



VIKMA05

Organizace znalostí

29. 11. 2013: Systémy organizace znalostí založené na přirozeném jazyce (P11)

FF MU, podzim 2013

Mgr. Josef Schwarz

126172@mail.muni.cz



Osnova

- JNK
- Praktická část
- Teoretická část
 - Pořádací systémy založené na přirozeném jazyce
 - Klíčová slova
 - Unitermy
 - Předmětová hesla
 - Tezaury
- Zadání 11. dílčího kroku DA



Praktická část

- Různé hierarchické stromy (kategorie)
- Dvě či více tříd pro jeden pojem



Teoretická část



Vybrané otázky

- Zdroj/Původce
- Účel
- Otázky struktury a morfologie
 - Prekoordinace / postkoordinace
 - Inverze



Klíčová slova

○ Původce

● Autor textu

- KS z názvu
- přidaná KS
 - Tagování (folksonomie)
Tagová mračna

● Program

● Informační profesionál



Klíčová slova

- Výhody
 - Otevřenost
 - Uživatelská srozumitelnost
- Nevýhody
 - Neřízenost
 - Nestrukturovanost



Předmětová hesla

○ Definice

- „Prekoordinovaný selekční informační jazyk používaný pro strukturní zpracování dokumentů nebo dat pomocí abecedně uspořádaných termínů nebo jejich kombinací s cílem umožnit předmětový přístup. Jeho lexikum představuje heslář (příp. pouze předmětový katalog), tj. řízený soubor lexikálních jednotek vybraných z přirozeného jazyka a užívaných k vyjádření témat dokumentů. Lexikální jednotkou je předmětové heslo.“ (TDKIV)

○ Základní charakteristiky „klasického“ předmětového hesla

- Strukturované (prekoordinované) lexikální jednotky
- Inverze sousloví
- (Částečně) řízený slovník



Předmětová hesla

- Příklady:

- malotraktory - charakteristiky technické - malovýroba zemědělská – studie ([NKP](#))
- řízení vnitropodnikové - chozraščot - metody - podniky socialistické ([NKP](#))
- zabezpečení počítačových sítí + počítačová kriminalita + internet ([NKP](#))
- Vnitropodnikové řízení – hospodárnost + Hospodárnost – vnitropodnikové řízení ([MZK](#))
- podniky + podnikový management (JVK ČB)
- průmyslové podniky + vnitropodnikový chozrasčot (MVKO)



Unitermy

- Uniterm = „nejmenší významový prvek selekčního jazyka použitý k vyjádření specifického pojmu v rámci systému koordinovaného indexování“ (ČSN ISO 5127-6/3.4.2-10)
- První verze: 1951 (Mortimer Taube) (jednoslovné lexikální jednotky, žádné vztahy)
- Rozšířená verze: 1955
- Konečná verze: přelom 50. a 60. let (částečně řízený slovník, gramatika, prekoordinace)
- Důvody vzniku: překonání omezení MDT a PH



Unitermy

- Výhody
 - Princip postkoordinace
 - Uživatelská přívětivost (ve srovnání s MDT a PH)
- Nevýhody
 - Neřízený slovník
 - Syntaktický rozklad
 - Nestrukturovanost



Tezaury

○ Definice

- „ Řízený a měnitelný slovník deskriptorového selekčního jazyka uspořádaný tak, že explicitně zachycuje apriorní (paradigmatické) vztahy mezi lexikálními jednotkami. Podle jazykového zaměření se rozlišuje jednojazyčný a vícejazyčný tezaurus, podle šíře tematického zaměření speciální, polytematický a univerzální tezaurus, podle struktury fasetový a tematický tezaurus, podle funkce tradiční, indexační a vyhledávací tezaurus. Dále se vyčleňuje cílový a zdrojový tezaurus.“



Tezaury

- Základní charakteristika
- Výhody X nevýhody