

Rešeršní činnost

Rešeršní strategie a taktiky

14. 3. 2008

komb. 21. 3. 2008

přednášející: Silvie Kořínková Presová

presova@phil.muni.cz

Kabinet informačních studií a knihovnictví, FF MU

Searching vers. browsing

CHU, 2007:81-90

SEARCHING – záměrná, cílená, systematická aktivita,
specifický dotaz - cílem je najít to, co se shoduje s řešerším
dotazem

BROWSING – hledání a výběr informací prohlížením,
postupným procházením, oproti searching není třeba
specifikovat problém

Vyhledávání v online rešeršních systémech (častý ekv. online databáze)

- možnosti se liší dle rešeršního systému a jeho mechanismu – chybí standardizace způsobu aplikace jednotlivých vyhled. technik

stejná databáze → různí poskytovatelé
(např. ProQuest, EBSCO, Dialog) → nabídka
vyhledávacích technik se může lišit

- základní a pokročilé vyhledávací techniky
- různé pokročilosti uživatelského rozhraní
- odlišnosti dle typu databáze (dokumentograf., faktografická, plnotextová db) aj.

*studie Retrieval features for online databases :
common, unique, and expected (Othman, 2004)*

- ☞ analýza řešer. možností u 25 db od 12-ti různých poskytovatelů
- ☞ snaha postihnout obvyklé (nabízeny více než 5-ti db) a neobvyklé/specifické řešer. možnosti (nabízeny méně než 5-ti db)

studie *Retrieval features for online databases : common, unique, and expected* (Othman, 2004)

obvyklé rešer. možnosti

- ☞ booleovs. operátory
- ☞ fráze
- ☞ vyhledávání dle polí
- ☞ formální omezení/limity
- ☞ krácení, zást. znaky, stemming
- ☞ ukládání rešerše a historie
- ☞ proximit. vyhledávání
- ☞ užití řízených slovníků
- ☞ věcné vyhledávání aj.

specifické rešer. možnosti

- ☞ prohlížení časopisů a obsahů jednotl. titulů
- ☞ expanze dotazu („query expansion“)
- ☞ navrhování řízených termínů
- ☞ dotaz příkladem
- ☞ automatický překlad
- ☞ odkazy na plný text prostřednictvím jiné služby, odkazy na www, napojení na katalog
- ☞ vyhledávání pomocí notací SSJ aj.

Rešeršní strategie – širší pojetí

- ☞ **systematický plán pro uskutečnění rešerše**
(jak si rešeršér poradí s nalezením neoptimálnější cesty k zodpovězení rešeršního požadavku)
- ☞ **plánování celého rešeršního procesu**
(analýza reš. požadavku – identifikace klíčových pojmů; výběr db, rešeršní dotaz, samotné vyhledávání)
 - na základě analýzy předmětu formulujeme optimální metody a následnost procesů, s jejichž pomocí dosáhneme zodpovězení konkrétního rešeršního požadavku realizovaného v IS
- ☞ **analýza a formulace optimálních metod**, kterými se zajistí uspokojení rešeršního požadavku prostřednictvím vyhledávacího IS (volně podle Makulová, 2005, s. 82)

Rešeršní strategie – užší pojetí

- ☞ tzv. **realizační přístup**, který spočívá v konkrétním výběru vyhledávacího nástroje a komunikaci se systémem (volně podle Makulová, 2005, s. 83)
- ☞ podceňuje se příprava vyhledávání a analýza rešeršního požadavku

Rešeršní strategie versus taktiky

- ☞ rešeršní strategie jako přístup k vyhledávání položek v IZ k jednomu požadavku (MATOUŠOVÁ, 1988, s. 69)
- X použité logické kombinace
- X **rešeršní taktiky** – souvisí s modifikací původní formulace dotazu
 - jsou úkonem, kterým se mění stávající postup vyhledávání (nová orientace dalšího postupu) nebo formulace dotazu
 - různé možnosti použitelnosti taktik – na počátku vyhledávání, v kterékoliv fázi, pouze v pokročilém stadiu práce na rešerši

Taktiky 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

Taktiky 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 termínům (deskriptorům, předmětovým heslům) – ty naleznete v příslušných řízených slovnících
- ☞ obecné termíny, tj. s vysokým výskytem
- ☞ zrušení předběžných omezení

Vyhledávací techniky pro zvýšení přesnosti

- ☞ použití operátoru AND
- ☞ použití operátoru NOT
- ☞ „case sensitive“
- ☞ proximitní operátory - vyhledání slov v přesném pořadí, např. v ProQuest oper. PRE
- ☞ „weighted searching“
- ☞ „field searching“ – omezení na pole

Vyhledávací techniky pro zvýšení úplnosti

- ☞ použití operátoru OR
- ☞ krácení, zástupné znaky
- ☞ fuzzy vyhledávání
- ☞ „query expansion“
- ☞ paralelní vyhledávání – „multiple database searching“

Rešeršní strategie - typy

- typy (Charles Bourne), Matoušová, 1988, Chu, 2007:84-87
 - možnosti, jak postupovat při vyhledávání
 - výzkum práce referenčních knihovníků a rešeršérů
- *strategie stavebních kamenů, vyhledávání pomocí nejspecifičtější fasety, strategie rostoucí perly, 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
- ☞ lze předpokládat, že budeme dotaz zužovat

Jaké vyhledávací služby na internetu aplikují SSJ?

- **vyhledávací služby:** vyhledávací nástroje, vyhledávače, vyhledávací stroje, portály, předmětové katalogy, OIB, virtuální knihovny, předmětové adresáře, tématické brány, předmětové brány, katalogové vyhledávací nástroje
- **SSJ:** klasifikační schémata, klasifikace, kategorizační schémata, klasifikační systém, fasetový systém, enumerické schéma
- **internet:** net, síť sítí, WWW

POJEM 1

vyhledávací služby OR vyhledávací nástroje

POJEM 2

klasifikační schémata OR klasifikace OR
kategorizační schémata OR klasifikační
systém

⇒ dotaz (vyhledávací služby OR vyhledávací
nástroje) AND (klasifikační schémata OR
klasifikace OR kategorizační schémata OR
klasifikační systém)

Příklad

Vytvořte dotaz v db LLIS

rešeršní požadavek: *vzdělávání dospělých v knihovnách se zřetelem na zlepšení jejich informační gramotnosti*

☞ *adult education OR lifelong learning*

☞ *information literacy OR information skills*

☞ *library*

☞ **vyzkoušejte si různé taktiky pro zúžení**, např. omezit na deskriptory, proximitní vyhledávání, chronolog. zpřesnění, výběr pole vyhledávání (v titulu, abstraktu apod.)

Vyhledávání pomocí nejspecifičtější fasety

- ☞ 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 řešeršního požadavku

- ☞ pokud je výsledek uspokojivý, nemusí být do řešerše zahrnuta další dílčí hlediska

Příklad

Vytvořte dotaz v db LLIS

rešeršní požadavek: „*předmětová kategorizace IZ v oborových informačních branách*“

- ☞ categorization, classification, scheme
- ☞ quality-controlled subject gateways, subject gateways, portal, gateways
- ☞ zahájení vyhledávání na základě nejužšího pojmu „quality-controlled subject gateways“

rešeršní dotaz ("quality-controlled subject gateways") <in> ALL

příklady taktik

- ☞ vyhledání pod autorem Koch, Traugott
- ☞ po projítí záznamů z 2. kroku, identifikace vhodného předmětu (su)
(web portals) <in> SU AND classification <in> ALL
- ☞ rozšíření na vyhledávání v keyword, tj. vyhledávání v Subject, Title, Abstract, Author, Journal name (web portals OR gateway) <in> KW
AND (categorization OR classification OR scheme) <in> KW
- ☞ **vyzkoušejte v db** Library and Information Science Abstracts (dále LISA)

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.

Cílem je alespoň jeden záznam

- 👉 zjištění použitelných selekčních termínů
- 👉 úprava formulace rešeršního dotaz

Strategie osekávání

- ➡ první formulace dotazu - **širší formulace, tj. pomocí obecného pojmu** – cílem je vyhledání více záznamů (hitů)
- ➡ **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

Vytvořte dotaz v db LLIS

rešeršní požadavek: *Fenomén Web 2.0 a jeho vliv na knihovny*

rešeršní dotaz:

Web 2.0 and Library

rešerše v LLIS, zajímají nás dokumenty řešící problematiku služeb Webu 2.0, které umožňují sdílet uživateli vytvářené popisky webových dokumentů

☞ identifikace klíčových pojmů:

☞ rešeršní dotaz:

☞ zužte dotaz pomocí nabízených deskriptorů

omezte na plnotextové dokumenty

vyzkoušejte v jiných db – LISA, ProQuest

Úkol na příští hodinu

Prostudovat nápovědu

- 👉 Library and Information Science Abstracts (LISA)
- 👉 ProQuest 5000

Co znát:

- 👉 jak se vyhodnocuje basic search (default operátor), relevance, co jednotlivá pole pro vyhledávání znamenají (např. All-Smart Search – v jakých polích se vyhledává), proximitní operátory, truncation, wildcards, základní booleovské operátory – AND, OR, NOT, Advanced Search, query expansion

Rešeršní strategie - praktické rady

(Makulová, 1993)

Bud'te flexibilní

- ☞ berte 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 termíny 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

- ☞ v českých knihovnách zejména klasifikování podle znaků MDT
- ☞ pomocí klasifikací vyhledáte většinou mnoho záznamů, proto se hodí jejich využití při strategii osekávání

Rešeršní strategie - praktické rady

Využívejte krácení - truncation

Využívejte zástupných znaků – wild cards

POZOR na používání NOT

- ☞ radikální snížení záznamů na výstupu
- ☞ vyloučení i těch záznamů, které obsahují žádané informace

Přizpůsobte rešeršní strategii vyhledávacímu systému, v němž vyhledáváte

Vytěžujte výhody databází

- ☞ reformulace dotazu
- ☞ taktiky pro rozšiřování a zužování výsledné množiny

Používejte akronymy

- ☞ chcete-li dosáhnout co nejúplnějšího vyhledávání, zadávejte zkratky, které se v daném oboru běžně používají
- ☞ ověřte, zda jsou zkratky zahrnuty do řízeného slovníku

Rešeršní strategie - nejčastější chyby

Logické chyby

- ☞ chybné kombinace výrazů pomocí booleovských operátorů (AND místo OR a naopak)

Při psaní slov

- ☞ záměna O x 0, písmeno versus číslice

Neuvědomění si odlišností dotazovacích jazyků jednotlivých databází

- ☞ rozdílné dotazovací jazyky a s tím spojená odlišná formulace rešeršního dotazu

Ignorování rozdílů mezi různými vyhledávacími systémy

- ☞ je nutné si uvědomit, že databáze, elektronické katalogy a vyhledávací nástroje internetu jsou odlišné vyhledávací systémy a liší se charakterem vyhledávání pomocí nich a typem vyhledaných dokumentů

Povinná literatúra

- ☞ Chu, H. *Information representation and retrieval in the digital age*. Medford : Information Today, 2007. 248 s.
- ☞ Makulová, S. *Vyhľadávanie informácií v internete : problémy, východiská, postupy*. 1. vyd. Bratislava : EI & T, 2002.
- ☞ Othman, R. *Retrieval features for online databases : common, unique, and expected*. *Online Information Review*, 2004, vol. 28, iss. 3, s. 200 – 210.

Doporučená literatura

- ☞ Makulová, S. *Úvod do informačných technológií*. 2, Databázové technológie a bázy dát : databázový systém CDS/ISIS. Bratislava : STIMUL-Centrum infromatických a vzdelávacích služieb, 1993. 92 s.
- ☞ Matoušová, M. *Rešerše : úvod do teorie a organizace rešeršních služeb*. Praha : ÚVTEI : Ústředí vědeckých, technických a ekonomických informací, 1988. 227 s.