Rigorous facet analysis as the basis for constructing knowledge organization systems (KOS) of all kinds.

Leonard Will

Abstract

Construction of any kind of knowledge organization system starts from the fundamental building blocks, which are concepts. Once a representative set of concepts has been chosen and each one clearly defined, their hierarchical relationships can be established. Fundamental facets emerge from these hierarchies. In a post-coordinate scheme, such as a thesaurus, associative relationships between concepts in different facets can be defined. In a pre-coordinate scheme, such as a classification, concepts from two or more facets can also be combined, following rules of citation order, to express compound subjects. Citation order has to take account of the role which each concept plays in the compound as well as the facet to which it belongs, and these roles and facet names should be explicitly shown in schedules to distinguish such combinations from hierarchical subordination. The first-cited facet need not be discipline, or field of activity, but can, for example, be “things” or “phenomena”. Hierarchies which are not based on genus-species relationships need special consideration; whole-part relationships, for example, are best restricted to concepts which are individual instances rather than classes.

1. Introduction

In May 1955 the Classification Research Group (1955) issued a manifesto entitled *The need for a faceted classification as the basis of all methods of information retrieval* and discussed the development of a general classification system based on faceted principles. The theme was picked up 51 years later by Vanda Broughton (2006, 2012), and this paper builds on that review, combining it with other approaches expressed by Claudio Gnoli (2011) in his discussion of the León Manifesto (2007) and the later thoughts of Brian Vickery (2009).

Many systems have been constructed on faceted principles, but in recent years the meaning of “faceted” or “facet analysis”, has been given many interpretations so that confusion arises in discussion. This paper tries to resolve this confusion and show that one clearly-defined meaning can be used as the basis for building the two distinct kinds of knowledge organization system – post- and pre-coordinate – exemplified by thesauri and classifications respectively. Although it is desirable that the principles of facet analysis should be understood and applied consistently, they do not provide a completely objective and mechanical framework for the construction of knowledge organization systems and there are several stages at which judgement and experience are called for.

2. Concepts

The building blocks of all knowledge organization systems are *concepts*, which may be defined as *units of thought.* These include not only physical things, such as shoes or ships, but also materials such as sealing-wax, living organisms such as cabbages, people such as kings, activities such as flying and abstract concepts such as temperature, taxes, corruption, hunger, and love. As is emphasized in a recent standard for thesauri (ISO 25964-1, 2011) there is a clear distinction between a concept and the labels which we use to identify it, be they words or symbolic notations. The *OWL Web Ontology Language Guide* (2004) contains the statement "A label is like a comment and contributes nothing to the logical interpretation of an ontology". Sometimes there is a single word or phrase which adequately conveys the scope of the concept to anyone using it, but in many cases it is necessary to specify the

scope separately by means of a scope note or by describing its relationship to other concepts.

The purpose of a knowledge organization system is to define its concepts clearly, so that they are interpreted in the same way by indexers and searchers, and to build them into a logical structure which allows indexers and searchers to find the most appropriate concepts to express what they find, or what they are seeking, in a collection of information resources. In general a thesaurus deals with individual concepts while a classification scheme additionally provides rules for expressing compound subjects by combining concepts together. There is no clear rule as to what constitutes an “individual concept”, distinguishing between simple and compound concepts, but thesaurus standards do give guidance on when it is appropriate to split a concept into constituent simpler ones.

Once we have assembled a list of the concepts which will form the building blocks of a knowledge organization system of any kind, the next stage is to group them into fundamental categories such as objects, materials, people, actions, properties, abstract concepts, places and times.

We can then establish hierarchical relationships between concepts, with the stipulation that generic/specific relationships require that each pair of related concepts must be in the same category. It is invalid to create such a relationship between objects, such as *onions* and an activity such as *cooking*, because neither of these concepts is a “kind” of the other. Ontologies also enforce this rule, specifying criteria for class membership such that membership of a subordinate class entails membership of the superordinate class, as expressed in the “is-a” relationship. In this process we consider each concept on its own merits, according to a definition of what it *is,* following the Aristotelian principle of *genus et differentia,* naming the broader class to which a concept belongs and specifying the differences which distinguish this species from other members of that broader class. In this process it is not necessary to specify the disciplinary context within which the concept may be studied or used, except for concepts which are themselves disciplines.

A single concept may be distinguished from its parent by a “characteristic of division”, and sibling concepts which have different values of that characteristic may be grouped into arrays, introduced by a node label specifying the characteristic, such as <foods by origin> or <vegetables by colour>. These subdivisions may be applied in parallel, at the same level, or sequentially, so that one array is subordinated to another, and a single concept such as “carrots” can occur in more than one array, as shown in the following simplified example:

Figure 1: Example of three facets, showing arrays, some introduced by node labels

materials	activities	people
<*materials by use*>	artistic activities	<*people by age*>
building materials	drawing	babies
cosmetics	painting	children
foods	cooking	adults
<*foods by origin*>	<*cooking by process*>	old people
animal foods	boiling	<*people by occupation*>
vegetable foods	braising	cooks
<*vegetable foods by part of plant*>	frying	farmers
grains, seeds	steaming	information scientists
leaf vegetables	information processing	librarians
cabbages	cataloguing	students
lettuces	indexing	
root vegetables	manufacturing activities	
carrots	boatbuilding	
parsnips	knitting	
turnips	recreational activities	
<*vegetable foods by colour*>	boating	
green vegetables	swimming	
cabbages		
lettuces		
orange vegetables		
carrots		
pumpkins		

The building of hierarchical trees and the sorting of concepts into categories is an iterative process, as checking the validity of a hierarchical relationship tests whether the two concepts concerned have been correctly allocated to the same category. It may also require us to adjust the definitions of some concepts. The process is thus generally not purely top-down or bottom-up, but a combination of these.

It is worth noting that some thesauri, such as the AAT, introduce another type of label into hierarchies, representing concepts which are needed in order to complete the logical structure of a hierarchy but which are not considered useful for use as indexing terms for assignment to documents. The AAT uses the expression "guide terms" to refer both to these and to node labels, as described above. This is confusing, and I suggest that these terms, which represent actual concepts rather than being labels showing how concepts have been organized, should be given a different name, such as "non-index terms" and be shown differently in displays. They may be annotated as "Do not use for indexing" or, as in the original TEST thesaurus (Engineers Joint Council, 1967), "Use of a more specific term is recommended – consult the terms listed below".

3. Facets

If this process is carried as far as possible, we can consider the top terms of such trees to be facet names, such as *objects*, *materials* or *activities*, and each set of concepts which has been linked hierarchically constitutes the respective facet.

This process does not give an absolute definition of facets, because some judgement is required at the top level. It would be possible to have a single top term such as “concepts”, to which everything could be linked, but it is usual to stop at a level below this. Even then there is scope for subjective judgement, for example when deciding whether *living things* and *inanimate things* should be treated as separate facets or as members of a *things* facet, whether *people* should stand alone or be subordinated to *living things*, or whether it is

desirable to make a distinction between *cabbage* as a foodstuff *material* and *cabbages* as individual *objects*. Abstract concepts present particular problems in assigning them to facets; this and other problems are discussed by Svenonius (2000).

This is also in accordance with the current ISO standard for thesauri, ISO 25964-1 (2011), which defines a facet as a "grouping of concepts of the same inherent category"; I think that this is the most useful definition and that we should restrict its meaning to this. Facets defined in this way are sometimes called "fundamental facets", but the distinction between these and "non-fundamental facets" is not clear enough to be helpful. ISO 25964-1 expands on the definition with the following examples and note:

> • *EXAMPLE 1 Animals, mice, daffodils and bacteria could all be members of a living organisms facet.*
>
> • *EXAMPLE 2 Digging, writing and cooking could all be members of an actions facet.*
>
> • *EXAMPLE 3 Paris, the United Kingdom and the Alps could all be members of a places facet.*
>
> *NOTE Examples of high-level categories that can be used for grouping concepts into facets are: objects, materials, agents, actions, places and times.*

The Art and Architecture Thesaurus (AAT) (2013) also adopts this approach, defining a facet as "a homogeneous class of concepts, the members of which share characteristics that distinguish them from members of other classes". and choosing the following facet names:

> associated concepts, physical attributes, styles and periods, agents, activities, materials, objects, brand names.

The AAT calls the first-level subdivision of a facet a "hierarchy", but I think it would be better to use the term "sub-facet" for such subdivisions, rather than using that term as a synonym of "array" as is sometimes done.

Because the concept of faceted search is currently popular, the definition of what a facet is has been stretched beyond the one outlined above, and the term is used loosely to mean any system in which terms are selected from pre-defined groups at the time of searching.

3.1 "Facets of a document"

In this interpretation, a facet is taken to mean any attribute of a document (or other resource), including its creator, distributor, title, date, size, copyright status or other pieces of information, in addition to its "subject" – what it is *about*. These are also known as "elements of metadata" or "fields of the catalogue record". Some of them will be considered to be "access points", by which the document might be retrieved, while others are part of the catalogue description, telling you more about the document once you have found it. As we have other names for these attributes, it is confusing to use these as one meaning of the word "facets".

3.2 "Facets of an object"

Many e-commerce applications allow you to narrow down the specification of an object by choosing one or more criteria. Vanda Broughton illustrates this with her example of socks, which can be characterised by colour, material, pattern, size or application (Broughton 2004). Similarly, Amazon allows you to select cameras by a combination of attributes such as number of pixels, zoom range, display size, price and so on (Amazon 2013). Though often called "facets", these attributes just serve to group things into arrays within a single "objects" facet, each of which can be introduced by a node label specifying a characteristic of division, such as <cameras by zoom range> or <socks by colour>.

3.3 Facets within disciplines

A more restrictive approach is to define facets only within the scope of a particular discipline, as described by Broughton (2006):

> "... In both the Ranganathan and CRG models, the universe of discourse is the discipline. Neither method attempts a unitary application of the fundamental categories to the whole of knowledge; although there has been some considerable discussion within the group regarding the feasibility of this, in practice it is seen to be unworkable (or at least very difficult). The faceted general classification is a series of subject classifications, each with its own facet structure and facet formula or citation order. The primary facet is not therefore "discipline" as stated in some sources (Priss, 2000); although there is an initial division into disciplines, this is external to the application of the facet analysis proper."

In this approach the object or purpose of each discipline is identified as its core element, being the equivalent of Ranganathan's "personality" facet, and other facets are then constructed in relation to this. The non-overlapping general classes of concepts which we have called "facets" are in this case called "categories" or "fundamental facets". These general facets have been recognized in development of the UDC, for example, which in addition to the long-standing tables of place and time now has "common auxiliary" schedules for properties, materials, relations, processes, activities and persons, though these are still restricted to being used in subordination to a disciplinary facet.

3.4 Non-disciplinary facets

A discipline-based approach is not essential, however, and, despite the reservations expressed above, discipline *can* be treated as a separate facet which need not be the primary one. Work on this approach by the Classification Research Group (1969) is referred to in paragraphs 6.2 onwards of the introduction to the second edition of the Bliss Bibliographic Classification (BC2; Mills and Broughton 1977), which calls these non-disciplinary facets "phenomena classes", a term which encompasses concepts such as properties, operations, processes and entities. It then goes on to say:

> "6.214.2 These phenomena classes are designed to take that literature on a given concept (entity, attribute, process) which treats it from the viewpoint of several or all disciplines. An example would be a work on the horse, treating it from the zoological, equestrian, agricultural, military, artistic, etc. viewpoint; or, a work on colour, treating it from the viewpoints of optics, biology, photography, painting, decoration, etc. ...
>
> 6.214.4 The BC takes the view that provision should be made now for the classification of such material from the phenomenon rather than the discipline view. Since the growth of such literature is not to be disputed, the sensible theory appears to make comprehensive provision for all the forms it might take – i.e. for all phenomena, not just a rather arbitrary selection based on the current situation."

Although the phenomena classes in BC2 have not yet been developed, recent work by Gnoli has produced a draft classification based on this approach, called the *Integrative levels classification (ILC)* (Gnoli et al. 2004). This organizes phenomena according to the theory of integrative levels, from the simplest to the most complex, and allows these phenomena to be combined with each other freely in any order. Disciplines are considered as one type of phenomenon, at the integrative level of knowledge, and if these are cited first a traditional discipline-based sequence can be generated.

Rather than treating all concepts as phenomena, as in the *Integrative Levels Classification*, Vickery suggested having at least two independent parallel sequences, one being phenomena – physical or abstract elements of the external world – and the other being activities, including the traditional disciplines as well as theories, methods, approaches, cultural perspectives, and intended applications (Gnoli 2011). This seems to be a useful distinction, though further subdivision into additional categories, as discussed above, would seem to be helpful.

When creating a thesaurus, building hierarchies of concepts within independent and self-consistent facets is as far as we need go, since a user can choose concepts from any of these to apply separately when indexing documents or constructing search statements. This is illustrated in Figure 6 of ISO 25964-1. It is helpful to create associative links between concepts, especially those in different facets, but these are additions which do not affect the structure.

It may also be helpful to create a display in which subsets of facets, containing concepts specific to particular disciplines, are shown grouped within each discipline, with the full facet, including general multi- and non-disciplinary concepts listed separately as shown in Figure 7 of ISO 25964-1. These subsets are represented in the ISO data model (NISO 2013; Will 2012a,b) as "concept groups", and should not be confused with facets or arrays.

4. Ordering and combining concepts

4.1 In a post-coordinate system

When working with a post-coordinate system, such as a thesaurus, we generally start by selecting a specific concept; displaying the record for this shows relationships which can be used for navigation. Choosing broader or narrower terms shows arrays of siblings at each level.

The order in which sibling concepts are arranged within an array is not generally of major importance, because arrays are typically quite short and can be scanned at a glance, especially if the display system allows the display of lower-level concepts to be suppressed. Alphabetical arrangement is common, unless the characteristic of division on which the array is based indicates some other sequence, such as size or chronology.

When seeking to retrieve items on a compound subject, we need to construct a search statement by choosing any number of independent concepts from one or more facets. These may be combined either by some default operator such as an implicit AND or by using a more complex Boolean search statement using AND, OR, NOT and parentheses to group concepts. This post-coordinate type of system is very flexible, but it can lead to ambiguities because the relationships between concepts are not normally specified.

4.2 In a pre-coordinate system

In a pre-coordinate scheme, such as a classification, on the other hand, we seek at the time of indexing to compile subject statements which express compound subjects and which arrange these compound subjects in a logical sequence, such that related items are near one another and dissimilar items are separated. By bringing related topics together it allows an enquirer to see a "map" of a subject area and to have attention drawn to adjacent topics which would not be retrieved by a specific search. It allows documents to be organized in a useful order, whether as an output list or as an on-line navigational structure.

If pre-coordinated strings of concepts are to be built consistently, there has to be agreement on the order in which concepts from separate facets are to be included. The CRG developed a "standard" citation order of 13 categories, which they consider applies in most cases, though acknowledging that it may need to be modified when applied to some disciplines, as follows:

> Thing – kind – part – property – material – process – operation – patient – product – by-product – agent – space – time

This list cannot be used directly with the type of "fundamental facets" discussed above, because the elements listed are not determined by what each concept *is* but rather what *role* it plays in relation to the elements around it. For example a "patient" or an "agent", being respectively the entity acted on or the entity carrying out an action, may both be concepts drawn from a people facet or an object facet. A *thing, part, product* and *by-product* may all be

concepts drawn from an objects or materials facet. Ranganathan recognized the fact that a given facet may occur at more than one place in a compound string, specifying each occurrence as a new “round”. The CRG sequence effectively does the same thing, though if we accept the definition of facet given above we should not say that it specifies a sequence of facets but rather a sequence of roles which concepts from each facet can take.

The *Integrative levels classification* uses explicit role indicators to show relationships as each element is added to the string. (Unfortunately it compounds the confusion by calling these role indicators “facets”, which is yet another meaning of that word, so allowance has to be made for this in understanding the scheme.)

While a pre-coordinated classification scheme could in theory consist purely of separate hierarchies of individual facets and rules for combining them, in practice such schemes normally show some pre-built combinations, either as examples of how the synthesis should work or to ensure the consistent placing of compound subjects. These are usually displayed as the hierarchies of the first-cited facet, with concepts from other facets interpolated where appropriate. For example, if we choose a phenomenon such as “materials” as the first-cited facet, we could have a structure such as shown in Figure 2. At each point where a different facet is introduced, we insert a label, which we may call a “facet label”, such as *(activities)* to indicate the facet from which the following concepts have been drawn. Where necessary, these labels can also indicate the role that these concepts play in relation to the concepts of the preceding facet, such as *(agents : people)*. In the following example, the *thing* and *kind* roles are drawn from the *materials* facet, and these are combined with concepts from the *activities* and *people* facets to express the roles of *process* and *agent* in the two compound subjects “Steaming cabbages by children” and “Roasting parsnips by cooks”.

Figure 2: Example of combining concepts from different facets in a classification scheme, showing facet labels in round brackets and node labels in angle brackets.

(people)
people
  <*people by age*>
  babies
  children
  adults
  old people
  <*people by occupation*>
  cooks
  farmers
  information scientists
  librarians
  students
 	
(activities)
activities
  cooking
  <*cooking by process*>
  boiling
  braising
  roasting
  steaming
  artistic activities
  drawing
  painting
  information processing
  cataloguing
  indexing
  manufacturing activities
  knitting
  recreational activities
  boating
	(materials)
	materials
	  <*materials by use*>
	  building materials
	  cosmetics
	  foods
	  <*foods by origin*>
	  animal foods
	  vegetable foods
	  <*vegetable foods by part of plant*>
	  grains, seeds
	  leaf vegetables
	  cabbages
	  (activities)
	  cooking
	  <*cooking by process*>
	  steaming
	  (agents : people)
	  people
	  <*people by age*>
	  adults
	  children
	  root vegetables
	  carrots
	  parsnips
	  (activities)
	  cooking
	  <*cooking by process*>
	  roasting
	  (agents : people)
	  people
	  <*people by occupation*>
	  cooks

It is important to recognize that in a classified sequence of this kind there is no hierarchical relationship between the concepts that precede and follow a facet label. There is no hierarchical relationship between cooking and cabbages or between roasting and cooks. The relationship across a facet label is one of combination, not subordination, so subsequent concepts have to be understood as representing compound subjects in which they are combined with the preceding concept from another facet.

This is quite different from the relationships between the concepts that precede and follow a node label such as <*cooking by process*>, where *steaming* and *roasting* are kinds of *cooking*. These three concepts are necessarily in the same facet.

5. Notation and facet indicators

The preceding examples of a classified arrangement have not been given any symbolic notation. Although notation is common and often useful, it is not of the essence of a classification, especially when, as above, each concept is given a unique verbal label, a preferred term, in accordance with thesaurus practice. A classification scheme without a symbolic notation is equivalent to a scheme of alphabetical subject headings, so long as each heading contains enough elements to show the string of concepts that have been combined to make it up, such as *foods : cabbages : cooking : by children*. This approach was fully developed in the PRECIS subject indexing system (Austin 1974, 1984; Bakewell 1975) and in the scheme of subject headings developed by Eric Coates for the British Technology

Index (Singleton 1974); sadly these are no longer used, but the underlying principles are still valid.

When a notation is used, it can represent a more compact representation of compound subjects and can allow a sequence of items to be sorted automatically into the classification sequence. A distinctive notation may be used for each facet, and facets may be combined and their roles shown by using symbolic “facet indicators” to match the facet labels used in the verbal representation, such as the punctuation signs used in UDC or the codes beginning "0" in the Dewey Decimal Classification. Computer searching on notation symbols may be more efficient than searching on words, but this does require the additional step of converting search expressions into the corresponding symbols.

For example, the Integrative Levels Classification uses many distinct facet indicators, in the following general categories:

0 under aspect
1 at time
2 in place
3 through process
4 made of element
5 with organ
6 from origin
7 to destination
8 like pattern
9 of kind

The specific interpretation of these relationships may vary according to the nature of the concepts which they link.

6. Hierarchical relationships

6.1 Polyhierarchy

A single concept may appear in more than one place in the hierarchical structure of a facet, as it may have more than one parent concept or may be included in more than one array according to different characteristics of division. For example, *carrots* are both *root vegetables* and *orange vegetables* as well as being a sub-species of *Daucus carota.* In order to avoid the generation of multiple strings when such concepts are built into compounds in a classification scheme, it is necessary to designate one location as being the primary one, to be used in compounds. This will normally be “the place of unique definition”, where the array is based on an attribute which defines the concept, rather than other attributes which it happens to possess, so for carrots the botanical placing would normally be the primary one.

6.2 Instances

In the foregoing discussion we have assumed that the hierarchies within facets would be based on generic (genus-species) relationships. At the lowest level we may define concepts as “classes of one”, where only a single instance exists. These are generally labelled with proper names, for example *John Milton* being an instance of the concept of *poets* and *Titanic* being an instance of the concept of *ships.* As these represent individuals rather than classes, they cannot have lower-level generically-related terms, and thus form leaf nodes of such a hierarchy. They do, however, inherit all the attributes of the classes to which they are subordinated.

6.3 Subdivisions of compound classes

Broughton (2010) has drawn attention to the problem of dealing with instances of compound classes, for example specific named books of Jewish scripture, where the general concept of *books of Jewish scripture* is represented by the synthesis of *Judaism* and *sacred texts*. To

avoid the inconsistency of enumerating specific concepts under a compound when these concepts do not occur in the parent class, it appears necessary to enumerate them under the most appropriate constituent element of the compound. In this example it would be under *sacred texts*, perhaps in an array headed by the node label <*sacred texts by religion*>. This array could then be "brought down" under the compound concept if necessary. There would be some redundancy of expression, but this cannot be avoided.

Another example is shown in Figure 3. Here there is some redundancy, in that *arthritis* means "inflammation of joints", and therefore is equivalent to the compound concept *joints : inflammation* to which it has been subordinated in the right-hand column. If the citation order used specifies that part of the body precedes disease, then the location in the right-hand column would be used to express the compound. The array in the left-hand column would allow a search to be done for all aspects of inflammation in any part of the body.

Figure 3: Example of "bringing down" concepts under a compound concept such as "*joints : inflammation*" from their enumeration in a single preceding facet.

(abstract concepts) ... health-related concepts diseases inflammation <inflammation by body part> arthritis osteo-arthritis rheumatoid arthritis appendicitis cystitis dermatitis	(living organisms) ... body parts bones joints (diseases) inflammation arthritis osteo-arthritis rheumatoid arthritis skin (diseases) inflammation dermatitis

6.4 Parts

It is commonly understood that hierarchies can also be built on partitive, or part-whole relationships. This can cause problems, because a part does not generally inherit the attributes of the whole to which it belongs. *Wheels* cannot be hierarchically related to *bicycles*, because wheels can form parts of many other wholes. Even if restricted to *bicycle wheels*, they do not have the attributes of *bicycles.* It is better to give them the broader concept *bicycle components*, which can in turn be given the broader concept *bicycles and bicycle components* if it is desired to keep these together.

It is therefore arguable that the part-whole relationship should be considered to be a type of associative relationship rather than a hierarchical one, as was suggested by Mills and Broughton (1977) in the introduction to the Bliss Classification, where they write:

> "Related terms: these are terms in any relationship other than the generic relationship (that of genus/species). Only one such relation is widely recognized and this is the Whole/Part."

ISO 25964-1 warns that whole-part relationships should be restricted to a few special cases, but even these are questionable. They are:

(a) *Systems and organs of the body.* It would be logical to treat this in the same way as parts or components of any other assembly, by giving them a generically broader concept *body parts*, analogous to *bicycle components.* The AAT does this, subordinating them to the concept *animal or human components*.

(b) *Geographical locations.* These are all instances, identified by proper names, rather than classes. A whole-part hierarchy consisting solely of such instances can be used, but it

should not be combined with a hierarchy built on genus-species relationships, to avoid misleading assumptions of inheritance of attributes. *France* may inherit the attributes of *countries*, being an instance of that class, but though we may say that *Paris* is a part of *France* it is not a country and does not inherit the attributes of *countries* or *France.* In an ontology context this relationship is described as non-transitive.

(c) *Disciplines or fields of discourse.* These can be considered as having generic hierarchical relationships rather than partitive ones. *Physics* can be considered as a kind of *science* rather than a part, and *painting* can be considered a kind of *art* rather than a part of art.

(d) *Hierarchical social structures.* The example usually given for this is the division of *armies* into *corps, battalions, regiments* and so on. This is questionable, because not all armies are subdivided in this way, and a *regiment* may not always and inherently be part of a *battalion,* for example. This is therefore the same situation as that of *wheels* and *bicycles*. It might be better to use a concept such as *organizational units*, as in the AAT, under which there can be an array of <*organizational units by size*>. If specific named units are to be used, such as the *Fifth Regiment of Foot*, then these are instances, not classes, and are subject to the points made under (b) above.

7. Conclusion

Facet analysis is an important technique which should underlie the construction of any type of knowledge organization system. Starting with concepts, we can successively build facets, thesauri and classification schemes. We are thus building the classification scheme from the thesaurus, rather than deriving the thesaurus from the classification scheme as is sometimes suggested (Aitchison and Dextre Clarke, 2004; State Records of South Australia, 2002, Broughton, 2008). The widespread popularity of the faceted approach has led to different interpretations, some of them misleading or confusing. We cannot lay down completely prescriptive rules for its application, but it is important that we should work to bridge the gaps of usage and terminology that exist between different communities so that developments and applications in each area can benefit from the experience of others.

8. References

Aitchison, Jean and Dextre Clarke, Stella (2004). The thesaurus: a historical viewpoint, with a look to the future. In: *The thesaurus: review renaissance and revision. Sandra K. Roe and Alan R. Thomas, editors.* – Binghampton: Haworth. (Co-published simultaneously as *Cataloging and Classification Quarterly*, vol.37, no.3/4, 2004.) p.5-21.

Amazon (2013). Amazon.co.uk: cameras. <www.amazon.co.uk> with search for "cameras". (Accessed 2013-06-11).

Art & Architecture Thesaurus Online (2013) / The J. Paul Getty Trust. – Revised 27 March 2013. <http://www.getty.edu/research/tools/vocabularies/aat/about.html>

Austin, Derek (1974). The development of PRECIS: a theoretical and technical history. *Journal of Documentation*, vol.30, no.1, p47-102.

Austin, Derek (1984). PRECIS: a manual of concept analysis and subject indexing. 2nd edition. – London : British Library Bibliographic Services Division.– isbn 978-0-7123-1008-6 397p.

Bakewell, K.G.B. (1975). The PRECIS indexing system. *The Indexer,* vol.9, no.4, p.160-166. <http://www.theindexer.org/files/09-4/09-4_160.pdf> (Accessed 2013-06-11).

Broughton, Vanda (2001). Faceted classification as a basis for knowledge organization in a digital environment; the Bliss Bibliographic Classification as a model for vocabulary management and the creation of multidimensional structures, *New review of hypermedia and multimedia.* vol.7, no.1, p.67-102.

Broughton, Vanda (2004). Essential classification. – London : Facet Publishing. ISBN 1-85604-514-5. – 324p.

Broughton, Vanda (2006). The need for a faceted classification as the basis of all methods of information retrieval. *Aslib Proceedings,* vol.58, no.1/2, p.49-72.

Broughton, Vanda (2008). A faceted classification as the basis of a faceted terminology: conversion of a classified structure to thesaurus format in the Bliss Bibliographic Classification, 2nd ed. *Axiomathes*, vol.18, no.2, p.193-210.

Broughton, Vanda (2010). Concepts and terms in the faceted classification: the case of UDC. *Knowledge organization* vol.37, no.4, p.270-279.

Broughton, Vanda (2012). Brian Vickery and the Classification Research Group : the legacy of faceted classification. In: *Facets of knowledge organization : proceedings of the ISKO UK Second Biennial Conference, 4th-5th July 2011, London.* – Bingley : Emerald. – isbn 978-1-78052-614-0. – p.315-326. <http://www.iskouk.org/conf2011/papers/broughton.pdf> (Accessed 2013-06-23).

Classification Research Group (1955). The need for a faceted classification as the basis of all methods of information retrieval / memorandum of the Classification Research Group, May 1955. – Reprinted in: From classification to "knowledge organization" : Dorking revisited or "past is prelude" / edited by Alan Gilchrist. – The Hague, International Federation for Documentation, 1997. – (FID publication ; 714). – isbn 92-66-00714-5. – p.1-9.

Classification Research Group (1969). Classification and information control: papers representing the work of the Classification Research Group during 1960-1968. – London : Library Association, 1969. – 130p.

Engineers Joint Council (1967). Thesaurus of engineering and scientific terms : a list of engineering and related scientific terms and their relationships for use as a vocabulary reference in indexing and retrieving technical information. – New York : EJC. – 690p.

Gnoli, Claudio et al. (2004). Integrative levels classification. < http://www.iskoi.org/ilc/>. Dated 2004.06.03 – 2011.08.25. (Accessed 2013-06-11).

Gnoli, Claudio (2011). Vickery's late ideas on classification by phenomena and activities. In: *Facets of knowledge organization : proceedings of the ISKO UK Second Biennial Conference, 4th-5th July 2011, London.* – Bingley : Emerald, 2012. – ISBN 978-1-78052-614-0. p.11-24.

ISO 25964-1 (2011). Thesauri and interoperability with other vocabularies. Part 1: Thesauri for information retrieval. – ISO, 2011. – 152p.

León manifesto (2007) / [Claudio Gnoli, Rick Szostak]. *Knowledge Organization*, vol.34, no.1, p.6-8. (Also available at <http://www.iskoi.org/ilc/leon.php>, accessed 2013-07-13).

Mills, J. and Broughton, Vanda (1977). Bliss bibliographic classification. Second edition. : introduction and auxiliary schedules. – London : Butterworths, 1977. <http://www.blissclassification.org.uk/Class1/introduction.pdf>

NISO (2013). ISO 25964 – the international standard for thesauri and interoperability with other vocabularies. <http://www.niso.org/schemas/iso25964/> (Accessed 2013-06-11).

OWL Web Ontology Language Guide (2004). W3C Recommendation 10 February 2004. <http://www.w3.org/TR/owl-guide/> (Accessed 2013-07-13).

Priss, U. (2000). Comparing classification systems using facets. In: *Dynamism and stability in knowledge organization. Proceedings of the 6th International ISKO Conference, Toronto, Canada, 10-13 July, 2000* / edited by C. Beghtol, L.C. Howarth, and N.J. Williamson. – Würtzburg : Ergon. (*Advances in Knowledge Organization*, vol. 7.)

Singleton, Alan (1974). Technical indexing at B.T.I. *The Indexer*, vol. 9 no. 2, October 1974, p.37-49. <http://www.theindexer.org/files/09-2/09-2_037.pdf> (Accessed 2013-06-11).

State Records of South Australia (2002). Developing a thesaurus : guideline. Version 1.3, 12 January 2012. 19p.
<http://www.archives.sa.gov.au/files/management_guidelines_developingthesaurus.pdf>

Svenonius, Elaine (2000). The intellectual foundation of information organization. – Cambridge, Mass. : MIT Press. p.139-144. (Chapter available at <http://polaris.gseis.ucla.edu/gleazer/462_readings/svenonius2000b.pdf>, accessed 2013-06-22)

Vickery, Brian (2009). The structure of subject classifications for document retrieval – <http://www.iskoi.org/ilc/vickery.php>. (Originally published (2008-2009) in Brian Vickery's personal website <www.lucis.me.uk> closed after the author's death.)

Will, Leonard (2012a). The ISO 25964 data model for the structure of an information retrieval thesaurus. *Bulletin of the American Society for Information Science and Technology*, volume 38, issue 4, April/May 2012, p.48-51. Available at: <http://www.asis.org/Bulletin/Apr-12/AprMay12_Will.pdf>

Will, Leonard (2012b). The ISO 25964 data model for the structure of an information retrieval thesaurus. In: *Categories, contexts and relations in knowledge organization. Proceedings of the Twelfth International ISKO Conference, 6-9 August 2012, Mysore, India*. Würzburg : Ergon, 2012, p.284-290. ISBN 978-3-89913-902-0. Also available at: <http://www.willpowerinfo.co.uk/Will_ISKO2012_paper.doc>