



VIKMA06

# Vyhledávání informací

13. 10. 2017: Přednáška P4+K2:

Objekty procesu vyhledávání informací (informační zdroje)

Metodika vyhledávání

FF MU, podzim 2017

Mgr. Josef Schwarz


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

# Informační objekty

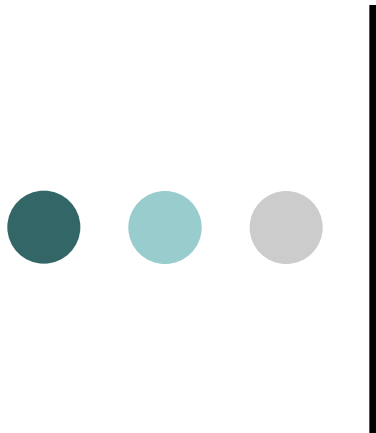


- informace
- dokumenty
- informační zdroje

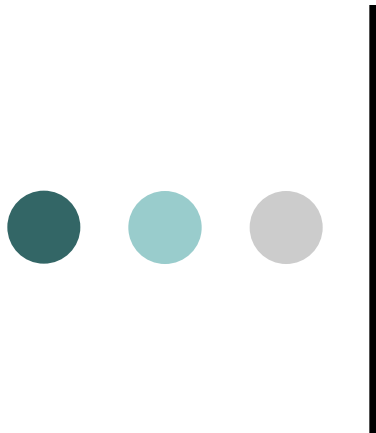
# Data, informace, znalost

- 
- data
    - záznam
  - informace
    - sdělení
      - →určení významu, smyslu údajů
  - znalost
    - kontexty, interpretace
      - →zasazením do širších souvislostí vzniká znalost a jsou generovány nové poznatky

# Dokument

- 
- informační objekt
  - typy (podle typu informace)
    - primární dokument
      - výzkumná zpráva, norma, disertace, článek, monografie, encyklopedie, výroční zpráva
    - sekundární dokument
      - bibliografie, rešerše
      - →*odkazuje na primární dokument*
    - terciární dokument
      - bibliografie bibliografií, databáze databází
      - →*odkazuje na sekundární dokument*

# Dokument (pokr.)

- 
- typy (podle informačního cyklu)
    - primární dokument
      - příklady: výzkumná zpráva, disertace, článek, výroční zpráva
      - *→ dokumenty, ve kterých jsou zveřejňovány nové informace vzešlé z vědeckého výzkumu, archivního bádání, činnosti organizací aj.*
    - sekundární dokument
      - příklady: kompendia, patentová literatura, referátová literatura
      - *→ přejímá informace z primárního dokumentu*
    - terciární dokument
      - příklady: monografie, učebnice, příručky, slovníky, encyklopedie
      - *→ přejímá informace ze sekundárního, resp. primárního dokumentu*

# Dokument (pokr.)



- typy (podle typu inf.)
  - textové
  - netextové
    - obraz
    - zvuk
    - video
  - smíšené

# Informační zdroje



## ○ charakteristika

### ● typ

- bibliografické IZ
- plnotextové IZ
- faktografické IZ
- adresářové IZ

### ● obsah

- oborové (polytematické)
- podle typu dokumentů

# Informační zdroje (pokr.)



- charakteristika

- producent, poskytovatel

- cena

- dostupnost

- veřejné IZ

- nekomerční IZ

- komerční IZ


- neveřejné IZ



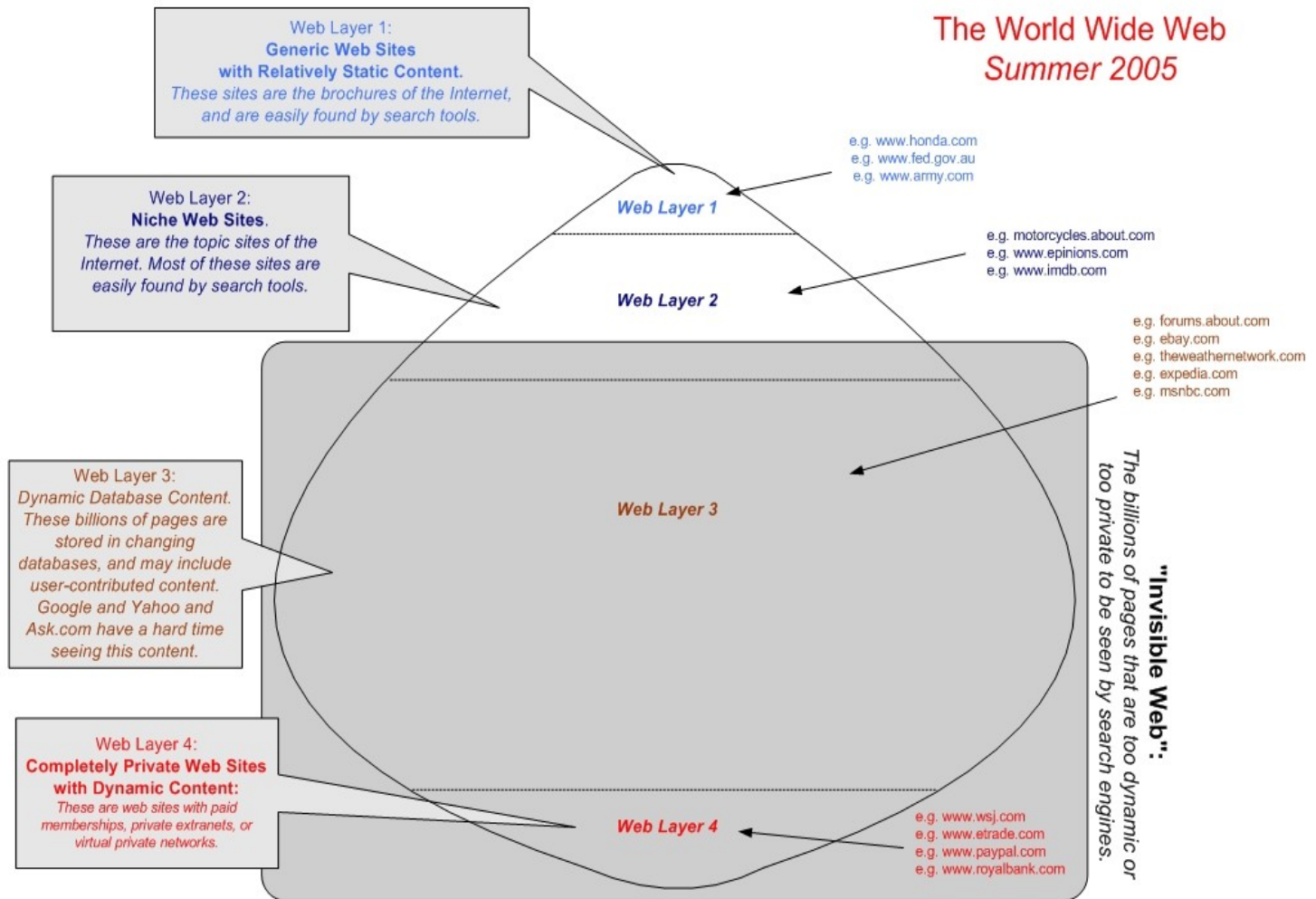
# Informační zdroje (pokr.)

- 
- kvalita
    - rozsah
    - úplnost
    - retrospektiva
    - aktuálnost
      - stárnutí informace
    - důvěryhodnost

# Informační okruhy

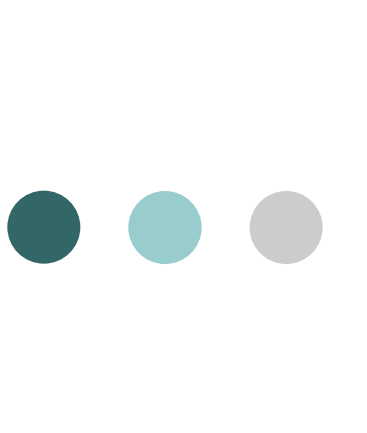
- 
- masmédia
  - veřejná správa
  - ekonomika a finance
  - věda a technika
  - kultura a umění
  - archivní fondy
  
  - a co internet?
    - hluboký web

# The World Wide Web Summer 2005

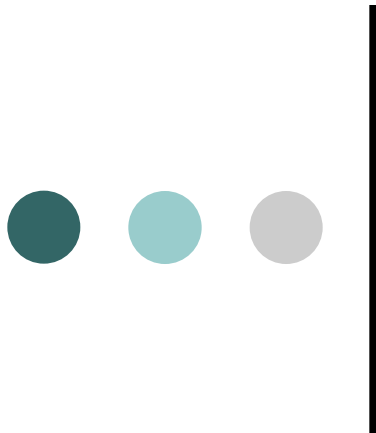


(Zdroj: [About: internet for begginers](#))


# Informační metazdroje

- 
- referenční příručky ([příklad](#))
  - databáze databází
  - bibliografie bibliografií (příklad: International Bibliography of Bibliographies in Library and Information Science and Related Fields)
  - portály ([JIB](#))

# Informační instituce

- 
- paměťové instituce
    - knihovny
      - národní knihovny
      - veřejné knihovny
      - odborné knihovny
      - akademické a univerzitní knihovny
    - galerie
    - muzea
    - archivy
  - informační střediska
  - zpravodajské agentury
  - databázová centra
    - Příklady: [STN International](#)

# Identifikace IZ

- 
- zjištění existence IZ
    - pomocí sekundárních a terciárních zdrojů
  - zjištění charakteristik IZ
    - vč. možností vyhledávání
  - volba IZ
    - obsahové aspekty
  - zjištění (a zajištění) dostupnosti IZ



# Metodika vyhledávání

- typy vyhledávání
- nástroje vyhledávání
- formulace řešeršního dotazu

# Typy vyhledávání

## ○ Podle hledané informace:

### ● identifikační vyhledávání

- (známe některé údaje o hledaném dokumentu nebo položce)
- vyhledávací výrazy: formální údaje - osobní jméno, název, nakladatel, rok, místo vydání, název časopisu, ISBN, ISSN, datum (konání konference, vydání, narození aj.) apod.
- příklad: [NTK](#), telefonní seznam, [Obchodní rejstřík](#)

### ● věcné vyhledávání

- (neznáme požadovaný dokument, hledáme určité téma)
- vyhledávací výrazy: věcné údaje - klíčová slova z názvu, předmětová hesla, klíčová slova, deskriptory tezauru, klíčová slova z textu dokumentu (redukovaného nebo plného), klasifikace ([MDT](#), [CZ-NACE](#), [NAICS](#)) apod.

### ● faktografické

- (chceme zjistit konkrétní informaci)
- vyhledávací výrazy: údaje podle obsahu a struktury zdroje





# Typy vyhledávání

- hledání (seeking)
- vyhledávání (searching)
- prohlížení (browsing)
- filtrace (filtering)
- data mining





# Typy vyhledávání

- nestrukturované (freetextové)
  - celý záznam dokumentu
- strukturované
  - metadata
    - selekční obraz dokumentu
    - redukovaný text
  - vazby dokumentů
    - citační vazby
    - formální vazby (FRBR)
- plnotextové



# Typy vyhledávání

- nestrukturované vyhledávání
  - základní, jednoduché vyhledávání
    - [KNAV](#), [PubMed](#), [Google](#)
- strukturované vyhledávání
  - pokročilé, podrobné vyhledávání
    - [KNAV](#), [PubMed](#), [Google](#)
  - řízený slovník (tezaurus, seznam předmětových hesel nebo klíčových slov apod.)
    - není dostupný: [KNAV](#)
    - je dostupný samostatně: [NTK](#)
    - je dostupný při vyhledávání: [NKP](#)
- plnotextové (fulltextové) vyhledávání
  - invertovaný rejstřík
  - sekvenční vyhledávání



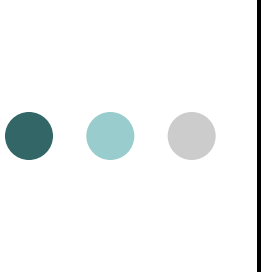
# Nástroje vyhledávání

- vyhledávací (dotazovací) jazyky
  - složky
  - standardizace (CCL)
  - tendence ke (kvazi)přirozenému jazyku
- selekční jazyky
  - věcné
  - identifikační (authority)
  - sémantické sítě
- uživatelské rozhraní
  - příkazový řádek
  - GUI
- algoritmy vyhledávání



# Formulace rešeršního dotazu

1. pojmová analýza
2. synonyma a související pojmy
3. převedení na výrazy řízeného slovníku
4. aplikace (booleovských) operátorů
5. aplikace dalších vyhledávacích technik



# Pojmová analýza

- identifikace klíčových pojmů
- reprezentace pojmů (substantiva a adjektiva, slovesa nahrazena operátory)



# Synonyma a související pojmy

- vytvoření seznamu synonym a dalších příbuzných výrazů
- využití seznamu:
  - výběru vhodného vyhledávacího výrazu
  - převod na výraz věcného SJ
  - rozšiřování a zužování tématu



# Převedení na výrazy řízeného slovníku

## Varianty

1. výraz v seznamu je shodný s výrazem ŘS
2. výraz v seznamu je synonymem/ekvivalentem výrazu ŘS
3. pro výraz v seznamu existuje pouze širší výraz ŘS
4. pro výraz v seznamu existují pouze specifictější/podřazené výrazy ŘS





# Aplikace (booleovských) operátorů

Určení vztahů mezi pojmy

- **operátor AND** – spojení významově odlišných výrazů
- **operátor OR** – spojení synonym a příbuzných výrazů
- **operátor NOT** – vyloučení nežádoucích výrazů



# Aplikace dalších vyhledávacích technik

- škála možností závisející na konkrétním informačním zdroji
  - krácení, zástupné znaky
  - proximitní operátory
  - vyhledávání podle polí
  - rozšiřování a úprava dotazu (query expansion – relevance feedback)
  - vyhledávání ve více databázích (multiple database searching)

# Query Expansion

Manual  
Query Expansion

Automatic  
Query Expansion

Interactive  
Query Expansion

Based on  
Search Results

Based on  
Knowledge Structures

Collection  
Dependent

Collection  
Independent



# Vyhledávací techniky

## obvyklé možnosti

- booleovské operátory
- fráze
- vyhledávání podle polí
- formální omezení
- krácení, zást. znaky, stemming
- ukládání rešerše a historie
- proximitní vyhledávání
- užití řízených slovníků

## specifické možnosti

- prohlížení časopisů a obsahů jednotlivých titulů
- rozšiřování dotazu
- navrhování výrazů ŘS
- dotaz příkladem
- automatický překlad
- odkazy na plný text prostřednictvím jiné služby, odkazy na web, napojení na katalog
- vyhledávání pomocí notací SSJ