



VIKMA05

# Organizace znalostí

3. 3. 2017: Hlediska KO, doménová analýza (P2)

FF MU, jaro 2017

Mgr. Josef Schwarz

[126172@mail.muni.cz](mailto:126172@mail.muni.cz)



# Úvod

- Výsledky 1. dílčího úkolu
- Celkový pohled na KO – přístupy
- Doménová analýza
- Zadání 1. a 2. kroku DA



# Smysl KO

- obsahová analýza dokumentů  
(informačních objektů)
  - *subject X content X text X information analysis*
- pořádání informací
- vyhledávání informací



# Východiska KO

- vyjádření obsahu (*content*)
  - *co je obsah?*
  - *jaký je vztah obsahu k autorovi dokumentu?*
- vyjádření (informační) potřeby (*need*)
- informační potenciál (*informative potential*)



# Východiska KO

- ad hoc klasifikace
- pragmatická klasifikace
- vědecká klasifikace



# Hlediska KO

(HJØ08)

1. Tradiční přístupy
2. Manažerské přístupy
3. Fazetová analýza
4. Počítačově orientované přístupy
5. Bibliometrie
6. Uživatelsky a kognitivisticky orientované přístupy
7. Doménová analýza



# 1. Tradiční přístupy

- Základní charakteristika
  - Enumerace
  - Vědecké disciplíny
  - Pozitivistický pohled na svět
  - Bez specifické klasifikační metodologie a teorie
  - Nároky na předmětnou znalost



# 1. Tradiční přístupy

## ○ Principy

- Řízený slovník
- Ch. A. Cutter – princip specifičnosti (1876)
- E. W. Hulme – princip literární základny (1911)
- Princip řazení od obecného ke specifickému





# 1. Tradiční přístupy

## ○ Příklady:

- DDT – manažerský přístup
- LCC – klasifikace (nejen) pro jednu knihovnu
- MDT – sirotek jednoho velikášství
- BC (Bliss 1) – snaha vystihnout řád přírody a vědy



## 2. Manažerské přístupy

- Základní charakteristika
  - Důraz na praktickou použitelnost, pragmatismus
  - Snaha o standardizaci pořádání
  - Institucionální orientace



## 2. Manažerské přístupy

- Příklad
  - DDT



## 3. Fazetová analýza

- Základní charakteristika

- Východiska v logice, formální pojmová analýza
- Analyticko-syntetický přístup
- Explicitní pravidla (metodologie)



## 3. Fazetová analýza

- Příklady

- Dvojtečková klasifikace (S. R. Ranganathan - PMEST)
- Classification Research Group 1952-1968



## 3. Fazetová analýza

- Omezení

- „Atomizace“ témat
- Chybějící kritická analýza fazetového principu
- Ranganathanova snaha najít jediný nejlepší klasifikační systém



# 4. Počítačově orientované přístupy

- Základní charakteristika
  - Algoritmizovatelné postupy
  - Automatické nebo poloautomatické zpracování textu (přirozeného jazyka)
    - Předpoklad: téma dokumentu je určeno jím samým a lze jej objektivně zjistit
  - Empirické testování na základě měr úplnosti a přesnosti
  - Nestrukturované vyhledávání
  - Transformace dotazu do množiny dokumentů



## 4. Počítačově orientované přístupy

- Cranfield I (1957-1962)
  - Unitermy (82,0 %)
  - Předmětová hesla (81,5 %)
  - MDT (75,6 %)
  - Fazetové klasifikace (73,8 %)





## 4. Počítačově orientované přístupy

- Omezení



# 5. Bibliometrie

- Základní charakteristika
  - Citační vazby
  - „Sociální“ forma organizace znalostí



# 5. Bibliometrie

- Výhody
  - Citují odborníci
  - Hloubka „indexace“ je dána počtem citací
  - Dynamická forma reprezentace znalostí
  - Kontextová interpretace



# 5. Bibliometrie

- Nevýhody

- Vztah mezi citací a jejím obsahem je nejasný
- Neexistuje logická struktura a soustavná klasifikace
- Omezení na (převážně) vědecké články



## 6. Uživatelsky a kognitivisticky orientované přístupy

- Základní charakteristika
  - Informace o uživateli nebo od uživatele (uživatelské průzkumy)
  - Uživatel jako konečný cíl informačního procesu
  - Aktivní role uživatelů v organizaci znalostí - „samoorganizace“ (folksonomie)



# 7. Doménová analýza

## ○ Základní charakteristika

- Metateoretický nástroj
- Ontologická, epistemologická a sociologická východiska
- Doménově chápaná organizace znalostí
- Diskurzní komunity
- Zohlednění potřeb dané skupiny uživatelů nebo účelu pořádání
- Klasifikace dokumentů není objektivní nebo neutrální (univerzální)



# 7. Doménová analýza

## ○ Teoretické předpoklady

1. Klasifikace je pořádání entit na základě určitých vlastností
2. Vlastnosti těchto entit nejsou dané, ale určené pomocí jejich popisu a předporozuměním těmto entitám; popisem nelze vyčerpat všechny vlastnosti entity
3. Reprezentace (popis) entity je daná vlastnostmi entity i subjektem, který tento popis provádí; reprezentace entity je více či méně ovlivněná účelem popisu a teoretickými východisky
4. Výběr vlastností entity, určené pro klasifikaci, musí zohlednit účel klasifikace; neexistuje žádný univerzálně platný výběr vlastností entity pro klasifikaci



# 7. Doménová analýza

## ○ Teoretické předpoklady

5. Předpoklad, že existují objektivní kritéria pro klasifikaci, lze označit jako pozitivismus (popř. empirismus), předpoklad, že klasifikace jsou účelové, jako pragmatismus
6. Různé obory (domény) mohou používat různé popisy a klasifikaci entity na základě jejich specifických potřeb. Kritéria pro klasifikaci entit jsou proto doménově předurčena.
7. V každém oboru (doméně) existují specifické teorie, přístupy, zájmy a paradigmaty, která ovlivňují způsob, jakým jsou entity popisovány a klasifikovány
8. Jakákoliv klasifikace je vždy výsledkem určitého pohledu nebo přístupu k entitám, které jsou klasifikovány





# 7. Doménová analýza

## ○ Teoretické předpoklady

9. Dokumenty, které mají být klasifikovány, jsou vždy více či méně průnikem různých domén a různých přístupů, teorií nebo pohledů. Tyto přístupy mohou být v dokumentu vyjádřeny explicitně nebo patrné pouze implicitně (v takovém případě je lze určit další analýzou).
10. Klasifikaci, která nezohledňuje různé účely a přístupy, vyjádřené v dokumentech, je možné označit jako pozitivistickou. Kritéria pro klasifikaci by měla vycházet z porozumění specifickým cílům, hodnotám a zájmům dané domény. Nelze je stanovit předem, ale pouze na základě kvalifikovaného průzkumu literatury.



# 7. Doménová analýza

## ○ Doménová analýza může zahrnovat:

1. Tvorbu a hodnocení referenčních příruček a oborových portálů
2. Tvorbu a hodnocení oborových klasifikací a tezaurů
3. Výzkum v oblasti indexace a vyhledávání informací ve specifických oblastech
4. Empirické uživatelské studie
5. Bibliometrické studie
6. Historické studie o organizaci znalostí a informačních službách v dané oblasti
7. Studie o typech dokumentů a žánrů v dané oblasti poznání
8. Epistemologické a kritické studie různých paradigmat, východisek a zájmů v doméně
9. Oborovou terminologii, jazyk a diskurz
10. Strukturu a instituce odborné komunikace
11. Oborové poznání a reprezentaci znalostí



# Doménová analýza

- Výběr domény
- Informační zdroje
- Pojmy a jejich charakteristika
- Pojmy a jejich struktura
- Notáčn<sup>í</sup> systém
- Formální struktura dokumentu
- Obsahová analýza a věcné zpracování
- Ontologie
- Sémantika
- Vizualizace



# DA – dílčí krok č. 1 a 2

zadání kroku 1

zadání kroku 2