

# VIKBA11 Selekční jazyky 1

## Úvod do problematiky

Přednáška č. 4 (denní studium) – 7.12.2007

*Filozofická fakulta Masarykova Univerzity, Kabinet knihovnictví -  
Ustav české literatury a knihovnictví  
podzim 2007*

**Mgr. Josef Schwarz**  
**[schwarzjv@seznam.cz](mailto:schwarzjv@seznam.cz)**



# Tvorba a správa SJ

# Metodika SJ

- ◆ metodika tvorby (implementace) SJ
- ◆ metodika správy (aktualizace) SJ
- ◆ metodika využívání (indexace, vyhledávání) SJ

# 4. úkol

## ◆ Zadání

- ◆ Pro vypracování tezauru lze využít program MultiTes.
  - Program je dostupný na <http://www.multites.com/download.htm> a lze si ho stáhnout pro zkušební účely jako Trial Software. Podmínkou je vyplnění stručného formuláře, na jehož základě obdržíte na vámi uvedenou mailovou adresu heslo pro přístup do programu.

# Obecná metodika implementace SJ

1. Rozhodnutí o implementaci SJ
2. Volba typu SJ
3. Základní vymezení tematického rozsahu SJ
4. Průzkum existujících SJ
5. Volba způsobu implementace
  - tvorba vlastního SJ
  - modifikace existujícího SJ
  - převzetí existujícího SJ

# Metodika tvorby SJ

1. Podrobné vymezení tematiky
  - vymezení jádra SJ
  - vymezení hraničních oblastí SJ
2. Volba systému třídění, faset, tematických oblastí
3. Volba charakteru pořadacích znaků a slovníku SJ
4. Excerptce LJ
  - ◆ deduktivní metoda
  - ◆ induktivní metoda
5. Zpracování excerptce a tvorba PZ
  - ◆ pre/postkoordinace, specifičnost, homonymie, synonymie
6. Strukturace PZ
7. Prezentace SJ

# Převzetí existujícího SJ

1. Volba SJ
2. Cizojazyčný SJ: Překlad/převod
  - problémy se strukturou
  - problémy s lexikem

# Specifické otázky tvorby SJ



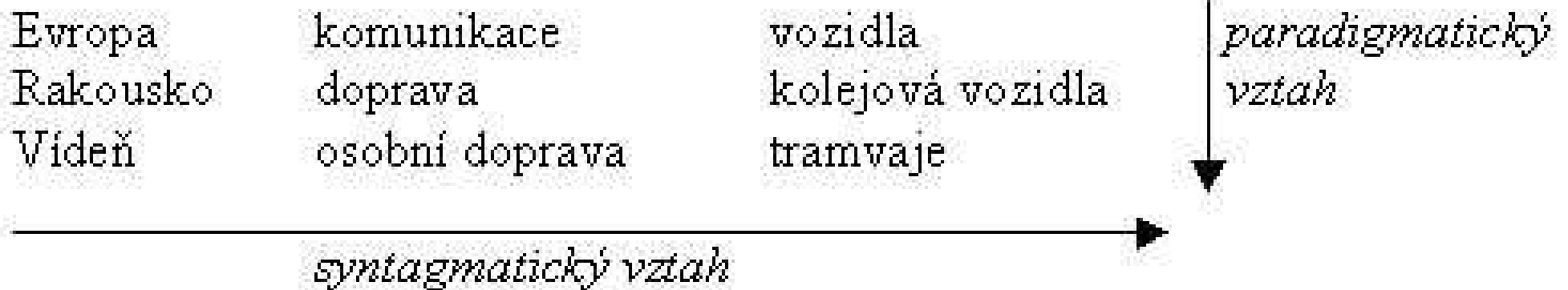
# Prameny

- ◆ ČSN 010193
- ◆ ČSN 010172
- ◆ Aichison, J., Gilchrist, A., Bawden, D.  
*Thesaurus construction and use : a practical manual.* 4th ed. Chicago : Fitzroy Dearborn Publishers, c2000. 218 S. (v knihovně FF MU)

# Vztahy mezi pojmy

- ◆ aposteriorní (syntagmatické)
- ◆ apriorní (paradigmatické)

*Příklad:*



# Typy pojmů

## ◆ konkrétní entity

- předměty, věci a jejich části (př.: opice, stavby, květináč)
- materiály (př.: cement, dřevo, bavlna)

## ◆ abstraktní entity

- činnosti a události (př.: evoluce, dýchání, lyžování)
- vlastnosti věcí, materiálů n. činností (př.: síla, přesnost, viskozita)
- obory n. vědní disciplíny (př.: fyzika, psychologie, informační věda)
- měřicí jednotky (př.: kilogram, metr, ampér)

## ◆ individuální jednotky (identifikátory)

- geografická jména
- jména osob
- jména organizací
- jména výrobků

# Tvar LJ

## ◆ substantivum n. substantivní slovní spojení

### ■ adjektivní

- ◆ modifikátor (adjektivum), základ (substantivum) *(př.: ženská práva, elektronická média, autodráhy)*

- s předložkovou vazbou *(př.: kurzy pro pokročilé, školy pro nevidomé děti)*

## ◆ adjektiva, příslovce, slovesa

- ◆ do ŘS se obvykle nezařazují
- ◆ výjimky

# Tvar LJ (pokr.)

## ◆ singulár/plurál

### ■ plurál:

- ◆ počítatelná substantiva (př.: *planety, děti, akvária*) (ale: části těla sg. (př.: *ústa, hlava*))

### ■ singulár

- ◆ nepočítatelná substantiva (př.: *sníh, pára, nikl*) (ale: třídy s více členy pl. (př.: *cukry, jedy, mouky*))
- ◆ abstraktní pojmy (př.: *autorita, vodivost, kapitalismus*) (ale: třídy s více členy pl. (př.: *biologické vědy, chemické reakce*))

- vyjádření odlišných pojmů pomocí sg./pl. (př.: *cukr/cukry, mýtus/mýty*)

## ◆ homografy, polysémy

- relátor (př.: *tiskárna (stroj) x tiskárna (budova); nafta (surovina) x NAFTA (organizace)*)
  - ◆ relátor se uvádí spolu s deskriptorem

# Poznámky

◆ **poznámka o rozsahu** *(př.: režijní náklady - Výdaje vynaložené podnikem, které nejsou přímo ztotožnitelné s náklady na výrobu.)*

- omezení významu pro účely daného systému

◆ **indexační poznámka** *(př.: mezinárodní nástroj - Uveďte ve spojení s danou organizací)*

◆ **historie** **LJ** *(př.: deskriptor zařazen v roce 1998)*

◆ **zdroj** **LJ** *(pramen použitý při výběru lexikální jednotky)*

# Sousloví

◆ **sémantický rozklad – nedoporučuje se** (př.: *barometr*  
= *tlak + měření + přístroj*)

◆ **syntaktický rozklad** (př.: *knihovnické časopisy = knihovnictví + časopisy*)

◆ **sousloví ponechaná beze změny:**

- **sousloví je běžné a užívané** (př.: *informační zdroje, invalidní vozík*)
- **rozklad vede ke ztrátě významu** (př.: *filozofie dějin, letadlová loď*)
- **sousloví obsahuje vlastní jméno** (př.: *Bradfordův zákon, Planckova konstanta*)
- **modifikátor ztratil původní význam** (př.: *lehký průmysl*)
- **modifikátor bez přímého vztahu** (př.: *stromová struktura*)
- **sousloví se „středním členem“** (př.: *benzinový motor*)
- **synkategorematická substantiva** (př.: *umělé květiny, čokoládová vejce, světelný rok*)

# Sousloví (pokr.)

## ◆ Sousloví rozložená

- Základ: vlastnost/část x modifikátor:nositel/celek (př.: *lodní motor = lodě + motory; knihovní personál = knihovny + personál*)
- Činnost/objekt (př.: *signování knih = signování + knihy; indexace článků = indexace + články; stěhování ptáků = stěhování + ptáci*)

## ◆ Příklad pravidel pro sousloví

- Appendix A – AAT compound term rules. In Aichison, J., Gilchrist, A., Bawden, D. *Thesaurus construction and use : a practical manual*. 4th ed. Chicago : Fitzroy Dearborn Publishers, c2000.



# Vztahy mezi lex. jednotkami

- vztah ekvivalence
- vztah hierarchie
- vztah asociace

# Vztah ekvivalence

Do vztahu ekvivalence jsou uváděny **synonymní lexikální jednotky** - termíny, které se liší formou, ale jejichž obsah je identický (označují stejný denotát).

- preferovaný termín versus nepreferované termíny

Záhlaví *práva žen*

Odkaz. forma *ženská práva [o]*

# Volba deskriptoru/nedeskriptoru

- ◆ pravopis
  - používání nejrozšířenějšího popř. kodifikovaného pravopisného tvaru
  - využití autoritních slovníků
  - dublety (př.: *sirup x syrob; salnitr x sanytr*)
- ◆ zkratky a akronymy (př.: *UNICEF, OSN, ČEDOK, LIS*)
  - preferuje se rozepsaná podoba - mezinárodní desetinné třídění X MDT [o]
  - výjimky v dané oblasti u široce používaných zkr. a akr.
- ◆ přejatá slova (př.: *software x programové vybavení; perestrojka x přestavba; gender studies x genderová/rodová studia*)
  - novotvary z cizího jazyka – běžně užívané lze zařadit do tezauru
  - volba mezi přejatým slovem a překladem – rozšířenost
- ◆ slangové termíny a žargon (př.: *hippies, íčko, koks*)
  - lze zařadit, pokud neexistuje pro nový pojem jiná, vhodná alternativa
  - jde-li o alternativní výraz k existujícímu, ustálenému termínu, ustálený je preferovaný t.

# Volba deskriptoru/nedeskriptoru

- ◆ **obecné názvy a obchodní názvy** (př.: *polytetrafluorethylen x teflon*)
  - preferované jsou obecné názvy jako deskriptory, je-li obchodní název široce používán zavede se jako nedeskriptor
- ◆ **běžné názvy a vědecké názvy** (př.: *bubo bubo x výr velký; arthritis uratica x dna*)
  - volba preferovaného termínu se odvíjí od uživatelů systému
- ◆ **místní názvy** (př.: *Vídeň x Wien; Holandsko x Nizozemí/Nizozemsko; Česko x Česká republika*)
  - oficiální x hovorový název, přeložený x originální název
  - volba deskriptoru se odvíjí od potřeb uživatelů systému

# Kvazisynonyma

◆ **Antonyma** (*př.: anotovaný záznam x neanotovaný záznam;  
zákonnost x nezákonnost*)

◆ **Hierarchizace vztahu ekvivalence**

- *KOMPAKTNÍ DISKY*

**EKV** *CD*

*CD-R*

*CD-I*

*CD-R*

*CD-RW*

- *HORNINY*

**EKV** *břidlice*

*čedič*

*žula*

# Vztah hierarchie

Nastává u lexikálních jednotek téhož sémantického okruhu a vyjadřuje poměr nadřazenosti a podřazenosti.

Je vytvářen pouze u preferovaných termínů.

Podřazená jednotka musí náležet k témuž základnímu typu pojmu jako její nadřazená lex. jednotka – věc x činnost x vlastnost.

- Generický vztah
- Partitivní vztah
- Kauzální vztah

# Generický vztah

Vymezuje souvislost mezi určitou třídou nebo kategorií a jejími členy nebo druhy.



➤ rodo-druhový vztah

# Partitivní vztah

„Sémantický hierarchický vztah mezi dvěma pojmy, z nichž jeden vyjadřuje celek a druhý jednu jeho část“ (TDKIV).

## 4 kategorie

- tělesné systémy a orgány (*př.: kostra -> lebka -> nadočnicové oblouky*)
- geografické lokality (*př.: Evropa -> Rakousko -> Vídeň*)
- disciplíny (*př.: jazykověda -> obecná jazykověda -> gramatika*)
- hierarchické společenské struktury (*př.: univerzita -> fakulta -> katedra -> kabinet*)



# Kauzální vztah

Vazba mezi kategorií věcí nebo událostí vyjádřenou obecným substantivem a individuálním případem dané kategorie (jednotková třída označena vlastním jménem).

pohoří

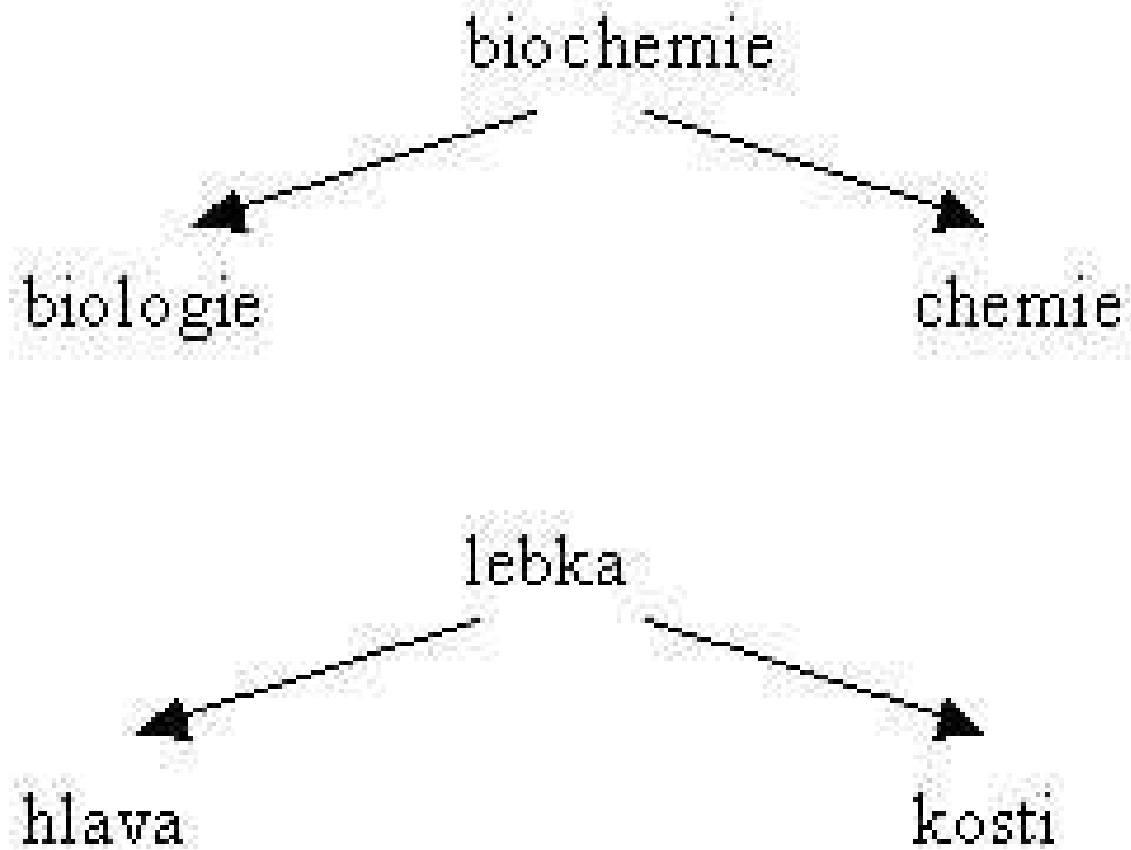
Alpy

Himaláj



individuální případy uvedené  
obecné kategorie

# Polyhierarchie



# Vztah asociace

Vazba mezi dvěma lex. jednotkami, které

- nejsou ve vztahu ekvivalence,
- nejsou spolu v hierarchickém vztahu.

Použití: „*Lexikální jednotky jsou mentálně natolik asociovány, že je účelné vyjádřit jejich vazbu v tezauru*“.

Pravidlo pro stanovení vztahu: Při indexaci jednou jednotkou je vyvolána potřeba indexace druhou asoc. jednotkou.

Záhlaví                      *versologie*

Asoc. termín                *verš*

# Asociace

- ◆ **Disciplína/objekt** (př.: *informační věda -- informace*)
- ◆ **Úkon, proces/konatel, nástroj** (př.: *katalogizace – katalogizační pravidla*)
- ◆ **Činnost/výsledek** (př.: *katalogizace – katalogizační záznam*)
- ◆ **Činnost/předmět** (př.: *indexace -- dokument*)
- ◆ **Pojem/vlastnost** (př.: *informace -- pravdivost*)
- ◆ **Pojem/původ** (př.: *Francouzi -- Francie*)
- ◆ **Kauzální souvislost** (př.: *nemoc -- léčení*)
- ◆ **Věc/agens působící proti ní** (př.: *rostliny -- herbicidy*)
- ◆ **pojem/jednotka měření** (př.: *informace -- bit*)
- ◆ **synkategorematické výrazy** (př.: *čokoládová vejce – vejce; umělé květiny -- květiny*)

# Časté chyby v úkolu č. 3

## ◆ chybný odkaz/ekvivalent

- Za ekvivalent považován termín vyjadřující širší/nadřazený pojem.

Do vztahu ekvivalence jsou však uváděny synonymní lexikální jednotky,

např. *miniházená* (navržený termín) EKV *míčové sporty* (termín dán do ekvivalentního vztahu), správně *miniházená* ND *míčové sporty*

◆ *miniházená* (navržený termín) → *házená*,  
*míčové sporty* (termíny dány do vztahu a jsou již součástí SVA)

**miniházená**

ND 1 házená

ND 2 míčové sporty

# Časté chyby v úkolu č. 3

## ◆ chybný odkaz/ekvivalent

- Za ekvivalent považován termín, který je již součástí SVA buď jako termín autoritní, nebo ekvivalentní,

např. *grunge* (navržený termín) →

EKV *rocková hudba* (termín dán do ekvivalentního vztahu, ale je již ekvivalentem k termínu *rock*),

*integrační informační zdroje* (navržený termín) →

EKV *seriálové publikace* (termín dán do ekvivalentního vztahu, ale jde o autoritní termín)

# Časté chyby v úkolu č. 3

- ◆ stejný nadřazený deskriptor a asociovaný termín

- např. *miniházená* (navržený termín)

*AD házená ND házená*

(termín dán do vztahu asociace, ačkoliv jde o nadřazený termín)

Nelze asociovat s termínem, který je v hierarchickém vztahu.

# Časté chyby v úkolu č. 3

## ◆ chybný podřazený deskriptor

- Za podřazený deskriptor považován deskriptor vyjadřující širší pojem, např. *Eagleheart* (navržený termín) PD *metalové skupiny* (termín navržený jako podřazený termín)
- Jde o kauzální vztah, Eagleheart je název metalové skupiny
- správně:  
*Eagleheart*  
ND *metalové skupiny*



# Časté chyby v úkolu č. 3

## ◆ chybný podřazený deskriptor

- *skupinová výsadba* (navržený termín) PD *okrasné dřeviny* (termín navržený jako podřazený termín)

Je každá okrasná dřevina součástí skupinové výsadby? či Je každá okrasná dřevina druhem skupinové výsadby?

Pokud zní odpověď NE, nelze podřadit.

# Časté chyby v úkolu č. 3

- ◆ nedodržení následující zásady:  
Podřazená jednotka musí náležet k témuž základnímu typu pojmu jako její nadřazená lexikální jednotka.
- ◆ *mapování porostů* (navržený termín – vyjadřuje činnost) → PD *dřeviny* (termín navržený jako podřazený termín – vyjadřuje konkrétní entitu)