

# Organizace znalostí I

PhDr. Jiří Stodola, PhD.

9. 10. 2020

# Požadavky

- **Prezenční studenti:**
  - 75 % účast na přednáškách
  - Úkol: jednoduchý model konkrétní domény
  - Písemný text: 60 % úspěšnost umožní přístup k ústní zkoušce
  - Ústní zkouška
- **Kombinovaní studenti:**
  - Místo účasti esej v rozsahu 5 normostran

# Osnova

- 01. Co je organizace znalostí?
- 02. Data - informace - znalost - moudrost
- 03. Doména a její analýza
- 04. Pojmy a sémantické vztahy
- 05. Znak (instrumentální)
- 06. Ontologická struktura reality
- 07. Kategorizace, klasifikace, fasetace
- 08. Indexace
- 09. Bibliografické univerzum jako ukázka ontologie
- 10. Dokument a jeho popis
- 11. Systematické selekční jazyky
- 12. Předmětové selekční jazyky

# Úvod – organizace poznání

- Užší pojetí
  - Popis dokumentů, indexace, klasifikace v knihovnách, bibliografických databázích a všech typech „paměťových institucí“
  - Dělají je knihovníci, archiváři, informační specialisté, oboroví specialisté, ale i počítačové algoritmy
  - Hlavní disciplínou v tomto pojetí je informační věda
- Širší pojetí
  - Provádí se všude tam, kde vznikají a šíří se znalosti
  - Univerzity, výzkumné instituce, věda a její disciplíny, obory a profese
  - Jde o sociální rozdělení mentální práce
  - Dělí se na sociální organizaci znalostí a kognitivní či intelektuální organizaci znalostí
  - Realitu zkoumají a třídí vědy jako chemie (periodická soustava prvků) či biologie (taxonomie)
  - V tomto pojetí k hlavním disciplínám patří sociologie vědění, jednotlivé speciální vědy a z univerzálních věd metafyzika a ontologie
- Nelze se zabývat pouze užším pojetím, aniž bychom vzali v potaz kontext širšího

# Teoretické přístupy k organizaci poznání

- Zpočátku praktická činnost bez teoretického podkladu, založená na „selském rozumu“, na analýze databází, případně znalostech oborových specialistů
- Rozvíjela se zejména syntaktická stránka organizace znalostí, v pozadí zůstávala sémantická
- Počítačovní vědci se snažili vytvářet systém OZ bez účasti lidí
- Knihovnictví a počítačová věda si v tomto konkurovali

# Teoretické přístupy

1. tradiční přístup: klasifikační systémy používané v knihovnách jako DDC, LCC, UDC
2. Fasetově-analytické přístup – Ranganathan
3. Přístupy založené na vyhledávání informací
4. Uživatelsky a kognitivně orientované přístupy
5. Bibliometrické přístupy
6. Doménově-analytický přístup
7. Jiné přístupy – sémiotický, hermeneutická, diskurzivně-analytický, žánrový

# Tradiční přístup

- **Melvil Dewey** – [DDC](#) –
  - chybí teoretický přístup,
  - výhodnější pro knihovníky více než pro uživatele,
  - založený spíše na sbírkách knihoven, než na klasifikaci vědních oborů a oblastí,
  - byrokratický, administrativní,
  - není založen na dedukci ani na společenském konsenzu
- **Henri Bliss**
- Přírodní řád > Vědecká klasifikace > Knihovní klasifikace

# Základní zásady tradičního přístupu

- Princip kontrolovaného slovníku (odstraňuje homonymii a synonymii, standardizace)
- Cutterovo pravidlo největší specifikace (využít nejspecifičtějšího termínu, který je dostupný)
- Hulmův princip literárního zdůvodnění (literary warrant) – existuje to, o čem je literatura, a to se má objevit v klasifikaci
- Princip organizace od neobecnějšího k nejspecifičtějšímu

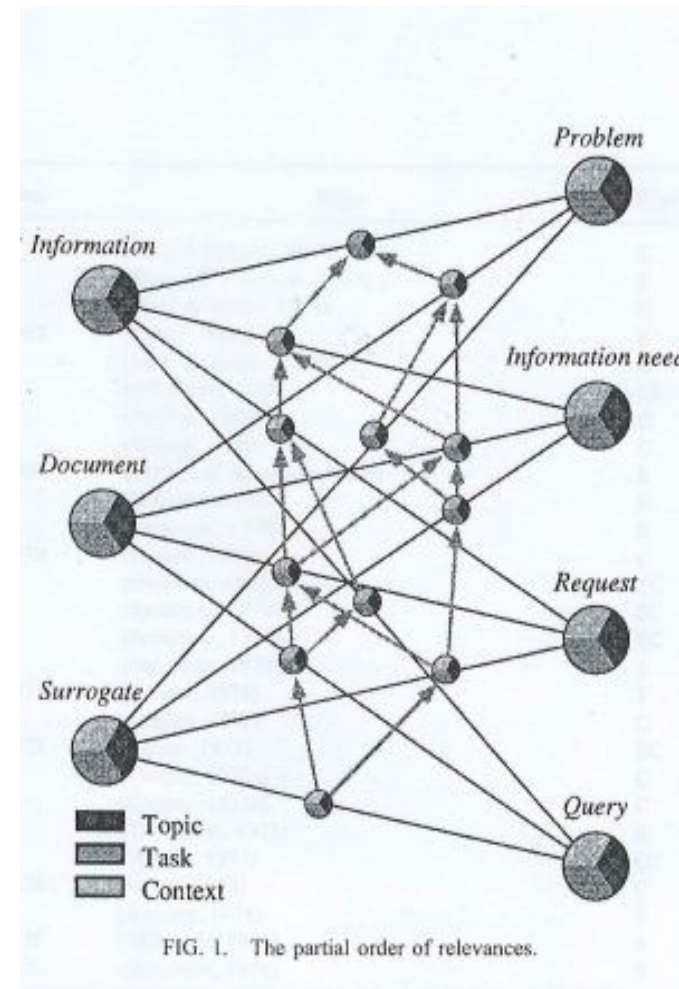


# Fasetově-analytický přístup

- **S. R. Ranganathan** – Colon Classification (1933)
- Analyticko-syntetická metoda: 1. rozdělení každého předmětu na základní pojmy, 2. složení předmětu, který máme v ruce, pomocí těchto základních pojmů.
- Nejzákladnější kategorie, na které lze každý předmět rozložit, jsou fasety.
- Fasety – PMEST
  - Osoba, osobnost – rozlišující vlastnost subjektu
  - Hmota – fyzický materiál, ze kterého se subjekt skládá
  - Energie – jakákoliv akce, kterou subjekt činí či trpí
  - Prostor – umístění subjektu
  - Čas – časové období spojené se subjektem
- Enumerativní systémy jsou podle Ranganathana povrchní a nedokážou předvídat nové poznatky, na rozdíl od fasetové dvojtečnové klasifikace

# Přístup spojený s vyhledáváním informací

- Vyhledávání informací a OZ jsou dvě odlišné oblasti LIS, prolínají se
- Dva přístupy ve vyhledávání informací
  - Systémově orientovaný (relevance je především vztah mezi vyhledanými záznamy a dotazem)
  - Kognitivně či uživatelsky orientovaný (zaměřuje se více na požadavek a potřebu), místo relevance se hovoří o pertinenci



# Relevance

- **Přesnost (precision)** = počet relevantních vyhledaných dokumentů k celkovému počtu vyhledaných dokumentů

	vybráno	nevybráno	
relevantní	<i>A</i>	<i>B</i>	<i>A + B</i>
irelevantní	<i>C</i>	<i>D</i>	<i>C + D</i>
	<i>A + C</i>	<i>B + D</i>	

- Koeficient přesnosti:

$$p = A / (A + C)$$

- **Úplnost (recall)** = počet vyhledaných relevantních dokumentů ku všem relevantním tedy i nevyhledaným

- Koeficient úplnosti:

$$R = A + B$$

# Uživatelsky orientovaný přístup

- Uživatelsky přívětivá OZ
- Tržně orientovaná OZ
- OZ založená na empirických studiích uživatelů (uživatelé dávají přednost systémům, založených na přirozeném jazyce)
- OZ dělaná uživateli (folkosonomie) - studování lidových a vědeckých klasifikacích
- Dává přednost studiu uživatelů před dokumenty, je protikladná k systémovému přístupu (je relevance vlastnost dokumentu nebo myslí uživatele?)

# Přístupy založené na bibliometrii

- Citační analýza
- Dokumenty na základě odkazů tvoří určitou síť, kterou knihovní katalogy nedokážou zatím zachytit
- Dá se využít k budování tezaurů, kde bibliometrické odkazy tvoří síť autorů a tezaury ontologii
- Blíže k sociální organizaci poznání než k intelektuální organizaci

# Doménově-analytický přístup

- Socio-epistemologický přístup
- Metodologický kolektivismus
- Sleduje vznik a práci s informacemi v různých oblastech, každá oblast má vlastní pojmy, termíny, diskurz
- **Doména se tedy pro nás skládá z předmětu zkoumání, paradigmatu dané disciplíny, sociální sítě odborníků a z odborného diskursu.**
- Doména je složitý agregát, který se skládá z
  1. objektů zkoumání
  2. subjektů poznání a tvůrců informačních objektů
  3. informačních objektů
  4. informačních a komunikačních vztahů mezi objekty zkoumání, subjekty a informačními objekty
  5. vztahů souladu či nesouladu mezi objekty, subjekty a informačními objekty, které vznikají na základě informačních a komunikačních vztahů.

# Co se organizuje

- Dokumenty (hmotný nosič + informace; to, co může vést ke znalosti, objektivní informace)
- Informace (sémantický obsah, přenositelný údaj o realitě pro někoho, subjektivní informace)
- Znalosti (vztah mezi subjektem a realitou, pravdivé zdůvodněné přesvědčení)

# Oblasti spojené s OZ

- Filosofie – metafyzika (ontologie), epistemologie, logika
- Lingvistika, sémiotika
- Počítačová věda