

DIGITÁLNÍ KNIHOVNA NÁRODNÍ LÉKAŘSKÉ KNIHOVNY – ARCHIVACE A ONLINE ZPŘÍSTUPNĚNÍ PLNÝCH TEXTŮ VĚDECKÝCH PUBLIKACÍ VE ZDRAVOTNICTVÍ

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

Abstrakt

Národní lékařská knihovna (NLK) provozuje Digitální knihovnu (DK) od roku 2008 v systému Kramerius, který umožňuje archivovat a zpřístupňovat plné texty různých typů dokumentů. Plné texty jsou automaticky propojovány s bibliografickými záznamy v portálu Medvik, takže je možné dostat se k plnému textu více způsoby. Dokumenty archivované v DK NLK jsou vytvářeny vlastní i externí digitalizací, případně se jedná o plné texty získané na základě licenčních smluv přímo od autorů a vydavatelů. Mezi dokumenty, které NLK zpřístupňuje přes DK, patří i závěrečné grantové zprávy Interní grantové agentury (IGA) MZ ČR. Tyto zprávy NLK shromažďuje, zpracovává, zpřístupňuje a uchovává v tištěné podobě od roku 1996. Svou povahou zprávy představují autorská díla, která dle autorského zákona řadíme mezi tzv. zaměstnanecká díla. V NLK je uchovávan pouze 1 exemplář, který je možné využívat pouze prezenčně v prostorách NLK. Pro efektivnější zpřístupnění zpráv přistoupila NLK k jejich digitalizaci a online zpřístupnění. Vytvoření digitálního úložiště, vybudování digitalizačního oddělení a online zpřístupnění dle pravidel autorského zákona je řešením projektu „Archivace a online zpřístupnění plných textů vědeckých publikací pro podporu rozvoje vědy a vzdělávání lékařských a nelékařských oborů“ podpořeného IGA MZČR. Výsledkem projektu bude převedení 3.500 závěrečných zpráv do elektronické podoby do konce roku 2013, včetně jejich volného zpřístupnění přes Internet. Ve spolupráci s IGA MZ ČR se připravuje též předávání závěrečných zpráv v elektronické podobě. NLK veškeré aktivity ohledně DK koordinuje s obdobnými projekty, řešenými v ČR (Národní úložiště šedé literatury – NUŠL, Národní digitální knihovna – NDK).

Klíčová slova

Digitální knihovny, digitalizace, lékařské knihovny, elektronické dokumenty, autorská práva, výzkumné zprávy

Úvod

Národní lékařská knihovna (NLK) od roku 2008 buduje Digitální knihovnu (DK). Digitální knihovna představuje systém pro získávání, zpracování, dlouhodobou archivaci a zpřístupnění elektronických dokumentů. DK je provozována v českém open-source systému Kramerius, který je vyvíjen Národní knihovnou (NK ČR), Knihovnou Akademie věd ČR a Moravskou zemskou knihovnou. Vývoj systému je zajišťován finančními dotacemi Ministerstva kultury ČR (MK ČR) a Ministerstva školství, mládeže a tělovýchovy ČR. Systém Kramerius slouží k zpřístupňování digitálních dokumentů v souladu s autorským zákonem.

Podporuje řadu formátů, např. XML, TXT, DjVu, JPG, PNG, PDF, JPEG2000, TIFF. Systém umožňuje replikaci dat mezi jednotlivými instalacemi.

Primárně byla Digitální knihovna NLK určena k ukládání digitálních dokumentů z fondu NLK digitalizovaných při řešení projektů v programu MK ČR Veřejné informační služby knihoven (VISK), podprogramy VISK 6 – Národní program digitálního zpřístupnění vzácných dokumentů Memoriae Mundi Series Bohemica a VISK 7 – Národní program mikrofilmování a digitálního zpřístupňování dokumentů ohrožených degradací kyselého papíru – Kramerius. Postupně začala být využívána i pro jiné výstupy, neboť obsah DK není omezen pouze na dokumenty vzniklé procesem digitalizace, ale je možné ho využít i pro ukládání dokumentů vytvořených již v digitální podobě. V roce 2009 NLK oslovila vybrané české nakladatele, zdali by měli zájem o službu dlouhodobé archivace a zpřístupnění elektronických verzí periodik na bezpečném úložišti včetně následného propojení s bibliografickou databází Bibliographia medica Českoslovac(BMČ). Pro ukládání a zpřístupňování elektronických verzí periodik v NLK bylo nutné získat povolení od nakladatelů. Smlouvu o poskytování elektronických dokumentů má NLK podepsanou s 27 nakladateli, celkem se jedná asi o 80 odborných periodik. V rámci smlouvy mohou nakladatelé zvolit časové embargo oproti tištěnému vydání a způsob zpřístupnění. Se 2 dodavateli došla spolupráce tak daleko, že se předávají nejenom plné texty článků, ale i bibliografická data ve strukturované podobě (Medicina sportiva Bohemica a Slovaca a časopisy České lékařské společnosti J.E.Purkyně zpřístupňované přes Meditorial).

Obsah Digitální knihovny NLK

Obsah DK NLK tvoří monografie, noviny, časopisy, konferenční sborníky, výzkumné zprávy, specializační atestační práce, výukové materiály, vzácné tisky, kvalifikační práce a závěrečné grantové zprávy. Závěrečné grantové zprávy Interní grantové agentury (IGA) MZ ČR NLK shromažďuje, zpracovává, zpřístupňuje a uchovává v tištěné podobě již od roku 1996. Vzhledem k tomu, že v NLK je uchovávan pouze 1 exemplář, je možné pouze jejich prezenční studium v prostorách NLK. Pro efektivnější využití v rámci celé ČR, se NLK rozhodla pro jejich převedení do elektronické podoby a zpřístupnění přes internet. Do konce roku 2013 by mělo být digitalizováno 3.500 závěrečných prací. Ve spolupráci s IGA MZ ČR se připravuje též předávání závěrečných zpráv v elektronické podobě. V DK se nyní nachází 1.120 dokumentů (tj. cca 150.000 stran). Z nichž 1/3 vznikla digitalizací (vlastní i externí), ostatní dokumenty (zejména periodika) získáváme na základě smluv přímo od nakladatelů nebo autorů.

NLK také plánuje rozšířit funkce DK o možnost tzv. autoarchivace, tj. umožnit samotným autorům, aby si do DK ukládali vlastní dokumenty (např. preprinty, příspěvky ze sborníků, výukové materiály).

Proces digitalizace

Digitalizace znamená převedení analogového obsahu do digitálního. Proces zahrnuje řadu činností, nejedná se pouze o čisté skenování, při kterém dochází

k převodu textového a obrazového materiálu do digitální podoby. Takto vzniklá data, obrazy se dále upravují (ořezy, narovnání), vytváří se metadata (popisná, administrativní, technická), provádí se OCR (optické rozpoznávání znaků), propojují se všechny části digitálního dokumentu. Výsledkem je digitální dokument skládající se z metadat a požadovaných digitálních souborů.

Pro digitalizaci NLK využívá knižní skenery Plustek OptiBook 4600 (formát A4), které pro tuto činnost ne zcela vyhovují. V letošním roce NLK plánuje zakoupit manuální knižní skener pro formát A2 s kolébkou a přítlačným sklem, s možností skenování v barevné a šedé škále, rozlišení min. 300 dpi. Požadavky na SW vybavení skeneru: ořez a vyrovnání, dělení stránek, OCR – čeština, angličtina, němčina (minimálně), výstupy: PDF/A, JPEG, TIFF, česká lokalizace a manuály.

Pro organizační zajištění procesů digitalizace v NLK bylo zřízeno nové pracoviště Centrum knihovnických a digitalizačních aplikací, které se bude problematikou digitalizace, uchovávání a zpřístupňování elektronických dat do budoucna zabývat.

Datové formáty a standardy

V DK NLK byly doposud používány formáty PDF a DjVu. Formát PDF bude i nadále využíván pro určité typy dokumentů, formát DjVu bude nahrazen formátem JPG případně JPEG2000. Pro potřeby dlouhodobé archivace (LTP – long time preservation) a zpřístupnění závěrečných zpráv IGA se jako nejvhodnější jeví formát PDF/A–1b, protože se jedná o otevřený standard (ISO 19005–1 Level B), obvyklý pro vědeckou komunikaci, tzv. “self–contained” formát – veškeré údaje potřebné pro zobrazení dokumentu jsou vloženy v samotném souboru. Pro PDF soubory již uložené v archivu nebo získané z externích zdrojů, které nejsou ve formátu PDF/A, bude po analýze typu souborů provedena normalizace na formát PDF/A–1b. V budoucnu bude možné využití formátu PDF/A–2 (ISO 19005–2), který je s PDF/A–1 zpětně kompatibilní.

Rozlišení při skenování bude nastaveno minimálně na 300 DPI, barevné předlohy (stránky s barevným obsahem) budou skenovány barevně (24-bit hloubka), černobílé předlohy v šedé škále (8-bit hloubka).

Právní ošetření zpřístupňování archivovaných dokumentů


Získávání a zpřístupňování dokumentů v DK se řídí Autorským zákonem č. 121/2000 Sb. Knihovna může převést dokumenty, které má ve vlastnictví (případně o ně přišla – ztráta, krádež, poškození), do elektronické podoby. Takto vzniklé elektronické kopie může volně zpřístupňovat přes internet, pokud již uběhlo 70 let od smrti autora/ů, případně získala povolení (licenci) k online šíření od nakladatele nebo autora případně dědiců autorských práv. V ostatních případech je možné dokumenty zpřístupňovat pouze na IP adrese knihovny. U sirotčích děl, kde není nositel práv k dílu určen, nebo i když je určen, není nalezen po důsledném vyhledávání nositele práv [cit. 9], je možné dílo zpřístupnit 100 let od vydání. U novin a časopisů, jejichž obsah je tvořen více autory se postupuje obdobně jako u sirotčích děl.

NLK si nechala v roce 2011 vyhotovit právní analýzu k problematice zpřístupňování šedé literatury. Mezi šedou literaturu řadíme dokumenty, které nejsou publikovány obvyklým způsobem a nejsou proto dostupné na běžném knižním trhu (např. diplomové a dizertační práce, výzkumné zprávy, interní dokumenty, oficiální publikace atd.) [cit. 3]. Závěrečné zprávy IGA MZ ČR patří též do šedé literatury, mají povahu autorských děl a patří dle autorského zákona mezi tzv. zaměstnanecká díla. „Není-li sjednáno jinak, zaměstnavatel vykonává svým jménem a na svůj účet autorova majetková práva k dílu, které autor vytvořil ke splnění svých povinností vyplývajících z pracovněprávního či služebního vztahu k zaměstnavateli nebo z pracovního vztahu mezi družstvem a jeho členem (zaměstnanecké dílo). Zaměstnavatel může právo výkonu podle tohoto odstavce postoupit třetí osobě pouze se svolením autora, ledaže se tak děje při prodeji podniku nebo jeho části. Smrtí nebo zánikem zaměstnavatele, který byl oprávněn vykonávat majetková práva k zaměstnaneckému dílu a který nemá právního nástupce, nabývá oprávnění k výkonu těchto práv autor.“ (AZ). U grantových zpráv za zaměstnavatele považujeme příjemce dotace. NLK může díla užívat pouze na základě svolení IGA. Do současné doby byl vztah mezi IGA a NLK řešen pouze ústní dohodou, což se ukázalo jako jeden z možných problémů v právní analýze. Do budoucna je proto nutné tento vztah ošetřit smluvně. Dle analýzy je možné povolit zpřístupnění již odevzdaných závěrečných zpráv v elektronické podobě na základě podepsání licenční smlouvy mezi IGA a NLK a mezi IGA a příjemci dotace (zaměstnavatelé). Do budoucna se doporučuje již do grantové dohody mezi IGA a příjemcem dotace vložit formulace ohledně zpřístupnění elektronických verzí v Digitální knihovně NLK. Volně přes internet budou moci být zpřístupněny pouze určité části závěrečných zpráv, a to věcná a ekonomická část. Doložená publikační činnost bude moci být zpřístupněna pouze v režimu neveřejné. Veřejné zpřístupnění publikační činnosti je možné, pouze pokud bude výsledek publikován v Open Access zdroji nebo pod licencí Creative Commons.

Co se týče již zmíněné problematiky autoarchivace, do Digitální knihovny NLK je možné vkládat autorská díla přímo autory, a to na základě využití licence Creative Commons. NLK však vždy ponese právní odpovědnost i za zpřístupnění těchto materiálů. Autoři mohou přes Digitální knihovnu zpřístupňovat konferenční příspěvky publikované ve sbornících, pokud nedošlo k uzavření smlouvy mezi autorem a vydavatelem sborníku. Totéž platí i pro preprinty. Výukové materiály, které vzniknou na základě dohody o provedení práce, patří též mezi zaměstnanecká díla.

Zpřístupnění

Data v DK NLK jsou zpřístupňována přes systém Kramerius. NLK pracuje na převedení dat do vyšší verze Kramerius 4 (K4), která je více orientována na fulltextové vyhledávání doplněné o fasetovou navigaci a časovou osu. Systém Kramerius je propojený s portálem Medvik pomocí linkovacího mechanismu, uživatel tak nemusí složitě zjišťovat, zda má dokument elektronickou podobu. Přímo z bibliografického záznamu v systému Medvik je směřován do elektronické verze.

- 4  **Zhodnocení vývoje, současného stavu a podmínek k zajištění zdravého tělesného, duševního a sociálního vývoje nové generace do roku 2010**



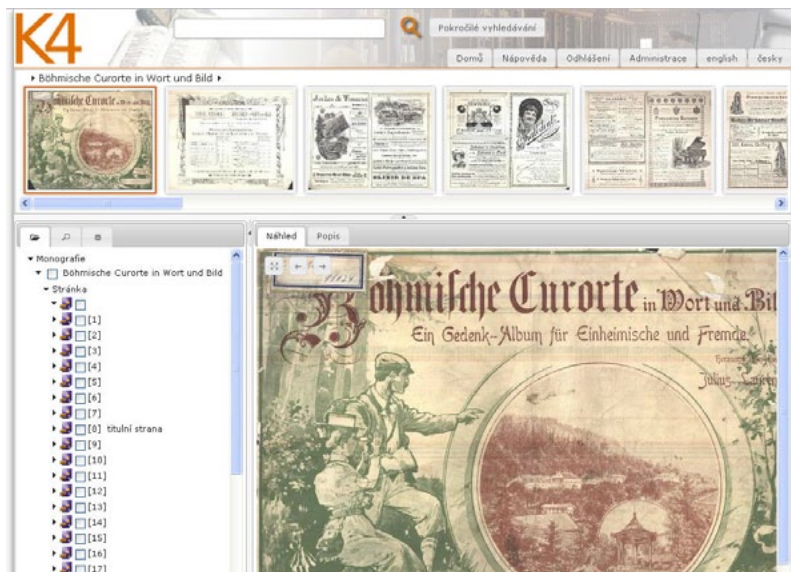
Pfeifer, Ivan, 1944- 

Praha : Iga MZ ČR, 2005. Přeřuš. str. : tab., grafy ; 30 cm.

[katalog](#) [detail](#) [Digitální archiv](#) 

Obrázek 1 — Zobrazení grantové zprávy v portálu Medvik

Z prostředí Java, v němž je provozován portál Medvik, je přístupováno pomocí ODBC konektoru přímo do databáze Krameria (PostgreSQL). Výsledkem je zobrazení odkazů na všechny úrovně metadat, která jsou v Krameriovi k dispozici – typicky tedy na článek, číslo, ročník a časopis. Řešení umožňuje propojení nejen na digital-born dokumenty, ale i na digitalizované dokumenty – zde však pouze maximálně na úroveň čísla. Vzhledem k odlišné architektuře systému Kramerius4 bude nutno upravit mechanismus propojení bibliografických záznamů v portálu Medvik a objektů v digitální knihovně.



Obrázek 2 — Zobrazení dokumentu v K4

Spolupráce

NLK veškeré aktivity ohledně DK koordinuje s obdobnými projekty, řešenými v ČR. Při digitalizaci využívá Registr digitalizace:

<http://www.registrdigitalizace.cz/>, ve kterém se dá ověřit, zdali dokument byl již digitalizován nebo se plánuje digitalizovat jinou institucí v ČR. Pokud ano, je možné zažádat o replikaci data a není nutné dokument znovu digitalizovat. Většina

institucí poskytuje tuto službu zdarma. NLK sleduje a koordinuje digitalizační aktivity s projektem Národní digitální knihovna NK ČR a spolupracuje s Národním úložištěm šedé literatury.

Závěr

NLK předpokládá, že digitalizací závěrečných grantových zpráv a jejich zpřístupněním širokému okruhu uživatelů přes internet dojde ke zlepšení dostupnosti informací o řešených projektech programů IGA MZ ČR a efektivnímu zviditelnění dosažených výsledků.

Článek byl podpořen grantem IGA MZ ČR NT12345.

Literatura:

- [1.] *Autorské právo a knihovny*. Dostupný z [www: http://www.nkp.cz/o_knihovnach/AutZak/Index.htm](http://www.nkp.cz/o_knihovnach/AutZak/Index.htm)
- [2.] *Cubr, L. Dlouhodobá ochrana digitálních dokumentů*. Praha: Národní knihovna ČR, 2010. 154 s.
- [3.] *Česká terminologická databáze knihovnictví a informační vědy (TDKIV)*. Dostupný z [www: http://aleph.nkp.cz/F/](http://aleph.nkp.cz/F/)
- [4.] *Hutař, J. Nové standardy digitalizace (od roku 2012)*. Dostupný z [www: http://www.ndk.cz/digitalizace/nove-standardy-digitalizace-od-roku-2011](http://www.ndk.cz/digitalizace/nove-standardy-digitalizace-od-roku-2011)
- [5.] *Kříž, F., Maixnerová, L. Digitalizace v Národní lékařské knihovně. Lékařská knihovna, 2009, č. 1–2*. Dostupný z [www: http://www.nlk.cz/publikace-nlk/lekarska-knihovna/2009/lk2009-1-2/digitalizace-v-nlk](http://www.nlk.cz/publikace-nlk/lekarska-knihovna/2009/lk2009-1-2/digitalizace-v-nlk)
- [6.] *Maixnerová, L., Kříž, F., Horsák, O. Zpřístupnění a archivace plných textů českých lékařských a zdravotnických časopisů. Knihovnický zpravodaj Vysočina, 2010, roč. 10, č. 2*. Dostupný z [www: http://kzv.kkvysoctiny.cz/archiv.aspx?id=1266&idr=10&idci=26](http://kzv.kkvysoctiny.cz/archiv.aspx?id=1266&idr=10&idci=26)
- [7.] *Maixnerová, L., Kříž, F., Horsák, O., Bouzková, H. Rozvoj digitálního archivu Národní lékařské knihovny ve spolupráci s vydavateli a autory vědeckých lékařských informací. MEDSOFT 2009. 2009, s. 97–103*.
- [8.] *Matušík, Z. K některým autorskoprávním otázkám činnosti knihoven v současnosti*. *Knihovna plus [online]*. 2010, č. 1. Dostupný z [www: http://knihovna.nkp.cz/knihovnaplus101/matusik.htm](http://knihovna.nkp.cz/knihovnaplus101/matusik.htm)
- [9.] *Matušík, Z. Osířelá díla a díla nedostupná na trhu a novela autorského zákona. 12. Konference AKM v digitálním světě, 30.11.2011*. Dostupný z [www: http://www.skipcr.cz/dokumenty/akm-2011/Matusik.pdf](http://www.skipcr.cz/dokumenty/akm-2011/Matusik.pdf)
- [10.] *Myška, M., Polčák, R., Šavelka, J. Zpřístupňování závěrečných zpráv IGA MZ ČR v Digitální knihovně Národní lékařské knihovny: opinio (stanovisko k právní otázce)*. 2011. 25 s.
- [11.] *Rosenthal, V., Blekinge–Rasmussen, A., Hutař, J. a kol. Průvodce plánem důvěryhodného repozitáře (PLATTER)*. Praha: Národní knihovna ČR, 2009. 51 s.

Kontakt:

Mgr. Lenka Maixnerová

Národní lékařská knihovna

Sokolská 54, Praha 2

maixnero@nlk.cz

Mgr. Filip Kříž

PhDr. Helena Bouzková

Mgr. Ondřej Horsák

Mgr. Adéla Jarolímková, Ph.D.

PhDr. Eva Lesenková, Ph.D.