

Nadstavbové služby portálu Medvik pro podporu evaluace vědy a výzkumu ve zdravotnictví

**Filip Kříž, Ondřej Horský, Lenka Maixnerová, Adéla Jarolímková,
Eva Lesenková, Helena Bouzková¹, Petr Lesný²**

¹ Národní lékařská knihovna, Praha

² Univerzita Karlova, 2. lékařská fakulta, Praha

kriz@nlk.cz

INFORUM 2013: 19. ročník konference o profesionálních informačních zdrojích
Praha, 21. - 22. 5. 2013

Abstrakt

Portál Medvik zpřístupňuje prostřednictvím centrálního indexu katalogy knihoven integrovaných v systému Medvik, bázi Bibliographia medica Českoslovacca (BMČ) a soubory personálních a věcných autorit (tezaurus MeSH) z biomedicínských oborů. BMČ je bibliografická, částečně plnotextová databáze, která je vytvářena v Národní lékařské knihovně (NLK). Architektura portálu Medvik je maximálně otevřená pro roboty vyhledávacích služeb, což dokládají vysoké statistiky přístupů a zesílená zpětná vazba od autorů článků indexovaných v BMČ v posledních 2 letech. NLK rozšířila nabídku služeb a nástrojů poskytovaných portálem Medvik za podpory projektu Interní grantové agentury (IGA) Ministerstva zdravotnictví ČR. Součástí řešení projektu je nástroj pro automatizované získávání a import záznamů z PubMed do BMČ. V portálu Medvik lze tedy využívat identifikátor PubMed (PMID) pro zobrazování počtu citací článku prostřednictvím API systémů Web of Science a Scopus. Současně lze propojovat záznamy časopisů a článků s grafem vývoje Impakt faktoru v bázi Journal Citation Reports a zobrazovat informace o možnostech archivace plných textů ze systému SHERPA/RoMEO. Záznamy báze BMČ byly doplněny o identifikátory projektů IGA, které umožňují vyhledávání publikačních výsledků a propojení na Centrální evidenci projektů (CEP). Součástí řešení je vývoj funkcionality pro auto-archivaci pre- a post-printů odborných článků v Digitálním archivu NLK s možnostmi stanovení úrovně zpřístupnění a nastavení veřejné licence Creative Commons. Tato funkcionality by měla autorovi poskytnout také možnosti opravy a případně obohacení ztotožněného záznamu osobní autority o identifikátory vědce (tzv. researcher identifiers) s následným propojením na příslušný registr.

Úvod

Vývoj portálu Medvik v posledních dvou letech přinesl nový přehlednější design, integraci zobrazování údajů o online dostupnosti časopisů z Portálu e-zdrojů, obohacení zobrazení záznamů o obálky a obsahy a také propojení na Google Books. K dispozici uživatelům je export záznamů v citačním formátu RIS pro import do citačních manažerů Zotero, EndNote atp. Bibliografické a autoritní záznamy v portálu jsou propojené množstvím hypertextových odkazů a jsou tak přístupné pro roboty vyhledávacích služeb. Díky možnostem procházení databáze prostřednictvím oblaků hesel MeSH se vyhledávače dostanou k záznamům, které jinak obvykle zůstávají skryté za vyhledávacími formuláři. V roce 2012 měl portál Medvik (www.medvik.cz/bmc) podle údajů nástroje Google Analytics 241 tisíc návštěv a 160 tisíc unikátních návštěvníků. V dnešní době je však nezbytné, aby portál poskytoval uživatelům nejen bibliografické informace, ale hlavně rovnou plné texty. NLK proto

buduje oborovou Digitální knihovnu (DK) v prostředí open source systému Kramerius. Obsah DK je získáván pomocí programů VISK, vlastní digitalizací, a také ve spolupráci s autory a vydavateli českých odborných časopisů a konferenčních sborníků, kteří na základě uzavřené licenční smlouvy s NLK poskytují plné texty. Aktuálně je licenčně ošetřeno 85 časopisů, jejichž plné texty jsou volně přístupné v DK (u některých titulů je embargo v řádu měsíců). BMČ se tedy stává plnotextovým zdrojem - celkem bylo dosud archivováno cca 45 000 článků, více informací o zpracování a pokrytí BMČ je uvedeno v [1]. V roce 2011 NLK získala grantovou podporu Interní grantové agentury Ministerstva zdravotnictví pro 3-letý projekt *Archivace a online zpřístupnění plných textů vědeckých publikací pro podporu rozvoje vědy a vzdělávání lékařských a nelékařských oborů* (NT12345 - dále jen 'projekt NLK'). Projekt NLK řeší nejen oblast získávání digitálních dokumentů - vlastní digitalizaci sbírky závěrečných grantových zpráv IGA, ale také jejich efektivní zpracování a legální zpřístupnění včetně poskytování potřebných souvisejících informací. Ukázka propojení záznamu článku BMČ s plným textem v DK je uveden na Obr. 1a.

Využití externích identifikátorů v záznamech Medvik

Závěrečná grantová zpráva se typicky skládá z vlastního textu zprávy a jejích příloh - kopií publikačních výstupů daného projektu. V rámci projektu NLK jsou tyto přílohy analyticky zpracovávány a propojovány se záznamy článků BMČ (aktuálně je propojeno 3 360 příloh). Do záznamu vlastní zprávy i do všech záznamů identifikovaných příloh je doplňováno oficiální číslo projektu podle Centrální evidence projektů (CEP). Závěrečná zpráva a všechny přílohy jsou tak vyhledatelné v portálu Medvik podle čísla projektu IGA a jsou propojené odkazem na stránku projektu v CEP (Obr. 1b). Pro potřeby pokrytí článků českých autorů ze zahraničních časopisů (bohemik) byl vytvořen v portálu NCBI rešeršní profil pro export XML záznamů z PubMed, takto získané záznamy jsou offline konvertovány na formát Marc21 včetně převodu hesel MeSH na českou verzi a importovány do báze BMČ. Aktuálně přes 7 000 záznamů článků obsahuje identifikátor PubMed ID (PMID), podle kterého jsou v portálu Medvik vyhledatelné a také jsou propojené na původní záznam PubMed (Obr. 1b). Zároveň je PMID používán pro získávání údajů o citovanosti prostřednictvím aplikačních rozhraní *Links Article Match Retrieval Service* (Web of Science, Journal Citation Reports) a *Scopus Search JavaScript API*. U konkrétního záznamu se tak veřejně zobrazuje odkaz na WoS nebo Scopus s uvedeným počtem citací v daném rejstříku a také odkaz na graf vývoje Impakt faktoru v JCR (Obr. 1c). V aktuální verzi konvertoru je importován také identifikátor DOI, pokud je v záznamu přítomen. U starších záznamů bohemik v BMČ je plánováno ztotožnění s PubMed záznamy pomocí NCBI E-utilities a doplnění záznamů BMČ o PMID a DOI.

Auto-archivace

Dalším cílem projektu NLK je zprovoznit možnost auto-archivace born-digital dokumentů samotnými autory v DK s možností volně zpřístupnit plný text pod licencemi Creative Commons. NLK tímto způsobem podporuje tzv. Zelenou cestu otevřeného přístupu (Green OA Self Archiving). Testování systému Kramerius 4 prokázalo, že je možné v tomto systému archivovat samostatné články ve formátech PDF. Vývoj auto-archivační funkcionality probíhá ve spolupráci se dvěma uživatelskými skupinami (vědci a knihovníci). Ze semináře uspořádaného pro tyto skupiny vyplynulo, že autoři většinou nemají povědomí o tom, jak nakládat se svojí kopií článku. Nástroj bude tedy muset být návodný a postup archivace co nejjednodušší. Musí být jasně uvedeno, za jakých podmínek a zda vůbec lze archivovat verzi plného textu, kterou má autor k dispozici. Problematika postupu při rozhodování autora je příkladně popsána v publikaci [2, kap. III]. Pomocníkem pro autora může být

informace o zdrojovém časopisu v systému SHERPA/RoMEO [3], nicméně vždy bude záležet na podmínkách uvedených v konkrétní smlouvě mezi vydavatelem a autorem. V portálu Medvik bylo zprovozněno zobrazování indikace možností archivace prostřednictvím SHERPA/RoMEO API rozhraní, a to jak u záznamu časopisu, tak u jednotlivých článků (Obr. 1d). Statistiky RoMEO [4] ukazují, že 69% vydavatelů umožňuje nějakou formu auto-archivace, nicméně v systému je zatím aktuálně zastoupeno pouze 7 českých vydavatelů (9 časopisů). Důležitý je také souhlas spoluautorů a zaměstnavatele v případě, že se jedná o zaměstnanecké dílo. V případech, kdy nebudou jasné podmínky auto-archivace (nebo bude přímo zakázána), je zvažována možnost archivace plného textu ale se zpřístupněním pouze formou zaslání autorské kopie, které bude vyžadovat povolení autora (tzv. request button).

Záznam dokumentu		Zdroj: bibliomedica
Název	TVT and TVT-0 for surgical treatment of primary stress urinary incontinence: prospective randomized trial / L. Krofta, J. Feyereisl, M. Otčenásek, P. Velebil, E. Kasíková, M. Krcmár	
Autor	Krofta, Ladislav <i>Institute for the Care of Mother and Child, Podolske nabrezi 157, 147 00, Prague 4, Czech Republic. ladislav.krofta@post.cz</i> Feyereisl, Jaroslav , 1954- Otčenásek, Michal Velebil, Petr , 1956- Kasíková, Eva Krcmár, Michal	
Zdroj	International urogynecology journal . 2010, roč. 21, č. 2, s. 141-148. ISSN: 0937-3462; 1433-3023 (elektronická verze).	
Datum vydání	2010	
Jazyk	angličtina	
Typ dokumentu	články	
číslo IGA	NR9309 (MZO)	b
Odkaz	Pubmed	
Signatura	NLK: sig ; neváz. sign	
Číslo záznamu	bmc12025472	
Persistentní link	http://www.medvik.cz/link/bmc12025472	
MESH deskriptor	stař ženské pohlaví lidé lidé středního věku prospektivní studie stresová inkontinence moči (chirurgie) výsledek terapie suburetrální pásky (škodlivé účinky)	
Publikační typ	srovnávací studie výzkum podporovaný jinými zdroji než vládou USA randomizované kontrolované zkoušky	

Schránka
 schránka: přidat +
Další akce
 kopie

Digitální knihovna
 plný text a
 zdroj i

Online přístup
 Health & Medical Complete (ProQuest)
 od 1997-01-01 do Před 1 rokem
 Hospital Collection (ProQuest)
 od 1997-01-01 do Před 1 rokem

Zobrazení záznamu
 standardní
 citace
 MARC
 Bibliomedica
 export EndNote, Zotero...

Citovanost
 Scopus : 14 c
 (přímý přístup)
 Web of Science @ : 10 i
 (přímý přístup)
 Impact Factor JCR @ i

Možnosti archivace d
 Sherpa/RoMEO
 (více informací)

Obr. 1 Ukázka náhledu záznamu BMC

Problém pro DK jako oborový repozitář představuje postup autorizace (ztotožnění) autora se záznamem autority, pro autory schůdný postup je v řešení. Součástí řešení auto-archivace bude systém podpory spolupracujících autorů, umožňující zpětnou vazbu pro případné doplňování a opravy bibliografických a autoritních záznamů; stejně tak možnost doplňování identifikátorů (DOI, jiné grantové identifikátory) a statistiky přístupů k archivovaným dokumentům. Pro autory je také zajímavá možnost uložit i článek/abstrakt ze sborníku, prezentaci příspěvku (jako PDF) nebo poster.

Připravuje se možnost zprovoznění rozhraní OAI-PMH pro dokumenty uložené v DK pod licencemi Creative Commons a registrace DK v registrech archivů s otevřeným přístupem OpenDOAR a ROAR, které umožní sklizení a šíření volně dostupných publikačních výstupů. Proveditelné se jeví i poskytnutí volně šiřitelného obsahu DK do systému OpenAIRE (Open Access Infrastructure in Europe) [5], který zatím archivuje pouze publikační výstupy projektů financovaných EU nebo jinou mezinárodní nebo národní agenturou, avšak časem umožní archivaci i jiných publikačních výstupů.

Profil autora - vědce

Pro autory vědeckých prací je v současnosti důležité budovat svou vědeckou identitu/profil v prostředí WWW. Pro autorizované uživatele - autory zpřístupňující publikace v DK - je v plánu možnost rozšíření zobrazení autoritního záznamu o další osobní identifikátory a prolinkování na příslušný registr, například ResearcherID (Thomson Reuters), Scopus Author Identifier nebo ORCID [6]. ORCID (Open Researcher and Contributor ID) je v roce 2012 spuštěný, otevřený a nekomerční systém, jehož identifikátor je založený na International Standard Name Identifier (ISNI). Podstatné je, že se jedná o nezávislou platformu, která však dokáže spolupracovat se vzájemně konkurenčními systémy WoS i Scopus. V případě, že je autor registrovaný ve Scopus i WoS, může sloučit pod svým ORCID profilem záznamy publikací z obou těchto platforem na jenom místě. První analýza API rozhraní ORCID ukázala, že bude možné vytvořit nástroj pro import záznamů publikací z portálu Medvik pro autora registrovaného v ORCID.

Závěr

Trvalé uchování a zpřístupnění hodnotných oborových dokumentů a materiálů v archivačním systému umožní trvalé zpřístupnění informací širokému spektru uživatelů vědeckých informací. Integrace Digitální knihovny s portálem Medvik poskytne uživatelům unikátní důvěryhodný zdroj. Systém persistentního propojení plných textů s článkovými metadaty BMČ umožní překonat problémy nefunkčních odkazů na externí informační zdroje, jednoznačně identifikovat autory a publikované dokumenty s možností budoucího využití identifikátorů např. v systémech hodnocení kvality. Veřejné zpřístupnění vědeckých publikací zvýší vyhledatelnost a dostupnost publikací v národním i mezinárodním kontextu a ve výsledku přinese zviditelnění výsledků českého výzkumu a vývoje.

Příspěvek byl podpořen projektem **IGA MZ ČR NT12345**.

Odkazy a literatura

1. Databáze bibliographia medica czechoslovaca jako zdroj informací o publikačních výstupech vědy a výzkumu ve zdravotnictví. In: MEDSOFT 2013: sborník příspěvků. Praha : Creative Connections, 2013. 2013, s. 131-139. ISSN 1803-8115. Dostupné z: http://creativeconnections.cz/medsoft/2013/Medsoft_2013_Maixnerova.pdf
2. MYŠKA, Matěj. Veřejné licence v České republice. V Tribunu EU vyd. 1. Brno: Tribun EU, 2012, 130 s. ISBN 978-80-263-0343-5. Dostupné z: <http://is.muni.cz/www/102870/Prirucka.pdf>
3. RYGELOVÁ, P. Služba SHERPA/RoMEO - další život vědeckého článku. Zpravodaj ÚVT MU [online]. 2009, XX, č. 2, s. 12-17. ISSN 1212-0901. Dostupné z: <http://www.ics.muni.cz/bulletin/articles/630.html>
4. SHERPA/RoMEO: RoMEO Statistics [online]. University of Nottingham, © 2006-2011. Dostupné z: <http://www.sherpa.ac.uk/romeo/statistics.php>

5. OpenAIRE [online]. The OpenAIRE Consortium. Dostupné z: <http://www.openaire.eu/cs/home>
6. ORCID [online]. Bethesda, MD: ORCID, Inc. Dostupné z: <http://orcid.org/about>