

Integrované vyhledávání v informačních zdrojích Národní lékařské knihovny - výzva 21. století

Filip Kříž, Ondřej Horský, Lenka Maixnerová, Jana Zindulková, Adéla Jarolímková, Eva Lesenková, Helena Bouzková

Národní lékařská knihovna, Praha

kriz@nlk.cz

INFORUM 2011: 17. konference o profesionálních informačních zdrojích
Praha, 24. - 26. 5. 2011

Abstrakt

Národní lékařská knihovna vyvíjí a provozuje portál Medvik (www.medvik.cz), který zajišťuje uživatelům přístup k bibliografickým a autoritním databázím provozovaným v systému Medvik (Medicínská virtuální knihovna). Hlavními bibliografickými databázemi spravovanými NLK jsou Katalog knihoven Medvik a Bibliographia medica Českoslovacca, autoritní báze představují soubory českých lékařských autorit (jména osob a korporací) a tezaurus Medical Subject Headings. Počet bibliografických a autoritních záznamů dosahuje téměř 900 tisíc. Řešení předchozí verze portálu nebylo optimální pro rychlé a uživatelsky přívětivé vyhledávání odpovídající očekáváním uživatelů ve 21. století. Bylo proto vytvořeno nové řešení, které automaticky integruje záznamy z uvedených bází do jedné společné databáze a umožňuje prohledávat všechny báze z jednoho místa. Uživatelům je v portálu k dispozici řada funkcí umožňujících rychlý a pohodlný přístup k poskytovaným informačním zdrojům a navazujícím službám, včetně plnotextového hledání, fasetové navigace a vizualizace výsledků hledání pomocí MeSH tag clouds. Od roku 2011 mají registrovaní uživatelé NLK k dispozici online službu federativního vyhledávání MedGate, která umožňuje lokálně i vzdáleně prohledávat vybrané licencované i volné elektronické zdroje.

Úvod

Důležitým předpokladem rozvoje informačních služeb knihovny je reagovat na očekávání uživatelů zvyklých na jednoduché vyhledávací postupy stávajících webových vyhledávačů. Současní uživatelé očekávají od vyhledávacího rozhraní jednoduchý formulář umožňující nejlépe plnotextové vyhledávání a to, pokud možno, v co největším počtu informačních zdrojů s přímo dostupnými plnými texty. Vedle vědecko-výzkumných pracovníků a knihovníků, kteří patří do kategorie uživatelů s vysokými nároky na vyhledávací systémy, poskytuje NLK služby pro široké spektrum zdravotnických pracovníků a laiků, kteří však nejsou ochotni zkoumat jiné složitější možnosti hledání nebo přístupu ke zdrojům a službám. Tradičně bohatě strukturované záznamy publikací zpřístupňované v online katalozích pomocí sofistikovaných formulářů tak mohou být současnými uživateli často opomíjeny a stejně tak i navazující služby knihoven. S přihlédnutím k potřebám pokročilých uživatelů nelze tedy při vývoji zavrhnout tradiční rešeršní nástroje jako jsou řízené slovníky, autoritní báze nebo možnosti tvorby rešeršní strategie. Řešením těchto problémů je vývoj rozhraní, která odpovídají očekáváním uživatelů a která zároveň dokážou využít naplno potenciál strukturovaných dat.

NLK vyvíjí a provozuje portál Medvik (www.medvik.cz), který zajišťuje uživatelům přístup k bibliografickým a autoritním databázím provozovaným v systému MEDVIK (Medicínská virtuální

knihovna). Hlavními bibliografickými databázemi spravovanými NLK jsou Katalog Medvik a Bibliographia medica Českoslovaca (BMČ), autoritní báze představují soubory českých lékařských autorit (jména osob a korporací) a tezaurus Medical Subject Headings (MeSH). Počet bibliografických a autoritních záznamů dosahuje téměř 900 tisíc.

Rozhraní portálu Medvik (www.medvik.cz) se příliš nezměnilo od svého spuštění v roce 2006. Od té doby přibyly do prostředí Medvik další systémy - Digitální archiv Kramerius, databáze BMČ, Portál e-zdrojů - které bylo třeba integrovat a jejich data bylo efektivně zpřístupnit uživatelům. Již při vytváření dílčích řešení - BMČ Beta, portlety zobrazující dostupnost online časopisů z Portálu E-zdrojů a dostupnost publikace v digitálním archivu atp. - začalo být zřejmé, že stávající architektura řešení portálu Medvik přestává být dostačující. Bylo proto vytvořeno nové řešení, které automaticky agreguje záznamy z uvedenýchází a systémů do jedné společné databáze. Důraz byl kladen zejména na jednoduchost rozhraní, možnost plnotextového hledání v integrovaných datech, fasetovou navigaci a využití a zviditelnění tezauru MeSH.

Nová verze portálu Medvik

Nová verze portálu pracující nad společnouází umožňuje efektivní přístup k uloženým informacím a navazujícím službám z jednoho místa. Základní hledání, dostupné vždy na všech stránkách portálu, je založené na plnotextovém přístupu a neklade tak na uživatele vysoké nároky při formulaci dotazu. Uživatel pomocí hlavního formuláře zadá nespécifikovaný dotaz, na základě kterého systém vrátí obvykle větší počet výsledků. Pomocí fasetové navigace (klastrů výsledků na základě různých kritérií) může uživatel seznam výsledků snadno dále zužovat a dostat se tak k relevantním výsledkům. Zvolené fasety může uživatel z dotazu snadno odstraňovat a má možnost vybrat více hledisek najednou pomocí lokálních filtrů.

Součástí úvodní stránky portálu je také blok rychlého hledání, který umožňuje rychlé zadání přesněji specifikovaných dotazů pro konkrétní potřeby - vyhledání dokumentu podle názvu, hledání časopisů, předmětů, autorů a odborných akcí. Zároveň však portál poskytuje nástroje pro pokročilé vyhledávání, které umožňuje provádět komplexní rešeršní postupy kombinováním dotazů s pomocí logických operátorů.

Seznam výsledků lze zobrazovat v různých formátech a řadit podle relevance, data vydání, názvu a autora. Důraz byl kladen na zobrazení nejdůležitějších údajů a využití vazeb mezi bibliografickými a autoritními záznamy. U každého zobrazeného záznamu je lišta s dostupnými navazujícími službami - odkazy na plné texty v Digitálním archivu NLK, objednání kopie, zobrazení detailu a abstraktu, objednání digitální kopie - služba E-books On Demand atd. Ukázka zobrazení výsledků hledání je uvedena na Obr. 1.

Úvod **Výsledky** Pokročilé hledání Autoři Témata Schránka

adiktologie

Hledej v: Katalogy BMČ-články vše

Portál E-Zdrojů
Digitální archiv Kramerius
Centrální vyhledávání MedGate

Upřesnit výsledky

Vydáno

Letošní rok	2
Loňský rok	90
Poslední dva roky	90
Posledních pět let	483
2000-2010	716

[filtrovat](#)

Jazyk

čeština	493
angličtina	233
slovenština	27

[filtrovat](#)

Země

Česko	716
-------	-----

[filtrovat](#)

Check Tag

lidé	302
ženské pohlaví	64
mužské pohlaví	59
mladiství	51
dospělí	43
srovnávací studie	28
dítě	27
více...	filtrovat

Výsledky dotazu:

událost	výsledky
adiktologie	716

[zpět k výchozímu dotazu](#) [o krok zpět](#) [přidat k dotazům](#)

Zobrazení: řadit dle: Výsledků:

Označ: vše stránka [do schránky](#) << < 1-10 z 716 >>

1 **Adiktologie** : časopis pro prevenci, léčbu a výzkum závislostí = Adictology
Tišnov: Sdružení SCAN, 2001-. sv.. [schránka](#): [přidat](#) [katalog](#) [detail](#) [Digitální archiv](#) [kopie](#)

2 **Editorial** : [adiktologie a její česká specifika]
[Radimecký, Josef, 1963-](#) ⁱ
Adiktologie. 2005, roč. 5, č. 3, s. 365-367. ISSN: 1213-3841.
[schránka](#): [přidat](#) [detail](#) [Digitální archiv](#): [ročník](#)

3 **Centrum adiktologie (CA)** : Sborník z XII. mezinárodní konference Společnosti pro návykové nemoci ČLS J.E. Purkyně a 45. mezinárodní konference AT sekce Psychiatrické společnosti ČLS J.E. Purkyně "Bio - psycho - socio - spirituální model závislosti", Měřín, 21.-25. května 2006. Abstrakt
[Šucha, Matúš, 1980-](#) ⁱ
Adiktologie. 2006, roč. 6, S2, s. 286. ISSN: 1213-3841.
[schránka](#): [přidat](#) [detail](#) [Digitální archiv](#): [číslo](#)

Omezit výsledky

Databáze

bibliomedica	692
katalogy Medvik	24

[filtrovat](#)

Vlastník

Národní lékařská knihovna	715
Oddělení vědeckých informací při MZČR	6

[filtrovat](#)

Typ dokumentu

články	444
abstrakta	248
knihy	23
monografie	22
recenze	9
časopisy / seriály	2

[filtrovat](#)

Autoři [seřadit](#)

Miovský, Michal, 1975-	48
Štastná, Lenka	26
Mravčík, Viktor	20
Kalina, Kamil, 1945-	19
Radimecký, Josef, 1963-	17
7áhranský, Tomáš	

Obr. 1 Ukázka zobrazení výsledků hledání

Náhled záznamu konkrétního titulu obsahuje persistentní odkaz na zobrazení záznamu a uživatel má pomocí AddThis pluginu možnost sdílet záznamy v sociálních sítích a odkazových službách. Zobrazení dostupných exemplářů pro objednání výpůjčky nebo rezervace bude zapracováno v další etapě vývoje portálu, aktuálně je pouze zobrazen odkaz na katalog ve staré verzi portálu. Stejně tak se připravuje zobrazení článků u záznamů zdrojových časopisů a sborníků. Výsledky lze snadno ukládat do schránky a exportovat záznamy ve zvolených formátech nebo pouze vybrané údaje z těchto záznamů.

Další funkce budou postupně doplňovány a portál dostane ještě novou grafickou podobu. Je pravděpodobné, že s novým designem dojde k optimalizaci rozložení jednotlivých částí portálu a zpřehlednění navigace. Připravuje se možnost tvorby uživatelských kolekcí - zařazování výsledků ve schránce do uživatelem definovaných skupin a jejich sdílení, autentifikace uživatele a možnost uložení některých nastavení, exporty v citačních formátech a podpora pro Zotero, zobrazení obálek publikací, speciální kolekce dokumentů, integrace dat z Portálu e-zdrojů, napojení na Google Books atd.

Využití tezauru Medical Subject Headings

NLK překládá do češtiny a zpřístupňuje americký tezaurus Medical Subject Headings (MeSH) pro indexaci bibliografických záznamů [1]. Většina záznamů v bázích Medvik je indexována s využitím hesel (deskriptorů) a podhesel (kvalifikátorů) tohoto tezauru. Význam používání hesel MeSH pro vyhledávání informací je dobře znám [2]. Uživatelé však často o existenci tohoto nástroje z různých důvodů nevědí nebo ho neumí při vyhledávání efektivně využívat. Proto byl při vývoji kladen důraz na zviditelnění a uživatelsky přívětivé využití tezauru MeSH při vyhledávání.

V novém portálu je MeSH prezentován více způsoby. Na úvodní stránce portálu má uživatel možnost procházet stromovou strukturou tezauru a následně si zobrazit bibliografické záznamy indexované příslušným heslem včetně použitých podhesel. Může také procházet celou databází dokumentů pomocí oblaku hesel MeSH (tzv. tag cloud), prohlížet a dále filtrovat seznamy souvisejících výsledků.

Po odeslání dotazu se pod seznamem výsledků zobrazuje dynamicky generovaný oblak hesel, který vždy odpovídá zadanému dotazu a který se aktualizuje při každé další úpravě dotazu. Implementace "tag cloud" byla inspirována řešením LigerCat [3]. Samotné zobrazení oblaku bylo vylepšeno o sémantické rozdělení jednotlivých hesel do souvisejících kategorií podle 16 hlavních větví stromové struktury MeSH, podobné řešení sémantického řazení tagů je navrhováno zde [4]. Velikost písma jednotlivých hesel je založena na statistickém výskytu daného hesla v aktuálním výsledku - množině bibliografických záznamů a je relativní pro každou kategorii.

Portál pracuje jak s českou tak anglickou verzí tezauru a je možno při vyhledávání používat oba jazyky a libovolně kdykoli zobrazení přepínat. Po kliknutí na konkrétní heslo MeSH je termín přidán do dotazu a dojde k aktualizaci seznamu výsledků a všech souvisejících klastrů fasetové navigace. V případě, že uživatel zadá obecný dotaz, pro který je plnotextovým vyhledáváním nalezeno málo výsledků, je mu nabídnuto použití konkrétních hesel MeSH s automatickým rozšířením dotazu. Uživatel se tak dostane k relevantním výsledkům pouhým klikáním, aniž by se musel zabývat identifikací a zápisem konkrétních hesel MeSH do vyhledávacího formuláře. Ukázka zobrazení oblaku MeSH hesel je uvedena na Obr. 2.

C - Nemoci

kardiovaskulární nemoci **diabetes mellitus** metabolismus - nemoci diabetes - komplikace diabetes mellitus 1. typu **diabetes mellitus 2. typu** hypertenze hypertyreoidismus hypotyroidismus **metabolický syndrom X** tělesná hmotnost zánět hladovění hyperaldosteronismus hyperglykémie hyperinsulinismus infarkt myokardu koronární nemoc ledviny - nádory **obezita** syndrom polycystických ovarií tělesná hmotnost - změny ateroskleróza hmotnostní přírůstek hmotnostní úbytek hyperlipidémie infertilita ženská **insulin - rezistence** hypercholesterolémie hyperlipoproteinémie

D - Chemikálie a léčiva

hormony laktony hormony ektopické ketolátky sulfonylmočovinné sloučeniny thyreoidální hormony triglyceridy angiotensinogen glukosa hypoglykemika látky ovlivňující centrální nervový systém leptin thyroxin antihypertenziva dehydroepiandrosteron krevní glukosa **látky proti obezitě** metformin psychotropní léky anorektika antipsychotika cyklobutany dehydroepiandrosteronsulfát hydrokortison insulin lipasa

E - Analytické, diagnostické a terapeutické techniky a přístroje

bariatrická chirurgie diagnóza diferenciální farmakoterapie žaludeční balónek cvičení a pohyb - techniky dietoterapie dlouhodobá péče gastroplastika analýza shluková dieta redukční epidemiologické studie faktorová analýza statistická **index tělesné hmotnosti** morbidita mortalita tělesná hmotnost prevalence riziko - zhodnocení rizikové faktory **obezita**

Centrální vyhledávání v online zdrojích - služba MedGate

Pro řadu uživatelů je problémem orientace v nabídce elektronických informačních zdrojů poskytovaných knihovnou a zvládnutí postupů práce v různých uživatelských rozhraních těchto zdrojů. Částečným řešením tohoto problému jsou systémy federativního vyhledávání, které umožňují přes jednotné rozhraní paralelně vyhledávat v těchto informačních zdrojích a dále pracovat s nalezenými výsledky. Jednotné rozhraní zároveň slouží uživatelům jako rozcestník pro orientaci v poskytovaných zdrojích. Uživatelé tak mohou snadno identifikovat konkrétní databáze, které obsahují informace relevantní k jejich požadavkům a případně mohou dále pokračovat ve vyhledávání již v rozhraní konkrétní databáze. V případě plnotextových zdrojů mohou z jednoho místa snadno zjistit, kde se nachází plný text požadovaného článku a rovnou se k plnému textu dostat bez nutnosti procházet jednotlivé databáze. NLK proto implementovala další online službu SerialsSolutions zvanou *360 Search*, která doplňuje stávající službu Portálu elektronických zdrojů NLK [5].

Nová služba se jmenuje *MedGate - centrální vyhledávání v e-zdrojích NLK* a je dostupná zde [6]. Protože však tato služba prohledává i licencované zdroje, je přístupná pouze pro registrované čtenáře NLK a to buď lokálně z počítačové sítě NLK nebo vzdáleně. Vzdálený přístup ke službě MedGate je možný přes URL adresu: <http://web.medvik.cz/han/medgate>. Aktuálně je prohledávatelných cca 40 licencovaných i volných online zdrojů. Dostupné zdroje jsou pro lepší orientaci rozdělené podle zaměření obsahu databází do skupin, pro vyhledávání plných textů je k dispozici samostatná skupina Plné texty. Uživatel má možnost před započítím samotného hledání vybrat buď konkrétní skupiny databází nebo konkrétní zdroje, případně může vyhledávat ve všech zapojených zdrojích.

MedGate nabízí dva způsoby hledání - základní a rozšířené, které umožňuje kombinovat více podmínek hledání pomocí logických operátorů. Podmínky hledání jsou následující: název, autor, plnotextové hledání, klíčové slovo, předmět, abstrakt, ISSN, ISBN a libovolné (ve všech polích).

Výsledky vyhledávání se zobrazují průběžně, jsou řazeny podle data publikování, lze je však řadit i podle názvu, autora či zdroje záznamu. Výsledky je dále možno omezit pouze na ty, u nichž je dostupný fulltext, či na peer-reviewed zdroje pomocí filtru. Výsledky jsou řazeny do skupin - klastrů podle tématu, data vydání, zdrojového časopisu a autora, je možno také procházet výsledky po jednotlivých zdrojích. Výsledky jsou automaticky deduplikovány v případech, kdy se obsahy zdrojů překrývají. Uživatelé mají možnost zobrazit abstrakt, přejít na dostupný plný text a exportovat (nebo si zaslat e-mailem) vybrané výsledky v různých citačních formátech (EndNote, ProCite). Ukázka zobrazení výsledků viz Obr. 3.



Upřesnit hledání | Nové hledání

Nápověda

Filtrovat dle:

plný text

peer-reviewed

Řadit výsledky dle: Datum | Název | Autor | Zdroj

1 až 25 výsledků z 394 dostupných (včetně 19 zdvojených) pro "klíčové slovo obsahuje atopic and dermatitis and adults" (celkem 36772)

1 - 2 - 3 - 4 - 5 - 6 - 7 - 8 - 9 - 10 >

Témata Datum Časopisy Autoři

Výsledky ve skupinách

- Veškerá témata**
- [-] Tacrolimus, Ointment (39)
 - [-] Allergic rhinitis (28)
 - [-] Atopic Eczema (20)
 - [-] Prevalence, Allergic (19)
 - [-] Eosinophil (18)
 - [-] Contact dermatitis (17)
 - [-] Infection (15)
 - [-] Atopic dermatitis patients (12)
 - [-] Onset (11)
 - [-] Patch test (11)
- Další

Výsledky dle zdroje

- [-] American Association for the Advancement of Science (4)
 - [-] Science Magazine (4)
- [-] Bentham Science Publishers (3)
 - [-] Bentham Science Publishers OA (3)
- [-] Bioline International (1)
 - [-] Bioline International (1)
- [-] BioMedCentral (179)

Vybrat veškeré záznamy: 3 vybraných záznamů: E-mail ▼ nebo Exportovat/Uložit ▲

Exportovat vybrané záznamy

Přímě do RefWorks

Ve formátu kompatibilním s EndNote, ProCite atp.

Jako HTML

Jako prostý text

- 1.69th Annual Meeting of the American Academy of Dermatology (1175-0561)
Fraser, Kathy.
2011-06-01. Vol.12,Vyd.3;str.215
Zdroj: Health & Medical Complete (ProQuest)
Zobrazit ve skupinách
- 2.Application of biological monitoring for exposure assessment following chemical incidents: A procedure for decision making (1559-0631)
Scheepers, Paul T J; Bos, Peter M J; Konings, Joke; Janssen, Nicole A H; Grievink, Linda.
2011-05-01. Vol.21,Vyd.3;str.247
Zdroj: Health & Medical Complete (ProQuest)
Zobrazit abstrakt | Zobrazit ve skupinách

Obr. 3 Ukázka zobrazení výsledků MedGate pro dotaz: *atopic dermatitis adults* (klíčové slovo)

Závěr

Domníváme se, že implementace nových vyhledávacích služeb, které odpovídají soudobým trendům, a které zároveň poskytují možnosti snadného a efektivního využití tradičních rešeršních nástrojů - tezaurů, kombinovaného hledání a souborů autorit, je správnou cestou pro zpřístupnění informačních zdrojů NLK a navazujících služeb. Nové rozhraní portálu Medvik tak může dobře sloužit jak uživatelům, kteří neovládají rešeršní postupy tak uživatelům pokročilým s jasnou vyhledávací strategií. Nabídka funkcí a služeb portálu Medvik se bude i nadále rozvíjet v reakci na zpětnou vazbu uživatelů. Plánujeme též integraci portálu Medvik do centrálního vyhledávání. Po dokončení integrace výpůjčních funkcí do nového rozhraní portálu Medvik bude stará verze portálu odstavena z provozu. Na uvedení nové verze portálu navazují školící aktivity pro uživatele, ze jejichž průběhu získáváme zpětnou vazbu. První ohlasy od uživatelů nové verze portálu jsou pozitivní. Věříme, že vyhledávač MedGate a nové rozhraní portálu Medvik zvýší uživatelský komfort v přístupu k poskytovaným informačním zdrojům a službám NLK.

Odkazy a literatura

1. Národní lékařská knihovna. Tezaurus Medical Subject Headings [online]. Praha : NLK, 2011 [cit. 2011-02-10]. Dostupné z WWW: <http://www.nlk.cz/informace-o-nlk/odborne-cinnosti/tezaurus-medical-subject-headings>
2. Chang AA, Heskett KM, Davidson TM. Searching the literature using Medical Subject Headings versus text word with PubMed. *Laryngoscope*. 2006;116:336–240

3. Sarkar IN, Schenk R, Miller H, Norton CN. LigerCat: using "MeSH Clouds" from journal, article, or gene citations to facilitate the identification of relevant biomedical literature. AMIA Annu Symp Proc. 2009 Nov 14;2009:563-7 [cit. 2011-02-10]. Dostupné z WWW: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2815376>

4. Hassan-Montero Y, Herrero-Solana V. Improving tag-clouds as visual information retrieval interfaces. International Conference on Multidisciplinary Information Sciences and Technologies, InSciT2006. Merida, Spain. October 25-28, 2006 [cit. 2011-02-10]. Dostupné z WWW: http://www.nosolousabilidad.com/hassan/improving_tagclouds.pdf

5. Národní lékařská knihovna. Portál elektronických zdrojů [online]. Dostupné z WWW: <http://rx4jt8nl5k.search.serialssolutions.com>

6. Národní lékařská knihovna. MedGate: centrální vyhledávání v e-zdrojích NLK [online]. Dostupné z WWW: <http://rx4jt8nl5k.cs.serialssolutions.com> [pouze pro registrované čtenáře NLK]