

Národní knihovna České republiky

Library Literature & Information Science Full Text

**Vyhledávání informací z online databáze
z databázového centra H.W. WILSON, Co.
v rešeršním systému WilsonWeb**

Metodická příručka

Verze 1.2

Zpracovala Eva Bratková

Ústav informačních studií a knihovnictví FF UK

Eva.Bratkova@ff.cuni.cz

**Praha
2004-11-08**

OBSAH

1	Databáze Library Literature & Information Science Full Text	3
1.1	<i>Bibliografický časopis a databáze Library Literature</i>	3
1.2	<i>Bibliografický časopis a databáze Library Literature & Information Science.....</i>	4
1.3	<i>Databáze Library Literature & Information Science Full Text</i>	4
2	Vyhledávání informací z databáze LLIS FTXT v režimu online	5
2.1	<i>Základní informace o rešeršním systému WilsonWeb.....</i>	5
2.2	<i>Základní principy vyhledávání a zobrazování informací v systému WilsonWeb</i>	6
2.2.1	Operátory dotazovacího jazyka rešeršního systému WilsonWeb	9
2.2.2	Selekční pole databáze LLIS FTXT	12
2.3	<i>Základní způsoby vyhledávání informací z databáze LLIS FTXT.....</i>	17
2.3.1	Pokročilé vyhledávání (Advanced Search)	17
2.3.2	Základní vyhledávání (Basic Search).....	18
2.4	<i>Vyhledávání informací z databáze LLIS FTXT pomocí rejstříků.....</i>	18
2.5	<i>Vyhledávání informací pomocí tezauru databáze LLIS FTXT.....</i>	19
2.6	<i>Operace s výsledky rešerší v databázi LLIS FTXT</i>	19
2.6.1	Změny zobrazování záznamů	19
2.6.2	Registrace vykonaných rešerší v dané relaci a ukládání formulací rešeršních dotazů	20
2.6.3	Ukládání, tištění a zaslání záznamů elektronickou poštou.....	21
	Použitá literatura a jiné informační zdroje	21

Tato metodická příručka byla připravena v rámci grantového projektu 1N04170 „*Informační zdroje pro knihovnictví a informační vědu*“ (2004-2008), který je finančně zajišťován z prostředků Programu „*Informační infrastruktura výzkumu*“ Ministerstva školství, mládeže a tělovýchovy ČR. Koordinátorem konsorciálního přístupu k vybraným online informačním zdrojům je Národní knihovna ČR.

© Eva Bratková, 2004-

1 Databáze Library Literature & Information Science Full Text

Mezinárodní oborová bibliografická databáze *Library literature & information science* (dále v textu bude užíváno také zkratky LLIS) a její další varianta - částečně plnotextová databáze *Library literature & information science Full text* (dále v textu bude užíváno také zkratky LLIS FTXT) patří, vedle databáze ISTA a LISA, k trojici nejvýznamnějších zdrojů oboru knihovní a informační vědy. Současná podoba databáze LLIS FTXT vznikla v roce 2000. Má však za sebou dlouholetý vývoj jednak v podobě tradičních bibliografických časopisů, jednak v podobě počítačové databáze zpřístupňované v mnoha formách.

1.1 Bibliografický časopis a databáze Library Literature

Bibliografická databáze LLIS má svého předchůdce v podobě klasického tištěného bibliografického časopisu *Library literature* (dále též jako LL), který ve 30. letech 20. století začala vydávat renomovaná americká vydavatelská společnost **H.W. Wilson Company** (<http://www.hwwilson.com>)

Library literature. 193?-1999. Bronx (N.Y.) : H.W. Wilson Co., 193?-1999. 6 x ročně. ISSN 0024-2373

Jak naznačuje název časopisu, jeho hlavní orientace byla po dlouhá desetiletí zaměřena na oblast **knihovnictví a knihovní vědy**. Předmětem registrace byla především americká a britská knihovnická literatura, v menším rozsahu byla zastoupena literatura z dalších zemí. Americká literatura byla registrována podrobně včetně **literatury regionálního významu**. Kromě časopiseckých článků bibliografie zachycovala knižní publikace, výzkumné zprávy a závěrečné práce z amerických knihovnických škol. Jinak kvalitní **bibliografické záznamy**, opatřované předmětovými hesly ze speciálního řízeného hesláře společnosti H.W. Wilson, nebyly v historii vývoje zdroje doplňovány referáty ani anotacemi. Až v současnosti se začínají v řadě záznamů již referáty objevovat.

Počítačové zpracování tohoto zdroje bylo zahájeno v roce 1984. Společnost H.W. Wilson databázi LL také od roku 1985 začala sama zpřístupňovat online z vlastního databázového centra (**WilsonLine**) v rešeršním systému **WilsonSearch** (doplňování databáze bylo čtrnáctidenní). Od roku 1988 začala společnost databázi vydávat také na nosiči CD-ROM:

Library literature [CD-ROM]. 1988-1999. Bronx (N.Y.) : H.W. Wilson Co., 1988-1999. 4 x ročně. ISSN 1076-7126

Vyhledávání z databáze na tomto nosiči bylo realizováno v aplikačním rešeršním systému **WilsonDisc**. Producent učinil také kontrakt se společností **SilverPlatter**, která tuto databázi začala zpřístupňovat v roce 1995 v rešeršním systému SPIRS:

Library literature [CD-ROM]. 1995-1999. Norwood (Mass.) : SilverPlatter, 1995-1999. 4 x ročně. ISSN 1093-7781

Databázi *Library literature* na základě kontraktu zpřístupňovala online také další databázová centra, k nimž patří zejména centum **DIALOG** (File 439) a **OCLC** (File Library Literature).

Zpřístupňování databáze *Library literature* v režimu online přes WWW rozhraní **WilsonWeb** zahájila společnost H.W. Wilson v roce 1997. První rešeršní systém, aplikovaný v rámci rozhraní WilsonWeb, byl adaptací rešeršního systému **WebSPIRS** společnosti SilverPlatter.

1.2 Bibliografický časopis a databáze *Library Literature & Information Science*

Společnost H.W. Wilson změnila na přelomu tisíciletí tradiční název svého bibliografického zdroje z oboru knihovnictví a knihovní vědy na název *Library literature & information science*. Tištěné sešity bibliografického časopisu vycházejí i nadále 6 x ročně, ale se změněným identifikátorem ISSN:

Library literature & information science. 1999- . Bronx (N.Y.) : H.W. Wilson Co. 1999- . 6 x ročně. ISSN 1528-0659

Změna názvu časopisu i databáze byla ovlivněna významnou proměnou praxe i teorie oblasti knihovnictví v posledních 10-15 letech. Do názvu se tak nutně dostal termín „information science“, tedy pojmenování oboru, který dnes zastřešuje také obor knihovní vědy. Časopis a databáze LLIS tak vykročily směrem k dalšímu sblížení obsahů s dalšími již výše jmenovanými zdroji - databázemi LISA a ISTA.

Počítačová podoba databáze LLIS je i nadále zpřístupňována databázovými centry **DIALOG** (File 439) a **OCLC** (File Library Literature).

1.3 Databáze *Library Literature & Information Science Full Text*

V roce 2000 zahájila společnost H.W. Wilson zpřístupňování databáze LLIS v nové verzi s úplnými texty v elektronické formě na CD-ROM:

Library literature & information science Full text [CD-ROM]. 2000- . Bronx (N.Y.) : H.W. Wilson, 2000- . Měsíčně. ISSN 1529-9880

Společnost postupně od roku 1997 na základě licencí s jednotlivými vydavateli získává, ve svém digitálním fondu ukládá a v rámci databáze LLIS FTXT zpřístupňuje autorizovaným uživatelům vybrané úplné texty časopiseckých článků, které jsou registrovány v záznamech. **Současná retrospektiva úplných textů jde k roku 1994.** Vlastní digitální fond společnosti postupně narůstá. Excerptovány jsou také články z volně dostupných elektronických časopisů a dalších seriálových publikací (*Ariadne, LIBRES, Information, Law and Technology, D-Lib Magazine, Annual review of OCLC Research, School Library Media Quarterly* aj.), jejich texty však nejsou systémem indexovány.

Společnost H.W. Wilson postupně rozšiřuje excerptční základnu zdrojových seriálů pro databázi LLIS / LLIS FTXT. V současné době (v roce 2004) je excerptováno cca **390 titulů seriálů** (s vyloučením titulů, které změnilý svůj název, jde o cca **300 titulů seriálů**), a to v celosvětovém záběru (<http://www.hwwilson.com/Databases/liblit.htm>). Přibližně ze **100 titulů časopisů jsou k dispozici úplné texty v elektronické formě**, tj. cca 25-30 % z celkového počtu excerptovaných časopisů. Ročně je do databáze uloženo cca **12 000 nových záznamů**. **Bibliografický popis** je veden v angličtině a je velmi podrobný. Využíván je také soubor jmenných autorit jmen osob a korporací. Mimořádná pozornost je věnována **věcnému zpracování pomocí oborového řízeného hesláře společnosti H.W. Wilson**, který je adaptací univerzálního hesláře Kongresové knihovny LCSH. Záznamy jsou opatřovány v průměru 2-6 předmětovými hesly. Nedostatkem záznamů databáze LLIS je **nepřítomnost referátů či anotací**, situace se však po roce 2000 mění.

Online zpřístupňování databáze LLIS FTXT je realizováno přes WWW rozhraní databázového centra společnosti H.W. Wilson v moderním sofistikovaném rešeršním systému **WilsonWeb** nové generace, který byl vyvinut na základě softwaru americké firmy **Verity, Inc.** Index databáze zahrnuje již reprezentaci úplných textů, ovšem jenom těch, které společnost vlastní v digitálním fondu (z automatické indexace unikají například články z

elektronických časopisů lokalizovaných pouze na WWW). Databáze zahrnuje v současnosti (k roku 2004) více než 250 000 záznamů.

Metodické pokyny, uváděné v dále v textu příručky, se týkají vyhledávání informací z databáze *Library literature & information science Full text* v režimu online z databázového centra společnosti **H.W. Wilson** v rešeršním systému **WilsonWeb**.

2 Vyhledávání informací z databáze LLIS FTXT v režimu online

2.1 Základní informace o rešeršním systému WilsonWeb

Rešeršní systém **WilsonWeb** společnosti H.W. Wilson je určen především k **online vyhledávání informací** z databází v prostředí WWW (přes WWW rozhraní). Jde o jeden z nejlepších rešeršních systémů moderní koncepce se sofistikovanými vlastnostmi. Zahrnuta je již funkce **vyhledávání z úplných textů** (ale jen z těch, které společnost vlastní). Rešeršní systém je vyvinut na základě softwaru významné americké firmy **Verity, Inc.** V rámci jeho současné verze lze kromě **vyhledávání informací bibliografického a plnotextového charakteru** také:

- Tisknout, ukládat nebo odesílat nalezené záznamy nebo jejich vybrané množiny elektronickou poštou
- Ukládat formulace rešeršních dotazů pro potřeby jejich dalšího využití
- Propojit se z nalezeného a zobrazeného záznamu dokumentu přímo na jeho úplný text (uložený ve vlastním digitálním fondu společnosti Wilson, volně dostupný na WWW nebo dostupný prostřednictvím jiného dodavatele), existuje-li; využívání služeb jiných dodavatelů úplných textů (EDD) je vázáno na další licence; systém je připraven na propojování záznamů pomocí technologie SFX
- Sledovat informace o fondech knihovny, které se vztahují k vyhledaným záznamům

The screenshot shows the 'Advanced Search' page of the WilsonWeb system. The search term 'LISA' is entered in the 'Search For' field. The 'Search In' section has 'Library Lit & Inf Full Text' selected. The 'Sort By' dropdown is set to 'Relevance'. The 'Limit Dates to' section has 'Any Year' selected. The 'Limit to' section has 'Full Text' selected. The 'Document Type' and 'Physical Description' dropdowns are both set to 'All'. The interface includes a sidebar with navigation options like 'Advanced Search', 'Basic Search', 'Browse', 'Search History', 'Thesaurus', 'Print Email Save', 'Help', and 'InfoCenter'. At the bottom, there are buttons for 'Customize Display', 'Clear', and 'Start'.

Obr. č. 1: obrazovka se standardním rozhraním pro komfortní pokročilé vyhledávání v databázi LLIS FTXT v rámci systému WilsonWeb (obrazovka zahrnuje formulaci dotazu hledání zkratky LISA s limitem hledání pouze článků s dostupným úplným textem) [získáno přes WWW 2004-10-25]

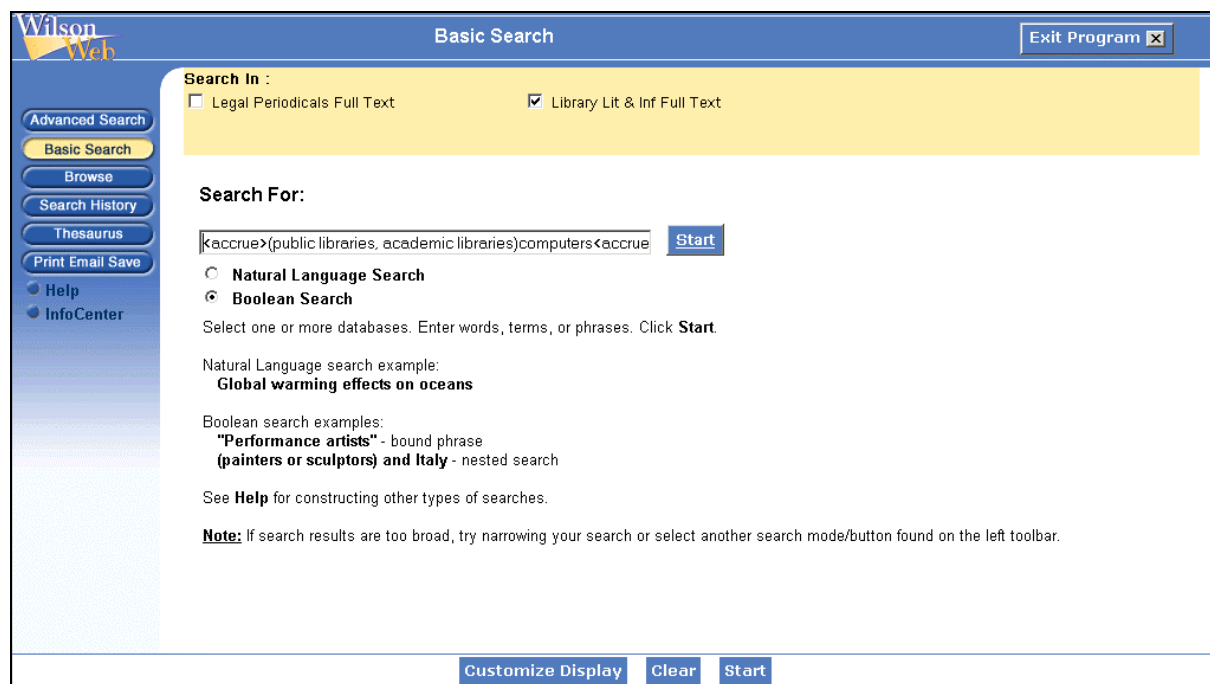
V rámci **konsorciálního přístupu** k databázi LLIS FTXT se uživatelé v ČR nemusejí do systému logovat. Přímý přístup ke standardní **úvodní stránce** rešeršního systému s rozhraním vyhledávání pro pokročilé (**Advanced search**), ze které je nutné před vlastním vyhledáváním zvolit databázi LLIS FTXT, je zajištěn dodavatelem služby. (Uživatelům různých univerzit v ČR mohou být nabízeny i další databáze). Po zaškrtnutí databáze LLIS FTXT může uživatel zahájit formulování rešeršního dotazu a vlastní vyhledávání. Počet přístupů do systému WilsonWeb je limitován (prozatím 4 přístupy v jednom čase). Po ukončení rešerší je nutné se odhlásit (**Exit Program**).

V rámci systému WilsonWeb může uživatel využívat texty základních manuálů v angličtině (tlačítko **Help**, viz obr. č. 1, levý rám nahoře) nebo může navštěvovat jednotlivé stránky Informačního centra webového sídla společnosti H. W. Wilsoson (tlačítko **InfoCenter**, viz obr. č. 2, rovněž levý rám nahoře).

2.2 Základní principy vyhledávání a zobrazování informací v systému WilsonWeb

Kromě standardního **rozhraní pro pokročilé vyhledávání (Advanced Search**, viz obr. č. 1) může uživatel přímo ze stránky tohoto rozhraní volit pomocí modrých tlačítek (jsou umístěná v levém rámu v horní polovině obrazovky):

- rozhraní pro **jednoduché vyhledávání (Basic Search**, viz obr. č. 2)
- rozhraní pro **prohlížení rejstříků (indexů) (Browse**, viz obr. č. 7)
- rozhraní pro **registraci vykonaných rešerší (Search History**, viz obr. č. 9) s návaznou funkcí **ukládání formulací rešeršních dotazů**
- rozhraní pro **vyhledávání z tezauru (Thesaurus**, viz obr. č. 6 a 8)
- rozhraní pro **tištění, ukládání a zasilání výsledků rešerší (Print, Email, Save)**






Obr. č. 2: obrazovka s rozhraním pro základní jednoduché vyhledávání v databázi LLIS FTXT v rámci rešeršního systému WilsonWeb (zapsán je složený booleovský dotaz s koncepčním operátorem ACCRUE) [získáno přes WWW 2004-10-25]


V rámci základních rozhraní pro vyhledávání z databáze LLIS FTXT jsou po vykonaných rešerších, registrovaných v dané relaci, systémem automaticky ohlašovány **počty nalezených záznamů** z databáze LLIS FTXT včetně kompletní formulace rešeršního dotazu. Systém zároveň automaticky **zobrazuje množinu záznamů** (standardně zkrácených citačních záznamů v úsporné formě „**in continuo**“, viz obr. č. 3). Název dokumentu je logicky hypertextově propojen s úplným záznamem. (Zkrácený záznam z pořadovým číslem 1 z obr. č. 3 je v úplné formě zobrazen na obr. č. 4)


The screenshot shows the 'Wilson Web' search results page. The title is 'Results: Brief Display'. The current search criteria are: 'LISA <in> ALL AND (FULL_TEXT = YES AND PAGE_IMAGE = YES)'. There are 773 records. The interface includes a sidebar with search options (Advanced Search, Basic Search, Browse, Search History, Thesaurus, Print Email Save, Help, InfoCenter) and a main results area. The results are listed with checkboxes for selection and icons for document types: Full Text (blue book icon), PDF (red document icon), and Peer Reviewed (black document icon). The results include titles like 'Is the Writing on the Wall for Dialog?', 'Evaluating libraryland topics in periodicals databases', 'Writing Novels and Having Fun: That's Amore', 'Guilty of Promoting Audiobooks', 'Flying Starts', 'In Profile', 'First Fiction Falls into Place', 'Meeting the new challenges at West and LexisNexis: post-SIAA Summit interviews with Michael Wilens and Lisa Mitnick', and 'Offshore Manufacturing'.

Obr. 3: obrazovka se seznamem zkrácených záznamů z databáze LLIS FTXT [získáno přes WWW 2004-10-25]

Vizualizace nalezených záznamů je na vynikající úrovni. Vlastní zkrácené záznamy předcházejí ikony, které reprezentují možnost přístupu k úplnému textu. Konkrétně jde o 3 ikony (viz také obr. č. 3):

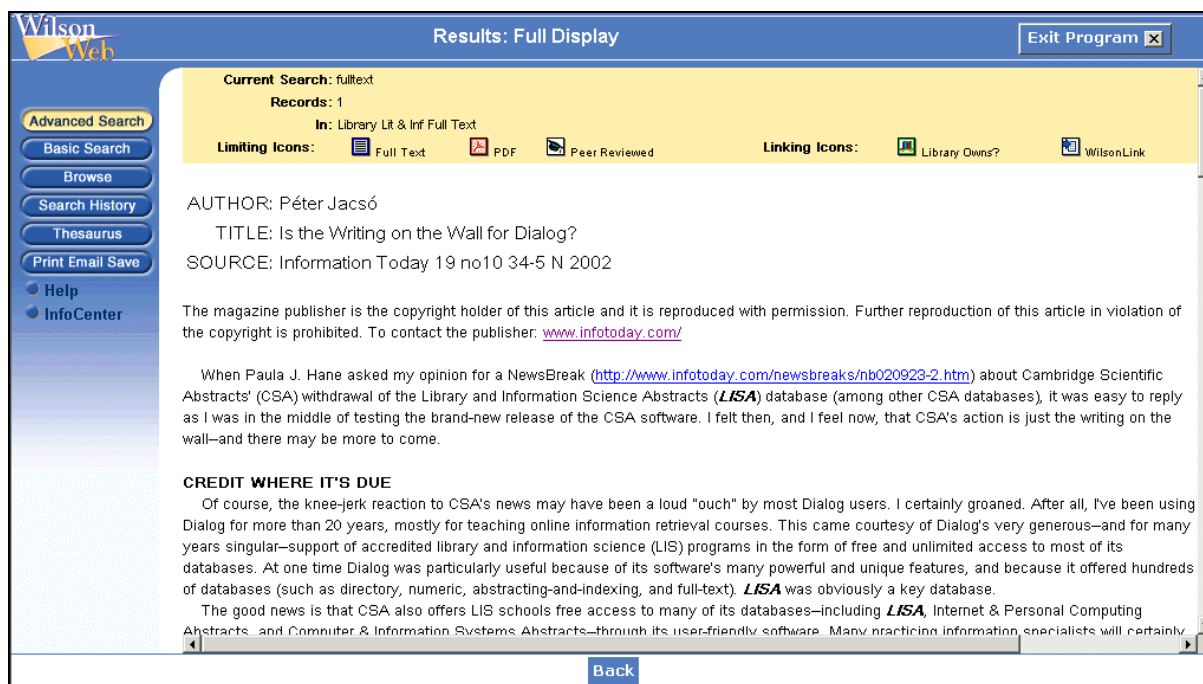
-  umožňuje bezprostřední získání úplného textu ve formátu HTML
-  umožňuje bezprostřední získání úplného textu ve formátu PDF
-  umožňuje propojení na jiné dodavatele úplných textů (licence je nutná)

Ikona  , která bývá umístěna na konci bibliografických záznamů, indikuje, že jde o dokument z kvalitního recenzovaného časopisu. Služby zpřístupňování **úplných textů od komerčních dodavatelů** jsou dostupné jen na základě dalších licencí či dohod (mohou být realizovány i na konkrétních českých univerzitách, nezbytné je však připravit odpovídající konfiguraci systému ze strany správce počítačové sítě).

Úplné bibliografické záznamy (viz obr. č. 4) jsou **hypertextově propojeny** z řady polí na specifické rejstříky databáze LLIS FTXT. Systém nabízí opět různé způsoby získání úplného textu a nabízí také **propojení do knihovního systému** (pomocí ikony ) k prověření existence dokumentů reprezentovaných v záznamech ve fondech knihovny.



Obr. 4: obrazovka s úplným záznamem z databáze LLIS FTXT (jde o záznam článku autora P. Jacsó, který je citován také v přiloženém seznamu použité literatury této příručky) [získáno přes WWW 2004-10-25]



Obr. 5: obrazovka s částí zobrazeného úplného textu ve formátu HTML z digitálního fondu společnosti H.W. Wilson (článek autora P. Jacsó) [získáno přes WWW 2004-10-25]

Základní typ vyhledávání informací (Basic Search) umožňuje také přímé formulování rešeršních dotazů do aktivního okna, které musejí být zapsány v souladu s principy dotazovacího jazyka systému WilsonWeb. Další informace jsou tudíž věnovány užívání

operátorů dotazovacího jazyka, vyhledávání variantních termínů a selekčním polím samotné databáze LLIS FTXT

2.2.1 Operátory dotazovacího jazyka rešeršního systému WilsonWeb

Rešeršní systém WilsonWeb je založen na booleovském modelu vyhledávání a nabízí ve stávající verzi velké množství booleovských operátorů rozdělených do několika **kategorií**. Využívání těchto operátorů je určeno pro zkušenější uživatele, méně zkušení mohou využívat jednodušší logické operátory nebo komfortní rozhraní pro vyhledávání s nabídkami operátorů.

V systému WilsonWeb lze fráze zapisovat jednoduše do uvozek: "**digital libraries**"

Ve formulacích dotazů s více logickými operátory se zapisují také kulaté závorky:

(bibliometrics or informetrics) and biology

Operátory **and**, **or**, **not** a **in** nevyžadují zápis ostrých závorek, jak je uvedeno níže v přehledu všech operátorů. Přímé formulování rešeršních dotazů s níže uvedenými operátory lze realizovat z obrazovky pro rozhraní základního vyhledávání (**Basic Search**).

Konceptní operátory (Concept Operators)

Tyto operátory kombinují význam selekčních termínů.

<accrue>

Příklad: **<accrue>(public libraries, academic libraries)computers<accrue>information**

V rámci uvedené formulace budou vyhledány záznamy obsahující alespoň jeden z uvedených selekčních termínů (čím více bude selekčních termínů, tím více bude výsledků)

<all>

Příklad: **<all>(public libraries and computers and readers)**

V rámci uvedené formulace budou vyhledány záznamy obsahující všechny selekční termíny, uvedené ve formulaci dotazu, záznamy ale nebudou řazeny podle relevance

<and>

Příklad: **<and>(public libraries and computers and readers)**

V rámci uvedené formulace budou vyhledány záznamy obsahující všechny selekční termíny; záznamy budou řazeny podle relevance

<any>

Příklad: **<any>(public libraries, academic libraries)<any>bibliography**

V rámci uvedené formulace budou vyhledány záznamy obsahující alespoň jeden z uvedených selekčních termínů; záznamy ale nebudou řazeny podle relevance

<or>

Příklad: **<or>(public libraries, academic libraries) or bibliography**

V rámci uvedené formulace budou vyhledány záznamy obsahující alespoň jeden ze selekčních termínů; záznamy budou řazeny podle relevance

Proximitní operátory (Proximity Operators)

Proximitní operátory specifikují relativní umístění slov v dokumentu. Specifikovaná slova musejí být ve stejné frázi, větě nebo odstavci v záznamu/dokumentu.

<in>

Příklad: **"information"<in>su**

V rámci uvedené formulace budou vyhledány záznamy obsahující selekční termín v poli předmětových hesel bibliografického záznamu

<near>

Příklad: **bibliography<near>library**

V rámci uvedené formulace budou vyhledány záznamy se všemi uvedenými selekčními termíny; záznamy budou řazeny podle relevance

<near/n>

Příklad: **patents<near/8>trademarks**

V rámci uvedené formulace budou vyhledány záznamy s uvedenými selekčními termíny, které budou od sebe o maximálně 7 slov; záznamy budou řazeny podle relevance

<phrase>

Příklad: **<phrase>(library users)**

V rámci uvedené formulace budou vyhledány záznamy, které obsahují danou frázi (slosoví)

Relační operátory (Relational Operators)

Relační operátory jsou určeny k vyhledávání údajů uložených v polích záznamu (Autor apod.); výsledné záznamy nebudou řazeny podle relevance; v polích se mohou vyskytovat písmena i číslice (u číslic lze uplatnit operátory =, !=, <, >, <=, >=)

<contains>

Příklad: **TI<contains>user**

PY<contains>2002

V rámci uvedené formulace budou vyhledány záznamy s uvedeným řetězcem v uvedených polích (název; rok vydání)

<ends>

Příklad: **TI<ends>cord**

V rámci uvedené formulace budou vyhledány záznamy s uvedeným řetězcem umístěným na konci pole název

<starts>

Příklad: **TI<starts>Know**

V rámci uvedené formulace budou vyhledány záznamy s uvedeným řetězcem lokalizovaným na začátku pole název

<substring>

Příklad: **JN<substring>journal of**

V rámci uvedené formulace budou vyhledány záznamy s uvedeným řetězcem, který bude součástí jiného řetězce (v poli název časopisu)

Evidenční operátory (Evidence Operators)

Následující operátory jsou určeny k základnímu nebo inteligentnímu vyhledávání slov

<stem>

Příklad: **<stem>film**

V rámci uvedené formulace budou vyhledány záznamy obsahující uvedené slovo (kmen) a všechny jeho varianty (films, filmable, filmed, filmcraft, filmdom, filmgoer aj.)

<typo/n>

Příklad: **<typo/3>search**

V rámci uvedené formulace budou vyhledány záznamy obsahující uvedené slovo a jemu podobná slova (založeno na bázi přibližného porovnávání vzoru slova; originální manuál nspecifikuje podrobnosti k užití tohoto operátoru)

<wildcard>

Příklad: **<wildcard>"bank[s,er,ing]"**

<wildcard>"gr[ae]y"

V rámci uvedené první formulace budou vyhledány záznamy obsahující slova banks, banker, banking, v rámci druhé formulace budou vyhledány záznamy obsahující slova gray a grey; (v současné verzi systému lze užívat tento operátor pouze v poli název TI); ve formulacích dotazů lze ale užit i běžných symbolů * (corp*) a ? (gr?y) bez uvádění ostrých závorek; symbol ? se může vyskytovat i na začátku řetězce (?an)

<word>

Příklad: **<word>delivery**

V rámci uvedené formulace budou vyhledány záznamy obsahující jeden nebo více výskytů přesně zapsaného selekčního termínu

Skórovací operátory (Score Operators)

Následující operátory mají vliv na uspořádání záznamů podle relevance; nedají se použít společně s modifikátory (viz dále)

<product>

Příklad: **<product>("library", "user")**

V rámci uvedené formulace bude vypočítána relevance nalezených záznamů pomocí násobení skóre každého z uvedených selekčních termínů

<sum>

Příklad: **<sum>("bibliography")**

V rámci uvedené formulace bude vypočítána relevance nalezených záznamů přidáním skóre selekčnímu termínu

<yesno>

Příklad: **<yesno>("loan")<and>"services"**

Uvedený operátor limituje vyhledávání na nalezení záznamů, ve kterých nebude vypočítáváno skóre uvedených selekčních termínů, jež má vliv na uspořádání dokumentů podle relevance

Operátory přirozeného jazyka (Natural Language Operators)

Tyto operátory umožňují vyhledávacímu systému pracovat se syntaxí přirozeného jazyka

<freetext>

Příklad: **<freetext>("interlibrary loans as the information services")and(2003)**

V rámci uvedené formulace budou nalezeny záznamy/úplné texty, které zahrnují uvedenou frázi a číslo; nalezené záznamy budou uspořádány podle relevance

Modifikátory (Modifiers)

Standardní chování operátoru se může změnit pomocí následujících operátorů

<many>

Příklad: **<many>information users**

V rámci uvedené formulace bude při výpočtu relevance přihlíženo k hustotě selekčního termínu

<not>

Příklad: **libraries not academic libraries**

V rámci uvedené formulace budou z výsledného počtu záznamů vyloučeny ty, které obsahují selekční termín uvedený po operátoru NOT

<order>

Příklad: **books<order>serials**

V rámci uvedené formulace se specifikuje pořadí selekčních termínů, ve kterém se musejí vyskytovat v dokumentu

2.2.2 Selekční pole databáze LLIS FTXT

Pro vyhledávání záznamů z **databáze LLIS FTXT** v režimu online z databázového centra společnosti H.W. Wilson lze využívat následujících selekčních polí (výčet jejich kódů i jmen je uveden v anglickém znění, níže komentovaný výčet zahrnuje český překlad; limitační selekční pole jsou uvedena s podtržením):

AB	Abstract
ABI	<u>Abstract Indicator</u>
AN	<u>Accession Number</u>
AU	Personal Author
BR	Book Reviewed
CA	Corporate Author
CP	<u>Country of Publication</u>
CSO	Chapter Source
DDC	<u>Dewey Decimal Number</u>
DT	<u>Document Type</u>
FI	<u>Fiction Indicator</u>
FR	Form Review
FTI	<u>Full Text Indicator</u>
ISBN	<u>ISBN</u>
ISSN	<u>ISSN</u>
JI	Journal Issue
JN	Journal Name
LA	<u>Language of Document</u>
LCC	<u>LC Classification</u>
LCCN	<u>LC Control Number</u>
NT	Notes
OT	Other Titles
PA	Pages
PD	<u>Physical Description</u>
PRI	<u>Peer Reviewed Journal</u>
PY	<u>Publication Year</u>
RS	Responsibility Statement
SO	Source
SU	Subject(s)
TI	Title
TXT	Text
UD	<u>Update Code</u>

Následuje přehled jednotlivých selekčních polí databáze LLIS FTXT s krátkou informací o jejich využívání. Kódy a jména selekčních polí jsou uvedena **tučně**. Limitační pole jsou uvedena navíc s podtržením:

AB Abstract (Referát)

Příklad: abstracting in ab

Pole **AB** zahrnuje referát. Objevuje se ale až v nových záznamech databáze LLIS. FTXT. Vyhledávat lze libovolné slovo nebo frázi v anglickém jazyce obsaženou v referátu.

ABI Abstract Indicator (Indikátor referátu) (Limitační pole)

Příklad: y in abi

Pole **ABI** indikuje, zda-li záznam obsahuje referát článku. Hodnota „y“ znamená, že abstrakt je přítomen.

AN Accession Number (Přirůstkové číslo [záznamu]) (Limitační pole)

Příklad: 200230506105015 in an

Selekční pole **AN** obsahuje jedinečné číslo přidělené producentem záznamu.

AU Personal Author (Personální Autor)

Příklad: christiani in au

Pole **AU** zahrnuje jméno (jména) autora (autorů) popisovaného dokumentu. Vyhledávat lze podle příjmení nebo podle křestního jména v libovolném pořadí. Případně lze využít jmenného rejstříku databáze.

BR Books Reviewed (Recenzovaná kniha)

Příklad: tenopir in br

Pole **BR** zahrnuje název a autora knihy, která je recenzována. Vyhledávat lze podle libovolného slova nebo fráze.

CA Corporate Author (Korporativní autor)

Příklad: wilson in au

Pole **CA** zahrnuje jméno korporativního autora (instituci, asociaci, konferenci, státní orgán apod.) dokumentu. V případě potřeby lze využít jmenného rejstříku databáze.

CP Country of Publication (Země publikování) (Limitační pole)

Selekční pole **CP** obsahuje zemi publikování časopisu nebo knihy. Hodnoty je vhodné ověřovat v rejstříku databáze.

CSO Chapter Source (Zdroj kapitoly)

Pole **CSO** zahrnuje název knihy, ze které je popisovaná kapitola nebo esej.

DDC Dewey Decimal Number (Třídník DDT) (Limitační pole)

Příklady: 028.5 in ddc
 ddc<starts>028

Pole **DDC** obsahuje třídíky DDT. Operátor <starts> lze použít, pokud hledáme záznamy knih, které patří do třídy začínající nějakým číslem.

DT Document Type (Typ dokumentu) (Limitační pole)

Příklady: feature article in dt

Pole **DT** obsahuje kategorie typů dokumentů. Je vhodné k limitování rešerší. Jde o níže uvedené hodnoty (nutno uvádět v angličtině):

Bibliography
Biography
Books
Book Chapter
Book Excerpt
Book Review
Corporate profile
Do It Yourself Work
Exhibit
Feature Article
Interview
Obituary
Product Evaluation
Speech
Symposium

FI Fiction Indicator (Indikátor Beletrie) (Limitační pole)

Příklady: fiction in fi
nonfiction in fi

V poli **FI** se objeví hodnota **nonfiction**, pokud recenzovaná kniha není beletrií, nebo **Fiction**, pokud beletrií je.

FR Form Review (Forma recenze)

Pole **FR** indikuje recenzi jiných dokumentů než knih, například opery, filmu aj.

FTI Full Text Indicator (Indikátor úplného textu) (Limitační pole)

Pole **FTI** indikuje, zda-li je k záznamu připojen úplný text. Hodnota „y“ znamená, že úplný text je přítomen (y in fti).

ISBN ISBN (ISBN) (Limitační pole)

Pole **ISBN** zahrnuje identifikátory ISBN.

ISSN ISSN (ISSN) (Limitační pole)

Pole **ISSN** zahrnuje identifikátory ISSN.

JI Journal Issue (Číslo časopisu)

Pole **JI** obsahuje údaj o čísle časopisu excerptovaného v databázi. K ověření čísel lze využít seznamu čísel časopisů v rámci specifického rejstříku databáze.

JN Journal Name (Rejstřík názvů časopisů)

Příklad: biblos in jn

Pole **JN** obsahuje názvy excerpovaných časopisů. K ověření názvů lze využít specifického rejstříku názvů časopisů.

LA Language of Document (Jazyk dokumentu) (Limitační pole)

Příklad: english in la

Selekční pole **LA** indikuje jazyk, ve kterém je napsán dokument. Většina dokumentů v databázi LLIS FTXT je v angličtině.

LCC LC Classification (Třídění LC) (Limitační pole)

Příklady: PZ7.H314975 in lcc

Pole **LCC** obsahuje třídění Kongresové knihovny (jde zároveň o signaturu dokumentu ve fondu LC).

LCCN LC Control Number (Identifikační číslo LC) (Limitační pole)

Příklad: 99-89390 in lccn

Pole **LCCN** obsahuje identifikátory záznamů z katalogu LC.

NT Notes (Poznámky)

Pole **NT** obsahuje poznámku o tom, že existuje jiný článek o určité osobě v databázi LLIS FTXT.

OT Other Titles (Jiné názvy)

Pole **OT** umožňuje doplňkové vyhledávání informací vztahujících se k názvu dokumentu. Může zahrnovat další slova rozšiřující název. Doplnuje je zpracovatel záznamu. Vyhledávat lze libovolné slovo nebo frázi.

PA Pages (Stránky)

Pole **PA** obsahuje údaj o počtu stran.

PD Physical Description (Fyzický popis) (Limitační pole)

Příklad: por in pd

Pole **PD** obsahuje údaje o vybavení dokumentu. Vyhledávat lze podle kódů nebo úplných termínů: bibl = bibliography, diag = diagram(s), chart = flow chart(s), f = footnotes, graph = graph(s), il = illustrations, por = portrait(s), tab = table(s)

PRI Peer Reviewed Journal (Recenzovaný časopis) (Limitační pole)

Příklad: y in pri

Pole **PRI** indikuje, zda-li je popisovaný dokument (článek) z recenzovaného časopisu. Hodnota „y“ znamená, že jde o recenzovaný časopis. Hodnota se v databázi LLIS FTXT uvádí od roku 1996.

PY Publication Year (Rok publikování) (Limitační pole)

Příklad: 2004 in py

Pole **PY** obsahuje rok publikování dokumentu ve formě 4 číslic. Lze vyhledávat i řadu dat publikování (s využitím operátoru AND):

(py>1990) and (py<1998)

Data publikování lze ověřovat ve specifickém rejstříku databáze.

RS Responsibility Statement (Údaj o odpovědnosti)

Pole **RS** zahrnuje jména osob, kteří přispěli ke vzniku díla (může jít o autory, editory, sestavovatele, ilustrátory, překladatele aj.)

SO Source (Zdroj)

Pole **SO** obsahuje informace, které popisují zdrojový dokument analytických popisných jednotek (především článků z časopisů). Základním údajem je název zdrojového seriálového dokumentu (časopisu aj.). Dalšími údaji v poli **SO** jsou **ročník, číslo, měsíc, rok a rozsah stran** (formální úprava těchto lokačních údajů se řídí pravidly zpracovatele). Vyhledávat lze podle slova nebo fráze. Informace o zdrojové knižní publikace se nachází v poli **CSO** (viz výše).

SU Subject(s) (Předmět(y))

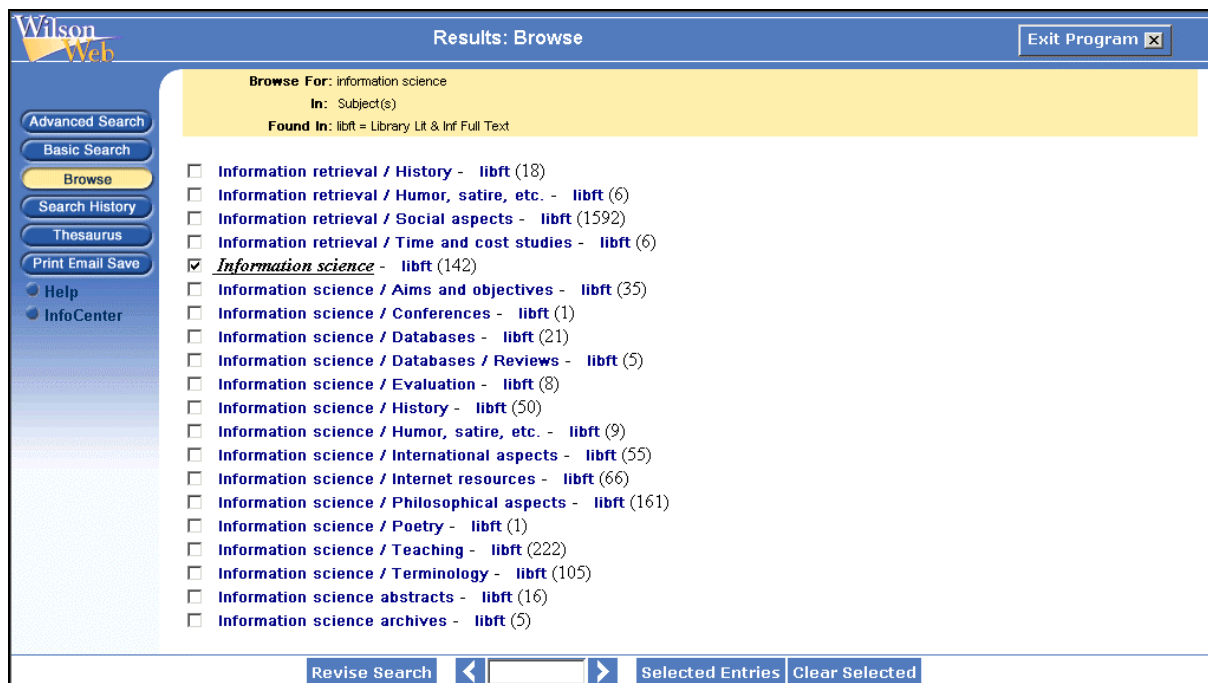
Příklad: DIALOG (information system) in su

Jde o jedno z **nejvýznamnějších selekčních polí databáze LLIS FTXT**. Zahrnuje autoritní hesla z řízeného oborového předmětového hesláře (v počítačové databázi je označen termínem tezaurus). Zahrnuta jsou **hesla personální, korporativní, geografická, tematická** (viz obr. č. 6 se seznamem předmětových hesel zobrazený v rámci rejstříku databáze LLIS FTXT) a také **unifikované názvy děl**.

TI Title (Název)

Příklad: library in ti

Pole **TI** obsahuje úplný název popisovaného dokumentu. Vyhledávání z pole **TI** lze realizovat jak na základě přesné fráze, tak na základě libovolného slova z názvu.



Obr. 6: obrazovka se seznamem předmětových hesel z řízeného hesláře databáze LLIS FTXT (zašknuto je heslo *Information science*) v rámci rešeršního systému WilsonWeb [získáno přes WWW 2004-10-25]

TXT Text (Text)

Pole **TXT** zahrnuje celý text, který je možné prohledávat.

UD Update Code (Kód aktualizace) (Limitační pole)

Příklad: 20041006 in ud
(ud>20030101) and (ud<20030630)

Pole **UD** obsahuje kód, který indikuje datum aktualizace databáze, tj. dobu, kdy do databáze LLIS přibyly nové záznamy. Kód je tvořený 8 číslicemi (4 číslice pro rok, 2 číslice pro měsíc a 2 číslice pro den). Pomocí operátoru **and** lze vyhledávat řadu dat.

URL Web Sites (Webové sídlo)

Pole **URL** zahrnuje URL adresy URL zmiňované autorem dokumentu, které se týkají tématu dokumentu.

V rámci některých limitačních polí, osahujících číslice, lze uplatnit specifické operátory. Příklady:

<all>(py<1993)
<all>(py>1993)
<all>(ud<=19931010)
<all>(ud>=19931010)

2.3 Základní způsoby vyhledávání informací z databáze LLIS FTXT

2.3.1 Pokročilé vyhledávání (Advanced Search)

Vícehlediskové vyhledávání se v databázi LLIS FTXT realizuje v rámci rozhraní pro **pokročilé vyhledávání** (viz obr. č. 1). V rámci velké plochy pravého rámu obrazovky uživatel může vyplnit postupně jednotlivé selekční údaje. Zapisovat lze slova nebo fráze do celkem 3 základních aktivních polí. Je-li polí více, je potřebné vybrat logický operátor. Z rozbalovacího menu lze vybírat způsob provedení rešerše.

Bude-li užito standardní nabídky **All-Smart Search**, nebude možné zapisovat symboly pro zkracování selekčních údajů ani žádné další speciální znaky. Rešeršní systém využívá v tomto případě hierarchii pravidel založených na indexaci společnosti Wilson s aplikací vyhledávacího algoritmu společnosti Verity. Výsledné záznamy budou standardně uspořádány podle relevance. V případě velké množiny získaných záznamů lze formulaci dotazu editovat v rozhraní pro registraci vykonaných rešerší (Search History). Rešerši lze také zúžit pomocí výběru jednoho ze selekčních polí z rozbalovací nabídky.

Vybírat lze také vyhledávání podle jmen autorů (**Authors**), podle klíčových slov (**Keyword**, zahrnuta jsou slova z různých polí záznamu včetně abstraktu) nebo lze vyhledávat v libovolném specifickém poli (**Searching in Specific Field**), která jsou v databázi definována. Před vykonáním rešerše lze ještě nastavit **další limity**: výsledné uspořádání výsledných záznamů (standardně podle relevance), časový limit rešerše, typ dokumentu, podrobnosti fyzického popisu a zobrazování záznamů pouze s úplnými texty, stranami s obrázky nebo dokumenty z recenzovaných časopisů. Konečná formulace dotazu se odesílá pomocí tlačítka **Start**.

2.3.2 Základní vyhledávání (Basic Search)

V rámci **základního vyhledávání** (viz obr.č. 2) se v systému WilsonWeb musí napřed zvolit způsob vyhledávání:

Natural Language search (vyhledávání v přirozené jazyce)

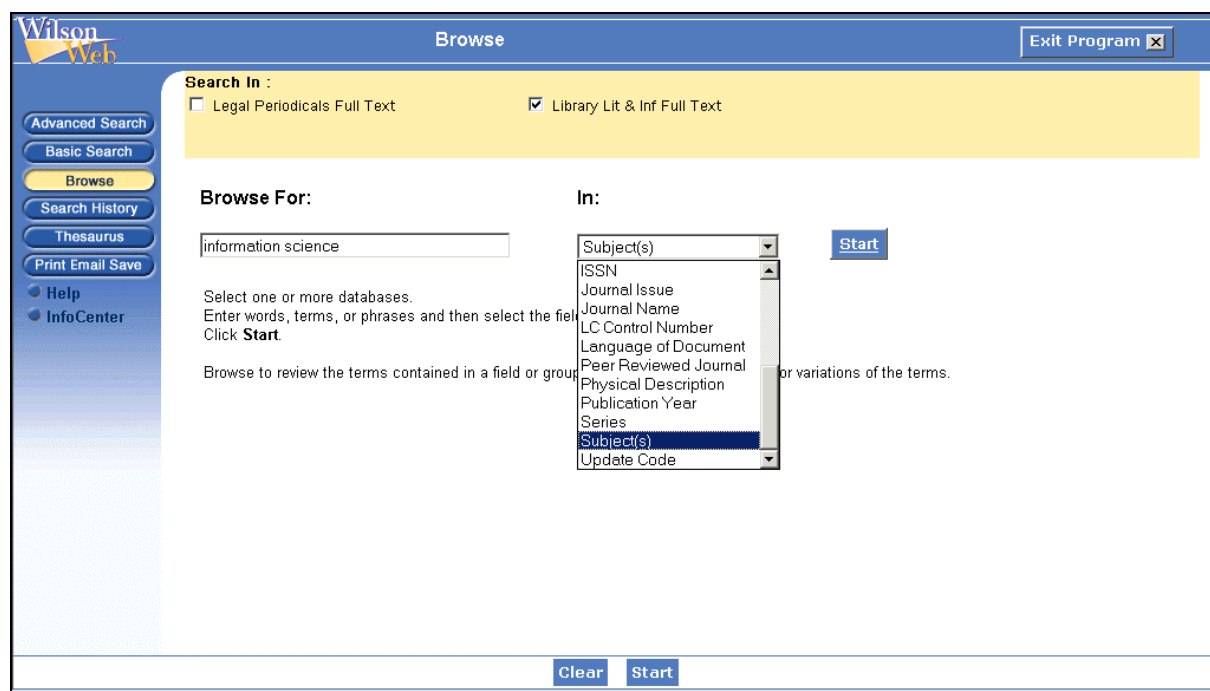
Boolean search (booleovské vyhledávání)

V rámci **vyhledávání v přirozeném jazyce** se zapisují fráze nebo úplné věty. Do vyhledávání nejsou zahrnuty některé obecné výrazy - stop slova (v angličtině například slovesný tvar **are**, člen **the**, předložka **for** aj.). V případě potřeby lze stop slova ale vyhledávat pomocí zápisu uvozovek (například: **library "and" information science**).

V rámci **booleovského vyhledávání** se užívá základních logických operátorů nebo specifických operátorů a syntaxe společnosti Verity (Topic). Uzavřené fráze se obklopují do uvozovek. Tlačítkem **Start** se opět provádí vyhledání záznamů.

2.4 Vyhledávání informací z databáze LLIS FTXT pomocí rejstříků

Rešeršní systém WilsonWeb nabízí uživatelům prohlížení a vyhledávání z řady **specifických rejstříků databáze LLIS FTXT**. Po otevření úvodní obrazovky s rozhraním pro prohlížení rejstříků (pomocí tlačítka **Browse**, viz obr. č. 7) si uživatel volí z rozbalovací nabídky potřebný rejstřík a do aktivního pole zapisuje hledaný selekční termín nebo jen několik jeho prvních znaků. Tlačítkem **Start** se vyvolá obrazovka s příslušnou abecední částí rejstříku (viz obr. č. 6). Jednotlivé položky seznamu jsou **navigovatelné** a klikáním se lze dostat přímo k záznamům (počet záznamů je uveden v kulaté závorce). V seznamu lze listovat po skupinách doprava i doleva. Na základě zaškrtnutých položek seznamu si lze pomocí funkce **Selected Entries** prohlížet pouze skupinu vybraných selekčních termínů.

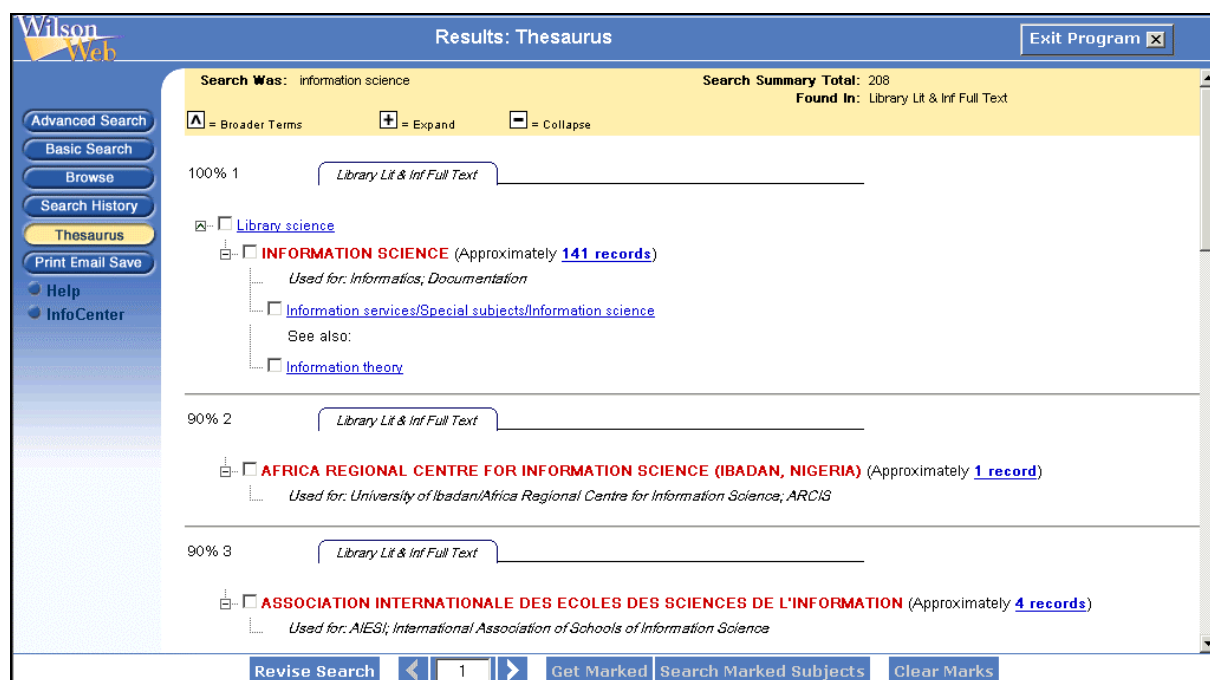


Obr. 7: obrazovka rozhraní pro prohlížení rejstříků v databázi LLIS FTXT (zapsán je termín *information science* a z nabídky je zvoleno je pole *Subject(s)*) v rámci rešeršního systému WilsonWeb [získáno přes WWW 2004-10-25]

2.5 Vyhledávání informací pomocí tezauru databáze LLIS FTXT

Tezaurem databáze LLIS je míněn oborový řízený heslář. Je vynikajícím nástrojem věcného vyhledávání v databázi.

Rozhraní pro vyhledávání v tezaurovi se vyvolá pomocí tlačítka **Thesaurus**. V tezaurovi lze vyhledávat podle deskriptorů i nedeskriptorů. Výsledné zobrazení deskriptorových odstavců zahrnuje všechny vazby na nadřazené, podřazené a související deskriptory i nedeskriptory (viz obr. č. 8).



Obr. 8: obrazovka se záznamy deskriptorových odstavců z tezauru (řízeného hesláře) databáze LLIS FTXT [získáno přes WWW 2004-10-25]

V tezaurovi lze snadno navigovat a zobrazovat záznamy, které obsahují dané deskriptory. Po vyhledávání systém zobrazuje v záhlaví obrazovky hledaný selekční údaj i celkový počet záznamů, které ho obsahují. Deskriptory, obsahující hledaný termín, jsou doplněny informací o přibližném počtu záznamů, zahrnujících daný deskriptor. Zaškrtnuté deskriptory lze pomocí tlačítka **Search Marked Subjects**, nacházejícího se v dolní části obrazovky, z databáze vyhledat.

2.6 Operace s výsledky rešerší v databázi LLIS FTXT

2.6.1 Změny zobrazování záznamů

V rámci obrazovek rozhraní pro základní typy vyhledávání informací z databáze LLIS FTXT lze pomocí funkce **Customize Display**, jejíž tlačítko se nachází na spodním okraji obrazovky (viz obr. č. 1 nebo 2), měnit parametry zobrazování záznamů. Po kliknutí na uvedené tlačítko se otevře obrazovka, v rámci níž lze měnit jednotlivé parametry: **typ zobrazování záznamu** (zkrácený nebo úplný), **uspořádání záznamů** (podle roku vydání, relevance nebo podle dalšího zvoleného údaje), **zařazení specifických polí do zobrazovaného záznamu** podle specifikace uživatele (pole se vybírají z nabídky) a **uživatelské nastavení parametrů** pro zobrazování, tištění e-mailování nebo ukládání záznamů.

Změna parametrů se potvrzuje tlačítkem **Save Changes**. Změny lze kdykoliv zrušit (**Cancel**). Standardní nastavení jednotlivých parametrů lze opět zavést pomocí funkce **Default settings**.

2.6.2 Registrace vykonaných rešerší v dané relaci a ukládání formulací rešeršních dotazů

Po vyhledání záznamů systém WilsonWeb v rámci jedné relace automaticky registruje jednotlivé rešerše v rámci rozhraní pro registraci vykonaných rešerší (**Search History**). Systém standardně pod titulkem **Active Searches** zobrazuje záznamy vykonaných rešerší (viz obr. č. 9). V záznamu je v aktivním okně zobrazována formulace rešeršního dotazu (Search String) podle dotazovacího jazyka systému WilsonWeb (Verity), kterou je možné eventuelně editovat, dále operátory pro případně kombinování jednotlivých rešerší mezi sebou, počet výsledných záznamů a datum posledního provedení rešerše. Nepotřebné rešerše lze označit a vymazat (Delete). Nové provedení rešerše se provádí po jejím označení pomocí funkce **New Results**.

Podstatnou funkcí tohoto rozhraní je funkce **Save**. Pomocí ní se **formulace rešerší** mohou **ukládat** na server databázového centra (až na dobu 6 měsíců) za účelem jejich opětovného využívání. V rámci dialogu je ale nejdříve systémem požadováno vytvoření **personálního ID** (v dolní polovině obrazovky nadepsané titulkem **Saved Searches**). Žádán zápis **e-mail adresy** uživatele a také **hesla**. Potom teprve lze, po označení záznamu zvolené rešerše a zápisu jména ukládané formulace této rešerše (do aktivního okna **Save as**) formulaci rešeršního dotazu uložit. K opětovnému převzetí uložené formulace dotazu a vykonání nové rešerše je nutné zapsat opět ID uživatele. V rámci dialogu se označí záznam uložené formulace a pomocí tlačítka **Load** se formulace převádí do obrazovky rozhraní **Search History**.

The screenshot shows the 'Search History' window in WilsonWeb. It features a sidebar with navigation options like 'Advanced Search', 'Basic Search', 'Browse', 'Search History', 'Thesaurus', 'Print Email Save', 'Help', and 'InfoCenter'. The main area is titled 'Active Searches' and contains a table with three rows of search results. Each row includes a checkbox, a search string, a radio button for logical operators (And, Or, Not), the number of search hits, and the last run date. Below the table is a 'Save As:' field. At the bottom, there is a 'Saved Searches' section with a 'Sign On' button and an ID input field.

	Search String	Search Hits	Last Run Date
1. <input type="checkbox"/>	<and>(<near>(LISA) <in> ALL,<yesno>(<and>(y <in>	1308	10/29/04
2. <input type="checkbox"/>	<and>(<near>(LISA) <in> SU,<yesno>(<and>(y <in> fti))	8	10/29/04
3. <input type="checkbox"/>	<and>(<near>(Library and Information Science Abstracts) <in> SU,<yesno> (<and>(y <in> fti))	4	10/29/04

Obr. 9: obrazovka se seznamem 3 vykonaných rešerší v databázi LLIS FTXT v jedné relaci [získáno přes WWW 2004-10-25]

2.6.3 Ukládání, tištění a zasílání záznamů elektronickou poštou

V rámci systému WilsonWeb lze množinu označených záznamů nebo řadu nalezených záznamů vybrat k ukládání na počítačové médium, tištění nebo zasílání elektronickou poštou. Rozhraní se vyvolává pomocí tlačítka **Print Email Save**. V rámci dialogu pro zasílání záznamů elektronickou poštou je potřebné uvést e-mail adresu, předmět záslky, označit množinu záznamů k zaslání a stanovit výstupní formát záznamů. Pro potřeby tištění záznamů je nezbytné stanovit označené záznamy nebo úplné texty a jejich formát. K ukládání záznamů nebo úplných textů je nutné rovněž stanovit formát, počítačové médium a jméno souboru.

Použitá literatura a jiné informační zdroje

- BRATKOVÁ, E. 1998. Bibliografické a plnotextové báze dat americké firmy H.W. Wilson pro společenskovědní a humanitní obory : vyhledávání informací v systému WilsonWeb. In *INFOMEDIA'98* [online]. Praha : Albertina icome Praha, 1998 [cit. 2004-10-25]. Dostupný z WWW: <<http://www.inforum.cz/infomedia98/pdf/wilson.htm>>
- H.W. Wilson Company. 2004. *H.W. Wilson Help : Documentation* [online]. Bronx (N.Y.) : H.W. Wilson Co., c2004 [cit. 2004-10-25]. Dostupný ze systému WilsonWeb na WWW: <<http://hwwilsonweb.com>>
- H.W. Wilson Company. 2004. *WilsonWeb Journal Directory* [online]. Bronx (N.Y.) : H.W. Wilson Co., c2004, updated October 25, 2004 [cit. 2004-10-25]. Dostupný ze systému WilsonWeb na WWW: <<http://hwwilsonweb.com>>
- JACSÓ, Péter. 2001. Electronic Shoes for the Cobbler's Children : Treatment of Digital Journals in Library and Information Science Databases. *Online* [online]. July 2001. Dostupný z WWW: <http://www.onlinemag.net/OL2001/jacso7_01.html>
- JACSÓ, Péter. 2002. Is the Writing on the Wall for Dialog ? *Information Today*. November 2004, vol. 19, no. 10, s. 34-35. Dostupný také z WWW: <<http://hwwilsonweb.com>>
- MATOUŠOVÁ, M; VONDRA, J. 1980. *Referátové časopisy a vybrané informační zdroje z informatiky*. Praha : Ústředí vědeckých, technických a ekonomických informací, 1980. 64 s. Metodické studie, sv. 139.