Indexing Languages – Development and Maintenance

Meeting of the Pilot Group of the Ad hoc Group of the Locarno Union
10th March 2009, Prague, Czech Republic

Peter Pálka and Linda Skolková

Institute of Information Studies and Librarianship, Faculty of Arts, Charles University, Prague

http://uisk.ff.cuni.cz/

Introduction

- It is necessary to consider design (especially industrial design) a very important field of human activity.

- Design is beneficial both for the manufacturer and the user.

- So the contributions of its authors must be protected (no matter who owns the contributions).

Defining industrial design

- There is a number of industrial design definitions which focus primarily on the author approach rather than covering a whole range of issues.

- Let us use the following definicion (by prof. ing. arch. Jaromíra Šimoníková, PhD., from Mendel University of Agriculture and Forestry in Brno):
  - Industrial design is a creative field attempting at reaching the optimal utility and esthetic features of products while respecting the technological and economic possibilities of the society.

Source: ŠIMONÍKOVÁ, Jarmila. *Počátek uplatnění zásad moderního designu v oboru nábytku : vývoj základních tvarů Thonetova nábytku z ohýbaného dřeva 1830-1930*. Praha, 1979. 2 vol. Dissertation. Czechoslovak Academy of Sciences. Institute for Theory and History of Arts.

Technology or art?

- The discussion about whether industrial design is more a technological or artistic field is not – from the point of view of our interest – productive.

Protection of authors' rights

- Approaches to the protection of authors' rights are not identical in all countries.

- The protection is based on *Berne Convention for the Protection of Literary and Artistic Works* from 1886 which was subsequently revised a number of times (http://www.wipo.int/treaties/en/ip/berne/trtdocs_wo001.html).

- In many countries (especially developing coucntries) the rights are not rigorously respected or respected at all.

World Intellectual Property Organization

- World Intellectual Property Organization (WIPO), now based in Geneva, Switzerland, was set up in Stockholm, Sweden, in July 1967 and operates from April 1970.

- WIPO's predecessor was the United International Bureaux for the Protection of Intellectual Property.

- URL: http://wipo.int/

6

WIPO's purpose

- WIPO's purpose is to support worldwide intellectual property protection and ensure cooperation among secretariats created on the basis of various agreements.

Industrial property

- Property related to the product

- Property related to the mark

- **Property related to industrial designs**

Industrial design

- Industrial design is typically a result of systematic creative efforts to create an object with a distinctive shape (and features).

- Definitions and rights may vary country from country.

Example – Czech Republic

- „[I]ndustrial design means the appearance of the whole or a part of a product resulting from the features of, in particular, the lines, contours, colours, shape, texture and/or materials of the product itself and/or its ornamentation,"

- „[P]roduct means any industrial of handicraft item, including parts intended to be assembled into a complex product, packaging, get-up, graphic symbol and typographic typefaces, except computer programs,"

Example – Czech Republic

- „An industrial design shall be protected, if it is new and has individual character.“

  - Source: *ACT No. 207/2000 Coll. of 21 June 2000, on the Protection of Industrial Designs and the Amendment to Act No. 527/1990 Coll., on Inventions, Industrial Designs and Rationalisation Proposals, as amended* [online]. Prague : Industrial Property Office [cit. 2009-02-23]. 17 p. URL: <http://isdvapl.upv.cz/pls/portal30/docs/FOLDER/PDF_DOKUMENTY/ZAKONY/207_2000_A.PDF>.

Imitations

- Imitations are:

  - *conscious* (cases when people are aware of the fact they imitate and act illegally on purpose),

  - or *uncounscious* (cases when people are not aware of it because of lack of information).

Plagiarius Award

- Besides legal action, some organizations also take other actions – award prizes - the example is Plagiarius Award in Germany.

- URL: http://www.plagiarius.com/e_index.html

Information as a necessary prerequisite

- In order to protect author and industrial rights to industrial designs efficiently **information** is necessary and indispensable as well as the possibility to search for them in such a way that guarantees completeness (a 100% recall) and use.

- Information loss is more serious than information noise (more risky, legal consequences).

Precision vs. recall in an ideal world
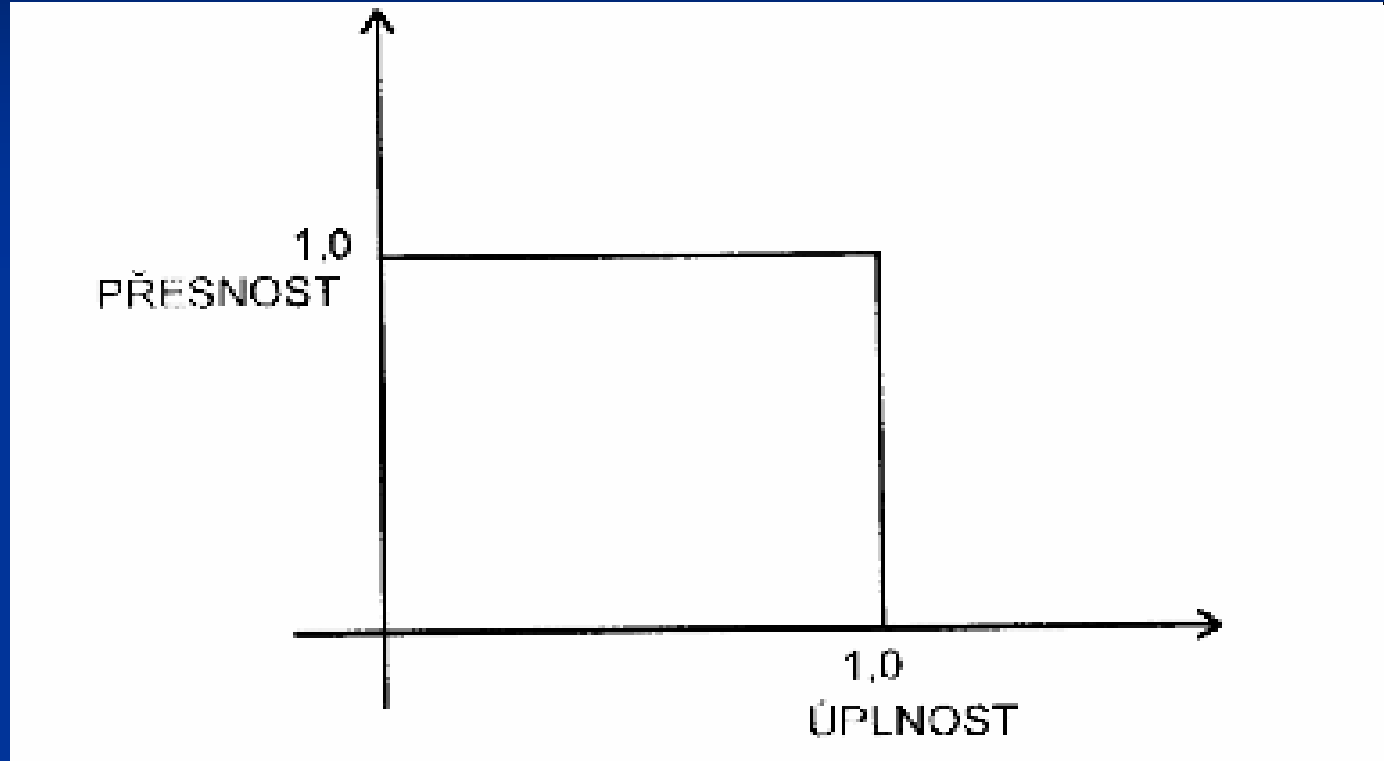
Precision vs. recall in practice
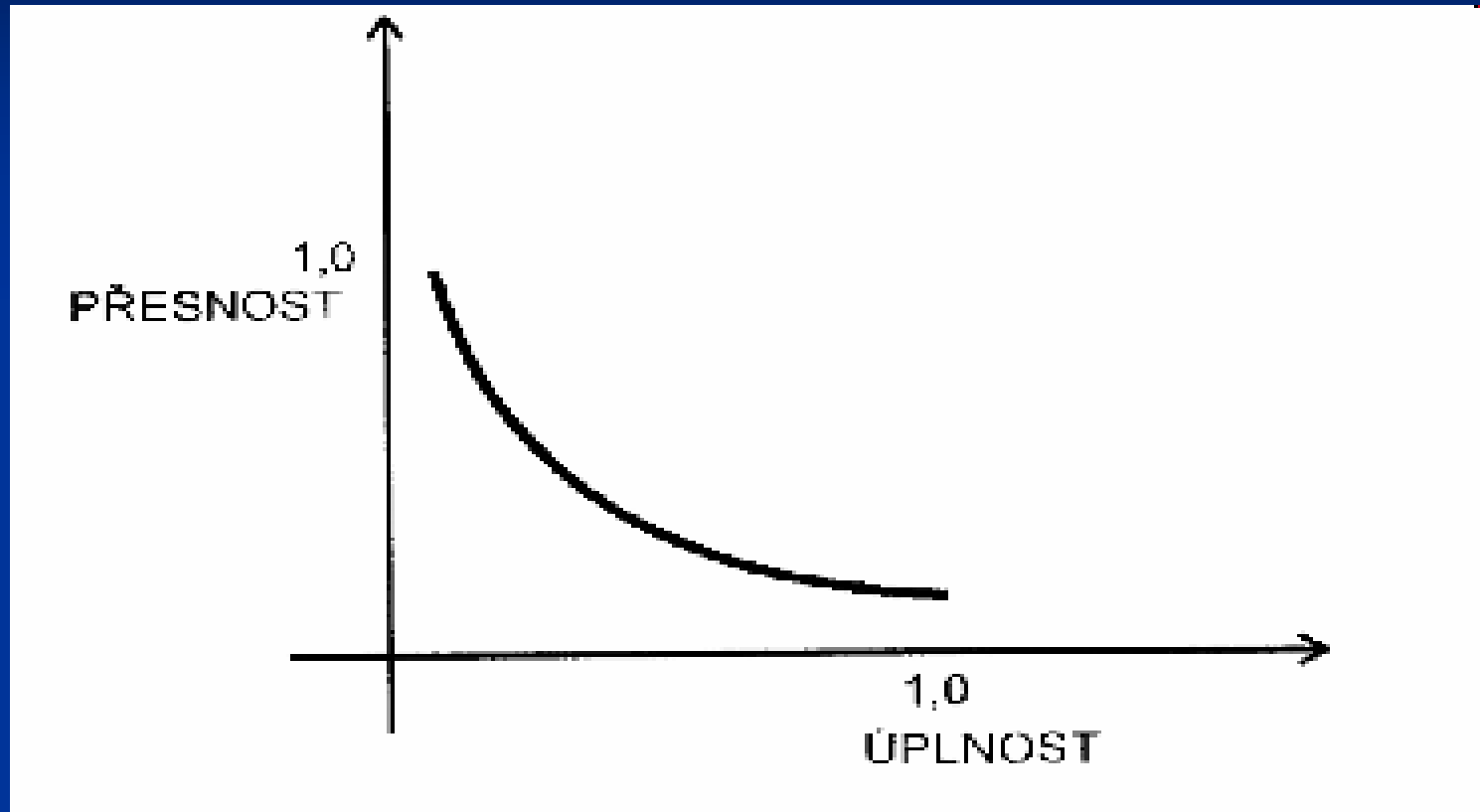
Information processing – indexing - research

- Industrial property applications serve as a basic source of information.

- The first obstacle is encountered when they are indexed as there is a high risk of **subjectivity** on the side of the indexer and therefore also unavoidable information losses while performing search (conducting research). The subjectivity is also present on the users' side.

Demands on the system and its users

- Demands on knowledge organization systems:
  - An ideal knowledge organization system should:
    - be **easy to understand** (user-friendly),
    - allow only the **right indexing** (optimally from multiple points of view),
    - and, at the same time, **be current**.
  - Don't forget **the users** -> TARGET GROUPS (professionals, laymen, both?)

Existing knowledge organization systems

- Currently, several systems are available which can be used to organize industrial designs (ordered from specialized to general systems):
  - Locarno Classification (International Classification for Industrial Designs under the Locarno Agreement)
    - URL: http://www.wipo.int/classifications/locarno/en/

Patent classifications

- **International Patent Classification**
  - URL: http://www.wipo.int/classifications/ipc/en/

- **Various national patent classifications, e.g., United States Patent Classification**
  - URL: http://www.uspto.gov/go/classification/

- **Other classifications, e.g. Derwent World Patent Index Classificaton System**
  - URL: http://scientific.thomsonreuters.com/support/patents/dwpiref/reftools/classification/

Mark classifications

- International Classification of Goods and Services (Nice Classification)
  - URL: http://www.wipo.int/classifications/nivilo/nice/index.htm

- International Classification of the Figurative Elements of Marks (Vienna Classification)
  - URL: http://www.wipo.int/classifications/nivilo/vienna/index.htm

Universal classifications

- **Library of Congress Classification**
  - URL: http://www.loc.gov/catdir/cpso/lcco/

- **Dewey Decimal Classification:**
  - URL: http://www.oclc.org/dewey/

- **Universal Decimal Classification:**
  - URL: http://www.udcc.org/

Which of these to choose?

- **None** of the above mentioned knowledge organization systems is fully suitable for our purposes.

Main drawbacks

- Single point of view

- Subjectivity – which then leads to information loss

- Enumeration – even classifications which tend to be enumerative are never complete -> classes like Others, Miscellanea, etc. (even in Locarno Classification)

Features of an optimal knowledge organization system

- Such a system should:
  - go into enough detail,
  - contain references (including image references),
  - have strong indexing rules (which limit indexing errors; assignment of one class vs. assignment of six classes; results of Cranfield experiments),
  - be sophisticated enough so that it provides tens of hits (results) at maximum,
  - periodically updated – extensible / hospitable / flexible – in compliance with updating/maintenance rules,
  - have a community of users who cooperate, provide feedback and help to develop the system (e.g. via an electronic discussion forum – even in social networks such as FaceBook, a mailing list)

Possible solutions

- Changing in currently used knowledge organization system – Locarno Classification

- Using facets and categories

- Indexing rules

- Validating indexes – using masks, classes, categories etc. as authorities

Possible solutions

- Using semi-automatic indexing

- Using content-based image retrieval (CBIR) as a supporting tool (not as the ONLY tool due to the experimental nature of CBIR)

- Using *shape (image) alphabet*

Shapes

Source: *Ottagono : rivista trimestrale di architettura arredamento industrial design.* Settembre 1971, anno 6. 22.

28
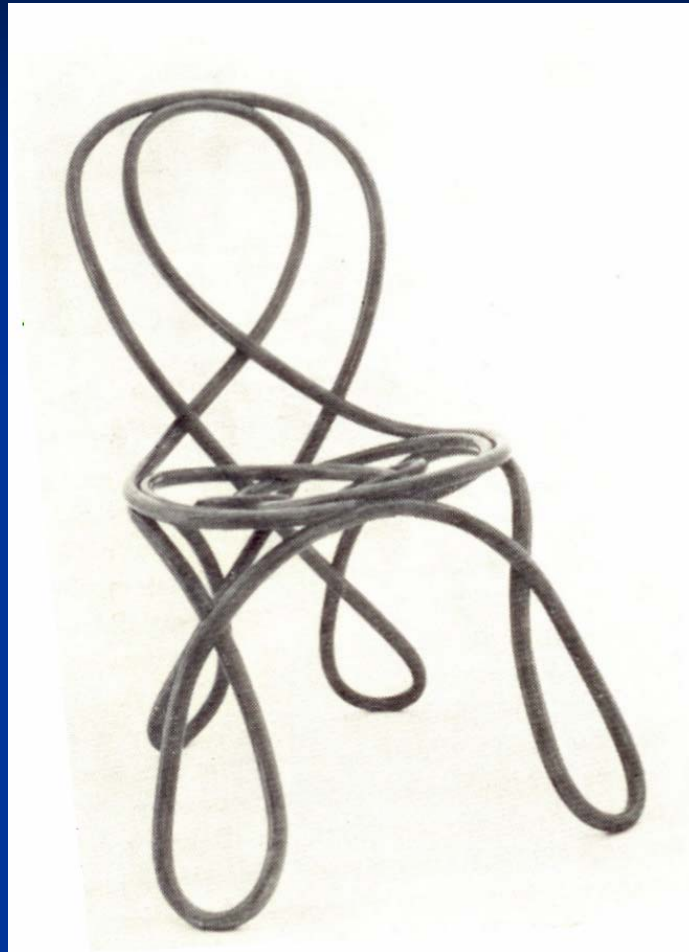
Shapes

Source: *Nábytek z ohýbaného dřeva : ze sbírek Okresního vlastivědného muzea ve Vsetíně.* Vsetín : Okresní vlastivědné muzeum, 1989. 62 pp. + annex.

Shapes

Source: *Ottagono : rivista trimestrale di architettura arredamento industrial design.* Settembre 1971, anno 6. 22.

Re-indexing

- Two options when a new system is introduced:
    - Re-indexing all items in the database using the new system (in some cases automated processing can be applied but it typically does not work in all cases)
    - Using the new system for newly inserted items only

Practical aspects of reindexing

- Let us count say **ten minutes** as an average time to index one industrial design (some designs can take more time to index, some less).

- Therefore, one indexer can re-index approx. 50 industrial designs per day.

- It is real to count with approx. 200 work days in a year.

- This means **one *highly skilled* indexer can re-index approx. 10,000 industrial designs per year**.

- If the European database consists of approx. 400,000 records representing industrial designs, then the work can be performed in a single year provided that **40 indexers work on the task full-time**.

- **So the question is: Is it feasible?**

Standards

- The use of standards is crucial for successful **automated data exchange** among various information systems.

- Examples:
    - SKOS – Simple Knowledge Organization System
        - URL: http://www.w3.org/TR/skos-primer/
    - MARC 21 Format for Authority Data:
        - URL: http://www.loc.gov/marc/authority/ecadhome.html
    - MARC 21 Format for Classification Data:
        - URL: http://www.loc.gov/marc/classification/eccdhome.html

Data validation

- Once the system is available in a suitable format which is not proprietary and can be used by more information systems, indexing data should only be **chosen from pull-down menus or validated so that the risk of error is minimized**.

- Searcher/user also needs to minimaze errors while performing the search – similar options should be offered.

Node labels (facet indicators)

Example:

cars
  by motive power
    diesel cars
    electric cars
  by purpose
    racing cars
    sports cars

Last but not least - interfaces

- Example:
  - EBSCOhost platform (EBSCO Information Services)

Homepage

Search results display

Thesaurus

Browsing: **Library, Information Science & Technology Thesaurus**

[   ] **Browse**

◉ Term Begins With   ○ Term Contains   ○ Relevancy Ranked

◀ **Previous** **Next** ▶   ◀ **Back to List**

Select term, then add to search using: [or ▼] **Add**   **Explode**

☐ ABSTRACTING   ☐

Scope Note	Here are entered works summarizing the main points of a piece of information. [EPC]
Broader Terms	☐ CONTENT analysis (Communication) ☐
	☐ DOCUMENTATION ☐
	☐ INFORMATION organization ☐
	☐ INFORMATION retrieval ☐
Narrower Terms	☐ AUTOMATIC abstracting
	☐ NEWSPAPERS -- Abstracting & indexing
Related Terms	☐ ABSTRACTING & indexing services
	☐ ABSTRACTS

Relationships between lexical units in the thesaurus

Browsing: **Library, Information Science & Technology Thesaurus**

[ ] **Browse**

( ) Term Begins With   ( ) Term Contains   ( ) Relevancy Ranked

◄ Previous Next ►   ◄ Back to List

Select term, then add to search using: [or ▾] **Add**   Explode

[ ] EXPERT systems (Computer science)   [ ]

Broader Terms   [ ] ARTIFICIAL intelligence   [ ]

[ ] COMPUTER systems   [ ]

Narrower Terms   [ ] KNOWLEDGE acquisition (Expert systems)

Related Terms   [ ] DECISION support systems

[ ] SEMANTIC networks (Information theory)

Used for   KNOWLEDGE-based systems (Computer science)

SYSTEMS, Expert (Computer science)

Page: Previous | Next

40

Tematres – a sample open-source tool for thesaurus creation
(http://www.r020.com.ar/tematres/index.en.html)

Source: *TemaTres* [online]. [cit. 2009-02-23]. URL:
<http://213.220.227.200:81/tematres/>.

Tematres – sample export format (SKOS)

```xml
-<rdf:RDF>
 -<skos:ConceptScheme rdf:nodeID="tematres">
    <dc:title>TemaTres</dc:title>
    <dc:creator>UISK FF UK</dc:creator>
    <dc:subject>Industrial design</dc:subject>
    <dc:description> </dc:description>
    <dc:publisher>UISK FF UK</dc:publisher>
    <dc:date>2008-10-12</dc:date>
    <dc:language>en</dc:language>
    <skos:hasTopConcept rdf:nodeID="tema15"/>
  </skos:ConceptScheme>
-<skos:Concept rdf:nodeID="tema25">
    <skos:prefLabel>Anorganické materiály</skos:prefLabel>
    <skos:inScheme rdf:nodeID="tematres"/>
    <skos:broader rdf:nodeID="tema18"/>
    <skos:narrower rdf:nodeID="tema73"/>
    <skos:narrower rdf:nodeID="tema67"/>
    <skos:narrower rdf:nodeID="tema79"/>
    <skos:narrower rdf:nodeID="tema68"/>
    <skos:narrower rdf:nodeID="tema66"/>
    <dct:created>2008-12-10 14:42:18</dct:created>
    <skos:subjectIndicator rdf:resource="http://localhost/tematres/index.php?tema=25"/>
  </skos:Concept>
</rdf:RDF>
```
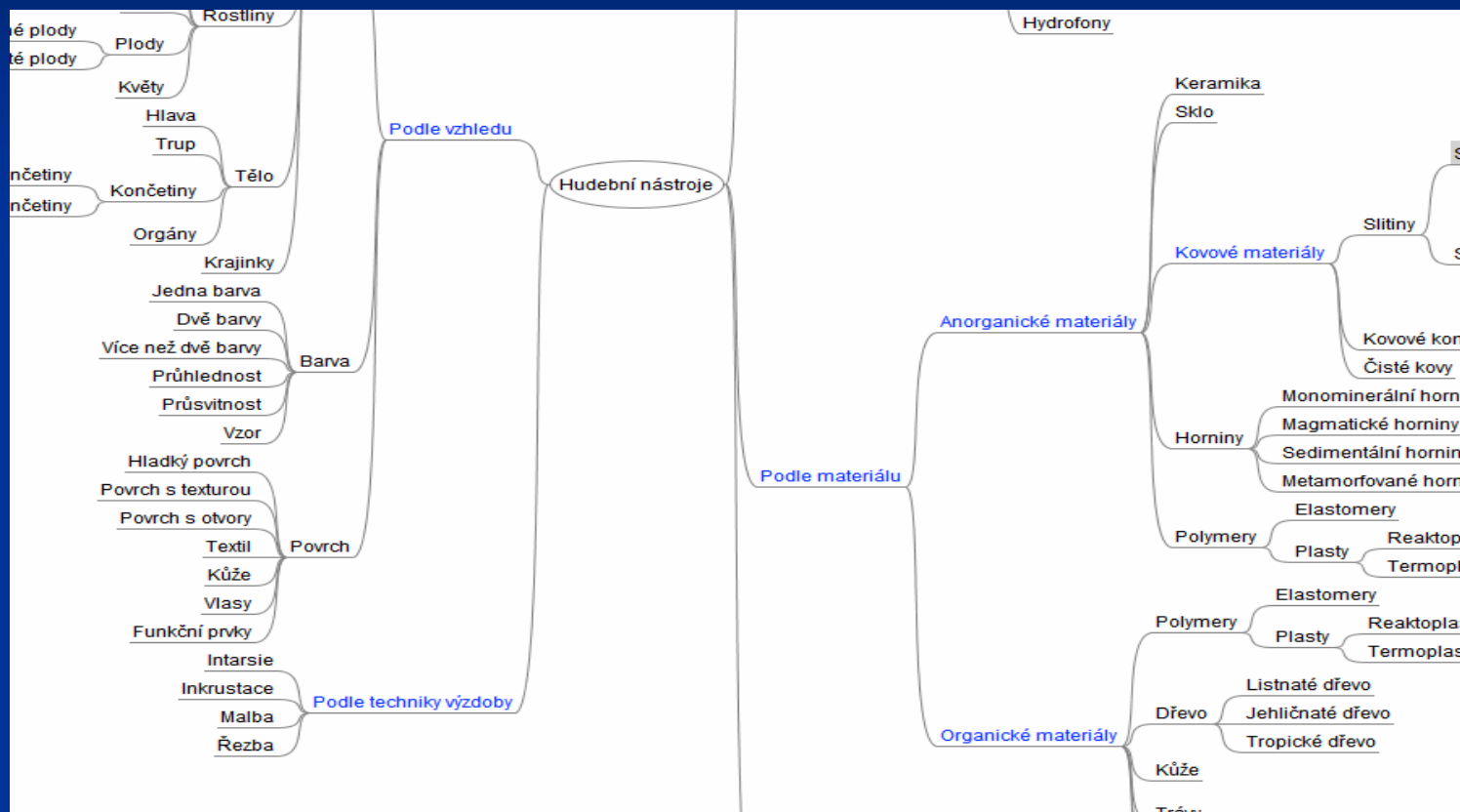
Source: http://213.220.227.200:81/tematres/, downloaded 2009-02-23

FreeMind – software for the creation of mind maps (http://freemind.sourceforge.net/)

Selected resources - organizations

- The American Society for Indexing:
  http://www.asindexing.org/site/index.html

- ISKO: International Society for Knowledge Organization:
  http://www.isko.org/

- IFLA Classification and Indexing Section:
  http://www.ifla.org/VII/s29/index.htm

- Networked Knowledge Organization Systems/Services:
  http://nkos.slis.kent.edu/

Selected resources – indexers and indexing

- ASI Indexer Locator:
  http://www.resourcenter.net/Scripts/4Disapi7.dll/4DCGI/resctr/search.html

- Indexing Evaluation Checklist:
  http://www.asindexing.org/files/public/checklist-book.pdf

- Getting Started in Indexing:
  http://www.asindexing.org/files/public/getting_started_in_indexing0604.pdf

Selected resources – policies, best practices

- French National Library – indexing policy: http://www.bnf.fr/pages/version_anglaise/cataloging/pol_ind_eng.htm

- Guide pratique du catalogueur: http://guideducatalogueur.bnf.fr/

- Best practices for subject access to national bibliographies: interim report by the Working Group on Guidelines for Subject Access by National Bibliographic Agencies: http://www.ifla.org/IV/ifla73/papers/089-Bourdon_Landry-en.pdf

Selected resources - interfaces

- Online Catalog Design Models: Are We Moving in the Right Direction?: http://www.ou.edu/faculty/H/Charles.R.Hildreth/clr-opac.html

- Interface to Classification: Some Objectives and Options: http://dlist.sir.arizona.edu/1621/01/slavic_E&C28.pdf

Selected resources – gateway to other resources in the field

- Sources of Information in the Field of Knowledge Organization: http://www.iskouk.org/links.htm#STANDARDS

Conclusion

- The problem mentioned is very complex and difficult to tackle.

- The optimal solution should have international dimension.

- No formal organization to solve the problem is needed (in our opinion).

- On the other hand, some financial and human resources are necessary.

Questions?

Thank you for your attention!

Contact information:

- Institute of Information Studies and Librarianship, Faculty of Arts, Charles University, Prague (http://uisk.ff.cuni.cz/), U Kříže 8, 158 00 Praha 5 - Jinonice

- e-mail addresses:
  - Linda Skolková: linda.skolkova@ff.cuni.cz
  - Peter Pálka: peter.palka@ff.cuni.cz