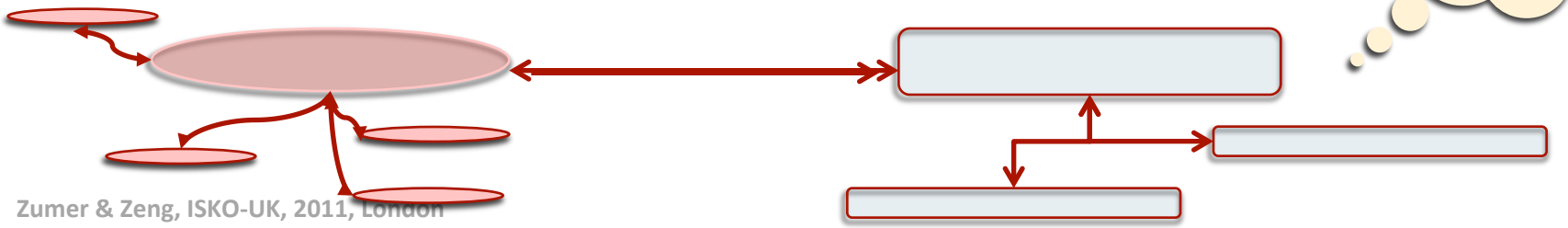


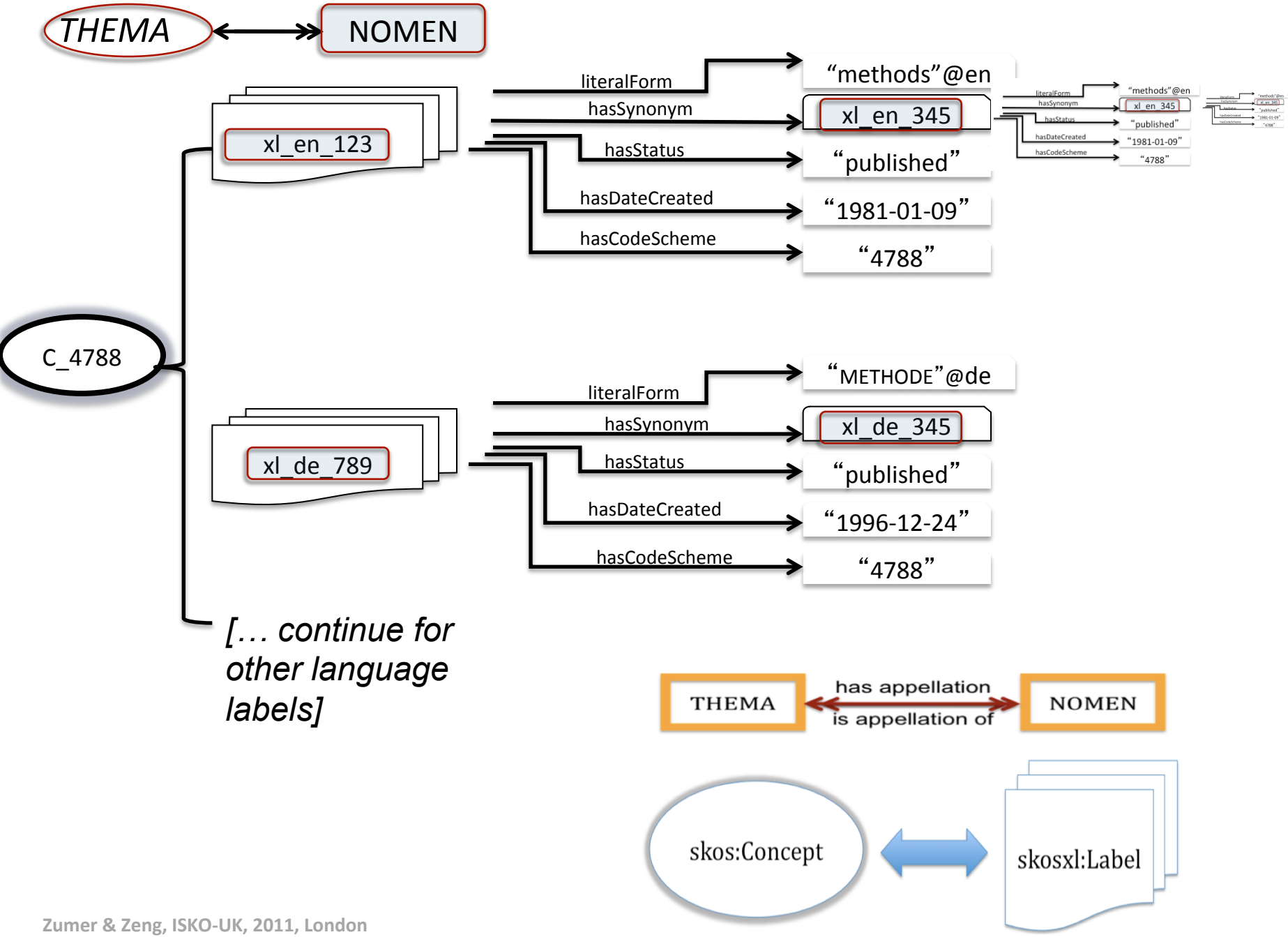
Comparing with model 1

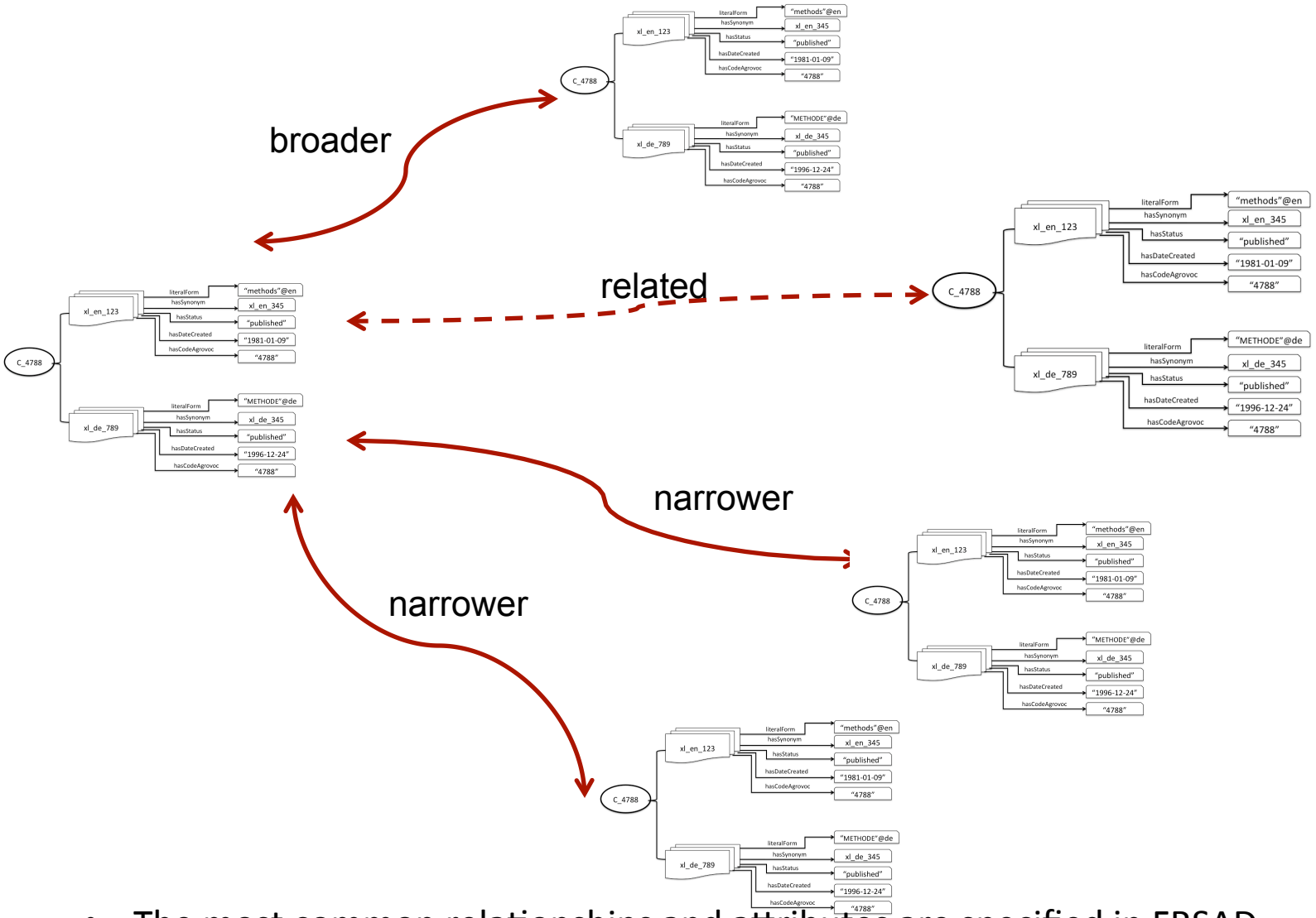


SKOS
(Core)

SKOS +
SKOSXL




THEMA
NOMEN
C_4788
xl_en_123
literalForm
hasSynonym
hasStatus
hasDateCreated
hasCodeScheme
"methods"@en
xl_en_345
"published"
"1981-01-09"
"4788"
xl_de_789
"METHODE"@de
xl_de_345
"published"
"1996-12-24"
"4788"
[… continue for other language labels]
THEMA
has appellation
is appellation of
NOMEN
skos:Concept
skosxl:Label



- The most common relationships and attributes are specified in FRSAD;
- Additional implementation- or domain-specific relationships and attributes can be added when needed.

Summary

- The two data models presented above demonstrate how the FRSAD model can be implemented。
- The power of the FRSAD conceptual model wields to meet the needs of both the conventional LIS environment and the emerging Linked Data environment.

References

- *AGROVOC Thesaurus* (Linked Data version 2011-). Food and Agriculture Organization of the United Nations. Retrieved from: http://aims.fao.org/website/Linked-Open-Data/sub
- Dextre Clarke, S. G. (2008). ISO 2788 + ISO 5964 + Much Energy = ISO 25964. *Bulletin of the American Society for Information Science and Technology*. 35(1): 31–33.
- *Functional Requirements for Subject Authority Data, A Conceptual Model* (2010). IFLA Working Group on Functional Requirements for Subject Authority Records (FRSAR). Retrieved from: http://www.ifla.org/en/node/1297.
- *Library of Congress Subject Headings* (Linked Data version 2009-). The Library of Congresss. Retrieved from: http://id.loc.gov/authorities/
- SKOS Core Vocabulary Specification. (2005). Eds. Miles, A. and Bechhofer, S. W3C Working Draft 10 May 2005. Retrieved from: http://www.w3.org/TR/2005/WD-swbp-skos-core-spec-20050510/
- SKOS eXtension for Labels (SKOS-XL). In: SKOS Simple Knowledge Organization System Reference (2009). Eds. Miles, A. and Bechhofer, S . W3C Recommendation. Appendix B. Retrieved from: http://www.w3.org/TR/skos-reference/ http://www.w3.org/TR/2009/REC-skos-reference-20090818/#xl
- SKOS Simple Knowledge Organization System Primer (2009). Eds. Isaac, A. and Summers, E. Retrieved from: http://www.w3.org/TR/skos-primer/
- SKOS Simple Knowledge Organization System Reference (2009). Eds. Miles, A. and Bechhofer, S . W3C Recommendation. Retrieved from: http://www.w3.org/TR/skos-reference/