Illuminating Chaos
Using Semantics to Harness the Web

Dagobert Soergel
Department of Library and Information Studies,
University at Buffalo

UDC Seminar, The Hague, October 28-29, 2009
Outline

• Overview of issues
  • Semantics for whom and for what
  • Representation to assist with query formulation
  • Representation for comprehension
  • Systems of representation
  • Support for finding: Indexing
  • Building KOS
  • How can it all get done

• Zeroing in on the conceptual foundation
Semantics, structure, meaning

- Classification
- Meaningful arrangement
- All kinds of relationships
Semantics for whom?

- **Semantics for computer systems**
  - inference
  - answers and solutions instead of lots of Web pages

- **Semantics for people**
  - assist users in creating meaning and making sense
  - structure for learning
Semantics for what

• Finding

• Comprehending

• To know what to look for, a user (a person or a system) must first comprehend something – a cycle

• Both finding and comprehending require navigating in an information space – need meaningful structure
Representation to assist with query formulation
### Problem clarification for search

<table>
<thead>
<tr>
<th>JG</th>
<th>prevention approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>JG10</td>
<td>individual-level prevention</td>
</tr>
<tr>
<td>JG10.2</td>
<td>individual- vs. family-focused prevention</td>
</tr>
<tr>
<td>JG10.2.2</td>
<td>individual-focused prevention</td>
</tr>
<tr>
<td>JG10.2.4</td>
<td>family-focused prevention</td>
</tr>
<tr>
<td>JG10.4</td>
<td>prevention through information and education</td>
</tr>
<tr>
<td>JG10.4.2</td>
<td>social marketing prevention approach</td>
</tr>
<tr>
<td>JG10.4.4</td>
<td>prevention through information dissemination</td>
</tr>
<tr>
<td>JG10.4.6</td>
<td>prevention through education</td>
</tr>
<tr>
<td>JG10.4.8</td>
<td>peer prevention</td>
</tr>
<tr>
<td>JG10.8</td>
<td>prevention through spirituality and religion</td>
</tr>
<tr>
<td>JG10.10</td>
<td>prevention through public commitment</td>
</tr>
<tr>
<td>JG12</td>
<td>environmental-level prevention</td>
</tr>
<tr>
<td>JG12.4</td>
<td>social policy prevention approach</td>
</tr>
<tr>
<td>JG14</td>
<td>multi-level prevention</td>
</tr>
</tbody>
</table>
Browse structure for search

- Make a table of contents for the entire Wikipedia using UDC
## Facet structure to guide search

<table>
<thead>
<tr>
<th>A Area of ability</th>
<th>combines with</th>
<th>B Degree of ability</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1 psychomotor ability</td>
<td></td>
<td>B1 low degree of ability, disabled</td>
</tr>
<tr>
<td>A2 senses</td>
<td></td>
<td>B2 average degree of ability</td>
</tr>
<tr>
<td>A2.1 vision</td>
<td></td>
<td>B3 above average degree of ability</td>
</tr>
<tr>
<td>A2.1.1 night vision</td>
<td></td>
<td>B3.1 very high degree of ability</td>
</tr>
<tr>
<td>A2.2 hearing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A3 intelligence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A4 artistic ability</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Examples

<table>
<thead>
<tr>
<th>Examples</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A2.1B1 visually impaired</td>
<td></td>
</tr>
<tr>
<td>A2.2B1 hearing impaired</td>
<td></td>
</tr>
<tr>
<td>A3B1 mentally handicapped</td>
<td></td>
</tr>
<tr>
<td>A3B3 intellectually gifted</td>
<td></td>
</tr>
</tbody>
</table>
Provide front-ends to assist users

- Elicit a query with a facet-based interfaces, then the system creates a free-text query.

- Create a structure that normalizes terms assigned through social tagging and arranges them in a meaningful structure. The user can then browse and select concepts. The system maps to all appropriate tags.
Problem space for diseases

Used by people or computer systems for search and arranging search output

Pathologic process

Body system affected

Cause (condition, organism, chemical substance, environmental factors)

Treatment
Representation for comprehension

A question of information representation (knowledge representation)

• For computer systems: formal representation
• For people: Text, images, graphical representation, visualization
• Transformations between representations, such as
  • from text to formal: information extraction
  • from text to a map showing the text structure
  • from a conventional thesaurus display to a concept map
Two representations

Text (for people)
High blood pressure is a serious disease often caused by being overweight. In kids 4 – 12 it can be treated highly effectively with Nystatin.

Formal representation (for computer system)
Causation (HighBloodPressure, Obesity)
Treatment (HighBloodPressure, {Human, [Age, 4-12y]}, Nystatin, [Effectiveness, 4])
Answering questions

Question
How can high blood pressure be prevented?

Answer
Loose weight?
Two representations

Text
Kids begin grazing independently from their mothers at three months

Formal representation
Separation (Mother, Child, {Goat, [Age, 3m]})
Information extraction

- Information extraction produces representations needed for the semantic Web
- Also useful for people if formal expressions are transformed into sentences that state the findings of a document as individual "bullets"
- Could arrange statements from one or more documents in UDC order as a kind of summary
- Information extraction needs rich KOS
ADHD In Children

What is the most effective treatment for ADHD in children? [Therapy question]

[Matching topic] - Symptom
- Inattention and hyperactivity/impulsivity
  Comorbid conditions, present in 65% of children with ADHD, were often poorly controlled.

[Method/Solution] - Medical treatment
- Stimulant medication therapy
- Side effect

[Condition] - Hindering factor
- Sometimes patients and parents are hesitant to take medication for ADHD. Likewise, children and adolescents may resist medication because of stigma or feeling unfairly labeled with a disease.

[Method/Solution] - Guideline

The American Academy of Pediatrics recommends that clinicians: 1) manage ADHD as a chronic illness, 2) collaborate with parents, the child, and school personnel to define specific desired outcomes, 3) use stimulant or behavioral therapy to improve these outcomes; if one stimulant is not effective at the highest feasible dose, try another, 4) reevaluate the diagnosis, treatment options, adherence, and possible coexisting conditions if treatment is not achieving the desired outcomes, and 5) follow-up regularly with parents, child, and teachers to monitor for progress and adverse effects.
Meaningful arrangement of terms assigned in social tagging
The Martyrdom of Saint Bartholomew
Matching topic (Direct)

- **Image content: Focal**
  - *Reference*
    - nude body
    - old man
    - Saint Bartholomew
    - executioner
    - knife
  - *Elaboration (Adj.)*
    - Bearded
    - physical anguish
    - profound emotion
      - {emotional}
    - luminous

- **Elaboration (Adv.)**
  - expressive hands
  - gestures
  - confronts
  - flayed alive
  - torture

- **Image theme**
  - martyrdom
  - mystical experience
  - biblical
  - religious

- **Image content: Peripheral**
  - *Elaboration (Adv.)*
    - lurking
Comparison

• By similarity: Metaphor / analogy
  • Christ’s sacrifice and crucifixion
    {Christ metaphor}

Context

• Biographic info: Artist
  • Jusepe de Ribera

• Biographic info: Time / period
  • 1634
  • 17th century

Cause / Effect

• Reaction or feeling
  • Intensity

• Effect / Outcome
  • Pulls the viewer into the scene
Support comprehension through links to KOS

• Map text term to concept in KOS, show definition, show place in hierarchical structure
Comprehension "in the large"

• Learning and sense making require comprehension across multiple sources
• Requires structure – can be supplied by KOS
• Require tools for the manipulation of external structures the learner / sensemaker builds
Representation systems
Representations need rules

• Formal representations need logical formalisms, such as full first-order logic or subsets (for ease of processing) or extensions (to be more expressive)

• Text needs rules of syntax and broader document structure

• Graphical representations need rules of design
Representations need names for entities

- Names for (abstract) concepts – classification
- Names for many different types of other entities, such as persons, places, buildings, events, currencies, … (named entities)
- Systems of such names – Knowledge Organization Systems, authority lists of personal names
- Mappings between such systems
Representations need relationships

- Relationships are used to connect entities, thus forming statements:
  obesity <causes> high blood pressure

- Need system of relationships:
  Many such systems exist (a type of KOS):
  Problem of mapping
Rhetorical relationships

• To map text structure
• To discern how a retrieved document, paragraph, statement, or image relates to the topic of a search
Topical relevance typology

**Function-based**

**Rhetorical structure**
- Matching topic
- Evidence (Indirect)
- Context
- Comparison
- Evaluation
- Method / Solution
- Purpose / Goal

**Argument structure**
- Grounds
- Warrants
- Claim

**Reasoning-based**

Generic inference
Comparison-based
Induction / rule-based
Causal-based
Transitivity-based

**Semantic-based**

(Taxonomy
Partonomy
Frame-based, etc.)

(Green & Bean, 1995)
### RST+ Functional Role

<table>
<thead>
<tr>
<th>Matching topic (Direct)</th>
<th>Cause / Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manifestation</td>
<td>Cause</td>
</tr>
<tr>
<td>Image content</td>
<td>Effect / Outcome</td>
</tr>
<tr>
<td>Image theme</td>
<td>Explanation (causal)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Evidence (Indirect)</th>
<th>Prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Context</td>
<td></td>
</tr>
<tr>
<td>Scope</td>
<td></td>
</tr>
<tr>
<td>Framework</td>
<td></td>
</tr>
<tr>
<td>Environmental setting</td>
<td></td>
</tr>
<tr>
<td>Social background</td>
<td></td>
</tr>
<tr>
<td>Time &amp; sequence</td>
<td></td>
</tr>
<tr>
<td>Assumption / expectation</td>
<td></td>
</tr>
<tr>
<td>Biographic information</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Condition</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Helping or hindering factor</td>
<td></td>
</tr>
<tr>
<td>Unconditional</td>
<td></td>
</tr>
<tr>
<td>Exceptional condition</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Purpose / Motivation</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Functional role: Comparison

Comparison

- By similarity vs. By difference (Contrast)
  - By similarity
  - Analogy & metaphor
  - By difference (Contrast)
- By factor that is different
  - Different external factor
  - Different time
  - Different place
  - Different participant
  - Different actor
  - Different subject acted upon
- Different act or experience
  - Different act
  - Different experience
Support for finding: Indexing

• Finding based on text:
  Knowledge-based expansion of query
  Front-end as discussed earlier

• Finding based on indexing:
  Semantically enriched documents
Reis et al. (2008)
**Impact of Environment and Social Gradient on Leptospira infection in Urban Slums** (doi:10.1371/journal.pntd.0000228).

<table>
<thead>
<tr>
<th><strong>Infectious disease studied:</strong></th>
<th>Leptospirosis</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pathogen (causative agent of disease):</strong></td>
<td>Leptospira spirochete</td>
</tr>
<tr>
<td><strong>Vector of disease pathogen:</strong></td>
<td>Rat (Rattus norvegicus)</td>
</tr>
<tr>
<td><strong>Pathogen host subjected to study:</strong></td>
<td>Human (Homo sapiens)</td>
</tr>
<tr>
<td><strong>Number of subject individuals in study:</strong></td>
<td>3,171</td>
</tr>
</tbody>
</table>

... 

**Purpose of study:** Quantify risk factors for leptospirosis ... 

**Principal finding 1:** Prevalence of Leptospira antibodies ... 

**Principal finding 2:** Disease risk ... open sewers ...
A semantically enriched document

Tag Trees of Individual Semantic Classes of Highlighted Terms

**disease**
- infectious diseases
- diarrheal disease
- childhood diarrhea
- dengue
- leptospirosis
  - human leptospirosis
- meningococcal disease
- pulmonary hemorrhage syndrome
- visceral leishmaniasis
- Weil's disease
- occupational disease
- zoonotic disease

ID = Infectious Disease Ontology
GO = Gene Ontology term used in ID

- ID:0000012 immunity
- ID:0000017 mortality
- ID:0000023 zoonotic
- ID:0000025 pathogenicity
- ID:0000034 endemic
- ID:0000038 parasite
- ID:0000056 host
- ID:0000057 carrier
- ID:0000063 vector
- ID:0000064 pathogen
- ID:0000066 infectious agent
- ID:0000069 primary pathogen
- ID:0000104 infection
ID = Infectious Disease Ontology   GO = Gene Ontology

**IDO:0000000 ! process**
- IDO:0000083 transmission
  - IDO:0000231 horizontal transmission (GO:0000031)
- IDO:0000104 infection
- IDO:0000084 pathogenesis
- IDO:0000221 ! infectious disease progression
  - IDO:0000100 ! pathogen evasion of host immune response
    - IDO:0000111 antigenic variation
      - IDO:0000115 genetic diversification
  - IDO:0000226 pathogen life cycle (GO:0000026)

**IDO:0000001 ! role**
- IDO:0000036 ! colonizer
  - IDO:0000038 parasite
- IDO:0000048 symptom
- IDO:0000056 host
  - IDO:0000057 carrier
- IDO:0000059 reservoir
- IDO:0000063 vector
- IDO:0000064 pathogen
  - IDO:0000066 infectious agent
    - IDO:0000069 primary pathogen
- IDO:0000200 mode of transmission (GO:0000000)

**IDO:0000002 ! quality**
- IDO:0000215 ! quality of host population
  - IDO:0000098 infectious disease
Semantically enriched documents

• Semantic enrichment supports semantic retrieval
  • Broad area of its own

• Many different forms
  • Explicit document structure
  • Concept and named entity tagging and identification
  • Assigning additional concepts or named entities
  • Assigning extracted propositions

• Closely linked with information extraction
  • IE produces elements of semantic enrichment
Need KOS

Needed for all this

- Large Knowledge Organization Systems
- Large knowledge bases with mappings
- Methods and procedures for developing KOS
How to get all this work done?

The forces that created the problem also support the solution

- Use automation
  - Automated information extraction gets better every day and also provides input to building KOS
  - Automated classification could be used for the UDC Wikipedia project

- Use Web-enabled collaborative work ("crowdsourcing")

- Use computer systems to assist people

- Use Web-based systems to collect and integrate results

- Bootstrap: The more knowledge is in formal systems, the more information extraction and structuring tasks can be automated
Example: Guided tagging

- Use facet structure to get taggers think a bit more out of the box
  For example, could ask
  What does this image remind you of

- Could assign some terms automatically, for example, extracting terms from text assigned to an image
DH June 2009

Instructions
Start tagging by entering terms into the General box. Press enter after each entry and tags will be added automatically to the categories below. If you feel a tag belongs in a different category, simply drag and drop it into the correct box. You may also apply tags by selecting them from a drop-down menu, clicking on them on the right, or entering them directly into the category boxes. See the Help section for more information on applying tags.

General
- surrealism
- peaceful
- blue

Image Content
- woman
- blue

Style
- surrealism

Facet

Interpretation
- peaceful
Surrealism, peaceful, blue

Disambiguation and Spelling Pop-up

1. [n] blue, blueness (blue color or pigment; resembling the color of the clear sky in the daytime) "he had eyes of bright blue"
2. [adj] blue (used to signify the Union forces in the American Civil War (who wore blue uniforms)) "a ragged blue line"
3. [adj] gloomy, grim, blue, depressed, dispirited, downcast, downcast, downhearted, down in the mouth, low, low-spirited (filled with melancholy and despondency) "gloomy at the thought of what he had to face"; "gloomy predictions"; "
Semantic analysis as the basis for everything
Mapping through a Hub

Dewey

<table>
<thead>
<tr>
<th>Dewey</th>
<th>Hub</th>
<th>LCSH</th>
</tr>
</thead>
<tbody>
<tr>
<td>387 Water, air, space transportation</td>
<td>Water transport</td>
<td>Shipping</td>
</tr>
<tr>
<td>386 Inland waterway &amp; ferry transportation</td>
<td>Inland water transport</td>
<td>Inland water transport</td>
</tr>
<tr>
<td>387.5 Ocean transportation</td>
<td>Ocean transport</td>
<td>Merchant marine</td>
</tr>
<tr>
<td>386.8 Inland waterway tr. &gt; Ports</td>
<td>Traffic station □ Water transport</td>
<td>Harbors</td>
</tr>
<tr>
<td>387.1 Ports</td>
<td>Traffic station □ Inland water tr.</td>
<td>Hafen</td>
</tr>
<tr>
<td></td>
<td>Traffic station □ Ocean transport</td>
<td></td>
</tr>
</tbody>
</table>

German
Outline

• Objective: Interoperability Plus
• KOS concept hub
• Method: Knowledge-based, computer-assisted creation of canonical representations of concepts
• Resulting knowledge base and applications
Objective

Improve semantic-based search across multiple collections in multiple languages.

- Interoperability between any two participating KOS (Knowledge Organization Systems)
- Support for search, esp. facet-based search
  - for any collection indexed by a participating KOS
  - for search based on free-text or free-form social tagging
- Assistance in cataloging (metadata creation) by catalogers or users (social tagging)
- Long-range goal: Web service where a KOS can be uploaded and mappings to specified target KOS are returned
KOS Concept Hub

• Interoperability is achieved by expressing concepts from all participating KOS as a canonical representation, such as a description logic formula using atomic concepts and relationships

• The backbone of the proposed system is a faceted core classification of atomic concepts together with a set of relationships

• Mapping from KOS to KOS is achieved by reasoning over these canonical representations
Mapping through a Hub

Dewey

387 Water, air, space transportation
386 Inland waterway & ferry transportation
387.5 Ocean transportation
386.8 Inland waterway tr. > Ports
387.1 Ports

Hub

Water transport
Inland water transport
Ocean transport
Traffic station ∩ Water transport
Traffic station ∩ Inland water tr.
Traffic station ∩ Ocean transport

LCSH

Shipping
Inland water transport
Merchant marine
Harbors
Hafen

German
Method: How to get DL formulas

Key: Efficient creation of canonical representations (DL formulas)

• Apply existing knowledge:
  Large knowledge base ➞ less effort for processing a new KOS

• Use knowledge of KOS structure for hierarchical inheritance

• Use linguistic analysis of terms and captions

• Eliminate redundant atomic concepts

• Check or produce mapping results from assignment of concepts to the same records

• Get human editors’ input and verification where needed through a user-friendly interface

• KOS “owners” may verify and edit data pertaining to their KOS
Knowledge base

Requires an ever larger classification and lexical knowledge base containing many kinds of data:

1. A faceted classification of atomic concepts
   Seeded from sources with well-developed facets such as UDC
   the Alcohol and Other Drug (AOD) Thesaurus
   the Harvard Business Thesaurus
   the Art and Architecture Thesaurus
   various systems called ontologies
Knowledge base 2

Requires an ever larger classification and lexical knowledge base containing many kinds of data:

2. Linguistic knowledge bases such as WordNet and mono-, bi-, and multi-lingual dictionaries and thesauri

3. Many KOS (Knowledge Organization Systems), such as LCC, UDC, DDC, DMOZ directory, LCSH, Gene Ontology, Schlagwortnormdatei

4. These will over time be fused into one large multilingual knowledge base with many terminological and translation relationships and relationships linking terms to concepts, with an increasing number of concepts semantically represented by a DL formula.
Examples of deriving DL formulas
<table>
<thead>
<tr>
<th>L00  Transportation and traffic</th>
<th>P00  Buildings, construction</th>
</tr>
</thead>
<tbody>
<tr>
<td>L10  Traffic system components</td>
<td>P23  Buildings</td>
</tr>
<tr>
<td>L13  Traffic facilities</td>
<td>P27  Architecture</td>
</tr>
<tr>
<td>L15  Traffic stations</td>
<td>P43  Construction</td>
</tr>
<tr>
<td>L17  Vehicles</td>
<td></td>
</tr>
<tr>
<td>L30  Modes of transportation</td>
<td></td>
</tr>
<tr>
<td>L33  Air transport</td>
<td></td>
</tr>
<tr>
<td>L37  Water transport</td>
<td></td>
</tr>
<tr>
<td></td>
<td>R00  Engineering</td>
</tr>
<tr>
<td></td>
<td>R30  Acoustics</td>
</tr>
<tr>
<td></td>
<td>R37  Soundproofing</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>T70  Military vs. civilian</td>
</tr>
<tr>
<td></td>
<td>T73  Military</td>
</tr>
<tr>
<td></td>
<td>T77  Civilian</td>
</tr>
<tr>
<td>HE Transportation</td>
<td>L00 Transportation and traffic ⊕ T77 Civilian</td>
</tr>
<tr>
<td>-------------------</td>
<td>---------------------------------------------</td>
</tr>
<tr>
<td>HE550-560 Ports, harbors, docks, wharves, etc.</td>
<td>Inherited: L00 Transportation and traffic ⊕ T77 Civilian</td>
</tr>
</tbody>
</table>

**Added by editor:**
- L15 Traffic stations ⊕ L37 Water transport

**Resolved to:**
- **L15 Traffic stations ⊕ L37 Water transport ⊕ T77 Civilian**
<table>
<thead>
<tr>
<th>NA6300-6307  Airport buildings</th>
</tr>
</thead>
</table>

*From database already established:*

Airport = L15 Traffic stations □ L33 Air transport
Buildings = P23 Buildings

*Added by editor T77 Civilian*

*Resolved to*

L15 Traffic stations □ L33 Air transport □
P23 Buildings □ T77 Civilian
From database already established:

Airplane = L17 Vehicles \sqcap L33 Air transport

Soundproofing = R37 Soundproofing

Added by editor: Nothing

Resolved to

L17 Vehicles \sqcap L33 Air transport \sqcap R37 Soundproofing
Aeroplanes-Soundproofing

<table>
<thead>
<tr>
<th>From database already established:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aeroplanes = Airplane [Spelling variant]</td>
</tr>
<tr>
<td><strong>Therefore</strong></td>
</tr>
<tr>
<td>Term is recognized as same as</td>
</tr>
<tr>
<td>Airplanes. Soundproofing</td>
</tr>
<tr>
<td><strong>Resolved to</strong></td>
</tr>
<tr>
<td>L17 Vehicles □ L33 Air transport □ R37 Soundproofing</td>
</tr>
<tr>
<td>Any class formed by geographical subdivision</td>
</tr>
<tr>
<td>---------------------------------------------</td>
</tr>
<tr>
<td>Such as</td>
</tr>
<tr>
<td>NA6300-6307 Airport buildings</td>
</tr>
<tr>
<td>NA6305.E3 Egypt</td>
</tr>
</tbody>
</table>
Examples from the resulting knowledge base
<table>
<thead>
<tr>
<th>Classification</th>
<th>Description</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>HE550-560</td>
<td>Ports, harbors, docks, wharves, etc.</td>
<td>L15 Traffic stations □ L37 Water transport □ T77 Civilian</td>
</tr>
<tr>
<td>NA2800</td>
<td>Architectural acoustics</td>
<td>P27 Architecture □ R30 Acoustics</td>
</tr>
<tr>
<td>NA6300-6307</td>
<td>Airport buildings</td>
<td>L15 Traffic stations □ L33 Air transport □ P23 Buildings □ T77 Civilian</td>
</tr>
<tr>
<td>NA6330</td>
<td>Dock buildings, ferry houses, etc.</td>
<td>L15 Traffic stations □ L37 Water transport □ P23 Buildings □ T77 Civilian</td>
</tr>
<tr>
<td>TC350-374</td>
<td>Harbor works</td>
<td>L15 Traffic stations □ L37 Water transport □ R00 Engineering □ T77 Civilian</td>
</tr>
<tr>
<td>TH1725</td>
<td>Soundproof construction</td>
<td>P23 Buildings □ P43 Construction □ R37 Soundproofing</td>
</tr>
<tr>
<td>TL681.S6</td>
<td>Airplanes. Soundproofing</td>
<td>L17 Vehicles □ L33 Air transport □ R37 Soundproofing</td>
</tr>
<tr>
<td>TL725-726</td>
<td>Airways (Routes). Airports and landing fields. Aerodromes</td>
<td>L13 Traffic facilities □ L33 Air transport □ Technical aspects</td>
</tr>
<tr>
<td>VA67-79</td>
<td>Naval ports, bases, reservations, docks</td>
<td>L15 Traffic stations □ L37 Water transport □ T73 Military</td>
</tr>
</tbody>
</table>
Aeroplanes-Soundproofing = L17 Vehicles □ L33 Air transport □ R37 Soundproofing

Airports-Buildings = P23 Buildings □ L15 Traffic stations □ L33 Air transport

Buildings-Soundproofing = P23 Buildings □ P43 Construction □ R37 Soundproofing

Ships-Soundproofing = L17 Vehicles □ L37 Water transport □ R37 Soundproofing
Mapping through a Hub

LCC

TL681.S6 Airplanes. Soundproofing
VM367.S6 Submarines. Soundproofing

Hub

L17 Vehicles ▇ L33 Air transport ▇ R37 Soundproofing

L17 Vehicles ▇ L37 Water transport ▇ R37 Soundproofing

L17 Vehicles ▇ L37 Water transport

R37 Soundproofing ▇ T73 Military ▇ Underwater

LCSH

Aeroplanes-Soundproofing
Ships-Soundproofing
Mapping user queries

**User query**
- Free text
- Combination of elemental concepts through facets (guided query formulation)
- Controlled term(s) from a KOS, possibly found through browsing a KOS

**Hub**
- Canonical form of query (DL formula)

**Final query**
- (Enriched) free text query
- Query in terms of a KOS
Query:
L17 Vehicles AND R37 Soundproofing

<table>
<thead>
<tr>
<th>L17 Vehicles AND R37 Soundproofing</th>
</tr>
</thead>
<tbody>
<tr>
<td>TL681.S6  Airplanes. Soundproofing</td>
</tr>
<tr>
<td>VM367.S6  Submarines. Soundproofing</td>
</tr>
<tr>
<td>Aeroplanes-Soundproofing</td>
</tr>
<tr>
<td>Ships-Soundproofing</td>
</tr>
</tbody>
</table>

- [L17 Vehicles ⊓ L33 Air transport ⊓ R37 Soundproofing]
- [L17 Vehicles ⊓ L37 Water transport ⊓ R37 Soundproofing ⊓ Military]
- [L17 Vehicles ⊓ L33 Air transport ⊓ R37 Soundproofing]
- [L17 Vehicles ⊓ L37 Water transport ⊓ R37 Soundproofing]
Examples from NALT and LCSH

- NALT  National Agricultural Library Thesaurus
- LCSH  Library of Congress Subject Headings
Air pollution laws

LCSH term

Air – Pollution – Laws and regulations


NALT terms

Air pollution


Laws and regulations

[isa] Legal rule

Mapping LCSH —► NALT

Air – Pollution – Laws and regulations —► Air pollution AND Laws and regulations

Interpretation for indexing and searching in both directions
Soil moisture vs. Soil water

LCSH term
Soil moisture
  [isa] Water [containedIn] Soil

NALT term
Soil water
  [isa] Water [containedIn] Soil

Mapping LCSH ➤ NALT
  Soil moisture ➤ Soil water
Greenhouse gardening

**LCSH term**
Greenhouse gardening


**NALT terms**
Home gardening


Greenhouse

[isa] Greenhouse

**Mapping LCSH ➔ NALT**

Greenhouse gardening ➔ Home gardening AND Greenhouse
Salad greens

LCSH term
Salad greens
  [isa] Green leafy vegetable  [usedFor] Salad

NALT term
Green leafy vegetables
  [isa] Green leafy vegetable

Mapping LCSH —► NALT
  Salad greens —► BT  Green leafy vegetables
Emerging diseases

**LCSH term**
Emerging infectious diseases

[isa] Disease [hasProperty] Infectious [hasProperty] Emerging

**NALT term**
Emerging diseases

[isa] Disease [hasProperty] Infectious ??? [hasProperty] Emerging

**Mapping LCSH —► NALT**

Emerging infectious diseases —► Emerging diseases
Emerging infectious diseases —► BT Emerging diseases
Distributed implementation

• A KOS on the Web could assign DL formulas to its concepts – let's call this a semantically enhanced KOS or SEKOS

• Could use any of a number of faceted core classifications or even several (using a unique URI for each elemental concept)

• Core classifications could be mapped to each other

• It is now a simple matter to map from any SEKOS to any other (somewhat dependent on the core classifications used)
Take-home message

Semantics gives powerful systems