

# Úvod do věcných SJ

**Předmět: Selekční jazyky**

2. 10. 2009

Přednášející: Mgr. Silvie Kořínková Presová

<http://kisk.phil.muni.cz/mgr-silvie-korinkova-presova-dis>

# Termíny výuky

Kombinované studium

**23.10., 13.11., 4.12.**

16:40 v učebně C14

Prezenční studium

**9. 10., 16. 10., 30. 10., 6. 11., 20. 11., 27. 11., 11. 12.**

15:00 (čas bude ještě upřesněn) v učebně C14

Test

**18.12.**

15:00 (čas bude ještě upřesněn) v učebně C14

# Ukončení předmětu

- způsob ukončení – písemná zkouška – test
- uzavřené otázky, počet otázek – 21
- klasifikační stupnice

A	21-20
B	19-18
C	17-16
D	15-14
E	13-12
F	11 - 0

# Témata přednášek

1. Úvod do problematiky selekčních jazyků
2. Systematické selekční jazyky – základní pojmy, enumerativní versus fasetové SSJ, využití SSJ, klasifikace, SSJ v online prostředí
3. Mezinárodní desetinné třídění, Skupiny Konspektu
4. Předmětové selekční jazyky – úvod do problematiky, deskriptorové selekční jazyky (tvorba tezauru)
5. Soubor věcných autorit NK ČR
6. Indexování – úvod do problematiky
7. Indexování – praktické aplikace
8. Věcné vyhledávání pomocí selekčních jazyků

# Selekční jazyk

- **Selekční jazyk (SJ)** (výklad z TDKIV) - Umělý informační jazyk používaný k vyjádření identifikačních nebo obsahových selekčních údajů za účelem pořádání, ukládání a vyhledávání dokumentů.
- Doplnění (Kovář, 1984, s. 14) - Selekční jazyk - **Soustava znaků přirozeného, formalizovaného nebo umělého jazyka** vytvořená pro účely pořádání, ukládání, vyhledávání, popř. i distribuce jak dokumentů samých, tak záznamů o nich.

# Selekční jazyk

**Identifikační SJ** (výklad z TDKIV) - Selekční informační jazyk sloužící pro popis formálních charakteristik dokumentu používaný v bibliografickém popisu a jmenné katalogizaci.

Více viz Kovář, 1984, 9-18  
např. jmenné autority NK  
ČR, nakladatelské údaje

**Věcný SJ** (výklad z TDKIV) - Selekční jazyk používaný pro zpracování dokumentů pomocí věcných údajů s cílem umožnit vyhledávání dokumentů podle obsahu.

- ang. ekv. - indexing language, subject retrieval language

# Příklady dalších výkladů věcného SJ

- An indexing language is a "language" used for subject classification or indexing of documents. (Not used about systems for descriptive cataloging or indexing). (Hjørland, 2006)
- Věcný SJ - **Selekční jazyk používaný při realizaci procesu věcného pořádaní informací jako jeho výrazový a vyjadřovací prostředek.** (Kovář, 1984, 15)
  - Věcné pořádaní vychází z obsahové analýzy, informace jsou při něm pořádaný podle obsahových údajů získaných inf. analýzou. (Kovář, 1984, s. 11)

# Co nás bude zajímat?

- věcný selekční jazyk  
*! Zkráceně ho budeme nazývat pouze selekční jazyk, SJ.*
- ale nejen to
- též věcné pořádání informací
- věcný SJ je nástrojem procesu věcného pořádání informací, je jeho výrazový a vyjadřovací prostředek



# Ukázky věcných SJ různých typů

Č Í M Z A Č Í T ?

autoritou nejvyšší

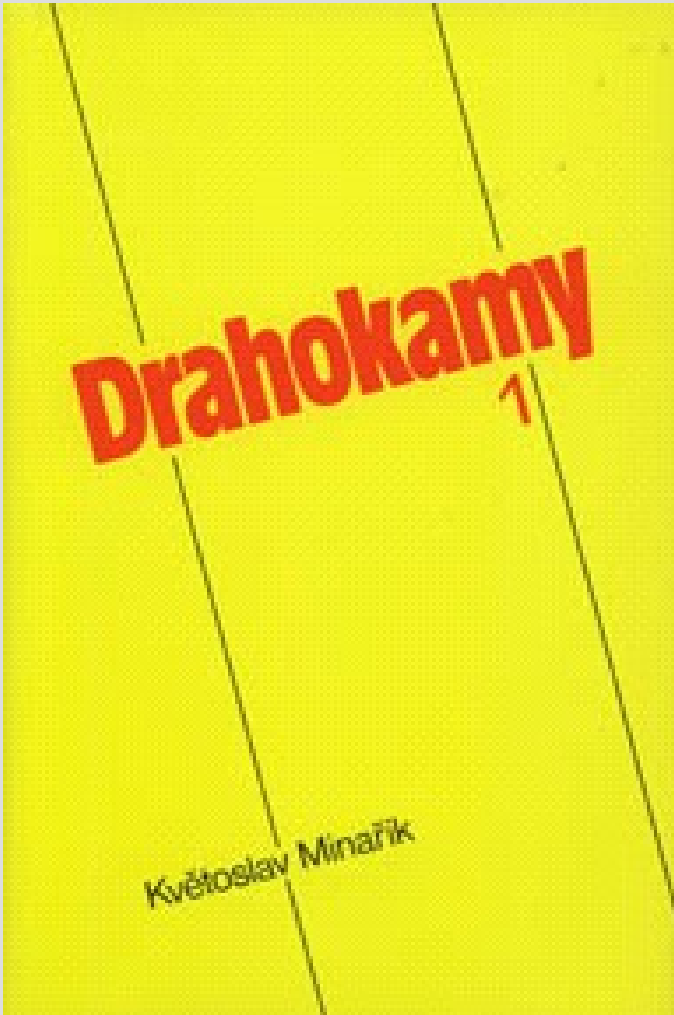
Selekčními jazyky v katalogu Národní knihovny ČR  
(NK ČR)

NEBOLI

Jaké druhy selekčních jazyků používá při indexaci  
NK ČR ?


Příklad

Hledáme knihu o drahokamech (nerost) pomocí  
klíčového slova z názvu - drahokamy



Pojednává kniha  
s tímto názvem o  
problematice, která  
nás zajímá?

Záznam v  
katalogu  
NK ČR  
(cit. 2008-16-4)

<b>Hlavní záhlaví</b>	● <a href="#">Minařík, Květoslav, 1908-1974</a>
<b>Název</b>	● <a href="#">Drahokamy / Květoslav Minařík</a>
<b>Vydání</b>	Vyd. 2.
<b>Nakl. údaje</b>	Praha : Canopus, 2004
<b>Popis (rozsah)</b>	2 sv. (319, 332 s.) ; 19 cm
<b>Roky/Sv.</b>	sv. 1-2
<b>EXEMPLÁŘE</b>	
<b>ISBN</b>	80-85202-15-8 (díl 1 : váz.) 80-85202-16-6 (díl 2 : váz.)
<b>Edice</b>	<b>Drahokamy</b> ; sv. 10, 11
<b>Předmět. heslo</b>	● <a href="#">Minařík, Květoslav, 1908-1974</a>
<b>Předmět. heslo</b>	● <a href="#">duchovní cesta</a> ● <a href="#">mystika</a> ● <a href="#">jóga</a> ● <a href="#">sebewýchova</a> ● <a href="#">sebepoznání</a> ● <a href="#">filozofie ducha</a> ● <a href="#">mystici -- Česko -- 20. stol.</a>
<b>Forma a žánr</b>	* pojednání * autobiografické vzpomínky
<b>Skupina Konspektu</b>	● <a href="#">13 - Filozofie ducha</a> ● <a href="#">929 - Biografie</a>
<b>MDT</b>	● <a href="#">2-584.5</a> ● <a href="#">2-587</a>


# Záznam v katalogu AV ČR

(cit. 2008-16-4)

## Úplné zobrazení záznamu

Zvolte formát: [Standardní Katalogizační záznam](#) [Citace](#) [Názvy polí MARC21](#) [Tagy MARC21](#)

Záznam 15 z 21

Hlavní záhlaví	• <a href="#">Minařík, Květoslav</a>
Název	• <a href="#">Drahokamy 1 / Květoslav Minařík</a>
Popis	286 s.
Vydání	1. vyd.
Naklad. údaje	• <a href="#">Praha : CANOPUS, 1993</a>
OBJEDNEJTE SI:	<a href="#">Všechny jednotky</a>
OBJEDNEJTE SI:	<a href="#">Orientální ústav</a> 
Edice	• <a href="#">Drahokamy ; Sv.2</a>
ISBN	• <a href="#">80-85202-20-4</a>

↳ Nezjistíme, že jde o knihu s odlišnou tematikou. Záznam obsahuje pouze identifikační údaje.

# Selekční jazyky používané NK ČR

**mystici - Česko - 20. stol.  
duchovní cesta**

**PHNK**                      **předmětový SJ**  
(předmětová hesla Národní knihovny)

**2-587-051**

**159.923.5**

**233-852.5**

**(437.3)**

**(0:82-94)**

**MDT**                      **systematický SJ**  
(Mezinárodní desetinné třídění)

- **Ukázka slovníků PHNK a MDT (soubor pdf Srovnání PSJ a SSJ)**

**133 - Okultismus**

**skupina Konspektu**      **hybridní SJ**  
blíží se k SSJ

# Příklady dalších možností uplatnění SJ

- **tematický rozcestník**  
<http://www.jib.cz/>



- **Catalog NCSU** <http://www.lib.ncsu.edu/catalog/> –
- **komerční databáze**, např. LLIS - Library Literature & Information Science Fulltext  
<http://library.muni.cz/ezdroje/>
- **kategorizační schémata v předmětových katalozích, portálech** např.:



<http://www.lii.org/>



<http://www.seznam.cz/>

# Seznamy různých druhů SJ

- **Taxonomy Warehouse**

<http://www.taxonomywarehouse.com/>

- **Beyond Bookmarks**

<http://www.public.iastate.edu/~CYBERSTACKS/CTW.htm>

*Oblast věcného pořádání informací a selekčních jazyků patří do širší problematiky pořádání informací a informačních jazyků (více viz Kovář, 1984, s. 9-18).*

## Pořádání informací

- ▶ v zahraniční problematika označující se jako **knowledge organization**, organization of knowledge, information organization, organization of information, organizing information
- ▶ **české** ekv **organizace znalostí**, organizace poznání pořádání znalostí, pořádání informací



# Pořádání informací

**Pořádání informací** (termín z TDKIV) ekv pořádání znalostí, termín anglicky **knowledge organization**

- Organizace a reprezentace zaznamenaných znalostí s cílem zajistit efektivní využívání údajů, informací a poznatků obsažených v informačních zdrojích.
- Pořádání informací zahrnuje vytváření organizovaných souborů identifikačních a obsahových informací získaných informační analýzou dokumentu podle předem stanoveného systému.
- Podle typu vstupních informací se vyčleňuje identifikační pořádání informací a věcné pořádání informací.

# Věcné pořádání informací

- **Věcné pořádání informací** (termín z TDKIV)  
ekv věcný popis, věcné zpracování, věcné zpracování informací, angl. ekv subject control

Specifický proces pořádání informací, při kterém jsou informace získané obsahovou analýzou dokumentu vyjádřeny pořádacími znaky věcných selekčních jazyků.



# Systematické pořádání informací

- Proces, při kterém jsou dokumenty, záznamy o nich nebo záznamy o jednotlivých informacích vřazovány na příslušné místo v rámci systematicky uspořádaného systému lidského poznání, kde se jejich postavení řídí v podstatě rodo-druhovými vztahy a slovní formulace obsahu dokumentu bývá většinou nahrazena znaky umělého jazyka (notacemi).  
(volně podle Kovář, 1984, s. 13)
- ekv třídění, klasifikace, systematické zpracování, systematická katalogizace
- vyjadřovacím prostředkem je systematický SJ

# Předmětové pořádání informací

- Proces, při kterém slovně vyjádřený obsah, téma či předmět dokumentu, jeho částí nebo jednotlivých v něm obsažených informací je převáděn do jednou více, jindy méně normalizovaných a formalizovaných hesel (deskriptorů), která jsou ve slovnících uspořádána abecedně. (volně podle Kovář, 1984, s. 13)
- ekv. heslování, předmětové zpracování, indexace, někdy nevhodně předmětové třídění, předmětová klasifikace
- vyjadřovacím prostředkem je předmětový SJ

# Účel věcného pořádní informací (Taylor, 2004, s. 242)

- Poskytnutí věcného přístupu k informačním objektům prostřednictvím vyhledávacího nástroje.
- Seskupení obsahově stejných nebo podobných informačních objektů.
- Logická lokace podobných informačních objektů.
- Šetří čas uživatele.

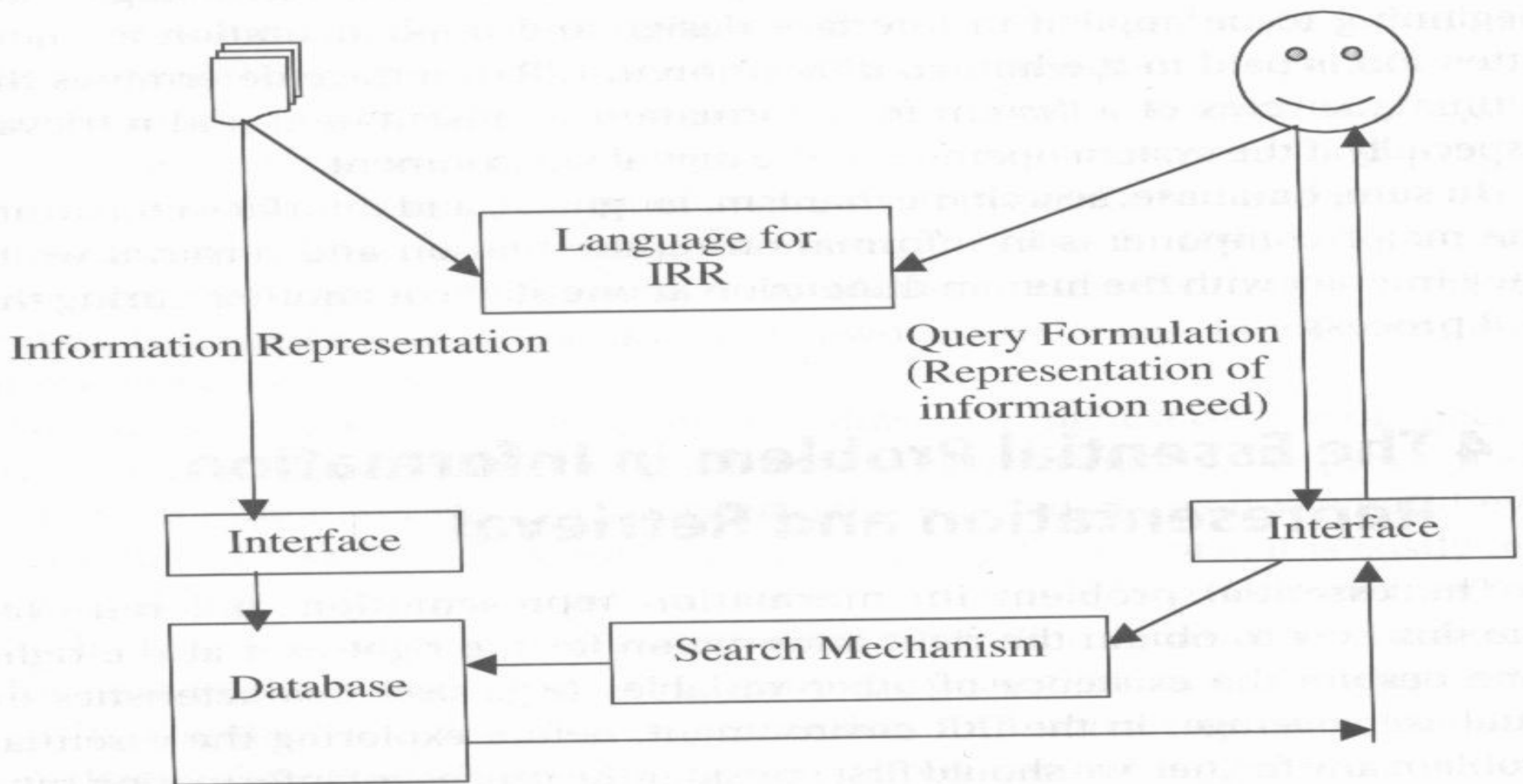
# SJ v informačním systému

SJ je součástí několika procesů v IS a současně subsystémů IS, vyplývá to z jeho funkce:

- výrazový a vyjadřovací prostředek procesů pořádání, ukládání a vyhledávání informací.

SJ je prvkem i prostředkem spojujícím pořádací, ukládací a vyhledávací subsystém IS. (více viz Kovář, 1984, s. 9-18)

# SJ v procesu reprezentace a vyhledávání informací (Chu, 2007)



**Figure 1.1 Process of Information Representation and Retrieval**

# SJ v procesu reprezentace a vyhledávání informací (Chu, 2007)

- Vedle přirozeného jazyka patří SJ mezi důležité složky tohoto procesu.
- Je výhodnější užít přirozený jazyk či SJ?
  - Obojí výhody i nevýhody (viz přednáška Věcné vyhledávání pomocí SJ).



# Charakter SJ z hlediska lingvistiky

- **Přirozený jazyk** (výklad z TDKIV) - Jazyk, který se vyvíjí, jehož pravidla vyplývají z úzu (nejsou definována formálně) a jenž je užíván konkrétním jazykovým společenstvím (např. národní jazyky).
- **Umělý jazyk** (výklad z TDKIV) - Jazyk vytvořený nebo řízený pomocí souboru předem stanovených pravidel s přesně definovanou syntaxí a sémantikou vytvořený pro určité praktické účely. Typem umělého jazyky je např. informační jazyk.
- SJ patří mezi informační jazyky, tj. umělé jazyky používané v rámci informačního systému.

# Charakter SJ z hlediska lingvistiky

Přirozený jazyk (dále PJ) má dvě základní složky:

- slovník (lexikum)
- pravidla, jak s tímto slovníkem nakládat, tedy gramatiku
  - morfologie: nauka o tom, jak se slova ohýbají (flexe) nebo tvoří (slovotvorba)
  - syntax: způsob, jakým se slova spojují do vět (zjednodušeně řečeno).

Podobně i SJ má tyto tři složky:

- slovník, morfologii a syntax

# Srovnání tří základních složek SJ a PJ: slovník

- Slovníku (lexiku) PJ odpovídají u SJ řízený slovník nebo (PSJ) klasifikační tabulky (SSJ).  
příklad:
  - Databáze národních autorit NK ČR  
<http://sigma.nkp.cz/F/>
  - Český teologický tezaurus  
[http://ckis.cuni.cz/F/2DL8NDU44PTBPE62VFU2X3KVDTA5I7CHUXNVXBTJ2FA74J3TBP-28409?func=find-a-o&local\\_base=etf11](http://ckis.cuni.cz/F/2DL8NDU44PTBPE62VFU2X3KVDTA5I7CHUXNVXBTJ2FA74J3TBP-28409?func=find-a-o&local_base=etf11)
  - Soubor MDT MRF  
online <http://aip.nkp.cz/mdt/>

# Srovnání tří základních složek SJ a PJ: slovník

## Pořádací znak

- Slovu nebo slovnímu spojení u PJ odpovídá u SJ pořádací znak.
- **Předmětový SJ** → **lexikální jednotka**  
**mystici - Česko - 20. stol.** - složené téma vyjádřené temat. autoritami (Db národních autorit NK ČR) a zapsané prekoordinovaně, základ přirozený jazyk
- **Systematický SJ** → **klasifikační znak**  
**2-587-051(437.3)“19“** - složené téma vyjádřené notací MDT, základ kód

- Pořádacími znaky přirozeného, formalizovaného nebo umělého věcného SJ mohou být slova převzatá z názvu nebo podnázvu dokumentu, ale také slova převzatá z jiného místa dokumentu, z redukce textu dok. (referát, anotace) nebo formulovaná informačním pracovníkem, a to zcela volně, anebo na podkladě nějakého řízeného slovníku. (volně podle Kovář, s. 11)

# Srovnání tří základních složek SJ a PJ: morfologie

- Morfologii (přesněji její části, slovo tvorbě, protože flexi nelze u SJ uplatnit) PJ odpovídají u SJ **pravidla pro tvorbu pořadacích znaků** (zde je jednou z podstatných otázek především míra prekoordinace/postkoordinace pořadacích znaků - viz dále)
- např. u **SSJ typ notačního systému** - MDT - desetinná notace

# Srovnání tří základních složek SJ a PJ: syntax

- Syntaxi (skladbě) PJ odpovídají u SJ **pravidla pro indexaci (klasifikaci) a formulaci dotazu při vyhledávání** (tedy pravidla pro spojování a kombinaci pořadacích znaků)
- **Větě** v PJ odpovídá v SJ **selekční obraz dokumentu** nebo **selekční obraz dotazu**.

# Typy SJ

Nejčastější dělení: podle uspořádání slovníku, podle charakteru pořádacích znaků SJ, šířky tematického zaměření.

- *Pořádací znaky lze ve slovníku SJ uspořádat abecedně, nebo tematicky, takže se SJ dělí na:*
  - **předmětové SJ** (více viz konkrétní přednáška) - abecední princip
  - **systematické SJ** (více viz konkrétní přednáška) - tematický princip



# Typy SJ

- *Charakter pořádacích znaků SJ je určen mírou postkoordinace, resp. prekoordinace, takže se SJ dělí na:*
  - **postkoordinované SJ**
  - **prekoordinované SJ**

# Prekoordinace versus postkoordinace

- Postkoordinace a prekoordinace jsou principy, které se uplatňují
  - při tvorbě pořadacích znaků ve slovníku SJ (morfologie),
  - při indexaci (klasifikaci) a vyhledávání (syntax).

Neexistuje čistě postkoordinovaný nebo čistě prekoordinovaný SJ. Postkoordinace a prekoordinace se uplatňuje v každém SJ; podle toho, který princip v daném SJ převažuje, označujeme jej jako prekoordinovaný nebo jako postkoordinovaný SJ.

# Prekoordinace versus postkoordinace

- **Na úrovni slovníku**

postkoordinace/prekoordinace vyjadřuje míru složení pořádacího znaku. Pokud jsou jako pořádací znaky použity jednoduché pojmy, hovoří se o postkoordinaci. Pokud jsou použity složené pojmy, hovoří se o prekoordinaci.

- **postkoordinovaný PSJ**

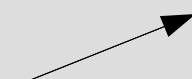
Dvořák, Antonín, 1841-1904

drama

televizní inscenace

bibliografie

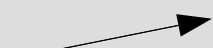
složené téma vyjádříme  
jednotlivými pořádacími znaky  
(deskriptory)



- **prekoordinovaný PSJ**

Dvořák, Antonín – hry divadelní a televizní – bibliografie  
(dramatik)

složené téma  
vyjádříme jedním  
pořádacím znakem  
(předmětové heslo)



# Prekoordinace versus postkoordinace

- Prekoordinovaný nebo postkoordinovaný charakter pořadacích znaků ovlivňuje samozřejmě i indexaci (klasifikaci) a vyhledávání.
  - U postkoordinovaného SJ je složený pojem popsán dílčími pořadacími znaky, které jsou k dokumentu přiřazovány samostatně a koordinovány (kombinovány) až při vyhledávání.
  - U prekoordinovaného SJ je naopak složený pojem vyjádřen jediným pořadacím znakem, který se použije jak pro indexaci, tak pro vyhledávání.

# Prekoordinace versus postkoordinace

## Příklad

- postkoordinovaná indexace v bibliogr. záznamu  
mořeplavci  
Itálie  
18.-19. stol.  
studie
- prekoordinovaná indexace v bibliogr. záznamu  
mořeplavci -- Itálie -- 18.-19. stol. - studie
- Pozn. Současné rešeršní systémy vyhledávají i v jednotlivých prvcích  
předmětového hesla. Rozdíl znatelný při prohlížení v rejstřících.

# Typy SJ

- Výsledkem tohoto třídění jsou celkem čtyři typy SJ:
  - prekoordinované systematické SJ
  - postkoordinované systematické SJ
  - prekoordinované předmětové SJ
  - postkoordinované předmětové SJ

# Příklady typů SJ

- **prekoordinované SSJ**
  - Mezinárodní desetinné třídění, Deweyho desetinné třídění, Třídění Kongresové knihovny, Cutterova expanzivní klasifikace
- **postkoordinované SSJ**
  - Dvojtečkové třídění (Ranganathan), Blissovo bibliografické třídění
- **prekoordinované PSJ**
  - předmětová hesla
- **postkoordinované PSJ**
  - deskriptorové SJ, volně tvořená klíčová slova, permutované (rotované) rejstříky (KWIC, KWOC)

# Šířka tematického zaměření

- Selekční jazyk lze klasifikovat také podle šíře univerza (oblasti), které popisují.
- **univerzální selekční jazyk** - univerzum selekčního jazyka pokrývá všechny oblasti lidského poznání a lidské činnosti
- **speciální selekční jazyk** - univerzum selekčního jazyka je omezeno na určitý obor (Pinkas, 2002)



# Použitá a povinná literatura

- Balíková, M. 2001 *Problematika věcného pořádní informací a jejich zpřístupnění*. *Národní knihovna*, 2001, roč. 12, č. 3, s. 175-186. Dostupné z WWW: <http://full.nkp.cz/nkkro103/0103175.html>
- Hjørland, Birger. 2006. *Indexing language* [online]. Last edited: 19-10-2006 [cit. 2009-09-29]. Dostupné z WWW: [http://www.db.dk/bh/lifeboat\\_ko/CONCEPTS/indexing\\_languages.htm](http://www.db.dk/bh/lifeboat_ko/CONCEPTS/indexing_languages.htm)
- Kovář, Blahoslav. 1984. *Věcné pořádní informací a selekční jazyky*. Praha : ÚVTEI, 1984. 1.1 *Úvod do problematiky věcného pořádní informací a selekčních jazyků*, s. 9-18.
- Pinkas, Otakar. 2002. *Zpracování informačních fondů*. Vyd. 1.V Praze : Vysoká škola ekonomická, 2002. Kap. 7 *Selekční jazyky*, s. 89-114.

# Použitá a doporučená literatura

- Chu, Heting. 2007. *Information representation and retrieval in the digital age*. Medford : Information Today, 2007. 248 s.
- *KTD : Česká terminologická databáze knihovnictví a informační vědy (TDKIV)* [online]. Praha : Národní knihovna České republiky, 2003. Dostupné z WWW: <http://sigma.nkp.cz/cze/ktd>
- Schwarz, Josef. 2007. *Selekční jazyky 1 : Úvod do problematiky* [ppt]. Přednáška č. 2 (kombinované studium). 9.11.2007.
- Taylor, Arlene G. 2004. *The organization of information*. Westport : Libraries Unlimited , 2004. 417 s. ISBN 1-56308-976-9 - dostupné u registračního pultu ÚK FF MU, složka Selekční jazyky