



VIKMA06

Vyhledávání informací

2. 11. 2018: Přednáška P6+K3: Rešeršní strategie

FF MU, podzim 2018

Mgr. Josef Schwarz

126172@mail.muni.cz

Rešeršní strategie

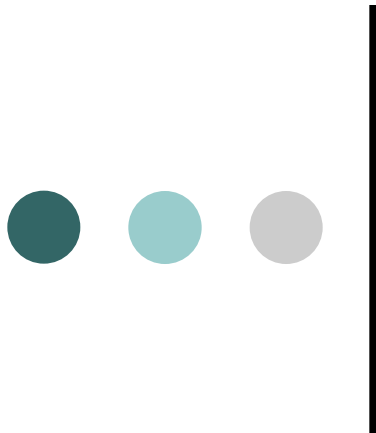


- širší pojetí

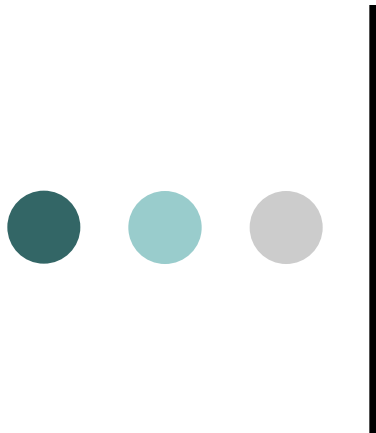
- užší pojetí

- výběr konkrétního vyhledávacího nástroje a komunikace se systémem

Cíle řešeršní strategie

- 
- relevance X pertinence
 - úplnost X přesnost

Úplnost a přesnost



	relevantní dok.	nerelevantní dok.
vyhledané dok.	a	b
nevyhledané dok	c	d

- úplnost (R)
 - $R = a / (a+c)$
- přesnost (P)
 - $P = a / (a+b)$
- vztah mezi úplností a přesností je nepřímo úměrný (vyšší přesnost znamená nižší úplnost a naopak)

Předpoklady řešeršní strategie

○ Předpokladem pro stanovení řešeršní strategie je znalost:

- informačního zdroje (databáze)
 - Obsah – jaké dokumenty, v jaké retrospektivě a úplnosti apod.
 - Struktura – podle jakých polí lze vyhledávat
- nástrojů
 - Řízené slovníky, hesláře, authority aj.
- algoritmů
 - Operátory dotazovacího jazyka, konvence pravostranného rozšíření atd.
- uživatelského rozhraní

Typy řešeršní strategie

-
-
- ○ strategie stavebních kamenů
- vyhledávání pomocí nejspecifičtější fazety
- strategie rostoucí perly
- strategie osekávání

Strategie stavebních kamenů

- samostatné dílčí dotazy vyjadřující ústřední pojmy původního rešeršního požadavku
- identifikace klíčových/významných pojmů
- množina výrazů vztahující se k pojmu: synonyma, kvazisynonyma, pravopisné formy, nadřazené, podřízené výrazy
 - OR, truncation (krácení podle slov. kořenů), stemming, wild cards (zástupné znaky)
- spojení dílčích formulací ve finální soubor
 - AND
- vhodné použít, když usilujeme o úplnost u úzce specifikovaných témat

Strategie stavebních kamenů

Příklad

- Informační dotaz...:
 - *Informační chování adolescentů v prostředí sociálních sítí při zjišťování informací o kosmetických produktech k redukci výskytu akné.*
- ... bude rozdělen na jednotlivé pojmové segmenty,....:
 - *informační chování*
 - *adolescenti*
 - *sociální sítě*
 - *zjišťování informací*
 - *kosmetické produkty*
 - *redukce výskytu akné*
- ... které budou řešeny jako samostatné dotazy a následně spojeny operátorem AND.

Vyhledávání pomocí nejspecifičtější fazety

- vztahuje se k vyhledávání složených témat – více aspektů
- uživatel musí znát všechny dílčí témata a musí být schopen určit, které téma je nejspecifičtější
- **Vyhledávání**
 - podle nejužšího pojmu z rešeršního požadavku
 - pokud je výsledek uspokojivý, nemusí být do rešerše zahrnuta další dílčí hlediska

Vyhledávání pomocí nejspecifičtější fazety

Příklad

Informační dotaz....:

- ☞ *Stanovení odolnosti legované oceli pomocí Bradfordovy stupnice tvrdosti za účelem zkujnění a sváření kovových profilů ve tvaru U*
- ☞ ... bude redukován na jediný pojmový segment, a to....:
 - ☞ *Bradfordova stupnice tvrdosti,*
- ☞ ... protože odolnost oceli se měří právě a jen touto stupnicí a nezáleží na tom, k jakému účelu se daná ocel používá.

Strategie rostoucí perly

Dotaz je postupně modifikován dle výsledků rešerše

- záznamy jsou postupně procházeny a zjišťovány relevantní termíny (řízené termíny, slova z názvů apod.), které jsou použity k revidování dotazu.

Prvotním cílem je alespoň jeden záznam

- zjištění použitelných selekčních termínů
- úprava formulace rešeršního dotazu

Strategie rostoucí perly

Příklad

Informační dotaz....:

Závislost míry nezaměstnanosti v regionech ČR na měsíčních srážkových úhrnech

... bude na začátku vyjádřen tímto rešeršním dotazem....:

nezaměstnanost AND počasí AND „Česká republika“,

... který bude generovat první relevantní záznamy, na jejichž základě pak bude dotaz obohacován o další prvky.

Strategie osekávání

- první formulace dotazu - **širší formulace, tj. pomocí obecného pojmu** – cílem je vyhledání více záznamů
- **postupná specifikace dotazu**
- uplatnění taktik pro zúžení záběru (AND, NOT, proxim. oper., field searching, formální omezení)
- formulace širší kategorie (obor, vědní disciplína), klasifikace
- náročnější na čas

Strategie osekávání

Příklad

Informační dotaz....:

- *Revize zaměřovacích bodů v nivě Moravy s ohledem na nové tachymetrické metody geodetického měření.*

- ... může být zpočátku poněkud obsahově nesrozumitelný, proto pro vyhledávání volíme nejprve obecnější pojmy, jako....:

- *tachymetrie, popř. geodézie*

- ... pomocí kterých najdeme širší množinu záznamů, kterou následně pomocí dalších klíčových slov omezujeme

Strategie pro zúžení záběru

- klíčová slova se kombinují s věcným selekčním jazykem
- omezení na určité pole záznamu
- využití proximitních operátorů
- omezení na určitý typ dokumentu
- operátor NOT pro vyloučení některých záznamů
- jazykové vymezení
- časové rozmezí
- kombinace množiny deskriptorů/hesel s podřazenými klíčovými slovy
- kombinace s množinou sel. údajů vyjadřující další pojem z dotazu, hledisko


Strategie pro rozšíření záběru

- uvedení synonym, tvarů slov, pravopisných variant (operátor OR, zástupné znaky, krácení podle slovních kořenů)
- uvedení jednotek věcného SJ jako klíčových slov (např. vyhledávání ve všech polích)
- dodatečné uvedení širších jednotek věcného SJ, tj. těch, které jsou nadřazeny použitým výrazům (deskriptorům, předmětovým heslům)
- obecné výrazy, tj. s vysokým výskytem
- zrušení předběžných omezení

Vyhledávací techniky pro zvýšení přesnosti

- použití operátoru AND
- použití operátoru NOT
- „case sensitive“
- proximitní operátory
- vážené vyhledávání („weighted searching“)
- omezení na pole („field searching“)

Vyhledávací techniky pro zvýšení úplnosti

- 
- použití operátoru OR
 - krácení, zástupné znaky
 - fuzzy vyhledávání
 - rozšiřování dotazu („query expansion“)
 - paralelní vyhledávání – „multiple database searching“

Rešeršní strategie - praktické rady

Bud'te flexibilní

- ber'te připravené kroky strategie orientačně
- přizpůsobujte další taktiky výsledkům rešerše
- nulový výsledek – hledání příčiny

Využívejte řízených slovníků

- využívejte souvisejících pojmů ke konkrétnímu řízenému termínu (nadřazené, podřazené pojmy)
- nikdy nespojujte výrazy s malou frekvencí výskytu (zjistitelné v katalogu) operátorem AND

Vytvářejte množiny termínů

- je velmi důležité k jednotlivým klíčovým slovům vytvářet množiny souvisejících termínů
- termíny v množině se spojují pomocí logického součtu – OR

Využívejte klasifikací

- pomocí klasifikací vyhledáte většinou mnoho záznamů, proto se hodí jejich využití při strategii osekávání

Literatura

● ● ● ○ kapitoly ze základní a doplňkové literatury

- CHU07, kap. 6, 9 (s. 81-96, 145-166)
- ING92, kap. 6 (s. 123-156)
- VIC04, kap. 7 (s. 180-209)