

ROZVOJ DIGITÁLNÍHO ARCHIVU NÁRODNÍ LÉKAŘSKÉ KNIHOVNY VE SPOLUPRÁCI S VYDAVATELI A AUTORY VĚDECKÝCH LÉKAŘSKÝCH INFORMACÍ

*Lenka Maixnerová, Filip Kříž, Ondřej Horsák,
Helena Bouzková*

Anotace

Národní lékařská knihovna (NLK) zprovoznila ve 4. čtvrtletí 2008 digitální archiv v Open Source systému Kramerius. Archiv primárně slouží pro ukládání digitálních dokumentů z fondu NLK digitalizovaných při řešení programu MK ČR VISK 7. Obsah archivu však není omezen pouze na dokumenty vzniklé procesem digitalizace, ale je možno ho využít i pro ukládání dokumentů vytvořených již v digitální podobě. Pro uživatele vědeckých informací jsou významné zejména odborné elektronické časopisy, jejichž trvalé a dlouhodobé uchování a zpřístupnění se stává stále naléhavějším úkolem. Důležitým zdrojem pro možné uložení a zpřístupnění v archivu NLK jsou další výstupy vědecké a vzdělávací činnosti – prezentace, sborníky z konferencí, postery atd. Tento příspěvek pojednává o možných způsobech spolupráce s vydavateli i autory vědeckých lékařských informací v oblasti trvalého uchování a zpřístupnění odborných zdravotnických informací v digitálním archivu NLK.

Klíčová slova

Informační služby, digitální archivy, Národní lékařská knihovna, uchování dokumentů, Creative Commons, digitalizace, zdravotnické knihovny, elektronické dokumenty

1. Úvod

Získávání, zpracování, trvalé uchování a zpřístupnění fondů a sbírek tradičních dokumentů je v prostředí knihoven relativně dobře a kvalitně zajištěno. Celosvětovým problémem je hrozba ztráty tištěných dokumentů způsobená degradací kyselého papíru, který se používal téměř 150 let a poškozením dokumentů častým používáním. Efektivním prostředkem pro uchování ohrožených publikací je jejich digitalizace a následné zpřístupnění v digitálních knihovnách, archivech nebo repositářích.

V souvislosti s rozvojem informačních a komunikačních technologií jsou knihovny nuceny se vypořádat se získáním, zpracováním, uložením a zpřístupněním nových typů dokumentů, které vznikly již elektronicky. Především se řeší problémy dlouhodobého uchování a zpřístupnění těchto digitálních dokumentů.

Potřeba a nutnost dlouhodobě uchovávat a zpřístupňovat kulturní dědictví a vědecké informace v tradičních i elektronických dokumentech je jasně formulována v iniciativě Komise evropských společenství *i2010: Digital Libraries Initiative*. Pro knihovny v České republice je určena *Koncepce trvalého uchování knihovných sbírek tradičních a elektronických dokumentů v knihovnách ČR* [1], která předpokládá vytváření oborových digitálních knihoven a archivů s návazností na systém Národní digitální knihovny [2].

Národní lékařská knihovna (NLK) zpracovala projekt na vybudování digitálního archivu (DA), jehož cílem je trvalé uchovávání a zpřístupnění tradičních i elektronických odborných zdravotnických informací. NLK zahájila spolupráci s českými nakladateli odborných biomedicínských periodik, která se soustřeďuje na možnosti získávání a zpřístupnění elektronických dokumentů v digitálním archivu. Neméně důležité je získávat obsah archivu přímo od autorů vědeckých informací a motivovat je k využívání DA pro ukládání a zpřístupňování výzkumných zpráv, konferenčních materiálů, prezentací, preprintů, výukových materiálů a dalších materiálů s odborným obsahem.

2. Kramerius

NLK zprovoznila na konci roku 2008 DA. Pro DA byl vybrán český systém Kramerius [3]. Kramerius je software s otevřeným zdrojovým kódem produkovaný firmou Qbizm Technologies, na jehož vývoji se podílí Národní knihovna ČR (NK ČR), Knihovna Akademie věd ČR a Moravská zemská knihovna v Brně. Vývoj systému je zajišťován finančními dotacemi Ministerstva kultury ČR a Ministerstva školství, mládeže a tělovýchovy ČR. Systém Kramerius slouží k zpřístupňování digitálních dokumentů v souladu s autorským zákonem. V současnosti jsou k dispozici DTD (dokument type definition = definice typu dokumentu) pro periodika a monografie, do budoucna se plánuje rozšíření i pro ostatní typy dokumentů. Podporuje formáty XML, TXT, DjVu, JPG, PNG, PDF. Systém používá fulltextové vyhledávání Convera a Lucene a umožňuje replikaci dat mezi jednotlivými instalacemi.

DA NLK je pro uživatele přístupný na adrese <http://www.medvik.cz/kramerius>. Archiv aktuálně obsahuje kompletní vydání 5 časopisů vytištěných na degradujícím kyselém papíře. Špatný fyzický stav těchto dokumentů již neumožňoval jejich další využívání. Jedná se o následující tituly:

1. Věstník Masarykovy ligy proti tuberkulóze, roky 1920–1929
2. Zdraví: list pro otázky zdravotní péče veřejné i soukromé, roky 1935–1940
3. Boj proti tuberkulóze, roky 1930–1950

4. Zdraví lidu, roky 1923–1940
5. Zprávy Československého Červeného kříže, roky 1920–1936

Všech 5 titulů podléhá ochranné lhůtě autorského zákona, proto je přístup možný pouze v prostorách NLK, ostatním uživatelům jsou přístupná pouze metadata (popisné údaje). Vlastní digitalizace byla částečně hrazena z dotace programu Ministerstva kultury VISK 7. Informaci o tom, že určitý dokument má elektronickou verzi v digitálním archivu NLK, nalezne uživatel v bibliografickém záznamu v katalogu systému Medvik s příslušným odkazem do systému Kramerius. Z digitalizovaných dokumentů je možné objednávat kopie v PDF formátu. Obrazové soubory mají uživatelé k dispozici ve formátu DjVu, pro jejich prohlížení je nutné nainstalovat plug-in DjVu od firmy Lizardtech (zdarma). Obsah archivu lze dále rozšířit o digitální kopie dokumentů z jiných instalací systému Kramerius za podmínky, že NLK tyto dokumenty vlastní nebo v minulosti vlastnila (ztráta, zničení apod.). Tímto způsobem bylo v únoru 2009 replikováno 17 monografií z NK ČR. Tyto monografie již nepodléhají ochranné lhůtě autorského zákona, jejich zpřístupnění v DA je tedy bez jakéhokoliv omezení. V letošním roce se NLK uchází o další finanční podporu z programu MK ČR VISK 7, a to na digitalizaci Zdravotnických novin z let 1952–2003.

3. Spolupráce s nakladateli odborných zdravotnických periodik

Obsah archivu není omezen pouze na dokumenty vzniklé procesem digitalizace, ale je možné ho využít i pro ukládání dokumentů vytvořených již v digitální podobě. V současné době má většina odborných tištěných časopisů i svou elektronickou verzi, některé již vychází pouze elektronicky. Trvalé a dlouhodobé uchování a zpřístupnění těchto elektronických verzí se stává stále důležitějším úkolem. NLK počátkem roku 2009 oslovila české nakladatele odborných biomedicínských časopisů, zda by měli zájem o službu dlouhodobé archivace a zpřístupnění elektronických verzí periodik na bezpečném úložišti NLK včetně následného propojení s bibliografickou databází Bibliographia medica Českoslovaca (BMČ). Z prvních poznatků je již zřejmé, že bude o tuto službu zájem, neboť nakladatelé primárně neřeší dlouhodobé a trvalé uchování elektronických verzí.

BMČ je národní registrující bibliografie, která v odvětví lékařství (biomedicíny) a zdravotnictví od roku 1947 zachycuje veškeré odborné písemnictví vydávané na území České republiky a práce českých autorů vydané v zahraničí. Do roku 1996 byla její součástí i lékařská literatura vycházející na Slovensku. BMČ je on-line zpřístupněna v systému Medvik na adrese <http://www.medvik.cz/bmc> a je také vydávána na DVD Bibliomedica (distribuci zajišťuje firma AiP Beroun). V počítačové podobě jsou zpracovány roky 1953–1956, 1968 a 1973 až do současnosti, ostatní ročníky

jsou k dispozici v knižní podobě. NLK postupně pracuje na jejich převodu do elektronické podoby. V roce 2008 došlo v NLK ke změně zpracování BMČ. Pro vytváření a zpřístupňování bibliografických záznamů BMČ se začal využívat systém Medvik. Dřívější dávkové zpracování bylo nahrazeno on-line zpracováním, bibliografické záznamy článků se uživatelům zpřístupňují hned po uložení do databáze. Při vytváření bibliografických záznamů se využívají elektronické verze časopisů, kopírují se názvy, autoři, abstrakta, URL adresy plných textů. Podmínkou je, aby elektronická verze byla k dispozici do 1/2 roku od vydání tištěné verze. Kopírováním základních údajů do zápisového formuláře BMČ dochází k urychlení zápisu a zamezuje se vzniku překlepů. Cílem těchto aktivit je, aby se po nalezení článku v bázi BMČ mohl uživatel dostat přímo do plného textu. Nevýhodou tohoto způsobu zápisu je to, že ze strany nakladatelů může dojít ke změně URL adresy, a tím se stávají odkazy z bibliografického záznamu na plný text nefunkční. U některých časopisů není možné odkazovat na konkrétní článek, neboť jednotlivé články nemají vlastní URL adresu, je možné odkazovat pouze na WWW adresu časopisu. Pokud je časová prodleva delší než 1/2 roku, není možné se zápisem do BMČ tak dlouho čekat. Možným řešením je archivovat elektronické verze periodik na datovém úložišti NLK, čímž by bylo zabezpečeno funkční prolínání bibliografických záznamů BMČ s plnými texty, v případě zániku nakladatele by nedošlo ke ztrátě elektronických dat. Nakladatelé, kteří budou mít o tuto službu zájem, uzavřou s NLK smlouvu o poskytování elektronických zdrojů podle § 51 zákona č. 40/1964 Sb., občanského zákoníku. Po vzájemné dohodě budou plné texty dostupné z portálu Medvik všem nebo pouze registrovaným uživatelům NLK, může být též stanovena doba zveřejnění elektronické verze oproti tištěné dle požadavků nakladatele.

4. Spolupráce s autory odborných zdravotnických informací

Neméně důležitým zdrojem pro možné uložení a zpřístupnění v archivu jsou další výstupy vědecké a vzdělávací činnosti: závěrečné grantové zprávy IGA MZ ČR, prezentace, konferenční materiály, postery, výukové materiály atd. V těchto případech postačuje, když autoři uzavřou s NLK licenční smlouvu o podmínkách užití díla ve smyslu zákona o právu autorském 121/2000 Sb. NLK oslovila několik autorů, o jejichž závěrečné grantové zprávy je velký zájem, zdali by byli ochotni tuto smlouvu s NLK podepsat, případně dodat i jejich elektronickou verzi (tištěné verze NLK archivuje). Většina oslovených autorů souhlasila. Do budoucna NLK bude jednat o tom, aby již při odevzdání závěrečné zprávy Interní grantové agentury MZ ČR, měl uživatel možnost tuto smlouvu podepsat, a kromě tištěné verze byla NLK odevzdána i verze elektronická.

5. Datové úložiště

Vybudování DA neznamena pouze výběr a zprovoznění systému pro archivaci, ale zároveň je nutné zajistit dlouhodobou ochranu digitálních dat, což samotný systém neřeší. ISO standardem pro potřebu digitálních úložišť je referenční model Open Archival Information System (OAIS) [4]. Tento model nabízí kompletní funkční a informační popis digitálního archivu a formuluje závazky, které musí organizace naplňovat, aby fungovala jako archiv „kompatibilní s OAIS“. V OAIS modelu by digitální archiv měl splňovat tyto minimální požadavky:

- Autenticita
- Integrita
- Spolehlivost a dostupnost
- Možnost opakovaného použití

Model neřeší pouze ochranu digitálních dat, ale též ochranu ostatních souvisejících prvků jako jsou metadata, identifikátory. Aplikace OAIS standardu do DA NLK je jedna z podmínek, aby byl archiv důvěryhodný a jeho uvedení do provozu by se nemělo oddalovat.

V rámci budování DA bude NLK též řešit trvalou identifikaci digitálních objektů [5]. Digitální objekt umístěný v DA by měl být označen persistentním identifikátorem (PI), který by byl jedinečný. PI je stále stejný, bez ohledu na to, kde se právě digitální objekt nachází. Mezi nejrozšířenější patří v současné době: DOI, Handle, URN:NBN.

V oblasti autorských práv by v letošním roce v ČR mělo dojít k zahájení používání licence „Creative Commons“ (CC) [6], jejíž anglická verze byla již přeložena a nyní je ve stadiu připomínkování k českému textu. CC je americká nezisková organizace, jejímž cílem je legální využívání a sdílení autorských děl. CC vychází z toho, že existují lidé, kteří nechtějí využívat všechna práva k duševnímu vlastnictví, která jim zaručuje zákon. CC nabízí různá licenční schémata, držitelé autorských práv si mohou vybrat, jaká z autorských práv k dílu si chtějí ponechat a jakých se naopak vzdát ve prospěch veřejnosti. Používání CC v ČR je umožněno díky novele Autorského zákona č. 216/2006 Sb. Pokud bude český text CC schválen, předpokládá NLK jeho využívání v rámci digitálního archivu.

6. Závěr

Digitalizace, archivace a zpřístupnění digitálních dokumentů v NLK je plně v souladu s projektem Národní digitální knihovny, jehož řešitelem je NK ČR v rámci Národní strategie digitalizace. Cílem tohoto projektu je urychlení digitalizace a dlouhodobá ochrana digitálních objektů a přístupu k nim.

Budování digitálního úložiště a zpřístupňování digitálních dat není pro

NLK jednoduchý úkol. Je nutné zajistit nejenom finanční a personální zabezpečení provozu digitálního archivu, ale též zvolit metody ochrany, archivace a zpřístupňování dat, metadata, vzory licenčních modelů, výběr SW a HW, vypracovat pracovní postupy. Z prvních ohlasů uživatelů se však domníváme, že se jedná o správné a důležité rozhodnutí, které usnadní komunikaci vědeckých informací.

Reference

- [1] Koncepce trvalého uchování knihovních sbírek tradičních a elektronických dokumentů v knihovnách ČR do roku 2010. Knihovna plus [online]. 2006, č. 1. Dostupný z [www: http://knihovna.nkp.cz/knihovna61/koncepce_uchovavani.htm](http://knihovna.nkp.cz/knihovna61/koncepce_uchovavani.htm)
- [2] Projekt „Národní digitální knihovna“ v širším kontextu. Dostupný z [www: http://www.ndk.cz/](http://www.ndk.cz/)
- [3] Kramerius. Dostupný z [www: http://kramerius.qbizm.cz/](http://kramerius.qbizm.cz/)
- [4] Reference Model for an Open Archival Information System (OAIS). Dostupný z [www: http://public.ccsds.org/publications/archive/650x0b1.pdf](http://public.ccsds.org/publications/archive/650x0b1.pdf)
- [5] E. Bellini, Ch. Cirinna, M. Lunghi. Trvalé identifikátory pro kulturní dědictví. Dostupný z [www: http://www.digitalpreservationeurope.eu/publications/briefs/cz_trvale_identifikatory.pdf](http://www.digitalpreservationeurope.eu/publications/briefs/cz_trvale_identifikatory.pdf)
- [6] Creative Commons. Dostupný z [www: http://creativecommons.org/](http://creativecommons.org/)
- [7] Zeman, M. Creative Commons – budoucnost copyrightu? Přístupné na: <http://www.lupa.cz/clanky/creative-commons-8211-budoucnost-copyrightu/>
- [8] Creative Commons. Dostupný z [www: http://cs.wikipedia.org/wiki/Creative_Commons](http://cs.wikipedia.org/wiki/Creative_Commons)
- [9] Stoklasová, B. Národní digitální knihovna. Dostupný z [www: http://www.ndk.cz/publikace/projekt-ndk-integrovaný-operacní-program-ndk-prezentace-phdr-bohdany-stoklasove](http://www.ndk.cz/publikace/projekt-ndk-integrovaný-operacní-program-ndk-prezentace-phdr-bohdany-stoklasove)
- [10] Vítejte na informačním portálu systému Kramerius! Dostupný z [www: http://kramerius-info.nkp.cz/](http://kramerius-info.nkp.cz/)
- [11] J. Handl. Digitalizace knihoven je běh na dlouhou trať. Dostupný z [www: http://www.lupa.cz/clanky/digitalizace-knihoven-je-beh-na-dlouhou-trat/](http://www.lupa.cz/clanky/digitalizace-knihoven-je-beh-na-dlouhou-trat/)
- [12] Autorskoprávní posudek Digitální knihovny AV ČR. Dostupný z [www: http://avi.lib.cas.cz/posudek_dknav](http://avi.lib.cas.cz/posudek_dknav)
- [13] Open Access and Institutional Repositories with EPrints. Dostupný z [www: http://www.eprints.org/](http://www.eprints.org/)
- [14] V. Pasqui. Ochrana digitálních dat a otevřené archivy. Dostupný z [www: http://www.digitalpreservationeurope.eu/publications/briefs/cz_otevrene_archivy.pdf](http://www.digitalpreservationeurope.eu/publications/briefs/cz_otevrene_archivy.pdf)
- [15] W. Nixon. „DAEDALUS: Initial experiences with EPrints and DSpace

at the University of Glasgow". Dostupný z www: <http://www.ariadne.ac.uk/issue37/nixon/>

- [16] B. Vorlíčková. Doporučení Evropské komise o digitalizaci, online zpřístupnění a digitální ochraně kulturního dědictví. *Ikaros* [online]. 2006, roč. 10, č. 10 Dostupný z www: <http://www.ikaros.cz/node/3629>
- [17] F. Vojtášek. Dlouhodobá archivace digitálních dokumentů. *Ikaros* [online]. 2000, roč. 4, č. 10. Dostupný z www: <http://www.ikaros.cz/node/675>

Kontakt:

Mgr. Lenka Maixnerová
Oddělení bibliografie a katalogizace
Národní lékařská knihovna
Sokolská 54
121 32 Praha 2
Tel: 296335925
e-mail: maixnero@nlk.cz
<http://www.nlk.cz>