

Building better controlled vocabularies

Guidance from ISO 25964

Jutta Lindenthal, Detlev Balzer

ISKO UK Conference, 2013

ISO 25964-1 can be regarded as a comprehensive guideline for building high-quality thesauri for information retrieval.

This presentation will look at a selection of quality issues addressed by the standard concerning the following:

- **Concept relationships** suitable for exploding search
- **Facets** and **node labels** to show multi-dimensionality
- **Concept groups** offering complementary navigation paths
- **Compound equivalence** to represent split compounds
- **Duplicate control** and **qualifiers** to avoid ambiguity
- **Top concepts** as seen by SKOS and the ISO standard

Overview

Real-world implementations:

Advanced features adopted by some existing thesauri

Quality issues that can still go unnoticed

Attempts to provide guidance to thesaurus developers:

(Mis)understanding thesaurus construction principles

Using ISO 25964 – A reference with latitudes

Implementation issues:

The ISO 25964 data model as an extension to SKOS

Ambiguities for implementers

Examples from three existing thesauri were selected to demonstrate the need for advanced thesaurus features:

- The [Art & Architecture Thesaurus](#)[®] contains highly structured hierarchy chains, making use of many features described by the standard.
- An example for splitting compounds is drawn from the [NAL Thesaurus](#) of the U.S. National Agricultural Library.
- An example from the [STW Thesurus for Economics](#) is included to show the application of a classified structure of subject fields for grouping thesaurus concepts by themes. These fields resemble what the ISO standard defines as concept groups.

While general considerations for thesaurus construction can be found in textbooks, more specific guidance with respect to current standards and specifications is still scarce.

- The EU-funded [Linked Heritage](#) project has attempted to fill this void with a [guideline](#) which, although mentioning some of the relevant questions, arrives mostly at wrong conclusions.

Concept relationships

. dairy products
.. milk
... fat milk

Thesaurus standards distinguish between two basic relationships between concepts: hierarchical and associative.

According to ISO 25964-1, hierarchical relationships hold *“between a pair of concepts when the scope of one of them falls completely within the scope of the other. It should be based on degrees or levels of superordination and subordination, where the superordinate concept represents a class or whole, and subordinate concepts refer to its members or parts”* [10.2.1]. **Fat milk** is a kind of **milk** and that in turn is a **dairy product**.

Associative relationships cover *“associations between pairs of concepts that are not related hierarchically, but are semantically or conceptually associated to such an extent that the link between them needs to be made explicit in the thesaurus, on the grounds that it may suggest additional or alternative terms for use in indexing or retrieval”* [10.3.1].

Building a thesaurus the hop way ;-)

. animal material
 ... <animal material by form or function>
 <excretions and secretions>
 milk
 cream (milk)
 ice cream
 cheese
 cheesecake
 cheesecake lifter
 handle of a cheese cake lifter

Source of blue labels: AAT® hierarchy of [ice cream](#)

**E
X
P
L
O
D
E
?**

The ISO standard defines criteria which hierarchical thesaurus relationships should meet. These recommendations cover

- decisions if a relationship is hierarchical or associative;
- generic hierarchies which are transitive;
- generic relationships as a pre-requisite for search explosion.

The above example is likely to be found in real-life thesauri. The terms in black are made up, but similar examples can actually be found in existing thesauri.

Once the logical hierarchy is broken, thesaurus maintainers may tend to carry on by adding arbitrary descriptors, ending up in a mixture of merely associated concepts. These „hierarchy chains“ are not suitable for a search explosion (expanded search over hierarchies), since the transitive closure becomes unpredictable.

Where transitivity fails (and notes are important)

.. materials (matter) ← **Note:** The matter or substance from which a thing is or may be made; the tangible substance that goes into the makeup of a physical object. [...]

.... animal material

..... <excretions and secretions>

..... milk ← **Note:** White or yellowish white fluid secreted by the mammary glands of female mammals.

..... ice cream

Source: AAT® Online, Hierarchy of [ice cream](#) and notes of the resp. concepts [2013-06-26]

The model of the ISO standard provides scope notes for concepts, though the standard does not mandate for the usage of scope notes or definitions. However, scope notes and/or definitions are crucial for the clarification of the intended meaning in almost all cases. The presence of such notes should be an element in quality metrics for thesauri.

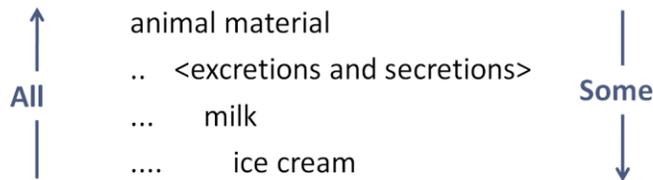
The scope note for **milk** in the above example restricts the meaning to glandular secretion, thus excluding dairy products.

An automatic procedure for detecting missing notes is included in Christian Mader's [qSKOS](#) test suite.

Is this relationship a generic one?

The ISO standard recommends an **all-and-some test** to determine if a hierarchical relationship is a generic one, thus being transitive.

So, let's perform an all-and-some test:



The all-and-some test fails in this case because the entailment „all ice cream is (a kind of) milk“ fails. Transitivity does not hold because the entailment „all ice cream is (a kind of) excretions and excretions“ is not true.

The standard distinguishes between types of hierarchies:

- **unspecified hierarchical relationship** of broader and narrower concepts (generalised role). This can only be tested for cycles.
- **specified hierarchical relationships** (specific roles)
 - the **generic relationship**. This type is transitive and holds if each pair of concepts passes the all-and-some test.
 - the **hierarchical whole-part relationship**. This type is transitive if the recommendations of ISO 25964-1 are followed.
 - the **instance relationship**. This type is not transitive.

Transitivity does not hold for hierarchy chains in which generic and whole-part relationships occur together.

Note that this test cannot be done by machines. A machine could, however, support this intellectual exercise by asking for confirmation of some entailments.

Are all these generic relationships?

The AAT® Online marks types of hierarchical relationships, namely the generic one (G). So called *guide terms*, displayed in angle brackets, are part of the hierarchy chain of generic relationships.

```

. materials (matter) (G)
.... <materials by origin> (G)
..... <biological material> (G)
..... animal material (G)
..... <animal material by form or function> (G)
..... <excretions and secretions> (G)
..... milk (G)
..... ice cream (G)

```

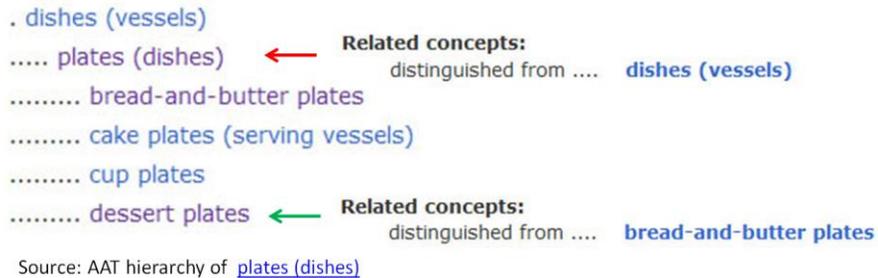
Source: AAT® hierarchical position of [ice cream](#) [2013-07-02]

The two AAT® guide terms, [<materials by origin>](#) and [<animal material by form or function>](#), correspond to node labels as defined by the ISO standard. As such, they are not allowed to have BT/NT relationships of any kind with concepts.

The two AAT® guide terms, [<biological material>](#) and [<excretions and secretions>](#), are different from node labels as defined by the standard in that they do not express a fundamental facet or characteristic of division. These are rather concepts which are not used for indexing, often because they do not occur in common language or because they lack literary warrant.

Check for inadmissible relationships

In this example the concepts labelled [plates \(dishes\)](#) and [dishes \(vessels\)](#) are hierarchically linked and linked as related concepts in the AAT®. The siblings [dessert plates](#) and [bread-and-butter plates](#) are associated as related concepts.



The ISO standard suggests validation checks in order to prevent inadmissible relationship combinations such as concepts that are connected by more than one of the basic relationships: Concept A must not be linked to concept B by a hierarchical and an associative relationship.

In the above example the relationships between [dishes \(vessels\)](#) and [plates \(dishes\)](#) violate the standard. Should this rule also apply if the „associative relationship“ is specialised?

Associative relationships between sibling terms, such as the one between [bread-and-butter plates](#) and [dessert plates](#) are not inadmissible, though redundant in many cases.

An automatic procedure for detecting "relation clashes" is included in Christian Mader's [qSKOS](#) test suite.

Facets and node labels

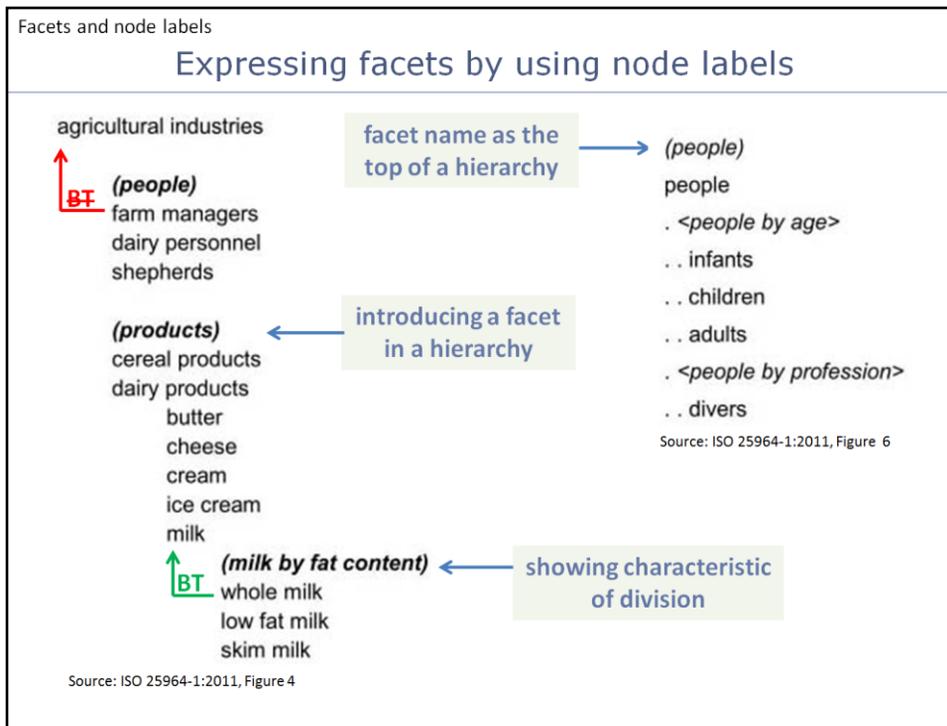
```
products
.. dairy products
... milk
.... (milk by fat content)
..... whole milk
```

ISO 25964-1 models a class **ThesaurusArray** to group concepts which are hierarchically linked. A thesaurus array is indicated by a **node label** showing how the concepts have been arranged.

A node Label “contains one of two different types of information:

- a) the name of a facet to which following terms belong; or
- b) the attribute or characteristic of division by which an array of sibling concepts has been sorted or grouped.” [ISO 25964-1, 2.38]

The class *ThesaurusArray* is mapped to *skos:Collection*.



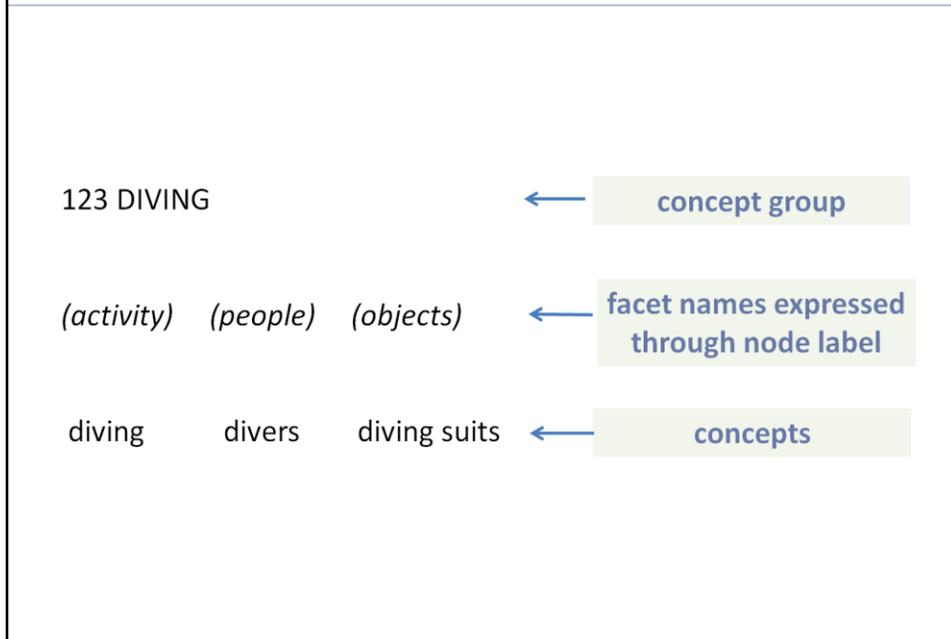
ISO 25964-1 addresses three cases in which node labels can be used to display facets or sub-facets:

1. They label a facet as the top of a hierarchy.
2. They are inserted in a hierarchy to introduce a new facet by which the subordinate concepts are arranged.
3. They are inserted as node labels to elicit the characteristic of division by which sibling concepts (member of array) are grouped. All subordinate concepts are narrower concepts of the superordinate concept, in this case, „milk“.

The model of the ISO standard does not explicitly distinguish between these different types of node labels. However, it is important to consider that a true hierarchical relationship holds between sibling concepts grouped by a characteristic of division, e.g. „whole milk“, and its superordinate concept, e.g. „milk“. In contrast, this is not true for concepts which are grouped by node labels showing facet names. Thus it should be taken into consideration to introduce a type distinction for *ThesaurusArray*.

Distinguishing types of node labels would allow for different views. In some cases it can be useful to omit node labels showing characteristics of division. Omitting these is possible without compromising the semantic coherence of the hierarchy chain, while omitting facet names would lead to implausible groupings.

Displaying node labels in context



According to the ISO data model, node labels cannot be directly related to concept groups. Views such as the one above can, however, be constructed algorithmically.

Concept groups

ice cream

CC P.20.02 Milk and Dairy Products

In true generic hierarchies, thematically related concepts are often scattered across many branches of the hierarchy. Associative relationships allow for lateral connection of concepts, but are rarely suitable for synoptic views of thematically related concepts.

Concept groups allow for compiling concepts from different facets and hierarchies under a common name.

Neither membership in a concept group, nor the nesting of concept groups can be equated to a BT/NT relationship.

„Many thesauri group concepts using a classification structure that exists in parallel to the hierarchies of thesaurus concepts based on BT/NT relationships. Groups created by the classification are often based on disciplines, subject areas or areas of business activity. They are sometimes called "subject categories", "themes", "domains", "groups", "subsets" or "microthesauri". The model provides for all of these by providing the classes `ConceptGroup` and `ConceptGroupLabel` and the specific type may be indicated by the attribute `conceptGroupType`. In general, there is not a BT/NT relationship between a `ConceptGroup` and the concepts that it contains.” [ISO 25964-1, 15.2.18]

The class `ConceptGroup` is mapped to `skos:Collection`.

Concept scheme, microthesaurus, array, ... ?

My terminology (ex: Architecture)	Thesaurus	Athena Format	Explanation
Micro-Thesaurus ex: Architecture	Micro-thesaurus ex: Architecture	skos:ConceptScheme (class) skos:hasTopConcept (property) ex: Architecture	If your terminology has a micro-thesaurus on Architecture, you can describe it as a concept Scheme according to the SKOS model.
Group of terms ex: Buildings	Thesaurus Array ex: Buildings	skos:Collections (class) ex: Buildings	If your terminology has thematic or other specific groups of terms, the SKOS Collections class allows you to reproduce these groups of terms.

Source: <http://www.linkedheritage.eu/getFile.php?id=244>, p. 108

The Linked Heritage guideline recommends to use elements like „Micro-Thesaurus“ and „ThesaurusArray“ to group concepts (cf. <http://www.linkedheritage.eu/getFile.php?id=244>, p. 108). This recommendation is misleading in several ways. Neither is the notion of „Micro-Thesaurus“ clarified nor is any advice given how to connect microthesauri to each other. It is not evident why a „Micro-Thesaurus“ is mapped to `skos:ConceptScheme`.

The guideline recommends microthesauri such as „Monuments“, „Habitations“, and „Architecture“ for structuring the vocabulary. If each of these microthesauri were represented as `skos:ConceptScheme`, semantically connected concepts would be scattered across many different schemes.

The relationship between a „Micro-Thesaurus“ and a so called „Group of terms“ remains unclear as well. Furthermore, the example given here for a „Group of terms“ is neither a fundamental facet nor does it show a characteristic of division; thus a mapping to „ThesaurusArray“ is wrong.

According to the standard, „Group of terms“ and „Micro-Thesaurus“ would each be modelled as *ConceptGroup*, the latter of type „microthesaurus“.

Subject Categories as example for concept groups

Ice cream 

Speiseeis (german)

Subject Categories

- P.20.02 Milk and Dairy Products ▼
- P.20.07 Luxury Foodstuffs and Tobacco ▼

P Commodities

P.20 Food and Tobacco

P.20.02 Milk and Dairy Products

P.20.07 Luxury Foodstuffs and Tobacco

Source: STW Thesaurus entry [Ice cream](#)

The STW Thesaurus for Economics assigns concepts to „Subject Categories“, e.g. [Ice cream](#) to the categories [P.20.02 Milk and Dairy Products](#) and [P.20.07 Luxury Foodstuffs and Tobacco](#). Note that concepts can be assigned to multiple categories. The categories are nested and thus form a chain of superordinate and subordinate categories.

Subject categories can be represented as *concept groups* as defined by the standard. The concept groups are nested by a *hasSubgroup/hasSupergroup* relationship in the ISO model.

The *hasSubgroup* relationship can be mapped to *skos:member*.

Concept scheme or groups and sub-groups?

STW Thesaurus for Economics ← **concept scheme**

P Commodities ← **concept group**

hasSubgroup

P.20 Food and Tobacco ← **concept group**

P.20.02 Milk and Dairy Products

P.20.07 Luxury Foodstuffs and Tobacco

Distinguishing concept schemes within a thesaurus may be justified in cases where these can be used out of context of the vocabulary as a whole. In this case, a partial thesaurus would require its own set of metadata.

Declaring **P Commodities** e.g. to be a *concept scheme* may be justified in case the members of this group shall be used independently from the entire thesaurus.

Ephemeral groupings

G Ice Cream**G 1 Ice cream parlor**

ice cream parlor chairs

ice cream scoops

jukebox

G 2 Ice cream ingredients

milk

sugar

If you feel like collecting *ice cream scoops* and *ice cream parlor chairs* under an *ice cream* theme (as you would perhaps do in a museum exhibition), then a concept group is the tool of choice. You can even organise your *ice cream* theme in sub-themes such as *ice cream ingredients* and *ice cream parlors* by creating nested concept groups.

Evidently, such a hierarchy of concept groups cannot be expected to fulfill the transitivity criterion required for "exploded" searches.

Compound equivalence

contaminated milk
USE+ food contamination
USE+ milk

ISO 25964-1 enumerates cases in which splitting a complex concept should be taken into account:

- when the concept is quite specific and falls outside the core scope of the thesaurus;
- if very few documents are likely to be indexed with the proposed compound term;
- if the focus is qualified by more than one difference;
- if the focus represents a property, part or component of the difference. In this case it is likely that the compound would recur in almost infinite combinations throughout the vocabulary.

[cf. ISO 25964-1:2011, 7.3.3 and [Art & Architecture Thesaurus \(AAT\): An In-Depth Look](#)]

Decomposition of concepts – a misleading guidance

B3 DEFINE WITH PRECISION THE LABELS EXPRESSING CONCEPTS

Example

Your terminology has a concept “Musical instrument” in order to define with precision the labels of your concepts, you decompose this concept into two concepts: “Music” and “Instrument”

```
Ex:music rdf:type skos:Concept
skos:prefLabel « music »@en;
skos :prefLabel « musique »@fr.
ex :music skos :narrower ex :intrument.
```

Source: <http://www.linkedheritage.eu/getFile.php?id=244>.

The Linked Heritage guideline recommends: *“If there are compound terms in your terminology, try as much as possible to decompose them in order to get to a simple form.”* (Cf. <http://www.linkedheritage.eu/getFile.php?id=244>)

This example is erroneous and misleading in several respects. A decomposition of compound concepts is

- a) neither used to define labels
- b) nor is one of the component concepts a narrower one of the other.

Example of compound equivalence

milk**Persistent URI:**<http://lod.nal.usda.gov/nalt/631>**RDF/XML Format:**<http://lod.nal.usda.gov/nalt/631.rdf>**Used For**cow milk
milk, cow**USE FOR And Type**contaminated milk
contamination of milk
milk-borne diseases
milk contamination
milkborne diseases**Broader Term**dairy products
maternal milk**Narrower Term**low fat milk
skim milk
whole milk***contaminated milk*****USE And Type**food contamination
milk**Term Number**

153840

contamination of milk**USE And Type**food contamination
milk**Term Number**

143715

milk contamination**USE And Type**food contamination
milk**Term Number**

142454

This example from the NAL Thesaurus has three compound terms referring to the same combination of simple concepts.

Retrieving split compound concepts assumes that a retrieval system is aware of compound equivalences and either

- prompts the user to rewrite the query as suggested by the thesaurus, or
- rewrites the query transparently by evaluating the USE+ relationships.

Compound equivalence according to the standard

```
:coal_mining
  a iso-thes:CompoundEquivalence ;
  iso-thes:plusUF :t_coal_mining_en ;
  iso-thes:plusUSE :t_coal_en, :t_mining_en .

:coal
  a iso-thes:Concept ;
  iso-thes:hasPreferredLabel :t_coal_en ;
  iso-thes:splitAltLabel :t_coal_mining_en .

:mining
  a iso-thes:Concept ;
  iso-thes:hasPreferredLabel :t_mining_en ;
  iso-thes:splitAltLabel :t_coal_mining_en .

:t_coal_en
  a iso-thes:PreferredTerm ;
  xl:literalForm "Coal"@en .

:t_mining_en
  a iso-thes:PreferredTerm ;
  xl:literalForm "Mining"@en .

:t_coal_mining_en
  a iso-thes:SplitNonPreferredTerm ;
  xl:literalForm "Coal mining"@en ;
  iso-thes:plusUseTerm :t_coal_en, :t_mining_en .
```

This is an RDF/Turtle rendering of the „coal mining“ example from the ISO standard, using the proposed iso-thes specification as of June, 2012.

The resulting RDF graph has multiple paths between the source term and the target concepts, which can lead to decision problems when such structures are queried. It also requires each term to be uniquely identified.

Concept-based compound equivalence

Introducing a *CompoundConcept* class and a *useplus* relationship, we can render the „coal mining“ example as follows:

```
:coal_mining
  a :CompoundConcept ;
  iso-thes:hasPreferredLabel [ xl:literalForm "Coal mining"@en ] ;
  :useplus :coal ;
  :useplus :mining .

:coal
  a iso-thes:Concept ;
  iso-thes:hasPreferredLabel [ xl:literalForm "Coal"@en ] .

:mining
  a iso-thes:Concept ;
  iso-thes:hasPreferredLabel [ xl:literalForm "Mining"@en ] .
```

Since in indexing and retrieval, split compounds are usually converted to an intersection of the constituent concepts, we could also speak of compound **concepts**.

Our preliminary explorations suggest that a concept-based modelling of compound equivalence, as in the example above, is likely to meet all requirements addressed by the term-based model. Moreover, it can be expressed with fewer relationship types, a smaller number of nodes in the knowledge graph, and without multiple paths between the source term and the target concepts or terms.

If the need arises, a concept-based compound can be easily transformed into an indexing concept. Such transformation is much more complex in the ISO data model.

Resolving ambiguity

milk (dairy product)
milk (glandular secretion)

Clarifying the meaning of descriptors is an essential task of controlled vocabularies.

The ISO standard recommends that no duplicate terms should be entered for the same language, whether a preferred or a non-preferred term. Thus, a qualifier should be added to each homographic term.

“Homographs (sometimes referred to by the broader term “homonyms”) are words with the same spelling but

different meanings. [...] When homographs are needed as thesaurus terms, the meaning of each term should be clarified and the traditional way to do this is by adding to it a qualifier in parentheses. The qualifier should be as brief as possible, ideally consisting of one word. Often a broader term, the qualifier should indicate the context or subject area to which the concept belongs. It forms part of the term and does not serve as a scope note.” [ISO 25964-1, 6.2.2]

The issue of duplicate control is partly covered by the test battery in qSKOS:

<https://github.com/cmader/qSKOS>.

Qualifiers and duplicate control

Why Scope Notes are so important

- Homographs: are words that are spelled the same yet have different meaning.
> Example: The French term 'Bois' has two meanings, both 'Wood' and 'Antlers'

- Appearance of the same term more than ones in the thesaurus.

➤ Example:

Animal > Antler > Antelope

Animal > Skin > Antelope

(French)

Animal > Bois

Fossile > Bois (could both be Fossil wood or Fossil antlers)

Végétal > Bois

Source: <http://de.slideshare.net/EuropeanaLocal/roxanne-wyns-belgium-2009>

Terms which have two or more meanings (homographs) are common in natural language. Homographs cause problems in thesaurus maintenance, indexing and retrieval.

The Linked Heritage guideline is aware of the problem of homography, and recommends to solve ambiguity by providing scope notes. This recommendation does not take into account that a scope note usually is not processed.

Please note that the hierarchical relationships in the above example, drawn from <http://de.slideshare.net/EuropeanaLocal/roxanne-wyns-belgium-2009>, are erroneous as well.

Duplicate control

swell boxes

(<swell components>, <swells and swell components>, ...

zwekkasten

swells (keyboard instrument components)

(<swells and swell components>, ...

zwekkasten

Source: AAT search result for [zwekkasten](#)

- . <swells and swell components>
- swells (keyboard instrument components)
- <swell components>
- swell boxes
- swell shutters

Duplicate control is essential to avoid using the same term for two different concepts, as shown in the example [zwekkasten](#) from the AAT® Online.

Resolving ambiguity

Fuzzy clarification vs. controlled qualifiers

Id: 863

Label: **archives (buildings) (en)**

Type: Concept

Denkmale
Denkmäler

Geografie
Geographie

karolingisch-ottonisch
karolingisch, ottonisch

Musikinstrument
Musikinstrument

Flügel (Instrument)
Flügel (Musikinstrument)

Variants of qualifiers.

Basic data
Relations
Mappings
Notes
Forum

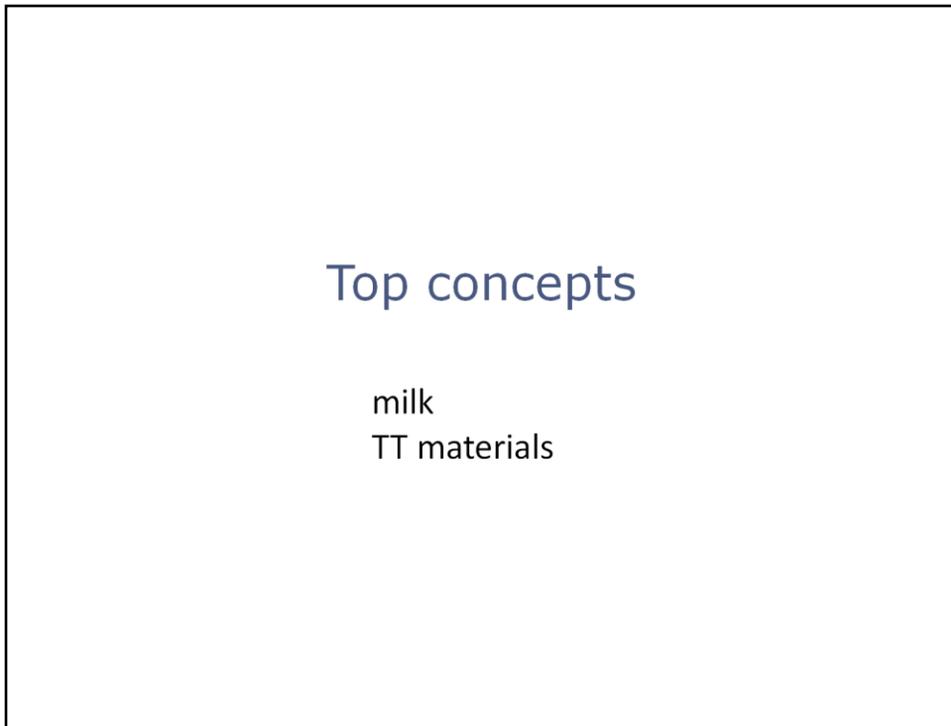
+ New term

Label	Qualifier
archives	<div style="border: 1px solid #ccc; padding: 2px;"> buildings <ul style="list-style-type: none"> buildings single built works structures symbols visual works </div>

Software: xTree,
digiCULT-Verbund eG, Kiel

According to the ISO standard, qualifiers are added to the homograph in parentheses, forming an integral part of the term. This method can lead to the accumulation of arbitrary variants of qualifiers as in the example above (right). Duplicates can go unnoticed if they are disambiguated by varying qualifiers with similar intent.

Spurious variation in qualifiers can be prevented e.g. by restricting these to terms from a controlled vocabulary, as in the above (left) example from the xTree thesaurus management platform.

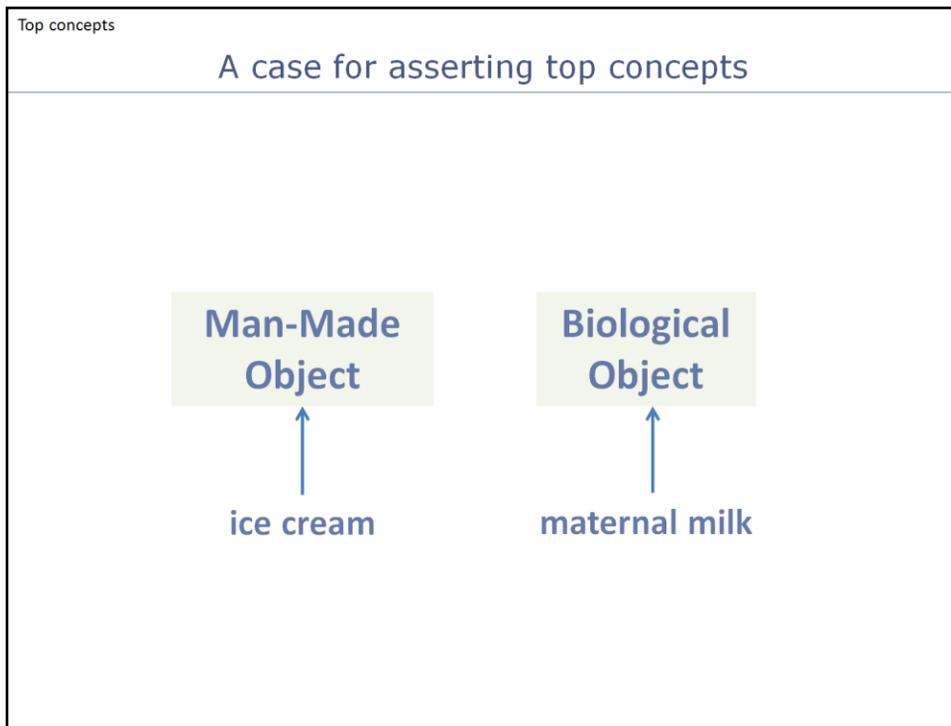


The ISO model defines a *hasTopConcept* relationship between thesaurus concepts and its inferred topmost concept of the hierarchy. The SKOS model instead defines a *hasTopConcept* relationship between a concept scheme and a concept, thus allowing arbitrary assertions about top concepts.

Manually assigning a **top concept** to individual nodes in the hierarchy is likely to produce errors that, even though they can be detected algorithmically, cannot be resolved without human intervention.

Under this assumption, a top concept relationship, if required, should only be inferred by following the BT axis within the hierarchy tree.

Providing *skos:hasTopConcept* as navigation aids should therefore not be considered a necessary quality criterion. Entries to facilitate browsing can be provided by **concept groups** (*skos:Collection*).



Assigning concepts to disjoint facets declared in advance to be *top concepts* could help avoid errors such as in the ice cream example. **Ice cream** could be assigned e.g. to a class **Man-Made Object** and **maternal milk** to a class **Biological Object** (these classes are defined by the CIDOC Conceptual Reference Model).

Since concepts which belong to different basic categories must not be connected in a broader/narrower hierarchy, such mismatches can be detected automatically.

In this case, facets are declared as a *skos:topConceptOf* and broader concepts should be refused by the software*. This method should be recommended for facets as top of a hierarchy only.

* This issue is covered by the test suite in qSKOS: <https://github.com/cmader/qSKOS>

What else the machine could do

palace
BT monument

monument
UF *construction*
UF *hut*
BT architecture
NT palace
RT building

[Palaces \(sh85096913\)](#) LCSH
BT Buildings

[Palast \(4044394-2\)](#) GND
[...]
BT Bauwerk

[Palais \(11937594\)](#) RAMEAU
BT Constructions

[palaces \(300005734\)](#) AAT
[...]
BT (single built works)

Example based on: <http://www.linkedheritage.eu/getFile.php?id=244>

Follow a mapping statement to determine if the mapped concept in another thesaurus has similar relationships. Allow for degrees of similarity, but warn if contradictions are found.

Outlook

Propose, examine, discuss and evaluate additions or modifications to the ISO 25964 data model, particularly in the areas of

- subtyping the **ThesaurusArray** entity
- specifying constraints for **concept relationships** in different roles
- finding sensible constraints for the usage of **ConceptGroup**
- exploring if a concept-based approach to the **CompoundEquivalence** relationship is a viable alternative to the term-based approach
- finding more options on **term disambiguation**
- clarifying the intent behind the **hasTopConcept** relationship in the ISO model and in SKOS
- examining if a class of **NonIndexingConcept** is useful for particular types of structural nodes

Additionally, for the techies,

- explore graph theoretical algorithms for finding computable quality metrics for more complex thesauri, drawing upon pioneering work on the SKOS data model by Mader et al. and the Finnish SeCo group.
- explore new ways in which thesaurus management tools can support intellectual quality control (complementary data views, "quizzing", etc.)
- explore "hybrid" approaches to quality control such as exploiting mappings to reference thesauri which machines can use to reveal contradictory connection paths between concepts.

References

- ISO 25964-1:2011: Thesauri and interoperability with other vocabularies
Part 1: Thesauri for information retrieval
http://www.iso.org/iso/catalogue_detail.htm?csnumber=53657
- Data model as UML diagram
http://www.niso.org/schemas/iso25964/Model_2011-06-02.jpg
- XML schema
http://www.niso.org/schemas/iso25964/iso25964-1_v1.4.xsd
- Test document of the XML schema
http://www.niso.org/schemas/iso25964/example_multi_lingual_08-09T15-21.xml
- Mader, Christian: Quality Issues
<https://github.com/cmader/qSKOS/wiki/Quality-Issues>
- Concept-based compound equivalence (LabNote by Detlev Balzer)
<http://www.jlindenthal.de/tmp/compound-equivalence-labnote-v1.pdf>

- ISO 25964-1:2011: Thesauri and interoperability with other vocabularies
Part 1: Thesauri for information retrieval
http://www.iso.org/iso/catalogue_detail.htm?csnumber=53657
- Data model as UML diagram
http://www.niso.org/schemas/iso25964/Model_2011-06-02.jpg
- XML schema
http://www.niso.org/schemas/iso25964/iso25964-1_v1.4.xsd
- Test document of the XML schema
http://www.niso.org/schemas/iso25964/example_multi_lingual_08-09T15-21.xml
- Mader, Christian: Quality Issues
<https://github.com/cmader/qSKOS/wiki/Quality-Issues>
- Concept-based compound equivalence (LabNote by Detlev Balzer)
<http://www.jlindenthal.de/tmp/compound-equivalence-labnote-v1.pdf>