

Koncipování obsahu digitální knihovny

Studijní text

Autor: Martina Smetánková

Konzultant: Václav Jiroušek

Ostrava 2019



MINISTERSTVO
KULTURY



VEŘEJNÉ
INFORMAČNÍ SLUŽBY
KNIHOVEN



Odborná způsobilost Koncipování obsahu digitální knihovny

Profesní kvalifikace Samostatný knihovník správce digitální knihovny
(kód: 72-018-R)

Kvalifikační úroveň 6

Klíčová slova digitální knihovna, digitalizace, repozitář, vyhledávací index, uživatelské rozhraní, webový klient, fulltextové vyhledávání, OCR, ochrana fondu, Kramerius, registr digitalizace, virtuální sbírka, vzdálený přístup, ochrana fondu, standardy pro digitalizaci, skeny, skenování, zpřístupnění, autorské právo, masová digitalizace, výběrová digitalizace, duplicitní záznamy, strategie digitalizace

Anotace Tento text řeší problematiku digitalizace, digitálních knihoven a vhodného koncipování jejich obsahu z pohledu paměťové instituce včetně uživatele. Text vás provede jednotlivými kroky přípravy digitalizace a pomůže vám s naplánováním vlastní strategie vedoucí k vytvoření atraktivního obsahu plánované digitální knihovny.

Tvorba studijních textů je realizována za finanční podpory Ministerstva kultury České republiky v rámci projektu Veřejné informační služby knihoven (VISK 1).

Zkouška z odborné způsobilosti **Koncipování obsahu digitální knihovny** má 3 části, zkoušený musí splnit tato kritéria hodnocení:

Kritéria hodnocení	Způsob ověření
a Zpracovat návrh obsahu digitální knihovny a prezentovat jej formou desetiminutové prezentace	Praktické předvedení
b Charakterizovat hlediska pro výběr dokumentů k digitalizaci	Ústní ověření
c Popsat základní funkce digitální knihovny	Ústní ověření

Tento text pro vás připravila Bc. Martina Smetánková, Moravská zemská knihovna v Brně, martina.smetankova@mzk.cz

Text konzultoval Václav Jiroušek, Národní knihovna ČR, vaclav.jirousek@nkp.cz

OBSAH

RYCHLÝ NÁHLED STUDIJNÍHO TEXTU.....	5
1 DIGITÁLNÍ KNIHOVNA	6
1.1 Digitální knihovna jako pojem	6
1.2 Komponenty digitální knihovny	7
1.3 Výhody používání digitální knihovny	8
1.4 Příklady digitálních knihoven	9
1.4.1 Příklady z českého prostředí	9
1.4.2 Příklady ze zahraničí	10
2 DIGITALIZACE	12
2.1 Digitalizace v českých knihovnách.....	12
2.2 K čemu slouží digitalizace?	13
2.3 Proces plánování digitalizace	14
2.4 Výběr dokumentů k digitalizaci	16
2.4.1 Kritéria pro výběr dokumentů	16
SHRNUTÍ STUDIJNÍHO TEXTU.....	20

RYCHLÝ NÁHLED STUDIJNÍHO TEXTU

ÚVOD

Tento text řeší problematiku digitalizace, digitálních knihoven a vhodného koncipování jejich obsahu z pohledu paměťové instituce včetně uživatele. Text vás provede jednotlivými kroky přípravy digitalizace a pomůže vám s naplánováním vlastní strategie vedoucí k vytvoření atraktivního obsahu plánované digitální knihovny.

Problematika zpřístupnění dokumentů ve vztahu k autorskému zákonu je zde zmíněna pouze okrajově, neboť jsou součástí jiných modulů tohoto knihovnického kurzu.

PO PROSTUDOVÁNÍ STUDIJNÍHO TEXTU BUDETE UMĚT

- vysvětlit termín digitální knihovna,
- popsat, co je digitalizace, proč se digitalizuje a jaké jsou její fáze,
- vytipovat vhodné dokumenty pro digitalizaci,
- vytvořit strategii digitalizace vhodnou pro vaši knihovnu.

KLÍČOVÁ SLOVA STUDIJNÍ OPORY

digitální knihovna, digitalizace, repozitář, vyhledávací index, uživatelské rozhraní, webový klient, fulltextové vyhledávání, OCR, ochrana fondu, Kramerius, registr digitalizace, virtuální sbírka, vzdálený přístup, ochrana fondu, standardy pro digitalizaci, skeny, skenování, zpřístupnění, autorské právo, masová digitalizace, výběrová digitalizace, duplicitní záznamy, strategie digitalizace

1 DIGITÁLNÍ KNIHOVNA

Digitální knihovna je služba, jejímž účelem je poskytnout přístup k digitálním či digitalizovaným zdrojům a dalším informacím o tištěných dokumentech uložených v knihovním fondu a zároveň je touto cestou chránit. Součástí procesu vytvoření digitální knihovny je ukládání datové zálohy v dlouhodobém úložišti dat. Budování digitální knihovny je však věcí dlouhodobé strategie¹.

Tato kapitola se zaměřuje především na definice a vysvětlení pojmů **digitalizace, digitální knihovna** a její **funkce**. Uvedeny budou příklady **české i zahraniční praxe**. Modul se věnuje problematice zpřístupnění dat prostřednictvím digitální knihovny. Problematice archivace zobrazovaných dat se budeme věnovat v samostatném modulu, proto je zde zmíněna pouze okrajově.

1.1 Digitální knihovna jako pojem

Je třeba zmínit, že termín “digitální knihovna” je značně problematický. Tento výraz zahrnuje komplexní vztah mezi elektronickými informačními zdroji a knihovnami jakožto institucemi, který může být vnímán jako řečnický rozpor z důvodu spojení dvou odporujících si slov “digitální” a “knihovna”. Definice pro termín digitální knihovna (z angl. *digital library*) se pak na základě tohoto paradoxu mohou různit. Pro české knihovnické prostředí a pro potřeby tohoto textu byly vybrány následující:

DEFINICE

- „Digitální knihovna je spravovaná sbírka informací spolu s odpovídajícími službami, přičemž informace jsou uloženy v digitální podobě a jsou dostupné prostřednictvím sítě.“²
- „Digitální knihovny jsou organizace, které poskytují zdroje (včetně specializovaného personálu) umožňující provádět výběr, strukturování a zpřístupnění sbírek digitálních prací, tyto práce dále distribuovat, udržovat jejich integritu a dlouhodobě uchovávat – a to vše s ohledem na snadné a ekonomické využití určitou komunitou nebo množinou komunit uživatelů.“³

¹ BRATKOVÁ, Eva. *Otevřený přístup a digitální knihovny v oblasti vědy a výzkumu: (vybrané systémy)* [online]. Verze 1.0. Praha: Ústav informačních studií a knihovnictví FF UK v Praze, prosinec 2008 [cit. 2015-11-24]. 126 s. Formát PDF. Elektronické studijní texty ÚISK. Dostupný z:

<http://texty.jinonice.cuni.cz/>

² ARMS, William Y. *Digital libraries*. Cambridge, Mass: MIT press, c2000, x, 287 s.

³ VAN DE SOMPEL, H. a P. HOCHSTENBACH. *Reference Linking in a Hybrid Library Environment. Part 3, Generalizing the SFX solution in the "SFX@Ghent*. 1999. Dostupné také z: <http://search.proquest.com/docview/57491073?accountid=35514>

- „Digitální knihovna je digitální obdoba klasické knihovny, která získává, katalogizuje, skladuje a ochraňuje digitální dokumenty v digitálním repozitáři a rovněž je zpřístupňuje.“⁴

DEFINICE

Výše uvedené definice mají několik společných znaků⁵:

- **digitalizace fyzického materiálu není klíčová**, jde spíše o způsob organizace elektronické sbírky a zajištění lepšího přístupu uživatelům,
- **informační zdroje** tvořící fond digitální knihovny mohou být **různorodé** (způsob uložení, organizace a správa uložení, použité platformy), **dynamické** (umožňují rychlé vkládání a vyřazování objektů) a **multimediální** (povaha dat),
- pro provoz digitální knihovny je nutné **propojení informačních prvků** (např. informací z katalogu), které musí být pro uživatele čitelné,
- **přístup** k relevantním digitálním informacím má být **jednotný** bez ohledu na jejich formu, formát, způsob a místo uložení.

Vývoj a nasazení digitálních knihoven je v dnešní praxi zpravidla výsledkem spolupráce dvou skupin: informačních profesionálů (knihovníci, nakladatelé a další poskytovatelé informačních služeb) a vývojářů aplikací. Výsledkem spolupráce těchto skupin je vytvoření zcela samostatné digitální knihovny s lokálním vymezením a fondem (specializované, muzejní, univerzitní, regionální, krajské, národní apod.). Další možnou variantou je sdružení více digitálních knihoven v rámci jednotného uživatelského rozhraní. Druhá jmenovaná možnost bývá formou federace nezávislých knihoven, které spojuje společné téma.

1.2 Komponenty digitální knihovny

Digitální knihovna se věnuje, podobně jako klasické knihovny, činnostem jako je sběr dokumentů a jejich třídění, popis a zpřístupnění. Tato činnost se liší pouze ve způsobu zprostředkování dat uživateli v digitální podobě. Skládá se z následujících částí:

- **digitální repozitář,**
- **vyhledávací index,**
- **uživatelské rozhraní.**

⁴ BARTOŠEK, Miroslav. *Digitální knihovny* [online]. Brno : Masarykova univerzita, 2001 [cit. 2015-12-04]. Dostupný z WWW: <http://www.ics.muni.cz/mba/dl-datakon01.pdf>

⁵ BARTOŠEK, Miroslav. *Digitální knihovny – teorie a praxe*. Národní knihovna, 2004, 15(4), s. 233-254. ISSN 0862-7487. Dostupné také z: <http://oldknihovna.nkp.cz/NKKR0404/0404233.html>

Digitální repozitář je uzavřené místo pro ukládání digitálních dat, která mohou být měněna pouze jeho správcem. Digitální objekty zde vložené mají své unikátní identifikátory. Systém eviduje datum a čas vložení a další změny, které usnadňují jejich vyhledání. Naopak k vyhledávání informací přímo v digitální knihovně slouží tzv. vyhledávací index.

Vyhledávací index lze popsat jako vrstvu mezi uživatelským rozhraním a digitálním repozitářem. V případě zadání uživatelského dotazu např. přes běžné vyhledávací pole (včetně zapojení filtrů) se tento dotaz neodešle přímo do repozitáře, ale do vrstvy nad ním. Ta obsahuje pouze vybraná a lépe strukturovaná data, která dokážou zprostředkovat rychlejší propojení uživatelského dotazu a přesného místa v repozitáři. Index je tedy klíčovým prvkem, neboť podstatně urychluje vyhledávání informací zadaných uživatelem v **uživatelském rozhraní**.

V tomto případě se myslí webové rozhraní, kudy uživatel (čtenář) přistupuje do digitální knihovny. Uživatelské rozhraní prezentuje čtenáři v graficky přátelské a intuitivní formě data uložená v repozitáři. Tímto způsobem si uživatel může procházet dokumenty, listovat v nich, prohlížet si obrázky, pročitat popisná metadata a v případě autorsky volných děl si data v různých formátech stahovat či tisknout.

1.3 Výhody používání digitální knihovny

Digitální knihovna poskytuje mnoho výhod jak pro uživatele, tak pro zprostředkovávající instituci. Z pohledu uživatele je velice atraktivní možnost využívat tzv. **vzdálený přístup**, pomocí něhož se může k fondu knihoven dostat odkudkoliv online. V rámci prostředí českých digitálních knihoven se většinou jedná o díla autorsky volná nebo taková, na která byla poskytnuta licence vlastníkem autorských práv. V rámci digitalizace a zpřístupnění dochází též ke koordinaci aktivit mezi knihovníky, aby se **zamezilo duplicitním dokumentům**. K jejich vyhledávání slouží nástroj **Registr digitalizace**. V rámci procesu digitalizace se provádí též optické rozpoznávání znaků či OCR (z angl. Optical Character Recognition), které uživateli umožňuje **fulltextové vyhledávání** v dokumentu a šetří tak značně jeho čas při vyhledávání, a to i přes určitou míru chybovosti. Další výhodou je **možnost současného přístupu více uživatelům** ke stejnému dokumentu, který knihovna často vlastní v jediném exempláři a není možné jej získat jiným způsobem. Pokud se jedná o autorsky chráněnou knihu, lze ji v rámci budovy knihovny zpřístupnit z více počítačů zároveň. Tato uživatelská výhoda je zároveň i výhodou pro knihovnu jakožto instituci, neboť unikátní či jinak ohrožené exempláře může digitalizace v praxi ochránit tím, že již není důvod je fyzicky půjčovat. Digitalizace a zpřístupnění dokumentů v digitální knihovně tak přispívá k **ochraně knihovního fondu** a zároveň

obohacuje knihovní služby. Dokumenty v digitální knihovně navíc mohou být volně organizovány do **tematických** (virtuálních) **sbírek**.

1.4 Příklady digitálních knihoven

V České republice i zahraničí existuje více druhů digitálních knihoven provozovaných různými institucemi na bázi komerčních i open source řešení, tedy počítačového software s otevřeným zdrojovým kódem, který za určitých licenčních podmínek smí použít kdokoli.

1.4.1 Příklady z českého prostředí

V České republice se nejčastěji pro prezentaci a uchování dat používá systém **Kramerius**⁶. Jedná se o softwarové open source řešení pro ukládání a prezentaci digitálních dokumentů s metadaty v souladu se standardy vydávanými Národní knihovnou ČR. Primárně je určeno pro zpřístupňování monografií a periodik z českých knihoven. Systém též umožňuje vkládat jednotlivé části dokumentů, jako jsou články či kapitoly a další **typy dokumentů**, např. mapy, grafiky, hudebniny, zvukové nahrávky, archiválie, rukopisy či born digital dokumenty, tedy dokumenty, které vznikly od počátku v elektronické podobě. Všechny uvedené typy dokumentů mohou být vkládány do jedné či více tematických virtuálních sbírek současně. Ty může dle potřeby zakládat kurátor či jiný správce digitální knihovny. Více informací o systému Kramerius a aktuálním webovém klientovi lze nalézt na vývojových stránkách projektu prostřednictvím služby Github⁷. Původně byl systém Kramerius zamýšlen především pro použití v knihovnách prostřednictvím dotačního programu **VISK 7**⁸. V současnosti je nicméně využívají pro prezentaci digitálních dokumentů i další instituce a nyní je známo přes 35 samostatných verzí systému Kramerius⁹.

Od počátku vzniku aplikace vznikl bezpočet verzí uživatelského rozhraní, které na území ČR stále využívají (podle vlastních technických možností) různé instituce. Aktuální verzi Krameria lze sledovat na webu **Digitalniknihovna.cz**¹⁰, který provozuje Moravská zemská knihovna v Brně. Kromě aktuální verze webového klienta (s plně responzivním designem pro mobily i počítače) aplikace umožňuje vstupovat a prohlížet data z více digitálních knihoven nejen v Česku na jednom místě. Prozatím však nelze prohledávat data z více digitálních knihoven současně v rámci jednoho vyhledávacího pole. Uživatel si musí vždy zvolit, kterou ze zapojených digitálních knihoven chce prohlížet.

⁶ <https://system-kramerius.cz/>

⁷ <https://github.com/ceskaexpedice/kramerius/wiki>

⁸ <https://visk.nkp.cz/visk-7>

⁹ <https://registr.digitalniknihovna.cz/>; uvedený stav ke dni 17. listopadu 2019

¹⁰ <http://www.digitalniknihovna.cz/>

Jednou ze zapojených knihoven je i projekt **Česká digitální knihovna**¹¹, která si klade za cíl agregovat všechny české digitální knihovny a tím jednotné vyhledávání umožnit. V současnosti je tato knihovna spravována Knihovnou Akademie věd ČR.

Systém Kramerius však není jedinou variantou pro zpřístupňování digitálních dokumentů. Příkladem velmi jednoduché digitální knihovny je např. digitalizovaný archiv časopisů, který provozuje **Ústav pro českou literaturu AV ČR**¹², který zpřístupňuje nejvyhledávanější česká periodika ve spolupráci s Divadelním ústavem, Masarykovým ústavem a Archivem AV ČR.

Dalším z úspěšných a velkých českých projektů je **Manuscriptorium**¹³, jehož administraci zajišťuje Národní knihovna ČR. Tato digitální knihovna je volně dostupná a vznikla pro účely prezentace výsledků dotačního programu VISK 6. Dnes slouží jako agregační prostředí pro rukopisy a staré tisky z mnoha významných institucí nejen z území Evropské unie. Na rozdíl od systému Kramerius její uživatelské prostředí umožňuje uživatelům z agregovaného obsahu zakládat osobní kolekce digitálních dokumentů, se kterými lze nakládat jako s vlastní virtuální knihovnou a sdílet výsledky své práce s dalšími uživateli. Manuscriptorium tedy slouží jako cenný nástroj především pro badatele. Pro systém Kramerius jsou tyto funkce ve fázi vývoje.

1.4.2 Příklady ze zahraničí

Jako příklady obsahově i uživatelsky inspirativních digitálních knihoven ze zahraničí jsme zvolili polskou národní digitální knihovnu **Polona**¹⁴ a finskou digitální knihovnu **Finna**¹⁵, která slouží jako agregátor veškerých digitálních dokumentů z finských knihoven. V obou případech je možné prohlížet digitalizované monografie, periodika, mapy, rukopisy, grafiky a další dokumenty v uživatelsky i vizuálně přívětivé webové aplikaci. Za zmínku též stojí **Europeana**¹⁶, která slouží jako kvantitativní agregátor digitalizovaných dokumentů z institucí napříč Evropou. Prohlížení dokumentů zde funguje na principu sklizení obsahu ze zapojených digitálních knihoven z různých zemí, který se společně s náhledy obrázků zobrazuje v jednotném uživatelském rozhraní. Europeana agreguje z českých projektů aktuálně dokumenty z Manuscriptoria a nově i České digitální knihovny. Podobným projektem na světové úrovni je např. **World Digital Library**¹⁷, která je provozovaná Kongresovou knihovnou ve Washingtonu. Ta se však od Europeany liší

¹¹ <https://www.czechdigitallibrary.cz/>

¹² <http://archiv.ucl.cas.cz/>

¹³ www.manuscriptorium.com

¹⁴ <https://polona.pl/>

¹⁵ <https://www.finna.fi/>

¹⁶ <https://www.europeana.eu/>

¹⁷ <https://www.wdl.org/>

kvalitativním přístupem ke sběru digitálních dokumentů, které pochází ze spolupracujících knihoven, archivů, muzeí, vzdělávacích institucí a mezinárodních organizací z celého světa. Tato digitální knihovna je přímo podporovaná UNESCO.

2 DIGITALIZACE

Digitalizaci lze popsat zjednodušeně jako *“proces, při kterém je dokument převáděn z analogové podoby do digitální, která je čitelná počítačem”*¹⁸. Jinými slovy se jedná o *souhrn činností, jejichž výsledkem je archivace či zpřístupnění dokumentů v digitální podobě*¹⁹. Digitalizaci je třeba chápat jako komplexní proces zahrnující **přípravu dokumentů, skenování, editaci skenů, OCR** (optické rozpoznání znaků), vytvoření popisných **metadat** a následně zajištění **zpřístupnění** těchto dat a jejich **archivaci v dlouhodobém úložišti dat** (LTP).

V následujícím textu bude dán prostor pouze digitalizaci v knihovnách či dalších paměťových institucích zaměřených především na ochranu fondů a všech stran originálu včetně ilustrací či čtenářských vpisů apod. Kromě běžných dokumentů na papírovém podkladu se v knihovnách též digitalizují audio i video záznamy či 3D objekty (např. sbírka glóbů v Mapové sbírce Přírodovědecké fakulty Univerzity Karlovy²⁰). Za zmínku též stojí digitalizace mikrofilmů.

2.1 Digitalizace v českých knihovnách

Největším digitalizačním projektem českého knihovnictví je nepochybně projekt **Vytvoření Národní digitální knihovny**²¹ (NDK), v rámci něhož probíhá spolupráce Národní knihovny ČR a Moravské zemské knihovny v Brně. Tento projekt digitalizuje všechny typy literatury zaměřené na významnou část **bohemikální produkce 19. až 21. století**, tj. knihy vydané na našem území v češtině nebo pojednávající o Česku. Cílem tohoto projektu je tuto produkci ve fondech NK ČR a MZK zdigitalizovat do roku 2040. V rámci projektu již bylo úspěšně digitalizováno přes 50 milionů stran. Digitalizované dokumenty se zpřístupňují prostřednictvím systému Kramerius v digitálních knihovnách NK ČR²² a MZK²³, archivní kopie se uchovávají v dlouhodobém úložišti dat (LTP) NK ČR s datovými zálohami na dvou místech v Praze a jedné v Brně. Projekt probíhal mezi lety 2010–2014 s následným pětiletým obdobím udržitelnosti a po uplynutí tohoto období se počítá i s další kooperací obou institucí.

Obdobně rozsáhlým projektem mimo NDK je **spolupráce společnosti Google a Národní knihovny ČR**. V rámci projektu se mezi lety 2011 až 2018

¹⁸ Oxford Dictionaries. Digitize [online]. © 2014 [cit. 2014-01-22]. Dostupné z: <http://www.oxforddictionaries.com/definition/english/digitize>

¹⁹ KRAVECOVÁ, Lenka; KRUTSKÁ, Markéta; RYCHTÁŘOVÁ, Pavlína: Úvod do digitalizace. In Moravská zemská knihovna v Brně. E-learningový kurz Digitalizace v knihovnách [online - e-learningový kurz]. Brno : Moravská zemská knihovna v Brně, 2019 [cit. 2019-11-17]. Dostupné z:

<https://www.mzk.cz/pro-knihovny/vzdelavani-knihovniku/e-learningovy-kurz-digitalizace-v-knihovnach>

²⁰ <http://www.mapovasbirka.cz/globy/>

²¹ <https://www.ndk.cz/>

²² <http://www.digitalniknihovna.cz/nkp>

²³ <http://www.digitalniknihovna.cz/mzk>

zdigitalizovala významná část fondu starých tisků až po 19. století společně se starším fondem Slovanské knihovny. Národní knihovna zajistila opravy a restaurování knih, katalogizaci, přidělení čárových kódů a vytvoření metadat. Společnost Google zprostředkovala dopravu knih do digitalizačního centra v Německu a sjednala pojištění. Digitální kopie knih jsou zpřístupněny v digitální knihovně Manuscriptorium a v aplikaci **Google Books**²⁴.

Kromě rozsáhlých digitalizačních projektů probíhá též **digitalizace v krajských knihovnách a krajských digitalizačních centrech**. Podoba těchto tzv. krajských digitalizací není jednotná. Zatímco v některých institucích je projekt financován a realizován ve spolupráci s krajem, jinde je přímo pověřenou institucí krajská knihovna a kraj vynakládá navíc pouze prostředky na techniku. V některých krajích se digitalizuje i formou tzv. **externí digitalizace** v jiném již zavedeném digitalizačním centru. Neznačená tedy, že knihovna provozující svou vlastní digitální knihovnu musí nutně provozovat vlastní digitalizační pracoviště. V krajských knihovnách se obvykle digitalizují lokální periodika a staré tisky společně s muzejními sbírkami či pro region unikátními a jinak důležitými dokumenty. Vše závisí nicméně na koncepci digitalizace v daném kraji a na zájmu jednotlivých paměťových institucí. Vybavení i počty úvazků pro realizaci digitalizace i výsledná čísla digitalizovaných stran se značně liší kraj od kraje.

Kromě již zmíněných projektů probíhá v rámci ČR i mnoho dalších **menších projektů**, které jsou obvykle financovány z dotačních programů **VISK 6 a 7**. V rámci těchto programů byla zřízena **samostatná digitalizační pracoviště** např. v Národní technické knihovně, Knihovně Akademie věd ČR a v univerzitních a dalších (nejen odborných) knihovnách. Financování českých digitalizačních projektů bylo v minulosti opakovaně realizováno též v rámci **Norských fondů**²⁵, např. v Národní knihovně ČR a Městské knihovně v Praze (**Hispra**²⁶). Další možností je zapojení knihoven do různých zahraničních projektů, jako je např. **eBooks on Demand**²⁷ (EOD), v rámci něhož se na objednávku čtenářů v zapojených knihovnách digitalizují autorsky volné knihy obvykle vydané před sto a více lety.

2.2 K čemu slouží digitalizace?

Digitalizace v knihovnách se provádí především z důvodu **ochrany fondů**. Nutnost ochrany spojená se zpřístupněním je typická především pro rukopisy,

²⁴ <https://books.google.com/>

²⁵ <https://www.eeagrants.cz/>

²⁶ MĚŘÍNSKÁ, Eva - VOJÍŘOVÁ, Kateřina: Projekt HISPRÁ - záchrana historických pragensií z fondu Městské knihovny v Praze (MKP). (Projekt digitalizace podpořený tzv. "norskými fondy"). In: *Knihovny současnosti 2008 : sborník z 16. konference, konané ve dnech 16.-18. září 2008 v Seči u Chrudimi /* Brno: Sdružení knihoven ČR, 2008 s. 27-33. [cit. 2019-11-26]. Dostupné z: <http://sdruk.mlp.cz/data/xinha/sdruk/2008-0-027.pdf>

²⁷ <https://books2ebooks.eu/>

staré tisky a noviny, které trpí degradací kyselého papíru a každé fyzické zapůjčení knihovní jednotky tak může napáchat nenapravitelné škody. Některé originály novin jsou dnes v již tak špatném stavu, že je nelze čtenářům vůbec půjčovat a digitalizace (v minulosti též mikrofilmování) je jediným způsobem, jak je čtenářům **zpřístupnit**. Dalším důležitým aspektem je **ochrana před ztrátou, krádeží a dalším poničením**. Digitalizace též přináší výhody v podobě **vzdáleného přístupu** odkudkoli²⁸, což je pro čtenáře výhodné zejména v případě potřeby studia rukopisů nebo dalších vzácných dokumentů, jako jsou např. staré mapy, které bývají dostupné pouze pro badatele a ve specializovaných studovnách s omezenou otevírací dobou. Čtenář se tedy nemusí teoreticky do knihovny vůbec vydat a má možnost dokumenty studovat z pohodlí domova. Výhodou je též, že dokument si může **ve stejnou chvíli prohlížet více čtenářů** zároveň a pořizovat z nich kopie. Texty navíc bývají, pokud je to možné²⁹, opatřené **OCR** umožňující **fulltextové vyhledávání**, které čtenáři šetří čas. Je třeba si uvědomit, že digitalizovaná strana dokumentu bez OCR je v podstatě obyčejný obrázek, který lze pouze prohlížet, ale nelze na něm vyhledávat. Zde je však nutné počítat možnými chybami, neboť kvalita fulltextu závisí vždy na vzhledu a kvalitě předlohy.

2.3 Proces plánování digitalizace

Proces plánování je možné rozdělit na oblasti výběru vhodných **dokumentů, architektonického řešení, archivaci a zpřístupnění dat a financování**.

Je třeba si rozmyslet, jaké **typy dokumentů** chceme digitalizovat a v jakém **objemu**. Je dobré si uvědomit **specifika** a **stav** předlohy, tj. jaké typy skenerů (včetně jejich počtů) budou potřeba, neboť z hlediska formátu je velký rozdíl mezi digitalizací např. starých map a monografií.

Další úvahy by měly směřovat k logické **kompletnosti dokumentů**, neboť není vhodné digitalizovat např. vícesvazkové dílo, z něhož knihovna nevlastní či nemá možnosti získat všechny svazky. Spolupráce knihoven je v tomto ohledu běžná. Jednotlivé dokumenty musí být opatřeny jedinečnými **identifikátory** (např. signatura, též příp. čárový kód³⁰) a je třeba mít k nim zpracované a zkontrolované kompletní **bibliografické záznamy** v katalogu. Exempláře samozřejmě musí být ve fyzickém stavu, který provedení digitalizace umožňuje.

Dále je nutné správně zvolit formu **architektonického řešení** a položit si základní otázky: Jakým způsobem bude probíhat samotná digitalizace? Bude se jednat o **vlastní řešení** (in the house) či knihovna zajistí **externí digitalizaci** přes jinou instituci nebo soukromou firmu? Z procesu plánování je nutné

²⁸ pro autorsky volná díla, pozn. autora

²⁹ s výjimkou rukopisů; v současnosti se nicméně vyvíjí software schopný rozpoznat ručně psané písmo a zajistit tak OCR i těch nejstarších dokumentů z českých knihoven, viz <https://pero.fit.vutbr.cz/>

³⁰ Jedná o pouze lokálně jedinečné identifikátory v rámci dané knihovny.

nevynechat rozhodnutí týkající se **kvality skenů** (např. volbu DPI, barevné škály apod.), jejich **formát** (včetně OCR) a jakým způsobem má být veden jejich **ořez**. Neméně důležitou částí je i volba **zápisu metadat**, tj. pouze na metadata popisná, strukturální, technická či administrativní a jaký formát zvolit. Je také nutné se rozhodnout, do jaké hloubky chceme dokumenty popisovat, tj. zda chceme u časopisů dojít na úroveň čísel nebo dokonce až na samostatné články, případně zda chceme v rámci skenů rozlišovat kapitoly u monografií. V rámci zkušeností s digitalizací v českých knihovnách a vzhledem k většinovému používání systému Kramerius důrazně doporučujeme před vlastním řešením prostudovat **standards pro digitalizaci pro dokumenty NDK**³¹, které vyhláší Národní knihovna ČR. Tyto standardy používá většina knihoven v ČR a jejich dodržování pak usnadňuje spolupráci mezi knihovnami a případnou replikaci dat mezi sebou.

Po úspěšné digitalizaci přichází neméně důležitá otázka **zpřístupnění a dlouhodobého uložení dat**. Na základě plánů ohledně objemu digitalizovaných dokumentů je třeba odhadnout, kolik místa na disku a ve zvoleném úložišti dat zaberou digitalizované dokumenty. Je třeba počítat s tím, že je potřeba uchovávat vždy minimálně uživatelskou a archivační kopii. V rámci zpřístupnění je nutné zvolit **metodu zpřístupnění** digitalizovaných dokumentů, tj. zda bude zvolena **digitální knihovna** (např. Kramerius) či nebo jiné řešení, např. zveřejnění dokumentů k prohlížení či volnému stahování na zvolených **webových stránkách**. Umožnění volného stahování v jakékoli podobě však znamená nutnost posouzení dokumentů na **autorská práva**, neboť nelze umožnit přístup a získání digitální kopie bez licence či uplynutí majetkových autorských práv. Zpřístupnění digitalizovaných dokumentů je možné v rámci **knihovní licence**³² pouze v budově knihovny na počítačích, kde dokumenty budou ošetřeny proti šíření rozmnoženiny v digitální podobě. Do autorsky chráněných dokumentů je možné v rámci licence v knihovnách registrovaným čtenářům pouze nahlížet a tisknout.

Závěrem je třeba zhodnotit celkovou **finanční náročnost** projektu a posoudit, zda je v silách knihovny zajistit jeho financování. Je nutné stanovit, jak dlouho bude práce na projektu probíhat, kolik **úvazků** na něj bude vyčleněno, zda má knihovna patřičné **prostory** a zda vlastní dostatečné množství **výpočetní techniky** na dostatečné výkonnostní úrovni. Finanční plán by měl kromě případného nákupu počítačů a další techniky též zahrnovat **nákup software**, pokud se knihovna nerozhodne celý proces digitalizace svěřit externímu dodavateli. Je nutné též počítat s **dlouhodobým finančním zajištěním uložení dat**, ať už se knihovna rozhodne jít cestou ukládání na server či disky nebo zvolí cloudové řešení.

³¹ <https://www.ndk.cz/standards-digitalizace>

³² § 37 zákona č. 121/2000 Sb.

2.4 Výběr dokumentů k digitalizaci

Současný trend v českých i zahraničních knihovnách začíná **upouštět od masové digitalizace** vlastních fondů a přechází spíše na **digitalizaci výběrovou**. Masová digitalizace fondů je z hlediska vynakládaných zdrojů velmi náročná a jen málo zohledňuje např. obsahovou kvalitu dokumentů. Je tomu tak z velké části z důvodu, že fondy bývají v knihovných skladech běžně poskládány podle formátu knih a nikoli tematicky. Problémy činí též (zejména u větších knihoven) nekompletní katalogizace starší produkce a rychlost retrokonverze. Starší fond též často nemá přidělené číslo **České národní bibliografie**³³ (ČNB) a čekání na jeho přidělení může trvat delší dobu. Digitalizační pracoviště jsou omezena kapacitně i technicky. Každý dokument vyžaduje speciální zacházení a nehodí se pro každý typ skeneru. Na kulturně či historicky hodnotné knihy a další vzácné kusy se tak může v rámci plánované digitalizace čekat i několik let. Masová digitalizace je tedy vhodná spíše pro zastřešující národní instituce, které jsou z historických důvodů příjemci povinného výtisku (např. NK ČR a MZK).

Pro výběrovou digitalizaci mluví v současné době i fakt, že při existenci a trvání velkých národních projektů, jako je např. Národní digitální knihovna, není takřka možné v běžném fondu nenarazit na dokumenty, které již byly digitalizovány jinde. České knihovny se pokouší **zamezit vzniku duplicitních záznamů** a z tohoto důvodu je nutné ve fázi plánování prověřit digitalizační plány dalších institucí a před samotnou digitalizací ověřit, zda je zvolený dokument skutečně možné v digitalizační lince zpracovat a zpřístupnit. Ke kontrole možných duplicit slouží českým knihovnám nástroj **Registr digitalizace**³⁴.

2.4.1 Kritéria pro výběr dokumentů

Na základě výše uvedených skutečností a na základě rešerše zahraničních zdrojů byla definována měřítko pro výběr dokumentů k digitalizaci, která mohou být v současnosti velmi dobře aplikována i na české prostředí, zejména pro menší knihovny.

³³ Číslo ČNB je unikátní identifikátor, který má za cíl jasně identifikovat českou národní produkci 19., 20. a 21. století. Číslo ČNB by měly mít veškeré publikované dokumenty vydané na území ČR od roku 1801 do současnosti, s výjimkou tzv. šedé literatury. Přiděluje jej Národní knihovna. Viz: <https://www.caslin.cz/caslin/spoluprace/sluzby-souborneho-katalogu-cr/cislo-cnb-v-sk-cr>

³⁴ <http://registrdigitalizace.cz>

Výběr dokumentů podléhá následujícím kritériím³⁵:

1. **kulturní, historická či informační hodnota dokumentu** a jeho unikátnost pro region či oblast, kde se knihovna nachází,
2. **využívanost dokumentu** (počet požadavků/výpůjček),
3. **dostupnost originálu** (např. umístění v externím skladu, pouze jedna kopie), stav originálu (ve smyslu ochranné digitalizace).

Kromě uvedených kritérií je vhodné přihlídnout i k dalším aspektům:

1. **(ne)možnost zveřejnění dokumentu** (preference autorsky volných dokumentů), příslušnost dokumentu ke sbírce nebo možnost ho do (virtuální) sbírky zařadit,
2. **využití digitalizovaného dokumentu** pro prezentaci knihovny vůči veřejnosti.

Knihovna by měla napláňovat na základě uvedeného klíče uskutečnitelnou **strategii digitalizace**, které by se měla v budoucnu pevně držet.

PŘÍPADOVÁ STUDIE

Příklad aplikování strategie digitalizace a následný výstup zpracovaných dat prostřednictvím digitální knihovny si představíme na případové studii v Moravské zemské knihovně v Brně:

PŘÍPADOVÁ STUDIE

Digitalizační centrum MZK se řídí v rámci digitalizace následujícími kritérii (v tomto pořadí priorit)³⁶:

1. **Ochranná digitalizace** dokumentů, kterým hrozí zánik (u seriálů je účelné zpracovat titul kompletně, totéž u vícesvazkových monografických děl apod.),
2. **Unikátní a vzácně dochované dokumenty** (nejen rukopisy, ale např. i regionalia – tj. tam, kde je málo dochovaných výtisků),
3. **Digitalizace na žádost** čtenářů a institucí,
4. **Digitalizace podle půjčovnosti** (podle kapacity vhodných skenerů Treventus, popř. Plustek),

³⁵ HRUŠKA, Zdeněk a kol.: Výběr dokumentů k digitalizaci dle zahraniční praxe. In: HRUŠKA, Zdeněk a kol.: Strategie digitalizace v Moravské zemské knihovně v Brně. Brno : Moravská zemská knihovna v Brně, 2017 [cit. 2019-11-17]. S. 12. Dostupné z:

https://www.mzk.cz/sites/mzk.cz/files/souboryMZK/pdf/strategie_digitalizace_mzk_02-2017_v1-0.pdf

³⁶ HRUŠKA, Zdeněk a kol.: Strategie digitalizace v MZK. In: HRUŠKA, Zdeněk a kol.: Strategie digitalizace v Moravské zemské knihovně v Brně. Brno : Moravská zemská knihovna v Brně, 2017 [cit. 2019-11-17]. S. 7. Dostupné z:

https://www.mzk.cz/sites/mzk.cz/files/souboryMZK/pdf/strategie_digitalizace_mzk_02-2017_v1-0.pdf

5. **Tematická digitalizace** (např. podle bibliografií, bohemikální tisky, staré hudebniny nebo podle sbírek, pokud jsou ve fondu identifikovatelné – např. dle signatury nebo druhu dokumentu),
6. **Digitalizace v rámci projektů** (např. EoD),
7. **Ochranná digitalizace** dokumentů, kterým hrozí zcizení,
8. **Destruktivní digitalizace**³⁷ (průběžně, podle kapacity skeneru a množství dostupných dokumentů),
9. **Systematická digitalizace** části fondu dle umístění.

Digitální knihovna Moravské zemské knihovny je pro čtenáře zpřístupněna na webu <http://www.digitalniknihovna.cz/mzk> a obsahuje následující typy dokumentů:

1. **noviny a časopisy** (např. starší vydání Lidových novin včetně článkové bibliografie, Národní listy, Rovnost atd.),
2. **rukopisy a staré tisky**,
3. **mapy**,
4. **archiválie**,
5. **hudebniny**,
6. **knihy** – výběrově novější česká i zahraniční literatura,
7. **obsahy CD z fondu MZK** (ve formátu PDF), najdete je mezi knihami,
8. **zvukové dokumenty** – nahrávky z gramofonových desek.

Fond digitální knihovny MZK je rozdělen na **tematické sbírky** obsahující vybrané dokumenty, které jsou k dispozici jak prostřednictvím sbírky, tak ve volném vyhledávání jednotlivých dokumentů.

Přístup k dokumentům je omezen autorskými právy. Dokumenty obsažené v digitální knihovně se dělí na:

- **veřejné** – volná díla, nevztahují se na ně autorská ani nakladatelská práva, jsou přístupné odkudkoliv,
- **neveřejné** – autorsky chráněná díla, která jsou chráněna autorskými či nakladatelskými právy; prohlížet a tisknout je možné pouze na počítačích v dané knihovně v jedné ze studoven; v tomto režimu není možné dokumenty stahovat ani kopírovat.

Autorsky chráněné dokumenty jsou přístupné registrovaným čtenářům po přihlášení do některého z počítačů ve studovnách v Moravské zemské knihovně. V aplikaci Digitální knihovna lze tyto dokumenty pouze prohlížet, či z nich tisknout. Vytváření a šíření digitálních kopií není možné. Registrace v knihovně je možná pro všechny občany České republiky po předložení občanského průkazu, pro občany žijící na území EU je nutné předložit platný

³⁷ Jedná se o metodu digitalizace, kdy může být knize odříznut hřbet knihy, následně je uvolněna vazba a celý knižní blok je rozdělen na jednotlivé stránky. Tato metoda se volí pro práci s automatizovanými a moderními skenery, které jsou navrženy pro automatický posun stránek. Knihy je možné po procesu digitalizace znovu sešít.

doklad totožnosti. Občané dalších států k získání čtenářského průkazu potřebují navíc vízum platným od data registrace nejméně na jeden rok.

SHRnutí STUDIjNíHO TEXTU

Digitalizace dokumentů patří mezi důležité knihovní procesy, které pomáhají chránit fondy knihoven a zároveň ve formě systematického budování digitální knihovny poskytují moderní službu, která čtenářům nabízí pohodlí při práci s dokumenty a zároveň šetří jejich čas. Kvalita služby se odvíjí od volby uživatelsky přívětivého a webového klienta optimalizovaného pro desktopová i mobilní zařízení a obsahu digitální knihovny jako takové.

Volba dokumentů vhodných k digitalizaci hraje klíčovou roli při budování digitální knihovny a měla by být součástí oficiální strategie digitalizace instituce. Ta by měla být v ideálním případě dostupná na webu knihovny, aby z ní mohly čerpat informace i další instituce a nedocházelo tak napříč (nejen) knihovnami k zbytečné duplikaci práce. Knihovny by měly využívat všech dostupných prostředků pro vyvarování se duplicitní digitalizace a vzájemně se informovat o vlastních digitalizačních plánech na další roky.

Každý dokument vkládaný do digitální knihovny by měl splňovat kritéria, která by měla být zvolena v souladu s digitalizační politikou instituce. Záleží však pouze na knihovně, jakou cestu pro jejich výběr zvolí. Každá knihovna by měla pečlivě zvážit, pro jaké uživatele bude obsah digitální knihovny komponovat, jaké k tomu má k dispozici prostředky a dokumenty ve fondu a digitalizační politiku tomuto přizpůsobit.

LITERATURA

ARMS, William Y. Digital libraries. Cambridge, Mass: MIT press, 2000.

BARTOŠEK, Miroslav. Digitální knihovny – teorie a praxe. Národní knihovna, 2004, 15(4), s. 233-254. ISSN 0862-7487. Dostupné také z: <http://oldknihovna.nkp.cz/NKKR0404/0404233.html>

BRATKOVÁ, Eva. Otevřený přístup a digitální knihovny v oblasti vědy a výzkumu: (vybrané systémy) [online]. Verze 1.0. Praha: Ústav informačních studií a knihovnictví FF UK v Praze, prosinec 2008 [cit. 2015-11-24]. 126 s. Formát PDF. Elektronické studijní texty ÚISK. Dostupný z: https://sites.ff.cuni.cz/uisk/wp-content/uploads/sites/62/2016/01/Otev%C5%99en%C3%BD-p%C5%99%C3%ADstup-a-digit%C3%A1ln%C3%AD-knihovny-v-oblasti-v%C4%9Bdy-a-v%C3%BDzkumu-vybran%C3%A9-syst%C3%A9my_Bratkov%C3%A1.pdf

DVOŘÁKOVÁ, Helena. Číslo ČNB v SK ČR. *Souborný katalog České republiky* [online]. 10.5.2012 [cit. 2014-01-23]. Dostupné z: <https://www.caslin.cz/caslin/spoluprace/sluzby-souborneho-katalogu-cr/cislo-cnb-v-sk-cr>

GREENBERG, Douglas. 1998. Camel drivers and gatecrashers : quality control in the digital research library. In HAWKINS, B.L.; BATTIN, P. (ed.). *The mirage of continuity : reconfiguring academic information resources for the 21st century*. Washington (D.C.) : Council on Library and Information Resources; Association of American Universities, 1998, s. 105-116.

KRAVECOVÁ, Lenka; KRUTSKÁ, Markéta; RYCHTÁŘOVÁ, Pavla: Úvod do digitalizace. In *Moravská zemská knihovna v Brně. E-learningový kurz Digitalizace v knihovnách* [online - e-learningový kurz]. Brno : Moravská zemská knihovna v Brně, 2019 [cit. 2019-11-17]. Dostupné z: <https://www.mzk.cz/pro-knihovny/vzdelavani-knihovniku/e-learningovy-kurz-digitalizace-v-knihovnach>

KRAVECOVÁ, Lenka; KRUTSKÁ, Markéta; RYCHTÁŘOVÁ, Pavla: Zpřístupnění. In *Moravská zemská knihovna v Brně. E-learningový kurz Digitalizace v knihovnách* [online - e-learningový kurz]. Brno : Moravská zemská knihovna v Brně, 2019 [cit. 2019-11-17]. Dostupné z: <https://www.mzk.cz/pro-knihovny/vzdelavani-knihovniku/e-learningovy-kurz-digitalizace-v-knihovnach>

MĚŘÍNSKÁ, Eva - VOJÍŘOVÁ, Kateřina: Projekt HISPRA - záchrana historických pragensí z fondu Městské knihovny v Praze (MKP). (Projekt digitalizace podpořený tzv. "norskými fondy"). In: *Knihovny současnosti 2008 : sborník z 16. konference, konané ve dnech 16.-18. září 2008 v Seči*

u Chrudimi/Brno : Sdružení knihoven ČR, 2008 s. 27-33. [cit. 2019-11-26].
Dostupné z: <http://sdruk.mlp.cz/data/xinha/sdruk/2008-0-027.pdf>

SIKORA, Radek. Vyhledávání v českých dokumentech pomocí Apache Solr .
Brno, 2012. Dostupné z: http://is.muni.cz/th/256499/fi_m/thesis_sikora.pdf .
Diplomová práce. Masarykova univerzita, Fakulta informatiky.

VAN DE SOMPEL, Herbert.; HOCHSTENBACH, Patrick. Reference Linking in
a Hybrid Library Environment. Part 3, Generalizing the SFX solution in the
"SFX@Ghent. 1999. Dostupné také
z: <http://search.proquest.com/docview/57491073?accountid=35514>

HRUŠKA, Zdeněk a kol.: Strategie digitalizace v Moravské zemské knihovně
v Brně. Brno : Moravská zemská knihovna v Brně, 2017 [cit. 2019-11-17].
Dostupné
z: https://www.mzk.cz/sites/mzk.cz/files/souboryMZK/pdf/strategie_digitalizace_mzk_02-2017_v1-0.pdf

Název: Koncipování obsahu digitální knihovny
Autor: Bc. Martina Smetánková
Jazyková korektura: Markéta Roupcová
Počet stran: 22