

Facet analysis as a fundamental theory for structuring subject organization tools

Some current work on faceted terminologies and the representation of relationships

The structure of the faceted classification:

- it uses 'simple' concepts, organized into a logical structure
- it uses a standard set of categories to analyse the concepts
- it has a standard 'syntax' for building compound descriptions
- it can be used to create structured classmarks, or
- it can provide thesaurus descriptors
- it can be used in a search interface

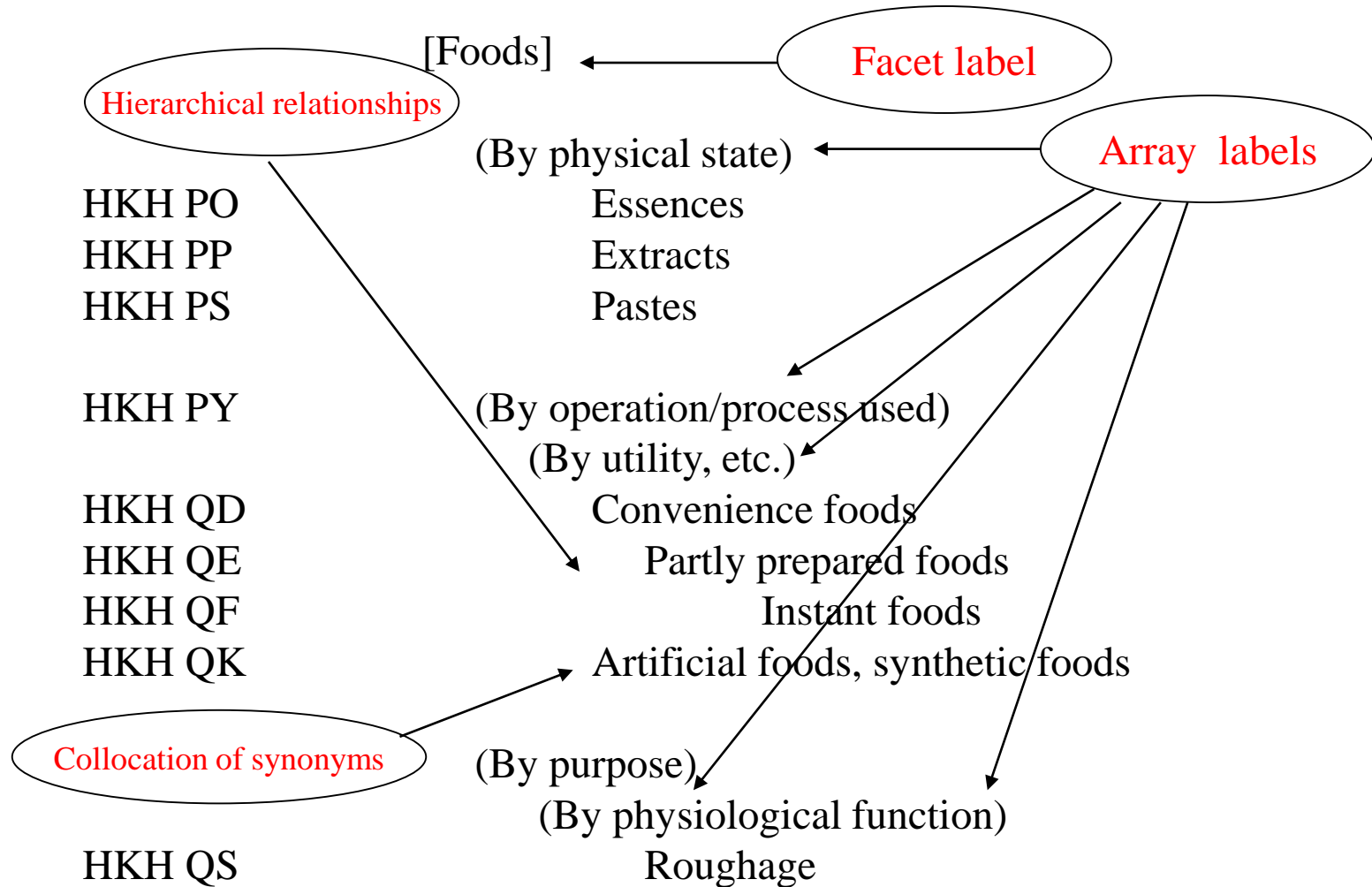
Standard categories:

- thing - kind - part - property - material - process - operation - agent - place - time
- any other useful or appropriate categories can be added
- concepts can be combined at will, but in a pre-determined order for physical arrangement
- this order can be repeated within a category if necessary

Facet analysis as a basis for classificatory structures:

- is a well established methodology
- organizes concepts in a domain into facets, and then into sub-facets (or arrays)
- within a facet, relationships of hierarchy are identified and visually displayed
- synonyms (or near synonyms) are collocated, and controlled by means of the notation

Basic classification structure:



Complex repeating structure derived from syntax rules:

HUQ W	Thymus gland (Physiology)
HUQ WH	(Pathology) (Hyperplasia)
HUQ WMD V	Lymphatism, status lymphaticus (Causal agents) (Symptoms) (Treatment) (Neoplasms)
HUQ WME	Thymomas (Products)
HUQ X	Thymus hormones (Molecular structure)
HUQ XS	Thymopoietins

[Compound terms pre-synthesized and added to published schedule]

[Examples of potential synthesized compounds]

Conversion to thesaurus format:

- all of the conceptual elements required to generate a thesaurus are implicit in the schedule
- BT/NT or intra-facet (paradigmatic) relationships (and some RTs) can be determined from the hierarchy
- other RTs can be identified from inter-facet (syntagmatic) relationships
- equivalence relationships are present in the synonym collocations

Intra-facet (paradigmatic) relationships in a basic schedule:

HKH PY

(By operation/process used)

(By utility, etc.)

HKH QD

Convenience foods

HKH QE

Partly prepared foods

HKH QF

Instant foods

HKH QK

Artificial foods, synthetic foods

Convenience foods

NT

Partly prepared foods

Partly prepared foods

BT

Convenience foods

NT

Instant foods

Convenience foods

RT

Artificial foods

Artificial foods

UF

Synthetic foods

Synthetic foods

USE

Artificial foods

Relationships in the faceted system:

- the vocabulary is very carefully and systematically analyzed
- as a result the relationships between terms are easily identifiable
- the thesaurus relationships can all be derived from a faceted classification
- a compatible thesaurus can be generated to a large extent by appropriate software

Automatic conversion from classification to thesaurus:

- BC2 has a suite of programs to generate schedule display and the A/Z index
- these have been extended to allow for automatic thesaurus generation
- some difficulties occur as a result of the schedule not having been written with the thesaurus in mind

BC2 source file markup for schedule display and indexing:

CLG	06Aluminium, aluminum
CLGLNM	07)Compounds with silicon & oxygen(
CLGLNMIFN	08Aluminium silicate
CLGM	07)Compounds with oxygen(
@	08)Salts(]IT
CLGMIFN	09Aluminates
CLGMJHN	08Aluminium oxide, alumina
@	07)Compounds with oxygen & hydrogen(]IT
CLGMKJHN	08Aluminium hydroxide, alumina trihydrate, hydrated aluminium oxide

BT/NTs inferred from the source file:

WWH P	Male voices	WWL ET	Dance forms
WWH R	Tenor	WWL ETM	Bourée
S	Baritone	WWL ETP	Chaconne
T	Bass	WWL ETS	Czardas
Male voices	NT Baritone	Dance forms	NT Bourée
	Bass		Chaconne
	Tenor		Czardas
Baritone	BT Male voices	Bourée	BT Dance forms
Bass	BT Male voices	Chaconne	BT Dance forms
Tenor	BT Male voices	Czardas	BT Dance forms

RTs derived from source file (but not inferred by software:

WWD G	Tonality
WWD GF	Overtones, intervals
WWD H	Scales
J	Diatonic, octave, 7 tone
K	Major scale
L	Minor scale
M	Chromatic scale
N	12 tone
O	Whole-tone scale
PD	Pentatonic scale
R	Modes

Major scale
Minor scale

RT Minor scale
RT Major scale

Scales
Modes

RT Modes
RT Scales

Equivalence relationships inferred from source file:

WWF O	Opera
WWF R	Operetta
S	Opera comique, opera buffa
T	Musical plays, musicals
Musical plays	UF Musicals
Musicals	USE Musical plays
WWG	Choral music
WWG C	Religious choral music
WWG F	Liturgical music, service music
Liturgical music	UF Service music
Service music	USE Liturgical music
WWQ Z	Bowed instruments
WWS B	Cello, violoncello
Cello	UF Violoncello
Violoncello	USE Cello

Problems with vocabulary control:

- there is no vocabulary control!
- authority control comes from the notation
- no thought was given to the idea of a preferred term
- therefore these have to be marked up manually
- some editorial guidelines must be introduced to manage this
- there are still some problems with formats of class headings

'Difficult' class headings:

Synonym sets:

WWG	Choral music
WWG C	Religious choral music
WWG F	Liturgical music, service music

Sibling sets:

WWG F	Liturgical music
WWG Q	Motets, anthems, hymns, etc

Inter-facet (syntagmatic) relationships in a complex schedule:

HUQ W	Thymus gland
HUQ WH	(Pathology)
	(Hyperplasia)
HUQ WMD V	Lymphatism, status lymphaticus
	(Neoplasms)
HUQ WME	Thymomas
	(Products)
HUQ X	Thymus hormones
HUQ XS	Thymopoietins

Thymus gland RT	Lymphatism
RT	Thymomas
RT	Thymus hormones

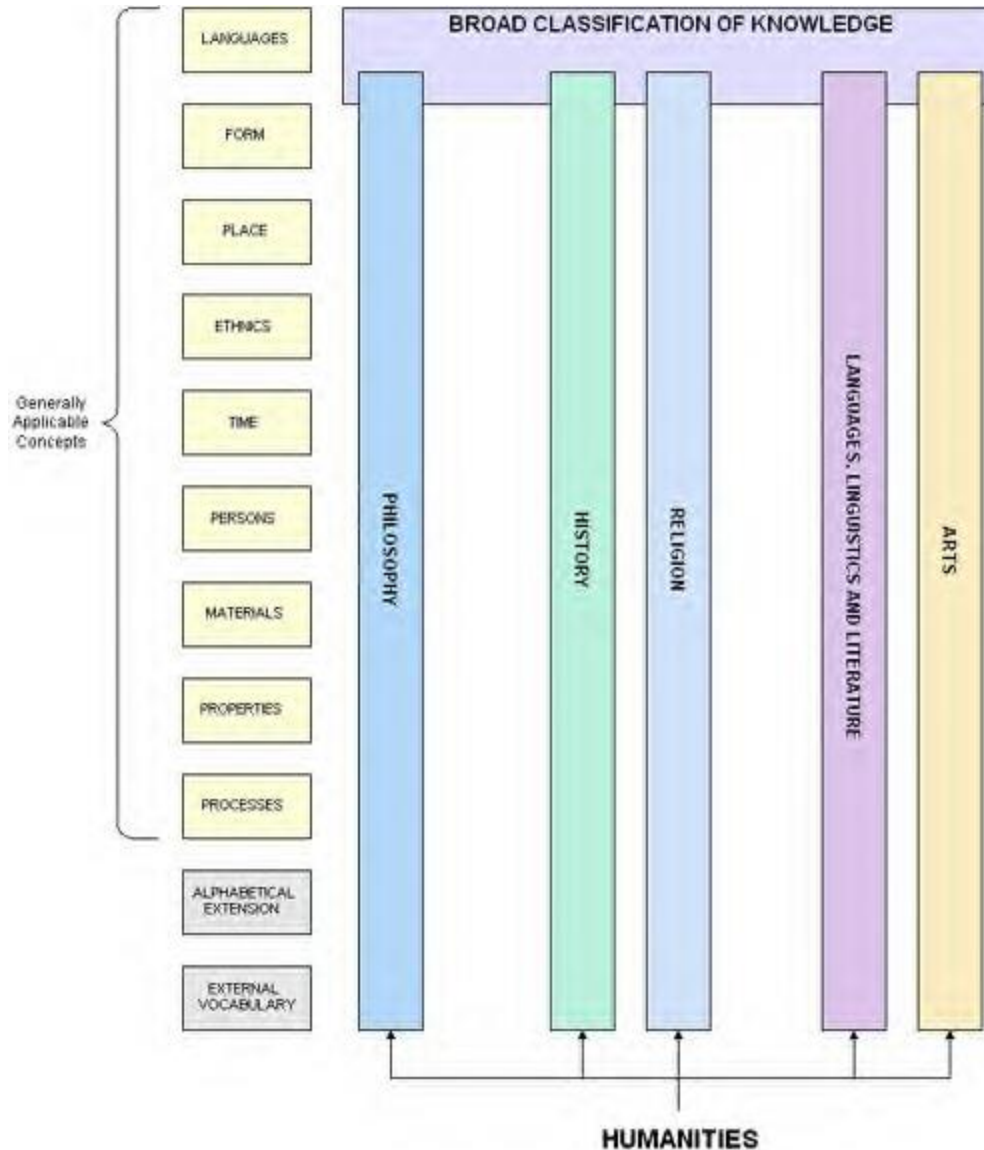
Syntagmatic relationships:

- in theory the relationship between a class and a sub-class created by combination with another facet = RT or associative term
- these can theoretically be identified by the presence of 'non-classes'
- in the BC2 terminologies these relationships are more precise than in current thesaurus practice
- e.g. entity-process, entity-product, agent-operation, etc.
- currently these cannot be inferred

FATKS: Facet analytical theory in knowledge structures:

- a project carried out at SLAIS
- an attempt to design a faceted classification that could be managed automatically
- a database was built to hold the classification data
- this included the hierarchical position of each class, and its containing category
- plus rules for combination and ordering of concepts

FATKS macrostructure:



Building compound and complex numbers

a) Combining concepts within the same facet

Notation:	Description:	Facet:
J15	Marriage and family	J Religious activities. Practice
J1477	Abstinence. Celibacy	J Religious activities. Practice
<hr/>		
J15J1477	Abstinence in marriage	

b) Combining concepts between facets

Notation:	Description	Facet:
5904	Buddhism	590 Religions and Faiths
E31	Originator, founder	E Agents. Subfacet: Persons as agents
A443	Physical form, appearance	A: Theory and Philosophy. Subfacet of God.
<hr/>		
5904 E31A443	Trikaya. Doctrine of the three bodies in Buddhism	

Notation:	Description:	Facet:
59033	Hinduism	590 Religion and faiths
5904	Buddhism	590 Religions and Faiths
5907	Christianity	590 Religion and faiths
J14247	Abstinence. Fasting. Prohibition	J Religious activities. Practice
<hr/>		
59033 J14247	Upavasa. Fasting in Hinduism	
5904 J14247	Abstinence. Fasting in Buddhism	
5907 J14247	Fasting in Christianity	

Untitled - SOFathum

File Edit Data View Help

Classmark

Classmark tag Fetch Browse

Component Terms Administration

Classmark component

Classmark Language

General category Change

Caption

Caption including

Heading Type

Scope Note

While this will differ from one religion to another, the concept of sacraments is inherent in many faiths. Individual sacraments are specified but the list is more comprehensive in coverage than will be required by any one faith. Use this location for general and comprehensive treatments of sacraments

Broader Category Change

Examples

Classmark	Description

Add Delete

Ready

Browsing Hierarchy

Search

Root Keyword Go

- 590J5 - Sacraments. Sacramentals
 - 590J52 - Rites of passage. Stages in life
 - 590J53 - Birth
 - 590J532 - Pre-natal. Antenatal
 - 590J533 - Peri-natal
 - 590J534 - Post-natal
 - 590J54 - Adolescence. Puberty. Transition to adulthood
 - 590J55 - Betrothal. Marriage. Wedding ceremonies
 - 590J555 - Procreation. Fertility. Childbirth
 - 590J56 - Old age
 - 590J57 - Death. Funeral ceremonies
 - 590J575 - Disposal ceremonies. Burial. Cremation
 - 590J58 - Stages in spiritual life
 - 590J582 - Initiation
 - 590J583 - Baptism. Naming ceremonies
 - 590J584 - Confirmation. Adult re-initiation
 - 590J585 - Ordination. Entry to priesthood
 - 590J586 - Beatification

Expand All Collapse All Print OK Cancel

What FATKS can do:

- represent hierarchical position
- represent categorical status
- support search and navigation of the vocabulary
- allow automatic synthesis through inbuilt syntax
- has potential to do more than this

FATKS: Facet analytical theory in knowledge structures:

- it can enable the creation of compound classmarks from extracted terms or keywords
- potentially it could also generate thesaural relations of the kind managed by the BC2 software
- in a populated system it also has the theoretical capacity to identify associative relations

Conclusion:

- all the structural elements of a thesaurus are implicit in a faceted classification
- vocabulary control is not well addressed and needs attention
- many of the relationships can be inferred automatically
- others have the potential to be recognized but need further identification in the source data
- currently it is not possible to represent the complexity of structure even though this is regular and predictable

References:

Aitchison J (1986) “Bliss and the thesaurus; the Bibliographic Classification of H. E. Bliss as a source of thesaurus terms and structure” *Journal of documentation* 42(3): 160-181

Aitchison J (2004) “Thesauri from BC2: problems and possibilities revealed in an experimental thesaurus derived from the Bliss Music schedule” *Bliss classification bulletin* 46: 20-26

BS 8723-1: 2005 (2005) Structured vocabularies for information retrieval. Part 1: Definitions, symbols and abbreviations. British Standards Institution, London

BS 8723-2: 2005 (2005) Structured vocabularies for information retrieval. Part 2: Thesauri. British Standards Institution, London

Broughton, V. (forthcoming) “A faceted classification as the basis of a faceted terminology” *Axiomathes*

Broughton, V; Slavic, A. (2007) “Building a classification for the humanities: principles and procedures” *Journal of documentation* 63(5)

Cordeiro, M. I.; Slavic, A. (2004) “Core requirements for automation of analytico-synthetic classifications”. In: *Knowledge organization and the global information society: proceedings of the Eighth International ISKO Conference, 13-16 July 2004, London, UK*. Ed. I. C. McIlwaine. Würzburg: Ergon Verlag. (Advances in Knowledge Organization; Vol 9) 187-192.

Facet analytical theory in knowledge structures www.ucl.ac.uk/fatks

Mills J, Broughton V (1977) *Bliss bibliographic classification. Second edition. Introduction and auxiliary schedules*. Butterworth, London