

# **Digitalizace a dlouhodobá ochrana fondů 2**

Distanční studijní text

**Dušan Katuščák**

**Opava 2019**



- Obor:** Knihovnictví, bibliografie, digitalizace, dokumentace, informační věda (sociální aspekty), Library and Information Science (LIS)  
Klasifikační kódy: 00, 022, 030, 0322, 038
- Klíčová slova:** *knihovnictví, bibliografie, dokumentace, dokumentologie, informační věda (sociální aspekty), digitalizace, uchovávání, konzervování, Library and Information Science (LIS), workflow digitalizace, digitální knihovna, on-line katalog, digitální repozitář, služby knihoven, digitální repozit, deskriptivní metadata*
- Anotace:** Předmět seznámí studenty s problematikou digitalizace knih a částečně i digitalizací zvuků a dokumentů. Důraz je kladen na analýzu a praktické využívání projektů digitalizace (Project Gutenberg, Google books, Manuscriptorium, Kramerius) a plánování a realizaci projektu (workflow) digitalizace v oboru knihovnictvo. V druhé části jsou studenti seznámeni s problematikou  
důrazem na plánování a realizaci  
procesu. Cvičení jsou zaměřena na praktické vyzkoušení si celého procesu digitalizace a uchovávání dokumentů.  
é digitalizační pracoviště.

## Obsah

<b>1</b>	<b><i>Významné digitalizační světové projekty - Project Gutenberg</i></b> .....	<b>8</b>
1.1	História .....	8
1.2	Záber zbierky.....	9
1.3	Ideály .....	9
1.4	Autorské práva.....	9
1.5	Kritika .....	10
1.6	Príbuzné projekty .....	10
<b>2</b>	<b><i>Významné digitalizační světové projekty - Google Books - Knihy Google</i></b> .....	<b>12</b>
2.1	Projekt Knihy Google .....	12
2.2	Spolupráce s knihovnami.....	13
2.3	Postup přispívání do Google Books.....	13
2.3.1	Výběr a ocenění .....	13
2.3.2	Katalogizace .....	13
2.3.3	Konzervování.....	14
2.3.4	Administrativa vývozu. Pojištění.....	14
2.3.5	Expedice .....	14
2.3.6	Vrácení. Přebírání digitálního obsahu.....	14
2.4	Soudní spory .....	14
2.5	Spolupráce Google s Národní knihovnou České republiky.....	15
2.6	Závěr .....	17
2.7	Odkazy .....	17
2.8	Související články .....	18
2.9	Literatura .....	18
<b>3</b>	<b><i>Významné digitalizační světové projekty - Europeana</i></b> .....	<b>19</b>
3.1	Europeana (Comité des Sages). Pozícia Európskej únie. ....	19
3.2	.....	20
3.3	Stimuly digitalizácie.....	21
3.4	Úloha Europeany.....	21
3.5	.....	22
3.6	Udržateľnosť .....	22
3.7	PPP projekty.....	23
<b>4</b>	<b><i>Významné digitalizační světové projekty – Europeana . Zbierky</i></b> .....	<b>24</b>
4.1	Ktoré dáta sa dávajú do Europeany? .....	26
4.2	Ktoré druhy dokumentov obsahuje Europeana? .....	26
4.3	Aké sú technické požiadavky .....	27
<b>5</b>	<b><i>Svetová digitálna knižnica</i></b> .....	<b>28</b>

<b>6</b>	<b><i>Digitální knihovna Manuscriptorium</i></b>	<b>29</b>
6.1	Abstrakt	29
6.2	Klíčová slova	29
6.3	O projektu. Počátky	29
6.4	Interoperabilita	30
6.5	Rychlý rozvoj spolupráce	30
6.6	Principy spolupráce	31
6.7	Distribuovaná obrazová úložiště nebo jediné centrální?	32
6.8	Vysoké náklady na implementaci a udržování centrálního obrazového repozitálu. 33	
6.9	Neochota partnera předat do cizí správy vlastní hodnotná obrazová data	33
6.10	Proč s námi spolupracovat	34
6.11	Zahájení spolupráce	34
6.12	Technické minimum	34
6.13	Administrativní úkony	35
6.14	Vlastní zahájení spolupráce	35
6.15	Nástroje pro koncové uživatele	35
6.16	Kde je Manuscriptorium dostupné	36
<b>7</b>	<b><i>Česká digitální knihovna Kramerius</i></b>	<b>37</b>
7.1	Abstrakt	37
7.2	Klíčová slova	37
7.3	Úvod	37
7.4		37
7.4.1	Úvod	37
7.4.2	Digitální knihovna KRAMERIUS	38
7.4.3	Registr digitalizace	38
7.4.4	—	39
7.4.5	Česká digitální knihovna — portál	39
7.4.6	Závěr	39
7.4.7	Reference	40
<b>8</b>	<b><i>Plánování a příprava procesu digitalizace v knihovně</i></b>	<b>41</b>
8.1	Praktické odporúčania	41
8.2	Projekt Minerva a Lundske zásady	41
8.3	Lundske zásady	41
8.4	Plánovanie projektu digitalizácie	42
8.5	Zdôvodnenie projektu	43
8.6	Ľudské zdroje	43
8.7	Výskum	44
8.8	Hrozby	45

<b>8.9</b>	<b>Výber zdrojového materiálu na digitalizáciu .....</b>	<b>45</b>
<b>8.10</b>	<b>Stanovenie kritérií výberu .....</b>	<b>45</b>
<b>8.11</b>	<b>Výber podľa kritérií .....</b>	<b>46</b>
<b>8.12</b>	<b>Príprava na digitalizáciu .....</b>	<b>47</b>
8.12.1	Hardvér.....	47
8.12.2	Softvér .....	49
<b>8.13</b>	<b>Proces digitalizácie .....</b>	<b>50</b>
8.13.1	Používanie skenerov.....	50
8.13.2	Používanie digitálnych kamier/fotoaparátov .....	52
8.13.3	Softvérové aplikácie na optické rozoznávanie znakov (OCR).....	53
<b>8.14</b>	<b>Metaúdaje (metadáta) .....</b>	<b>53</b>
8.14.1	Metaúdaje použité pre opis predmetu (deskriptívne metadáta) .....	54
8.14.2	Vyhovujúce štandardy pre metaúdaje .....	54
<b>8.15</b>	<b>Zverejnenie .....</b>	<b>55</b>
<b>8.16</b>	<b>Spracovanie obrazu .....</b>	<b>55</b>
<b>8.17</b>	<b>Trojrozmerné materiály a virtuálna realita .....</b>	<b>56</b>
<b>8.18</b>	<b>On-line zverejnenie .....</b>	<b>57</b>
<b>8.19</b>	<b>Práva duševného vlastníctva a copyright .....</b>	<b>58</b>
<b>8.20</b>	<b>Právne aspekty.....</b>	<b>58</b>
8.20.1	Autorské práva (Copyrights).....	58
8.20.2	Vydavateľské a licenčné zmluvy .....	58
8.20.3	Copyright v knižnici podľa amerického a austrálskeho práva .....	58
8.20.4	Praktický príklad .....	59
8.20.5	Dispenzácia s formálnym povolením .....	59
8.20.6	Vytvorenie copyrightu.....	59
8.20.7	Ochrana copyrightu.....	60
<b>8.21</b>	<b>Manažment projektov digitalizácie.....</b>	<b>61</b>
8.21.1	Riadenie procesu digitalizácie .....	61
8.21.2	Vytvorenie tímu.....	62
<b>8.22</b>	<b>Školenie zamestnancov .....</b>	<b>62</b>
8.22.1	Získavanie technického poradenstva od tretích strán .....	63
8.22.2	Spolupráca s projektovými partnermi a vytvorenie spoločného obsahu.....	64
<b>8.23</b>	<b>Náklady.....</b>	<b>64</b>
8.23.1	Orientačné ceny digitalizácie viazaných kníh .....	66
8.23.2	Ďalej uvedené informácie ilustrujú finančnú náročnosť digitalizácie tlačených viazaných dokumentov na základe informácií o službách komerčnej digitalizácie s použitím špičkovej technológie.....	66
8.23.3	Odhad nákladov na digitalizáciu slovacikálneho fondu kníh v Slovenskej digitálnej knižnici (SDK) .....	66
8.23.4	Náklady na digitalizáciu tlačených viazaných dokumentov v SDK.....	66
<b>8.24</b>	<b>Prostredie .....</b>	<b>67</b>
<b>9</b>	<b><i>procesu digitalizace.....</i></b>	<b>69</b>
<b>9.1</b>	<b>Digitálna knižnica a digitálny archív (DIKDA). Slovensko .....</b>	<b>69</b>
<b>9.2</b>	<b>Proces digitalizácie prebieha: .....</b>	<b>69</b>
<b>9.3</b>	<b>Digitalizačný workflow .....</b>	<b>70</b>

<b>9.4</b>	<b>Konzervácia dokumentov (konzervovanie)</b> .....	<b>76</b>
9.4.1	Reštaurovanie .....	76
<b>9.5</b>	<b>Integrácia konzervovania a digitalizácie</b> .....	<b>76</b>
9.5.1	Sterilizácia .....	77
9.5.2	Deacidifikácia .....	77
9.5.3	Lyofilizácia .....	77
<b>10</b>	<b><i>Plánování procesu dlouhodobého uchování digitálních děl v knihovnictví ....</i></b>	<b>78</b>
10.1	Uchovávanie originálneho digitálneho materiálu .....	78
10.2	Formáty súborov .....	78
<b>11</b>	<b><i>Postupy dlouhodobého uchování digitálních děl.....</i></b>	<b>80</b>
11.1	Výber médií.....	80
11.2	Migračné stratégie .....	81
<b>12</b>	<b><i>Práce s původními digitálními dokumenty .....</i></b>	<b>82</b>
12.1	Zaobchádzanie a práca s originálmi .....	82
12.2	Premiestňovanie a manipulácia s originálmi .....	82
<b>13</b>	<b><i>Práce s puvodními digitálními dokumenty .....</i></b> <i>Chyba! Záložka není definována.</i>	
13.1	Zaobchádzanie a práca s originálmi .....	<i>Chyba! Záložka není definována.</i>
13.2	Premiestňovanie a manipulácia s originálmi .....	<i>Chyba! Záložka není definována.</i>

<b>B-III – Charakteristika studijního předmětu</b>				
Název studijního předmětu	Digitalizace a dlouhodobá ochrana fondů 2			
Typ předmětu	Povinný, ZT (společný základ)	doporučený ročník / semestr	1/LS	
Rozsah studijního předmětu	13p+26s	hod.	39 h/sem.	kreditů 6
Prerekvizity, korekvizity, ekvivalence				
Způsob ověření studijních výsledků	Zápočet, zkouška	Forma výuky	Přednáška, seminář	
Forma způsobu ověření studijních výsledků a další požadavky na studenta	Zápočet: praktické prokázání dovedností věcného pořádkání informací Zkouška: písemný test			
Garant předmětu	Prof. PhDr. Dušan Katušák, Ph.D.			
Zapojení garanta do výuky předmětu	Vedení přednášek.			
Vyučující	Prof. PhDr. Dušan Katušák, Ph.D. (přednášející, 100 %)			
<b>Stručná anotace předmětu</b>				
<b>Cíl výuky</b>				
Předmět seznámí studenty s problematikou digitalizace knih a částečně i digitalizací zvukových a obrazových záznamů. Důraz je kladen na analýzu projektů digitalizace (Project Gutenberg, Google books, Manuscriptorium, Kramerius) a plánování a realizaci projektu digitalizace v oboru knihovnictvo. V druhé části jsou studenti seznámeni s problematikou dlouhodobého uchování výsledků digitalizace s důrazem na plánování a realizaci procesu. Cvičení jsou zaměřeny na praktické vyzkoušení si celého procesu digitalizace a uchování dokumentů. Součástí předmětu je i exkurze na významné digitalizační pracoviště.				
<b>Osnova – bloky</b>				
<ol style="list-style-type: none"> <li>1. Významné digitalizační světové projekty – projekt Gutenberg</li> <li>2. Významné digitalizační světové projekty – projekt Google books</li> <li>3. Významné digitalizační světové projekty – projekt Europeana. koncept</li> <li>4. Významné digitalizační světové projekty – projekt Europeana. sbírky</li> <li>5. Významné digitalizační světové projekty – projekt World Digital Library</li> <li>6. Významné digitalizační české projekty – projekt Manuscriptorium</li> <li>7. Významné digitalizační české projekty – projekt Kramerius</li> <li>8. Plánování a příprava procesu digitalizace v knihovně</li> <li>9. Řešení technologických aspektů procesu digitalizace</li> <li>10. Plánování procesu dlouhodobého uchování digitálních děl v knihovnictví</li> <li>11. Postupy dlouhodobého uchování digitálních děl</li> <li>12. Práce s původními digitálními dokumenty</li> </ol>				
<b>Studijní literatura a studijní pomůcky</b>				
<b>Povinná literatura</b>				
Studijní materiály, prezentace a učební text umístěné na webových stránkách Ústavu bohemistiky a knihovnictví FPF SU v Opavě – oddělení knihovnictví – webová adresa: <a href="http://knihovnik.fpf.su.cz/">http://knihovnik.fpf.su.cz/</a>				
KONVIT, M. <i>Digitization and Long-term Fund Protection I, II</i> . Opava: SU v Opavě, 2017.				
<b>Doporučená literatura</b>				
ČNK. <i>Digitalizace a digitální zpřístupnění dokumentů</i> . Online. Dostupné na <a href="https://www.nkp.cz/o-knihovne/odbornosti/sprava-a-ochrana-fondu/zzz_osof/digitalizace">https://www.nkp.cz/o-knihovne/odbornosti/sprava-a-ochrana-fondu/zzz_osof/digitalizace</a>				
<i>Archivy, knihovny, muzea v digitálním světě 2005</i> . Praha: Národní technické muzeum, 2006 - 121 s. ISBN 80-7037-149-8				
<i>World Digital Library</i> . Online. Dostupné na <a href="https://www.wdl.org/">https://www.wdl.org/</a>				
<i>Europeana Collections</i> . Online. Dostupné na <a href="http://www.europeana.eu/">www.europeana.eu/</a>				
<b>Informace ke kombinované nebo distanční formě</b>				
Rozsah konzultací (soustředění)				hodin
Popis systému kontaktu s vyučujícím				

## 1 VÝZNAMNÉ DIGITALIZAČNÍ SVĚTOVÉ PROJEKTY - PROJECT GUTENBERG

Projekt bol založený v roku 1971 a je najstaršou digitálnou knižnicou. **Projekt Gutenberg** (angl. **Project Gutenberg**, často sa skrakuje na PG) je dobrovoľnícka snaha digitalizovať, archivovať a distribuovať kultúrne diela. Väčšina diel sú plné texty kníh so statusom voľné dielo. Projekt sa snaží sprístupniť diela zo svojej zbierky tak slobodne, ako sa len dá, v dlhotrvajúcich otvorených formátoch, ktoré je možné používať takmer na každom počítači.

Oficiálna stránka digitálnej knižnice Gutenberg: <http://www.gutenberg.org>

Študenti získajú vedomosti o histórii vzniku projektu digitálnej knižnice, o tom, aké zbierky a dokumenty sa do tejto knižnice zaraďujú, o zámeroch projektu, o autorských právach v projekte a o rôznych podobných projektoch na národnej úrovni. Zoznámia sa prakticky s knižnicou, ktorá v roku 2019 obsahovala okolo 60 000 digitalizovaných diel v rôznych jazykoch.

### 1.1 História

Projekt Gutenberg začal Michael Hart v roku 1971. Hart, študent University of Illinois dostal prístup k veľkému počítaču Xerox Sigma V na univerzite. Vďaka tomu, že jeho priateľ bol operátorom, dostal k dispozícii účet s prakticky neobmedzeným množstvom výpočtového času; jeho hodnota sa odvodila odhadovala rôzne od 100 000 do 100 000 000 USD. Hart povedal, že sa chcel odvdáčiť za tento dar tým, že bude robiť niečo, čo by bolo hodné takej sumy peňazí.

Tento počítač bol zhodou okolností jedným z 15 uzlov počítačovej siete, ktoré sa stali internetom. Hart veril, že počítače sa jedného dňa stanú dostupnými širokej verejnosti a rozhodol sa slobodne sprístupniť literárne diela v elektronickej podobe. Náhodou mal v ruksaku Deklaráciu nezávislosti Spojených štátov a táto sa stala prvým e-textom Projektu Gutenberg. Projekt pomenoval podľa Johanna Gutenberga, nemeckého tlačiara z 15. storočia, ktorý spustil tlačiarenskú revolúciu pohyblivými znakmi. V polovici 90. rokov bol sídlom PG Illinois Benedictine College. K snahe sa pripojilo viac dobrovoľníkov. Väčšina textu sa vpisovala ručne, kým sa nezlepšili a nerozšírili scannery a softvér pre optické rozpoznávanie znakov (OCR). Hart neskôr uzavrel dohodu s Carnegie Mellon University, ktorá súhlasila so správou financií projektu. Ako narastal objem e-textov, dobrovoľníci postupne začali preberať každodenné úlohy, ktoré Hart vykonával.

V roku 2000 založil Charles Franks projekt Distributed Proofreaders, ktorý umožnil distribúciu kontroly naskenovaných textov medzi množstvo dobrovoľníkov prostredníctvom internetu. Táto snaha v obrovskej miere zvýšila množstvo a rôznorodosť textov, ktoré sa pridávajú do PG, ako aj zjednodušila začínajúcim dobrovoľníkom začať so svojimi prvými príspevkami. Pietro Di Miceli, taliansky dobrovoľník, vyvinul a spravoval prvý web Projektu Gutenberg a začal s vývojom online katalógu projektu. Počas desiatich rokov, kedy sa zhostoval tejto úlohy (1994 – 2004), web projektu vyhral niekoľko ocenení, často mu boli udelené označenia „najlepšie na Webe“, čo prispelo k popularite projektu.



Počínajúc v roku 2004, vylepšený online katalóg zjednodušil prístup k obsahu PG pre prezeranie, vyhľadávanie, prístup a odkazovanie naň.

Projekt Gutenberg je v súčasnosti hosťovaný na ibiblio na University of North Carolina na Chapel Hill.

## 1.2 Záber zbierky

V marci 2006 mal Projekt Gutenberg vo svojej zbierke viac ako 18 000 titulov, pričom priemerne pribúdalo viac ako 50 nových titulov mesačne. <sup>[2]</sup>

Sú to väčšinou literárne diela západnej kultúrnej tradície, okrem literárnych foriem ako romány, poézia, poviedky a dráma PG obsahuje aj kuchárky, referenčné diela a vydania periodík. Zbierka PG má tiež niekoľko netextových diel ako zvukové súbory a súbory s hudobným zápisom.

Väčšina diel je v anglickom jazyku, ale tiež významné množstvá sú v nemčine, francúzštine, taliančine, španielčine, holandčine, fínčine a čínštine a ko aj rastúce množstvá v iných jazykoch. Viac ako 50 diel je zdigitalizovaných v 13 jazykoch. Katalóg však neobsahuje ani jedno dielo v slovenčine a v češtine len diela 8 autorov.

Vždy, keď je to možné, sú diela dostupné v čisto textovom formáte, väčšinou za použitia znakového kódovania US-ASCII, ale často rozšíreného na ISO-8859-1. Iné formáty sa tiež vydávajú, keď ich dobrovoľníci zašlú, pričom najbežnejšie býva HTML. Formáty, ktoré nie sú ľahko upravovateľné, ako PDF, sa vo všeobecnosti nepovažujú za vhodné v rámci cieľov projektu, hoci niekoľko ich bolo pridaných do zbierky. Už roky prebieha diskusia o použití niektorého typu XML, hoci vývoj je pomalý.

## 1.3 Ideály

Michael Hart povedal v roku 2004: „Misia projektu Gutenberg je jednoduchá: 'Podporovať tvorbu a distribúciu e-textov'“.

Slogan projektu je „strhnúť mreže neznalosti a nevzdelanosti“, keďže niektorí dobrovoľníci šíria verejnú vzdelanosť a docenenie literárneho dedičstva tak, ako to robili verejné knižnice na začiatku 20. storočia.

Projekt Gutenberg je medzinárodne decentralizovaný. Napríklad neexistuje politika výberu, ktoré diela sa pridávajú. Jednotliví dobrovoľníci pracujú na dielach, ktoré ich zaujímajú alebo ktoré sú dostupné.

Zbierka PG má uchovávať diela dlhodobo, tak, aby sa nemohli stratiť pri žiadnej miestnej nehode. V snahe zabezpečiť to sa celá zbierka denne zálohuje a mirroruje na servery na rôznych miestach.

## 1.4 Autorské práva

Projekt Gutenberg dôsledne overuje status svojich e-textov podľa autorského práva Spojených štátov amerických. Materiál sa pridáva do archívu PG iba potom, ako dobrovoľníci potvrdia, že sa na dielo nevzťahujú autorské práva a tieto záznamy sa uchovávajú pre kontrolu do budúcnosti.

Na rozdiel od niektorých projektov digitálnych knižníc si Projekt Gutenberg nenárokuje na nové autorské práva na diela, ktoré publikuje. Naopak, podporuje ich voľné šírenie a distribúciu.

Väčšina diel zbierky sa distribuuje ako voľné dielo v zmysle autorského práva USA. Právny text, ktorý sprevádza každý e-text do istej miery obmedzuje, čo je možné robiť s textom (ako distribúcia v zmenenom tvare či pre komerčné účely) v prípade, že je použitá obchodná značka Projekru Gutenberg. Tiež je umožnené odstrániť túto hlavičku a samotný text používať ako voľné dielo bez obmedzení.

V zbierke existuje tiež niekoľko diel, ktoré Projekt Gutenberg distribuuje s povolením autora. Tieto podliehajú obmedzeniam, ktoré špecifikoval držiteľ autorských práv.

V roku 1998 Sonny Bono Copyright Term Extension Act predĺžil trvanie už existujúcich autorských práv retroaktívne o 20 rokov. To zabránilo Projektu Gutenberg pridať množstvo titulov, ktoré by inak v Spojených štátoch boli považované za voľné diela.

## 1.5 Kritika

Projekt Gutenberg býva kritizovaný za nedostatok bibliografických informácií v svojich e-textoch: napríklad neadekvátne podrobnosti o vydaní a nezverejnenie pôvodne publikovaných predslovov a kritický aparát. Zlepšenie v týchto oblastiach je badateľné pri porovnaní starších textov s novšími; väčšina e-textov zachováva informácie o vydaní a predslovy.

## 1.6 Príbuzné projekty

Všetky príbuzné projekty sú nezávislé organizácie, ktoré zdieľajú rovnaké ideály a bolo im udelené povolenie používať obchodnú značku *Project Gutenberg*. Často majú konkrétne národné alebo lingvistické zameranie.

Zlomok kníh českej proveniencie je v projekte Gutenberg obsiahnutá: pozri: <http://www.gutenberg.org/browse/languages/cs>

Na Slovensku existuje analogický dobrovoľnícky projekt, ktorý sa volá *Zlatý fond denníka SME*

[https://sk.wikipedia.org/wiki/Zlatý\\_fond\\_denn%C3%ADka\\_SME](https://sk.wikipedia.org/wiki/Zlatý_fond_denn%C3%ADka_SME)

- Project Gutenberg Australia hostuje množstvo textov, ktoré sú voľným dielom podľa Austrálskeho autorského práva, ale stále chránené (alebo v istom stave) v Spojených štátoch; so zameraním na austrálskych spisovateľov a knihy o Austrálii.
- PG-EU je sesterský projekt pracujúci v súlade s autorským právom Európskej únie. Jedným z jeho cieľov je zaradiť do Projektu Gutenberg čo najväčšie množstvo jazykov. Pracuje v Unicode, aby bolo zabezpečené, že všetky abecedy sú jednoducho a korektne reprezentovateľné.
- Project Gutenberg of the Philippines [4] „cieli sprístupniť čo najviac kníh čo najviac ľuďom, so špeciálnym zameraním na Filipíny a filipínske jazyky“.
- Project Gutenberg Europe [5] je projekt pod záštitou Project Rastko v Srbsku a Čiernej hore. Cieľom je stať sa Projektom Gutenberg pre celú

Európu. Začal zasielať svoje prvé diela v roku 2005. Zabezpečuje na svojich stránkach beh Distributed Proofreaders softvéru (oddelene od pôvodného).

- Project Gutenberg Luxembourg [6] publikuje väčšinou, ale nie výlučne knihy písané v luxemburštine.
- Project Gutenberg Consortia Center [7] je pobočka špecializovaná na zbierku zbierok. Tieto nemajú redaktorský dohľad či konzistentné formátovanie hlavného Projektu Gutenberg. Obsahuje tematické zbierky ako aj rozličné jazyky.
- Projekti Lönnrot [8] je projekt začatý fínskymi dobrovoľníkmi Projektu Gutenberg.

Hoci Projekt Gutenberg-DE dostal pred rokmi povolenie používať názov Gutenberg pred rokmi, nie každý ho považuje za príbuzný projekt, z dôvodov filozofických odlišností. Projekt Gutenberg-DE autorsky chráni svoje produkty a obmedzuje prístup na browsovateľné webové verzie svojich textov.

## 2 VÝZNAMNÉ DIGITALIZAČNÍ SVĚTOVÉ PROJEKTY - GOOGLE BOOKS - KNIHY GOOGLE

Internetový gigant Google buduje a zpřístupňuje svoji mezinárodní digitální knihovnu - **Knihy Google** (anglicky Google Books, starší označení **Google Book Search**, ještě starší **Google Print**) je vyhledávací služba, kterou poskytuje Google. Cílem je zpřístupnit naskenované (digitalizované) knihy knihoven. V počátcích tohoto projektu (2004) bylo naskenováno přibližně 15 milionů knih z největších amerických knihoven - *knihovna Michiganské university, Harvardská univerzitní knihovna, Standfordská Green library, Oxfordská Bodleian Library, Newyorská veřejná knihovna*.

Študenti získají poznatky o digitální knihovně Google Books jako projektu, pak o tvorbě jejího obsahu, o spolupráci s partnery, o postupu předávání dokumentu Google Books, autorskoprávních aspektech zpřístupňování digitálního obsahu a názorech na tuto knihovnu z hlediska konkurence. Studenti si osvojí praktické postupy využívání Google Books.

Podle našeho názoru je spolupráce s Google Books nevyvážená. Participaci s Google Books vidíme z hlediska národní digitální knihovny spíše jako marketingový krok. A to jednak proto, že v Google Books je jen torzo národní produkce a jednak proto, že *výběr, konzervování, copyright, administrace, pojištění transport, předávání a přebírání dokumentů, celá logistika* je na bedrech knihovny a ne Google Books. Google vlastně jen digitalizuje, přesněji skenuje, což je relativně nejjednodušší proces. Knihovna, která se zapojuje do Google Books snáší podstatnou část nákladů, jakož i administrativní a odborné práce, kterou by měla provádět pro vlastní národní digitalizační program.

Knihy mají různé způsoby zobrazení podle toho, jak to umožňuje autorský zákon:

- úplné zobrazení - knihy u kterých autoři, respektive vydavatelství, uvolnili knihy na zveřejnění anebo jsou autorská práva promlčená - staré knihy.
- omezený náhled - knihy, u kterých autoři, respektive vydavatelství, vlastníci autorských práv, neuvolnili knihy ke zveřejnění.

### 2.1 Projekt Knihy Google

Projekt Knihy Google se vytvářel několik let. Již dříve utvořený pracovní tým se v roce 2004 spojil s pěti univerzitními knihovnami, a sice s Harvardem, Michiganem, Stanfordinem, Oxfordem a New York Public Library. Knihovny poskytovaly knihy, které Google začal skenovat. Ve smlouvě oxfordské knihovny se společností Google<sup>[1]</sup> bylo naplánováno během následujících tří let zdigitalizovat více než milion tištěných knih, které poté mají být přístupné na internetu skrze službu Google a internetové stránky univerzity. Prozatím byly plánovány pro digitalizaci pouze ty knihy, kterým už vypršela ochrana autorským zákonem (tzn. vydané před rokem 1920). Pro každou knihu se měly vytvořit dvě digitální kopie, jedna z nich je pro Google a jedna pro Oxford. Kopie pro Google má být indexovaná a dohledatelná v jeho vyhledávací službě a kopie pro univerzitu bude spojena s odpovídajícím záznamem v knihovním katalogu.

Projekt Knihy Google čerpal inspiraci z předchozích projektů, jako je například Gutenberg, Universal Library a další. Digitalizační projekt Gutenberg

byl založen v roce 1971 a je nejstarší digitální knihovnou. Dobrovolnická skupina skenuje díla, na která už vypršel autorský zákon. Digitalizační projekt společnosti Google se skládá z více částí. Služba *Knihy Google Search* umožňuje vyhledávat knihy na základě fulltextového prohledávání jejich obsahu, nejen prohledáváním bibliografických údajů jako knihovní katalogy. Spolupracuje s *Projektem knihovna* (Library Project) a *Partnerským programem* (Partner Program). *Projekt knihovna* spolupracuje od roku 2004 s knihovnami, které poskytují knihy ke skenování. Automatizovaně se indexuje text, doplňují základní bibliografická metadata, vazby na jiné informační zdroje a další údaje. Pro rozpoznání textu se používá technologie OCR. Naproti tomu *Partnerský program* je zaměřen na autory a vydavatele. Ti mohou dodávat společnosti Google své knihy přímo v elektronické podobě nebo je dát ke skenování. Jako protislužbu Google nabízí zvýšení jejich prodeje umístěním odkazů na knihkupectví, nezobrazení celých textů knih a finanční podíl na kontextové reklamě.

Google knihy rozdělil do tří skupin podle míry zpřístupnění.

- První skupinu tvoří knihy, jimž už vypršela doba ochrany autorským zákonem. U nich je umožněno volné vyhledávání a zpřístupnění plných textů.
- Do druhé skupiny patří knihy, které jsou chráněné autorským zákonem, a majitel autorských práv nevyjádřil souhlas s jejich zveřejňováním. U těchto knih Google umožňuje pouze vyhledávat, ale nezobrazuje plný text, pouze několik úryvků, které souvisí s hledaným výrazem. Vlastník však může vyjádřit nesouhlas a kniha bude z databáze odstraněna nebo ani nebude skenována.
- U poslední skupiny vlastník souhlasil se zařazením svého díla do projektu Google Book Search a sám určí, které části knihy budou zobrazovány.<sup>[2]</sup>

## 2.2 Spolupráce s knihovnami

Počet knihoven, které se zapojují do digitalizačního projektu, stále narůstá. Nabídka společnosti Google je pro ně výhodná, protože Google velkou část prací financuje. Knihovna se sama musí postarat pouze o přípravu knih na digitalizaci. Od Google dostanou detailně rozpracovanou metodiku přípravných prací.

## 2.3 Postup přispívání do Google Books

### 2.3.1 VÝBĚR A OCENENÍ.

Výběr knih spočívá ve zhodnocení jejich fyzického stavu, velikosti, ceny a obsahu. Je třeba zjistit, které knihy již Google zdigitalizoval, je zbytečné jeden titul skenovat dvakrát.

### 2.3.2 KATALOGIZACE

Nejdříve se musí provést katalogizace knih. Knihovna musí vytvořit elektronický katalog podle požadavků Google.

### 2.3.3 KONZERVOVÁNÍ

Pak přichází na řadu konzervování a restaurování knih. To je důležité proto, aby knihy nemohly být skenováním poškozeny. Google vydal podrobný manuál, ve kterém popisuje, jaké parametry musí připravované knihy splňovat. Restaurátoři knihy mechanicky čistí, přichycují odpadávající části knižní vazby a fixují místa, která se při skenování mohou nejvíce poškodit. Také se pomocí speciálních testů musí provést kontrola mikrobiologické aktivity knih, aby nebylo mikroby zaneseno skenovací zařízení a ostatní skenované knihy.

### 2.3.4 ADMINISTRATIVA VÝVOZU. POJIŠTĚNÍ

Také je potřeba ocenit knihy pro pojištění a zajistit složitou administrativu spojenou s vývozem knih. Každá kniha musí mít osvědčení o vývozu do zahraničí ve čtyřech stejnopisech, v České republice se však dá uplatnit výjimka a použít hromadné seznamy knih.

### 2.3.5 EXPEDICE

Poté následuje samotná expedice knih do zahraničí spolu s předáním elektronických katalogových záznamů.

### 2.3.6 VRÁCENÍ. PŘEBÍRÁNÍ DIGITÁLNÍHO OBSAHU

Po naskenování se knihy vrací zpět do původní knihovny. Následně knihovna přebírá od Google digitální dokumenty a zařazuje do své digitální knihovny, elektronické zdroje.

## 2.4 Soudní spory

Projekt Knihy Google však má i své odpůrce, mezi něž patří především spisovatelé, kteří se bojí ohrožení svých autorských práv a dalším společnostem se nelíbí, že by měl mít Google v této sféře monopol. Ze soudních materiálů se ukazuje, že hlavní motivací Google k digitalizaci byl konkurenční boj.<sup>[3]</sup>

Negativní reakce se objevují i mimo USA, nejsilněji v Kanadě, Austrálii a Evropě. V Evropě s dohodou nesouhlasila především Francie, Itálie a Německo. Francie i Německo proti ní podaly stížnost, obávaly se rizika vytvoření monopolu. Velmi ostře se vyjádřil francouzský Národní syndikát nakladatelů, který reprezentuje více než 500 nakladatelů a 80% francouzského knižního trhu. Postupně se přidávali další držitelé práv již zdigitalizovaných děl se svými protesty. Nakladatelská skupina La Martinière podala na společnost Google žalobu u pařížského soudu. Požadovala stažení všech sporných digitálních děl a náhradu ve výši patnácti milionů euro. Do konfrontace se však dostalo francouzské a americké právo, přičemž obě strany mají na autorská práva jiný názor. Nakonec francouzský soud rozhodl ve prospěch skupiny La Martinière. Nicméně ta v červenci roku 2011 sama podepsala s Google dohodu, podle které Google bude digitalizovat díla chráněná autorským zákonem, která už jsou na trhu rozebraná. Další dohodu s Google uzavřela druhá francouzská nakladatelská skupina Hachette Livre. K této dohodě došlo také v červenci 2011 a Hachette podle ní dostalo

právo kontroly knih, které budou digitalizované, a tyto knihy může dále komerčně využít.

Se společnostmi Google uzavírali postupně smlouvy další evropští nakladatelé i knihovny. Do projektu se zapojily obě italské národní knihovny, Římská i Florentská, Národní knihovna Rakouska, Nizozemská královská knihovna, kantonální a univerzitní knihovna ve Švýcarském Lausanne a také Národní knihovna České republiky a Moravská zemská knihovna. O zapojení uvažovala i Národní knihovna Francie, která však zpočátku byla proti a chtěla si svá díla digitalizovat sama, ale kvůli nedostatku finančních prostředků přistoupila také k dohodě s Google.

## 2.5 Spolupráce Google s Národní knihovnou České republiky

Národní knihovna České republiky v roce 2011 zahájila spolupráci se společností Google na digitalizaci historických a vzácných fondů svých sbírek. Digitalizace je prozatím naplánována do roku 2016 a je možnost prodloužení smlouvy vždy o jeden rok. Cílem má být zpřístupnění historických knižních fondů do roku 1800 a některých mladších vzácných fondů. Celkem je naplánováno zdigitalizovat asi 200 tisíc svazků a zpřístupnit je prostřednictvím Google Books a také v rámci Manuscriptoria a Krameria. Podle dohody zdigitalizované knihy Google zpřístupní v rámci svého zdroje Google Books a také služby Google Book Search a spolupracující knihovna může digitální dokumenty archivovat a také zpřístupnit podle vlastního uvážení a začlenit je do svých služeb. Z toho vyplývá, že spolupráce s Google knihovny nijak nesvazuje a může vést k vytvoření otevřeného virtuálního prostředí pro jejich uživatele.

Pro národní knihovnu kvůli spolupráci s Google vyvstává i otázka autorského práva. Na staré tisky se již nevztahuje, ty jsou však úzce spojeny se sekundárními prameny, které se k nim vztahují, což jsou například komentáře, anotace, ale také edice pramenů. Je záhodno digitálně zpřístupnit i tyto dokumenty společně se starými tisky a to může být právě kvůli autorským právům problémem. Pracovníci Národní knihovny Hejnová a Uhlíř ve svém příspěvku<sup>[4]</sup> k této problematice navrhuje adaptaci autorského práva na nové digitální prostředí a odmítají trvání na starých principech, které fungovaly v tištěném prostředí.

Samotné digitalizaci musí předcházet přípravné práce. Ty Národní knihovna České republiky začala provádět už v průběhu roku 2011. V květnu následujícího roku se uskutečnilo setkání čtyř zástupců Google s pracovníky Národní knihovny, které mělo za cíl předat informace nutné k digitalizačnímu procesu a jeho přípravám. Důraz byl kladen na problematiku transportu historických a vzácných dokumentů na digitalizační pracoviště Google. Následně Google poslal Národní knihovně vzorový vozík, na který měly být nakládány knihy určené k převozu na digitalizaci. Problém způsobila jeho velikost, protože se vejde pouze do jediného výtahu v Klementinu, který navíc není určen pro transport nákladu. Těchto vozíků Klementinum dostalo 20 a muselo vyřešit, kam je uložit.

Další informace Národní knihovna získává od zahraničních partnerů, kteří své fondy také digitalizují v rámci spolupráce se společností Google. V říjnu 2012 se konal každoroční mezinárodní summit knihoven digitalizujících s Google, který pro ně společnost Google připravila tentokrát v Haagu v Národní

knihovně. Summitu se zúčastnil jeden zástupce Národní knihovny České republiky a přivezl informace o nových technologiích Google v oblasti zpřístupňování dat, jejich agregace a deduplikace. Ve své zprávě<sup>[5]</sup> píše, že v té době bylo v Google Books přes 23 milionů knih, při jejich zpřístupňování jsou respektovány rozdíly v autorských právech jednotlivých zemí a zdigitalizované dokumenty v angličtině již tvoří méně než 50% všech dokumentů.

Pro digitalizaci knihovna musí mít elektronický katalog podle zásad Google. Národní knihovna elektronický katalog starých tisků neměla, musí ho tedy vytvořit. Oddělení rukopisů a starých tisků připravuje k digitalizaci dokumenty z historického fondu z let 1501-1800. Během dvou let 2011-12 bylo externími pracovníky zkatalogizováno přes 40 tisíc záznamů. Slovanská knihovna připravuje na digitalizaci dokumenty vydané v 19. století. Během sledovaných dvou let bylo zkatalogizováno přes 8 tisíc záznamů.

Katalogizace však postupuje mnohem rychleji než opravy a restaurování fyzických dokumentů. Historický fond Národní knihovny není totiž v nejlepším stavu, protože v minulosti se s ním nezacházelo moc dobře. Staré tisky byly uloženy v nevhodných klimatických podmínkách. Průzkum fyzického stavu fondu tedy zjistil velmi vážná poškození dokumentů a jejich řešení je mnohem náročnější na čas než katalogizace. Z tohoto důvodu dochází ke skluzu restaurování za katalogizací. Z toho vyplývá, že Národní knihovna potřebuje mnohem více restaurátorů a pracovišť vybavených speciálními přístroji.

Ošetřené knihy se evidují v nově vytvořené databázi, kde jsou rozdělené do čtyř kategorií podle typu poškození. Restaurování knih spočívá zejména ve fixaci uvolněných listů, opravě potrhaného šití, trhlin papíru, opravy knižních hřbetů a výrobě ochranných obálek. Náročnější zásahy zahrnují rozešití knižního bloku, čištění papíru pomocí vodních systémů nebo dolévání papírových listů, opravu knižní vazby a doplnění chybějících knižních desek. Z digitalizace bylo za sledované období vyřazeno asi 540 knih, které nesplňovaly požadavky na fyzický stav určené společností Google. Tyto knihy byly nekompletní nebo zčásti nečitelné, vazba neumožňovala dostatečné otevření, listování knihou by ji poškozovalo apod.

Další činností, kterou musí Národní knihovna vykonat, je předávání samotných knih a metadat společnosti Google. Jednotlivé dávky knih musí být stejně velké, aby byla digitalizace plynulá a pravidelná. Společně s knihami se předávají i seznamy, ve kterých musí být knihy seřazeny stejně jako v jednotlivé dávce, kvůli usnadnění kontroly. Metadata musí knihovna předávat ve formátu MARC XML.

Od r. 2014 běží spolupráce obou partnerů rutinním způsobem a do podzimu 2016 bylo již digitalizováno na 105 000 knih, z toho asi 20 tisíc ze Slovanské knihovny (i z těch bylo cca 2 tisíce starých tisků). Knihy jsou ihned po digitalizaci zpřístupněny v Google Books a tam nacházející se digitální kopie také odkázány z online katalogu starých tisků Národní knihovny ČR a katalogu Slovanské knihovny. Knihy jsou také postupně zpřístupňovány v Manuscriptoriu.

Přebírání dat od společnosti Google probíhá v návaznosti na digitalizaci. Datové balíčky obsahují popisná metadata (původně dodávaná Národní knihovnou), strukturální metadata, obrazová data, textová data a pozice slov v obrazech. Obrazová data Google poskytuje ve formátu JPEG. U starých tisků je



datový balíček Google transformován do datového balíčku dle specifikace VISK 6, což je základní předpoklad pro zpřístupnění v Manuscriptoriu a předpoklad pro archivaci. Národní knihovna plánuje takto transformovaná data uložit do systému LTP, který je součástí Národní digitální knihovny.

Spolu s tímto projektem má Národní knihovna i vlastní digitalizační programy, z nichž starší je zaměřen na rukopisy a staré staré tisky a novější na periodika a na mladší knihy vydané po r. 1800.

## 2.6 Závěr

Masová digitalizace prováděná společností Google je přijímána dvojznačně. Akademické prostředí ji hodnotí spíše negativně, zejména proto, že je prováděna za účelem zisku. Naproti tomu knihovny ji vítají. Uvědomují si, že samy by digitalizace v takové míře nikdy schopny nebyly, jak z organizačního, tak převážně z finančního důvodu. Zejména u starých dokumentů, u kterých je nutné klást velký důraz na vyšší kvalitu, by knihovny se svými technickými možnostmi silně zaostávaly. Bez spolupráce knihoven se společností Google by byl přechod do digitálního prostředí mnohem pomalejší.

Digitalizace knih je pro knihovny výhodná i z důvodu ochrany historických fondů a starých tisků. Když budou dokumenty přístupny v digitální podobě, nebudou muset být čtenářům a badatelům půjčovány originály a tím se uchrání před dalším poškozením. Nevýhodou pro knihovny však je fakt, že po zdigitalizování většiny knih lidé přestanou mít zájem chodit do knihovny. Ta je však postavena před nový úkol, a sice uživatelům co nejvíce usnadnit vyhledávání knih v digitální podobě, rozšiřovat jednoduchá metadata, psát obsáhlé obsahy děl apod.

Nevyřešena však stále zůstává otázka autorského práva. Navíc doba autorskopravní ochrany díla může být v každé zemi různě dlouhá. V zemích Evropské unie a Spojených států amerických trvá ochrana od vytvoření díla až do 70 let po smrti jeho autora. U děl ve vlastnictví korporací je ve Spojených státech ochrana 95 let od vydání díla. Google bere v potaz autorskopravní ochranu země, ve které bylo dílo vydané. Přesto se vede kolem autorského práva stále mnoho diskuzí.

## 2.7 Odkazy

*V tomto článku byl použit překlad textu z článku Google Books na slovenské Wikipedii.*

1. ↑ CARR, Reg. Oxford-Google digitisation agreement. [online]. [cit. 2014-01-16]. Dostupné z: <http://www.bodley.ox.ac.uk/news/news58.htm> Archivováno 6. 4. 2014 na [Wayback Machine](#)
2. ↑ BARTOŠEK, Miroslav. Google Book Search. Inflow [online]. 26.4. 2009. [cit. 2014-01-16]. Dostupné z: <http://www.inflow.cz/google-book-search>. ISSN 1802-9736
3. ↑ <http://www.bookz.cz/wordpress/2012/08/07/digitalizace-knih-u-google-mirila-na-amazon-zadna-dobrocinnost-se-nekona/> - Digitalizace knih u Google mířila na Amazon, žádná dobročinnost se nekoná

4. ↑ HEJNOVÁ, UHLÍŘ. Digitalizace s Google a její vztah k NDK – IOP. Bibliotheca Antiqua 2012: sborník z 21. konference. [online]. Olomouc: Vědecká knihovna v Olomouci, 2012. ISBN 978-80-7053-298-0. [cit. 2013-12-03]. Dostupné z: <http://www.vkol.cz/data/soubory/import/konf21/Bibliotheca%20Antiqua%202012-Hejnova-Uhlir.pdf>
5. ↑ KNOLL, Adolf. Zpráva ze zahraniční služební cesty. [online]. [cit. 2014-01-16]. Dostupné z: [http://wwwold.nkp.cz/files/cz\\_haag\\_google2012\\_ak.pdf](http://wwwold.nkp.cz/files/cz_haag_google2012_ak.pdf)

## 2.8 Související články

- [Google Ngram Viewer](#)
- [Google Trends](#)

## 2.9 Literatura

- About Google Books. [online]. [cit. 2014-01-16]. Dostupné z: <http://books.google.com/googlebooks/about/>
- BARTOŠEK, Miroslav. Google Book Search. Inflow [online]. 26.4. 2009. [cit. 2014-01-16]. Dostupné z: <http://www.inflow.cz/google-book-search>. ISSN 1802-9736
- Google Books Library Project. [online]. [cit. 2014-01-16]. Dostupné z: [http://en.wikipedia.org/wiki/Google\\_Books\\_Library\\_Project](http://en.wikipedia.org/wiki/Google_Books_Library_Project)
- Google Books. [online]. [cit. 2014-01-16]. Dostupné z: [http://en.wikipedia.org/wiki/Google\\_Books](http://en.wikipedia.org/wiki/Google_Books)
- HEJNOVÁ, Miroslava. Digitalizace starých a vzácných tisků v Národní knihovně společností Google. Duha [online]. 2011, roč. 25, č. 3 [cit. 2014-01-16]. Dostupné z: <http://duha.mzk.cz/clanky/digitalizace-starych-vzacnych-tisku-v-narodni-knihovne-spolocnosti-google>. ISSN 1804-4255.
- HEJNOVÁ, UHLÍŘ. Digitalizace s Google a její vztah k NDK – IOP. Bibliotheca Antiqua 2012: sborník z 21. konference. [online]. Olomouc: Vědecká knihovna v Olomouci, 2012. [cit. 2013-12-03]. Dostupné z: <http://www.vkol.cz/data/soubory/import/konf21/Bibliotheca%20Antiqua%202012-Hejnova-Uhlir.pdf>. ISBN 978-80-7053-298-0.
- KNOLL, Adolf. Zpráva ze zahraniční služební cesty. [online]. [cit. 2014-01-16]. Dostupné z: [http://wwwold.nkp.cz/files/cz\\_haag\\_google2012\\_ak.pdf](http://wwwold.nkp.cz/files/cz_haag_google2012_ak.pdf)
- MACHOVÁ, Anna. Google Books - Google Livres: klikaté cesty digitalizace. Knihovna [online]. 2011, roč. 22, č. 1, s. 94-100. [cit. 2014-01-09]. Dostupné z: <http://knihovna.nkp.cz/knihovna112/11294.htm>. ISSN 1801-3252.

### 3 VÝZNAMNÉ DIGITALIZAČNÍ SVĚTOVÉ PROJEKTY - EUROPEANA

**Europeana** je digitálna knižnica, ktorá umožňuje širokej verejnosti prezerat' si digitálne zdroje z európskych múzeí, knižníc, archívov a audiovizuálnych zbierok. Portál Europeana.eu bol spustený 20. novembra 2008 Viviane Redingovou, eurokomisárkou pre informačnú spoločnosť a médiá. Pamiatkou na štart Europeany je pamätná tabuľa vo vestibule v Belgickej kráľovskej knižnici. Medzi zakladateľmi je aj Dušan Katuščák, ktorý bol v tom čase generálnym riaditeľom Slovenskej národnej knižnice v Martine a zabezpečil spoluprácu s Europeanou a budovanie slovenskej národnej digitálnej knižnice tak, aby bola súčasťou Europeany.

Projekt je realizovaný centrálnym tímom, sídliacim v národnej knižnici Holandska - Koninklijke Bibliotheek. Europeana je internetový portál, ktorý funguje ako rozhranie k miliónom kníh, obrazov, filmov, múzejných predmetov a archívnych záznamov v celej Európe. Do Europeany prispieva viac ako 2000 inštitúcií. (Louvre či British Library až po regionálne archívy a miestne múzeá Európskej únie). Umožňuje prístup k rôznym typom harvestovaného obsahu (digitálne objekty nie sú uložené na centrálnom počítači, ale zostávajú v jednotlivých inštitúciách). Europeana zhromažďuje kontextové informácie (metadáta) o položkách, vrátane malého náhľadového obrázku. Europeana obsahuje prístup k digitálnemu obsahu – *texty, zvuk, obraz, audiovideo*.

#### 3.1 Europeana (Comité des Sages). Pozícia Európskej únie.

”

, a p

, digitálneho priemyslu.

•

•

•

•

•

•

•

•

*Comité des sages*

3.2

*Kul*

*V  
poplatok za*

•

- 

### 3.3 Stimuly digitalizácie

- 

- 

- 

- 

, a

- 

- 

- *poskytovatelia.*

- 

, zvuku.

### 3.4 Úloha Europeany

- *Europeana by sa*

- 

, aby do roku 2

-

).

•

1

•

3.5

•

•

– ako je to te

•

•

•

a

•

*Europeanou.*

3.6 Udržateľnosť

*Europeany*

1

- 

*financovania.*

- 

- 

).

- 

### 3.7 PPP projekty

*partnerstiev (PPP)*

A

*podmienok:*

- 

*nerom.*

- 

-

- 

*vytvorenie stimulov*

- 

- 

- 

- 

*pre digital  
spolufinancovania s  
prostriedkami;*

- 

*dedi*

*renesanciu.“*

#### 4 VÝZNAMNÉ DIGITALIZAČNÍ SVĚTOVÉ PROJEKTY – EUROPEANA . ZBIERKY

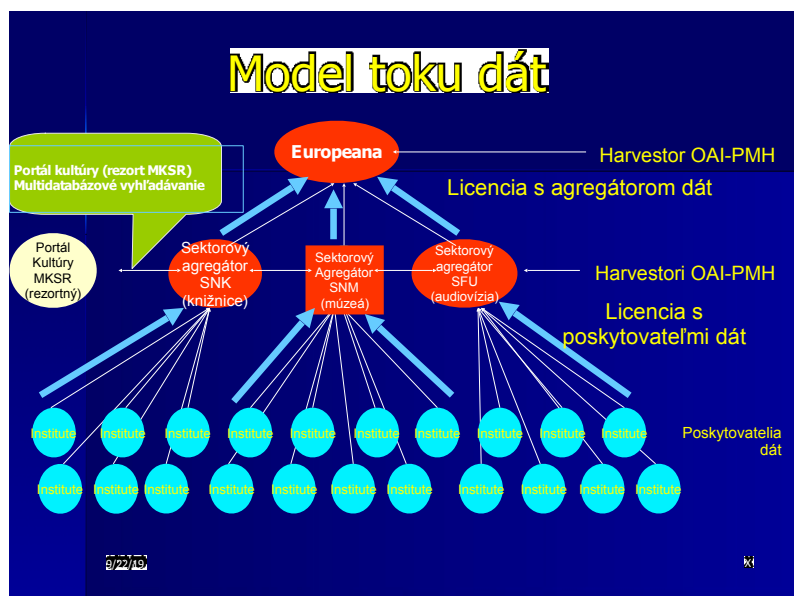
Do Europeany prispievajú po uzavretí (DEA) kontraktu,<sup>1</sup> najmä:

1. Knižnice,
2. Múzeá,
3. Archívy,
4. Filmové archívy,
5. Rozhlasové archívy,
6. Regionálni agregátori

---

<sup>1</sup> The Europeana Data Exchange Agreement (DEA). Kontrakt z roku 2018 je dostupný:  
<https://pro.europeana.eu/page/the-data-exchange-agreement>

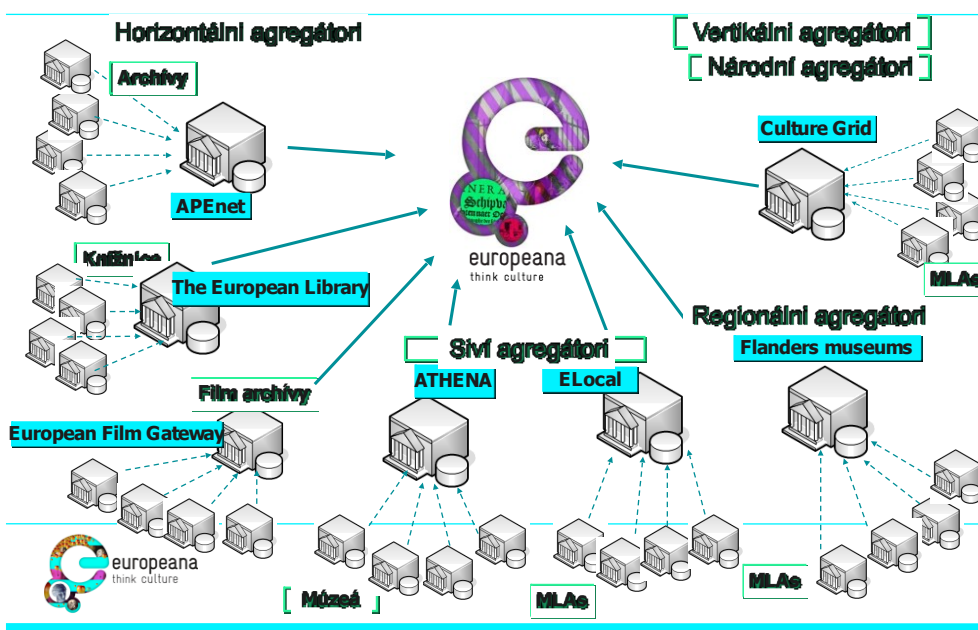




Kontrakt upravuje vzťah medzi Europeanou a poskytovateľmi údajov. Stanovuje dve hlavné pravidlá týkajúce sa práv duševného vlastníctva (IPR) v metaúdajoch a obsahu dodanom Europeane:

1. Všetky metadáta/metaúdaje predložené Europeane sa uverejnia ako otvorené údaje v súlade s podmienkami vyhradenia pre verejnosť (Creative Commons Zero Public Domain, CC0).
2. Každý digitálny objekt, ktorý obsahuje pridružený náhľad, thumbnail, ktorý je k dispozícii na portáli Europeana, musí obsahovať vyhlásenie o právach, ktoré opisujú stav autorských práv daného objektu. Ak je objekt vo verejnej doméne, musí byť označený ako *vo verejnej doméne*.

## Kto prispieva do Europeany?



#### 4.1 Ktoré dáta sa dávajú do Europeany?

Kooperanti poskytujú do Europeany:

1. Náhľad (Thumbnails)
2. Linky na digitálny obsah (umožňujú harvestovanie)
3. Metadáta (hlavne deskriptívne, popisné)

### Ktoré dáta môžem dať do Europeany?

The screenshot shows the BnF Gallica interface. The main content is a document titled "GARGANTUA" with the subtitle "Fêtes des Pauvres à Bordeaux, LES 2, 3, 4 MAI 1892." A green box highlights the document title and metadata area, with labels "1. Náhľad", "2. Metadáta", and "3. Linky na online digitálny obsah" pointing to different parts of the page. The left sidebar shows navigation options like "Rechercher dans ce document" and "Ajouter une étiquette". The bottom left shows "Actions: Add a tag, Share with a friend, Save to My Europeana, Login | Register".

#### 4.2 Ktoré druhy dokumentov obsahuje Europeana?

Europeana obsahuje dokumenty:

1. TEXT
2. OBRAZ
3. ZVUK
4. AUDIOVIDEO

### Ktoré druhy dokumentov idú do Europeany?

The screenshot shows the BnF Gallica interface with search results for "GARGANTUA" on the ina.fr platform. The main content is a video player showing a concert. A green box highlights the search results, with labels "TEXT", "OBRAZ", "VIDEO", and "AUDIO" pointing to different parts of the page. The left sidebar shows navigation options like "Rechercher dans ce document" and "Ajouter une étiquette". The bottom left shows "Actions: Add a tag, Share with a friend, Save to My Europeana, Login | Register".

### 4.3 Aké sú technické požiadavky

Pre nahrávanie digitálneho obsahu do Europeany je nevyhnutné:

#### 1. Metadáta mapované podľa špecifikácie ESE

Europeana semantic elements (ESE) – to je súčasný model dát v Europeane. Obsahuje minimálnych 8 metadát podľa Dublin Core. Použiť možno a ďalšie v štandarde Dublin definované Core (DC, je ich 15)) a ďalších 13 údajov, ktoré môžu byť dôležité pre funkcionality Europeany. Dostupné na: <https://pro.europeana.eu/page/ese-documentation>

#### 2. Online linkovanie na objekty u vlastníka objektu

Niekedy úplný dokument (zobrazenie objektov v Europeane – požiadavky. Tzv. „thumbnail“ objektu (malý náhľad)

#### Mapovanie metadát a pravidiel štandardizácie

Na správne mapovanie metadát pre Europeanu má poskytovateľ obsahu poznať a použiť tento zdroj: Zdroj je dostupné v dokumentácii na: <https://pro.europeana.eu/page/ese-documentation>

## 5 SVETOVÁ DIGITÁLNA KNIŽNICA

**Svetová digitálna knižnica** (angl. World Digital Library, *WDT*) je projekt medzinárodnej neziskovej digitálnej knižnice pod správou UNESCO a Kongresovej knižnice. Oficiálne začala fungovať 21. apríla 2009.

Úlohou knižnice je prezentácia medzinárodného a medzikultúrneho porozumenia a zároveň zväčšovanie množstva a rôznorodosti kultúrneho obsahu spoza hraníc západnej kultúry. Materiály umiestnené na stránkach knižnice sú bezplatné a dostupné v mnohých jazykoch. Medzi nimi sú staré rukopisy, mapy, vzácne knihy, notové záznamy, nahrávky, filmy, fotografie, architektonické skice a iné kultúrne hodnotné materiály.

V deň spustenia knižnica poskytla 1 170 súborov na stránkach v arabskom, čínskom, anglickom, francúzskom, portugalskom, ruskom a španielskom jazyku.

## 6 DIGITÁLNÍ KNIHOVNA MANUSCRIPTORIUM23

### 6.1 Abstrakt

*Manuscriptorium je volně dostupná digitální knihovna, která pomocí sofistikovaných vyhledávacích nástrojů poskytuje přístup k existujícím zdrojům v oblasti historických dokumentů. Cílem projektu je zpřístupnit veškerý existující digitální obsah tak, aby byl snadno dostupný koncovým uživatelům. Za tím účelem spolupracuje s řadou renomovaných světových institucí (23 zemí, 120 institucí, 46 000+ plně digitalizovaných dokumentů a 400 000+ popisných záznamů). V současnosti je projekt otevřen spolupráci s dalšími správci digitálního obsahu a je připraven shromáždit a prezentovat obsah dostupný v rámci kulturních a historických institucí zemí V4. Příspěvek se zaměřuje na popis způsobů spolupráce a její přínos badatelům, studentům, pedagogům a zainteresované více či méně laické veřejnosti.*

### 6.2 Klíčová slova

*Manuscriptorium, digitální knihovna, historické dokumenty, digitalizace, zpřístupnění, kooperace, interoperabilita*

### 6.3 O projektu. Počátky

Počátky projektu spadají do dávné minulosti (devadesátých let minulého století). Skupina badatelů a techniků cítila potřebu digitalizace vzácných historických materiálů. V dnešní době není potřeba vysvětlovat přínos takového druhu digitalizace (zachování fyzických dokumentů, lepší zpřístupnění jejich obsahu atd.). První digitalizovaný dokument byl v České republice vytvořen v roce 1993, rutinní digitalizace vzácných dokumentů běží od roku 1996.

Od prvopočátků se náš digitalizační projekt soustředil nejen na samotné pořizování digitálního obsahu, ale i na jeho dlouhodobé uchování a prezentaci koncovým uživatelům. Což, ač možná překvapivě, není tak běžné ani u některých soudobých projektů.

Bezpečného uchování se koncem minulého století se dosahovalo budováním a řízeným sledováním archivu CD-R disků pro kopie v nejvyšší kvalitě (Master Copy). Prezentace koncovým uživatelům se realizovala pomocí odvozených CD-R s uživatelskou kopií - typicky pro prezenční studium.

---

<sup>2</sup> Autor state je **Tomáš Psohlavec**, AiP Beroun s.r.o., jednatel a ředitel společnosti, narozen v roce 1978, vystudoval Českou zemědělskou univerzitu v Praze, obor Informatika. Původně programátor a spoluvývojce webových aplikací pro on-line publikaci odborných informačních zdrojů. Technický koordinátor evropského projektu ENRICH zaměřeného na agregaci digitálního obsahu z oblasti psaného kulturního dědictví, technický koordinátor evropského projektu ETNOFOLK zaměřeného na vznik, agregaci a zpřístupnění digitálního obsahu z oblasti lidové kultury centrální Evropy. Dnes vede společnost AiP Beroun, která se specializuje na šetrnou bezkontaktní digitalizaci vzácných dokumentů a na jejich on-line zpřístupnění.

<sup>3</sup> PSOHLAVEC, Tomáš. 2016. Digitální knihovna Manuscriptorium. In: *Libraries V4 in the Decoy of Digital Age. Proceedings of 6th Colloquium of Library and Information Experts of the V4+ Countries held from 31st May – 1st June 2016 in Brno.* – Brno : Moravská zemská knihovna v Brně, 2016. S.(cze) 367-374. – ISBN 978-80-7051-216-6 (brož.)

O něco později, s rozvojem webových služeb, byl zpřístupněn on-line katalog digitalizovaných rukopisů a inkunábulí s výběrem několika málo nejzajímavějších či nejdůležitějších stránek každého z dokumentů.

Konečně v roce 2003 byl tento katalog nahrazen plnohodnotnou *Digitální knihovnou vzácných a historických dokumentů*. V ten šťastný den, kdy byl projekt zahájen, jsme pro něj vybírali jméno. Na seznamu bylo mnoho zvukných a úderných jmen, ale z jakéhosi důvodu bylo vybráno jméno Manuscriptorium, a pod ním je tento projekt znám dodnes.

I když byl projekt zpočátku založen jako projekt národní, díky své výjimečnosti brzy nabral nový směr k *Evropské digitální knihovně písemného kulturního dědictví*.

#### 6.4 Interoperabilita

Technické řešení projektu bylo od samého počátku navrženo tak, aby bylo možno získávat obsah z různých zdrojů. Díky tomu bylo možné přizvat ke spolupráci první partnery - významné archivy, knihovny a další správce digitálního obsahu. V návaznosti na první úspěchy jsme si stanovili nové ambiciózní cíle:

- Poskytovat přístup k existujícím digitalizovaným dokumentům v oblasti historických fondů
- Vytvářet virtuální badatelské prostředí tak, aby byl existující obsah dostupný v jednotném uživatelském rozhraní

Tento přístup odlišuje Manuscriptorium od jiných příbuzných projektů, které fungují jako portálová řešení. Portály (jako např. Europeana) jsou samozřejmě užitečné a poskytují koncovým uživatelům nástroje k tomu, aby získali informace o dostupnosti konkrétních dokumentů v různých zdrojích. Aby uživatel mohl tyto informace využít, opouští portál a míří s pomocí odkazů do jednotlivých zdrojových aplikací, kde je nucen používat různá uživatelská rozhraní. To je podle našeho názoru málo pohodlné obecně a příliš komplikované pro seriózní badatelskou práci či studium.

Výhody přístupu Manuscriptoria jsou následující:

- *Žádná zjednodušení*: portálová řešení často při agregaci krátí dostupné informace, nebo v lepším případě zjednodušují jejich strukturu. To vede k tomu, že vyhledání relevantní informace není dostatečně přesné a účinné
- *Odhalování nových souvislostí*: pokud jsou související dokumenty zpřístupněny jednotným způsobem (dedikované indexační rutiny atd.), je mnohem snazší tuto vzájemnou souvislost odhalit
- *Obecně pohodlná práce v systému*: odpadá potřeba učit se pracovat v různých systémech
- *Specializované nástroje*: nástroje vyvíjené se zřetelem k potřebám konkrétních typů cílových uživatelů usnadňují studium a bádání

#### 6.5 Rychlý rozvoj spolupráce

Výše uvedené výhody projektu Manuscriptorium byly zřejmé řadě evropských producentů digitálního obsahu v dané oblasti. Proto bylo možné

ve spolupráci s nimi vytvořit konsorcium, které spolupracovalo v letech 2007-2009 v rámci programu eContent + na projektu ENRICH (<http://enrich.manuscriptorium.com>). V rámci tohoto projektu jsme využili Manuscriptorium jako platformu pro hromadnou agregaci a zpřístupnění obsahu a získaná podpora umožnila skutečně masivní urychlení procesu agregace. V projektu spolupracovalo 17 významných evropských institucí a během toho, jak projekt úspěšně šel k dosažení vytčených cílů, se k nám připojilo dalších 36 partnerů. Není proto divu, že v roce 2010 byl rozsah zpřístupněného obsahu skutečně velmi neobvyklý.

Dnes je v katalogu Manuscriptoria přes 400 000 popisných záznamů a více než 46 000 plně digitalizovaných dokumentů.

Obsah pochází z více než sto dvaceti institucí - nejvýznamnějších vlastníků a správců digitálního obsahu v oblasti historických fondů z více než dvaceti zemí.

Pokud jde o počty zpřístupněných zdrojů, je Manuscriptorium nejrozsáhlejší digitální knihovnou rukopisů na světě.

Financování z evropských fondů skončilo a my jsme samozřejmě museli za jistit pět let dalšího provozu systému. Ale na rozdíl od některých jiných dotovaných projektů, my zajišťujeme provoz systému dodnes - díky financování z rozpočtu Národní knihovny České republiky, která projekt stále vede.

Přirozeně je dnes financování projektu značně omezené v porovnání s rozpočtem projektů podporovaných evropskými fondy. Přesto je Manuscriptorium nejen udržováno, ale i dále rozvíjeno. Dnes máme k dispozici již třetí verzi systému a na čtvrté se pracuje (beta verze je plánována na začátek roku 2017).

Kromě toho je projekt stále otevřen pro nové poskytovatele obsahu a proto jsou všichni případní zájemci, kteří provozují vlastní digitalizační projekty, srdečně zváni ke spolupráci. Tímto způsobem jsou každý rok do Manuscriptoria připojeni jeden až dva noví partneři. (Dobrou zprávou pro případné zájemce je fakt, že Národní knihovna České republiky zatím kryje náklady na zavedení takové spolupráce.)

## 6.6 Principy spolupráce

Interním formátem Manuscriptoria je formát založený na TEI P5( *Recommendations of the Text Encoding Initiative* ). Staví tak na tradici a dvaceti letech zkušeností s kódováním textových informací různých druhů. Doporučení *TEI P5 Guidelines for Electronic Text Encoding and Interchange* pokrývají širokou oblast využití a představují de facto standardní formát pro elektronické texty v badatelsky orientovaných výzkumných projektech.

Samotné TEI je velmi obecné schéma, které umožňuje maximálně flexibilní přístup ke tvorbě záznamů. Lze je využít v mnoha různě zaměřených popisech nejen v oblasti rukopisů a vzácných dokumentů. Z těchto důvodů jsme během projektu ENRICH definovali podmnožinu tagů a využili dedikované slovníky tak, aby bylo vytvořeno schéma optimální pro oblast historických fondů. Tak vzniklo TEI P5 ENRICH schema a bylo plně implementováno do Manuscriptoria .

Je samozřejmé, že ne všechny spolupracující instituce jsou připraveny

přijmout stejný formát pro popis svých dokumentů. Přestože jsme tedy přesvědčení, že TEI P5 ENRICH schema je optimální volbou pro tvorbu detailních, vysoce strukturovaných analytických popisů vzácných dokumentů, uvědomujeme si, že poskytovatel obsahu využívá pro uložení svých digitálních informací jiné formáty. Často se sektáváme s „tradičnějšími“ formáty, jako je MARC XML, DC, MODS, METS a mnoho dalších.

A nejde pouze o metadata. Digitální prostředí v naší oblasti je velmi diverzifikované i pokud jde o formáty, způsoby uložení a organizace obrazových dat.

Jak je tedy možné, že se v tomto heterogenním prostředí sdružuje v projektu Manuscriptorium tolik institucí?

### **Konektory**

Manuscriptorium totiž využívá systém konektorů, který umožňuje adaptovat se na různé formáty a workflow partnerů. V případech, kdy přispívající partner nemůže nebo nechce měnit strukturu vlastních metadat, ale přesto by rád publikoval svůj obsah v Manuscriptoriu, pomohou tyto nástroje zajistit potřebnou kompatibilitu. To umožní partnerům vytvářet a spravovat svá data v původní podobě a nezávisle na prezentační platformě Manuscriptoria.

V Manuscriptoriu pro jednoho každého přispívajícího partnera vytváříme dedikovaný importní profil, který definuje:

- 1) způsob přenosu metadat do Manuscriptoria (OAI-PMH, ITP, web scraper a další)
- 2) pravidla mapování mezi původním formátem a TEI P5 ENRICH schéma
- 3) řadu dalších vlastností a pravidel ingestu podle potřeb konkrétní spolu - práce (jako např. periodu těžení atp.)

Právě díky systému flexibilních konektorů je možné dosáhnout vysoké míry interoperability a harmonizovat tak spolupráci i v existujícím vysoce heterogenním prostředí.

## **6.7 Distribuovaná obrazová úložiště nebo jediné centrální?**

Systém Manuscriptoria se v současné době skládá z jediné databáze agregovaných metadat a z celé řady distribuovaných obrazových repozitářů (data repositories). To znamená, že tam, kde to bylo možné, shromáždilo Manuscriptorium pouze metadata předávaných dokumentů, zatímco spolupracující partner zajišťuje dostupnost dat (obrazů) v rámci vlastní on-line služby.

Rozsah agregovaných metadat je typicky následující:

- *popisná metadata* - obvykle se skládají z jednoho nebo více následujících typů informací:
  - identifikace fyzického předlohy
  - popis intelektuálního obsahu



- předlohy
  - popis dochování předlohy
  - historie předlohy
  - administrativní informace
  - a další
- *strukturální metadata:*
  - logické mapy (seznamy kapitol a jiných logických jednotek, obsahy a další)
  - fyzické mapy (seznamy obrazů a jejich umístění)

Popisná metadata zpracováváme a plníme jimi katalog Manuscriptoria. Strukturální informace jsou využívány k tomu, abychom rekonstruovali strukturu dokumentu, naplnili ji obrazy a zobrazovali ji koncovým uživatelům.

Kde to bylo možné, bylo povinností partnera, aby zpřístupnil obrazy v rámci vlastní repository a umožnil jejich přenos do prohlžeče či aplikace koncového uživatele.

Potřeba distribuovaných obrazových repozitářů byla vyvolána následujícími omezeními, dnes již částečně překonanými.

## 6.8 Vysoké náklady na implementaci a udržování centrálního obrazového repozitáře.

V současnosti je celkový objem obrazových dat Manuscriptoria odhadován na přibližně 25 TB. Odhadujeme dále, že tento objem bude v roce 2019 až čtyřnásobný. Pokud vezmeme v úvahu, jak drahý by byl provoz podobného úložiště ještě před pár lety, je zřejmé, že provoz systému by nebyl udržitelný (přinejmenším s ohledem na dostupný rozpočet). Čili rozhodnutí jít cestou distribuce odpovědnosti (a nákladů) na zajištění dostupnosti obrazů bylo rozhodně správnou volbou, která významně přispěla k udržitelnosti projektu.

Dnešní vývoj cen na trhu IT technologií (a také významný vývoj v IT infrastruktuře Národní knihovny České republiky) umožňují tento přístup přehodnotit, což přináší možnost dalším potenciálním partnerům zapojit se do projektu. Zejména menší instituce mohou nyní využít možnost i poskytnout k agregaci nejen metadata, ale i data do centrálního datového úložiště.

## 6.9 Neochota partnera předat do cizí správy vlastní hodnotná obrazová data

Pro některé partnery není akceptovatelná představa, že svá obrazová data (byť jen jejich odvozené uživatelské kvality pro on-line prezentaci) předají do správy jinému subjektu. Ačkoliv z technického hlediska může být diskutováno o tom, že každý obraz přenesený ke koncovému uživateli tak jako tak opustil bezpečí mateřského repozitáře, z pohledu managementu a administrace obsahu se jedná o vcelku srozumitelný požadavek. Model distribuovaných repository pak takovému požadavku vychází vstříc.

## 6.10 Proč s námi spolupracovat

Jakmile jsou vaše data akceptována Manuscriptoriem, stanete se součástí nejrozsáhlejší digitální knihovny vzácných rukopisů na světě.

Pro své dokumenty získáte výkonné vyhledávací a zobrazovací nástroje - ať už jako hlavní platformu pro zpřístupnění, nebo jako paralelní zobrazení k Vaší hlavní digitální knihovně.

Nejen, že se vaše dokumenty budou využívat mnohem častěji, ale také míra návštěvnosti vaší digitální knihovny může vzrůst, protože Manuscriptorium své uživatele mimo jiné také odkazuje na původní umístění dokumentu.

Všechny dokumenty agregované v Manuscriptoriu se automaticky stávají součástí projektu EUROPEANA, se kterým Manuscriptorium spolupracuje jako subagregátor pro oblast historických fondů.

Dále se Vaše dokumenty stanou součástí dalších informačních zdrojů, protože Manuscriptorium je pravidelně těženo službami jako jsou The European Library SUMMON web-scale discovery service, EBSCO Discovery Service a dalšími.

Všechna data samozřejmě zůstávají ve vašem vlastnictví, Manuscriptoriu pouze udělujete právo zpracovávat je pro potřeby on-line zpřístupnění v digitální knihovně Manuscriptoria.

## 6.11 Zahájení spolupráce

Jednání o spolupráci probíhá ve třech úrovních. Nejprve je obsahovým správcem posouzen obsah dokumentů. Národní knihovna České republiky je obsahovým správcem projektu, který dohlíží na výběr dokumentů pro Manuscriptorium. Obecně platí pravidlo, že Manuscriptorium shromažďuje písemné kulturní dědictví do roku 1800 s primárním zaměřením na knižní dokumenty, sekundárně však zpřístupňuje i jiné typy písemných materiálů (listiny, mapy a další).

## 6.12 Technické minimum

Aby bylo možno přispívat dokumenty do Manuscriptoria, sestavili jsme souhrn minimálních technických požadavků. Za kompatibilní dokumenty se považují takové,

- jejichž metadata jsou vytvořena libovolnými nástroji tak, aby měla formu XML validního dle TEI PS ENRICH DTO či XSD schématu
- jejichž metadata obsahují alespoň potřebné minimum informací umožňujících automatizované zpracování v Manuscriptoriu (viz dále)
- jejichž obrazová data jsou dostupná on-line pomocí protokolu HTTP (přímo či prostřednictvím skriptu) na základě informací dostupných v metadatach
- jejichž obrazová data jsou dostupná ve formátech, které jsou přímo podporované nejrozšířenějšími soudobými

internetovými prohlížeči (aktuálně formáty JPEG, GIF a PNG)

Za povinné technické minimum informací obsažených v metadatech se aktuálně považuje:

- bibliografický popis v rozsahu elementární identifikace exempláře:
  - země a místo uložení
  - knihovna
  - signatura - jednoznačná v rámci knihovny
- seznam datových souborů (typicky obrazů) umožňující přístup k datům dostupným on-line

Informace tohoto rozsahu skutečně představují pouze souhrn technického minima. Optimální minimum, jež je vhodné dodržet, aby byl digitální document využitelný, by mělo zahrnovat podstatně širší soubor informací, jehož struktura se samozřejmě liší v závislosti na vlastnostech fyzické předlohy. Proto si Manuscriptorium neklade za cíl stanovit minimální rozsah informací nad hranicí technického minima.

### 6.13 Administrativní úkony

Mezi každým partnerem a Národní knihovnou České republiky jako provozovatelem Manuscriptoria je uzavřeno licenční ujednání, které specifikuje vše výše uvedené.

### 6.14 Vlastní zahájení spolupráce

Prvním krokem k zahájení spolupráce může být vyplnění připraveného on-line dotazníku. Ten nám umožní udělat si základní obrázek o obsahu kolekce nabízené k agregaci a o jejích základních technických parametrech. Formulář dotazníku je dostupný on-line na stránkách projektu Manuscriptorium.

### 6.15 Nástroje pro koncové uživatele

Veškeré nástroje v systému Manuscriptoria jsou dostupné uživatelům bezplatně, část z nich vyžaduje elektronickou registraci. Jen v rychlém souhrnu v příspěvku uvedeme přehled základních nástrojů, které mají registrovaní koncoví uživatelé v systému k dispozici:

- *Oblíbené položky*: celé dokumenty nebo jejich jednotlivé obrazy lze označit hvězdičkou a uživatel k nim pak má kdykoliv přímý přístup ze svého účtu
- *Kolekce záznamů*: jednotlivé dokumenty z katalogu lze sdružovat do pojmenovaných (tematických) kolekcí
- *Virtuální dokumenty*: z digitálních obrazů agregovaných v Manuscriptoriu lze složit vlastní virtuální dokument, v jednom

virtuálním dokumentu je možné shromáždit obrazy z více různých předloh

- *Oblíbené dotazy*: jednoduché i s ložitě dotazy nebo sekvence dotazů lze uložit a kdykoliv zopakovat jediným kliknutím
- *Anotační nástroje*: uživatelé mají možnost tvořit krátké poznámky nebo delší strukturované texty a připojovat je k jimi vytvořenému obsahu (virtuálním dokumentům, kolekcím, oblíbeným dotazům)

Naším cílem je vytvářet aktivní a kooperativní prostředí, a proto umožňujeme uživatelům vytvářený obsah publikovat také ve formě článků na webových stránkách Manuscriptoria, tak zvaných „Data articles“. Tento obsah je prohledávatelný a dostupný jak prostřednictvím website projektu, tak vzájemným prolinkováním „oficiálního“ agregovaného obsahu se souvisejícími Data articles.

### **6.16 Kde je Manuscriptorium dostupné**

Veškeré informace o psaném kulturním dědictví snadno naleznete na [www.manuscriptorium.com](http://www.manuscriptorium.com). Rádi s vámi budeme spolupracovat!  
Manuscriptorium

[www.manuscriptorium.com](http://www.manuscriptorium.com)

## 7 ČESKÁ DIGITÁLNÍ KNIHOVNA<sup>4</sup> KRAMERIUS

### 7.1 Abstrakt

ČDK slouží jako jednotné rozhraní pro koncové uživatele i jako primární národní zdroj dat pro další zastřešující

KNIHOVNY.CZ. V rámci projektu jsou vyvíjen

, je používán v několika knihovnách pro produkci digitálních dokumentů. S ohledem na archivační součást řešení jsou sledovány standardy pro dlouhodobou archivaci (např. model O AIS) a zajištěno propojení se systémem Archivematica.

### 7.2 Klíčová slova

komunikační protokoly, standardy, interoperabilita, úložiště, archivace

### 7.3 Úvod

řešení podporujících vytváření digitálních dokumentů, jejich zpřístupňování a archivaci jsou hlavní cíle projektu Česká

(například OAI-PMH, Z39-50, SRU/SRW) a formátů metadat (např Dublin Core, METS, MODS, mixu, PREMIS, FOXML).

### 7.4

#### 7.4.1 Úvod

. Bude sloužit jako jednotné rozhraní pro koncové uživatele i jako prim

---

<sup>4</sup> LHOTÁK, Martin. 2016. Česká digitální knihovna – agregace a zpřístupnění digitálního obsahu z českých knihoven. In: *Libraries V4 in the Decoy of Digital Age. Proceedings of 6th Colloquium of Library and Information Experts of the V4+ Countries held from 31st May – 1st June 2016 in Brno*. – Brno : Moravská zemská knihovna v Brně, 2016. S. (eng) 345-351; (cze) 352-356. – ISBN 978-80-7051-216-6 (brož.)

— jedná se o systémy Kramerius, ProArc a RDflow a částečně jsou využita již zavedená řešení používaná v České republice po mnoho let, mezi něž patří zejména Registr digitalizace a knihovní informační systémy (katalogy knihoven).

#### 7.4.2 DIGITÁLNÍ KNIHOVNA KRAMERIUS

dokumentům pocházejících z digitalizace. Je určen především pro zpř

. Systém poskytuje rozhraní pro koncové uživatele, které umožňuje vyhle

sbírkami digitálních dokumentů. Navazuje na funkčnost předchozí verze systému Kramerius,

— spuštěn jako webová aplikace v jakémkoli J2EE kontejneru (např. Apache Tomcat). Více než 30 knihoven v České republice (většina z nich patří mezi největší v ČR) používá systém Kramerius pro provozování své vlastní digitální knihovny. Všechny instalace dohromady o

knihovnami v České republice. Česká digitální k

knihoven - KNIHOVNY.CZ.

#### 7.4.3 REGISTR DIGITALIZACE

Registr digitalizace byl založen jako projekt Knihovny Akademie věd a Národní knihovny České republiky. Z hledí

, vlastníka a podl

dokumentů

, což vede k lepšímu využití finančních prostředků. V současné době Registr digitalizace obsahuje téměř 320 000 titulů. Registr digitalizace poskytuje zároveň řešení, které monitoruje digitalizační workflow. Spolupracuje přitom s knihovními katalogy a digitálními knihovnami a repozitáři. V rámci interoperability Registr komunikuje s knihovními informačními systémy a umožňuje dávkové importy ve fo

. Na konci celého procesu odešle Registr informací o digitalizovaném dokumentu do veřejného knihovního OPACu společně s odkazem na tento dokument v digitální knihovně. Informace jsou odeslány také do Souborného katalogu České republiky.

#### 7.4.4 PROARC —

#### IGITÁLNÍCH DOKUMENTŮ A JEJICH ARCHIVACI

. ProaArc umožňuje rychlé částečně automatizované vytváření standardních metadat. Skládají se ze strukturální, deskriptivní a archivní části a OCR. Z hlediska archivační součásti řešení jsou zohledněny standardy pro dlouhodobou archivaci, zejména ISO normy referenční model OAIS. ProArc e —

systemu Archivematica.

#### 7.4.5 ČESKÁ DIGITÁLNÍ KNIHOVNA — PORTÁL

bod k obsahu digit

. Cílem je zajistit přístup k digitálním dokumentům v knihovnách z jednoho místa. Česká digitáln

dynami

/původních digitálních knihovnách.

V ČDK

doplňován. Česká digitální knihovna je veřejně přístupná z <http://www.czechdigitallibrary.cz/>.

#### 7.4.6 ZÁVĚR

Česká digitální knihovna je základní infrastrukturou na národní úrovni, která zajišťuje přístup k digitálním dokumentům a zároveň poskytuje data dalším centrálním řešením na národní i mezinárodní úrovni.

. Hlavním cílem je umožnit vytváření, sdílení a opakované využívání digitálního obsahu snadno a v České republice pod jednou střechou.

#### 7.4.7 REFERENCE

- FOLTÝN, Tomáš. The Kramerius System — Open Source Solution for Digital Libraries. In: Proceedings of the Third Workshop on Very Large Digital Libraries. Glasgow, Scotland (UK) September 10, 2010, Pisa 2010, ISBN 978-88628015-7.
- FOLTÝN, Tomáš. Kramerius Information System as a Tool of Access to Digital Documents. In: PRESIDENTIAL LIBRARY COLLECTIONS, DigitalLibrary series. Issue 3, Coordination and Standardization in the Sphere of Creation and Use of the National Information Resources 2012, ISBN 978-5-905273-25-4.
- FOLTÝN, Tomáš. Has It Been Already Digitized? How to Find Information about Digitized Documents. In: Review of the National Center for Digitization, Faculty of Mathematics, Belgrade, 2013, Issue: 22, ISSN: 1820-0109.
- FOLTÝN, Tomáš. Registrdigitalizace.cz, ITL lib 2010/2, Bratislava 2010.
- LHOTÁK, Martin. Česká digitální knihovna a nástroje pro zajištění digitalizačních procesů. [The Czech Digital Library and Tools for the Management of Complex Digitization Processes.] [online]. INFORUM 2012: 18. ročník konference o profesionálních informačních zdrojích 2012, 1. ISSN 1801-2213- [INFORUM 2012: konference o profesionálních informačních zdrojích /18./. Praha, 22.05.2012-24.05-2012] [cit. 2016-04-24]. K dispozici na: <http://hdl.handle.net/H104/0218127>.
- LHOTÁK, Martin a Tomáš FOLTÝN. Registr digitalizace CZ. [Digitization Registry CZ.] [online]. INFORUM 2010. Praha: Albertina icome Praha, 2010. ISSN 1801-2213. [INFORUM 2010. Konference o profesionálních informačních zdrojích /16./. Praha (CZ), 25.05.2010-27.05.2010] [cit. 2016-04-24]. K dispozici na: <http://hdl.handle.net/11104/0194066>.



## 8 PLÁNOVÁNÍ A PŘÍPRAVA PROCESU DIGITALIZACE V KNIHOVNĚ

Kultúrne inštitúcie, ako knižnice, múzeá, archívy a organizácie venujúce sa prírodnému a environmentálnemu dedičstvu sú hlavnými dodávateľmi digitálneho obsahu, no musia sa mobilizovať. Ich úsilie je potrebné koordinovať, aby sa čo najlepšie využili existujúce technológie a prispelo sa k tvorbe, využívaniu a poskytovaniu lokálneho kultúrneho obsahu, ktorý vyhovuje potrebám všetkých občanov.

Prostredníctvom tohto aktualizovaného plánu je na realizáciu vízie Európskeho kultúrno-informačného priestoru sledovaných šesť cieľov, pričom sa berie do úvahy a ďalej rozvíja predchádzajúci *Lundských zásady*. Rovnako sa pri plánovaní digitalizácie využívajú aj zistenia a návrhy projektu Minerva a najlepšia prax inštitúcií, ktoré už majú s digitalizáciou praktické skúsenosti.

### 8.1 Praktické odporúčania

V tejto kapitole sú prezentované najdôležitejšie praktické vedomosti získané tímom pre najlepšie praktiky projektu Minerva. Zameriava sa na praktické pravidlá, ktoré by mali organizácie zobrať do úvahy pri začatí, realizácii alebo riadení digitalizačných projektov vo sfére kultúry. Odporúčania sú rozdelené do nasledovných oblastí, z ktorých každá odráža nejakú etapu projektu digitalizácie:

1. Plánovanie projektu digitalizácie
2. Výber zdrojového materiálu na digitalizáciu
3. Príprava na digitalizáciu
4. Zaobchádzanie a práca s originálmi
5. Proces digitalizácie
6. Uchovávanie originálneho digitálneho materiálu
7. Metaúdaje
8. Zverejnenie
9. Práva duševného vlastníctva a copyright
10. Manažment projektov digitalizácie

### 8.2 Projekt Minerva a Lundské zásady

Projekt Minerva, začal v roku 2002 pod vedením talianskeho Ministerstva kultúry (zmluva IST 2001-35461). Projekt zahŕňal zástupcov príslušných vládnych ministerstiev alebo centrálnych štátnych agentúr z mnohých členských štátov EÚ so spoločným cieľom presadzovania spoločného prístupu a metodológie digitalizácie európskeho kultúrneho materiálu. Projekt uznáva jedinečnú hodnotu európskeho kultúrneho dedičstva a strategickú úlohu, ktorú môže zohrávať v rastúcom priemysle digitálneho priemyslu v Európe. Uznáva aj hodnotu koordinácie úsilia národných vlád a kultúrnych organizácií, aby sa zvýšila úroveň syntézy a synergie digitalizačných iniciatív.

### 8.3 Lundské zásady

Dňa 4. apríla 2001 sa stretli predstavitelia a experti Európskej komisie a členských štátov v meste Lund vo Švédsku (v rámci švédskeho

predsedníctva), aby diskutovali o tom, ako koordinovať a pridať hodnotu národným programom digitalizácie na európskej úrovni. Stretnutie vyústilo do publikovania súboru všeobecných zásad, ktorými sa majú riadiť verejné digitalizačné iniciatívy a ich koordinácia. Tieto zásady (Lundské zásady) boli transformované do Lundského akčného plánu, ktorý stanovuje zoznam úkonov, ktoré majú vykonať členské štáty, Komisia, Komisia a členské štáty spoločne, aby sa vylepšili podmienky digitalizácie v Európe.

(1) **Zabezpečiť vedúcu úlohu národných pamäťových** inštitúcií v dynamickom a meniacom sa prostredí, v ktorom sa uskutočňuje rýchly technologický a ekonomický rozvoj.

(2) **Posilniť koordináciu** a presadzovať silnejšie prepojenie medzi digitalizačnými iniciatívami inštitúcií, systémov a projektov prostredníctvom riadiaceho grémiu zriadeného pri ministerstve kultúry (Rada pre informatizáciu a digitalizáciu)

(3) Pokračovať v snahe o **prekonávanie fragmentácie a duplikácie** digitalizačných aktivít a o maximalizáciu synergií sprístupnením štandardov, metodík a odporúčaní pre projekty digitalizácie

(4) Hodnotiť a **identifikovať vhodné modely**, spôsoby financovania a politické prístupy na udržanie rozvoja a stratégie dlhodobého uchovávanía.

(5) Presadzovať **kultúrnu a jazykovú rôznorodosť** prostredníctvom tvorby digitálneho obsahu.

(6) Zlepšovať on-line **prístup** ku kultúrnemu obsahu.

(7) Prispôbovať digitálny obsah pre špecifické **využívanie** (napr. v školách).

#### 8.4 Plánovanie projektu digitalizácie

Plánovanie projektu je prvým krokom akéhokoľvek projektu digitalizácie. Čas venovaný plánovaniu sa oplatí tak, že sa uľahčí riadenie a realizácia projektu. Za normálnych okolností je potrebné odpovedať na nasledovné otázky:

- **Čo** sa musí urobiť?
- **Kto** to bude robiť?
- **Kde** sa to bude robiť?
- **Kedy** sa to bude robiť?
- **Ako** sa to bude robiť?

Projekt digitalizácie musí mať jasne špecifikované ciele, pretože majú priamy vplyv na výber, copyright a zverejnenie materiálu. Do projektu by mal byť zapojený vhodný personál s vyhovujúcimi vedomosťami a zručnosťami. Mal by zahŕňať aj plán školení, aby sa zaručila dostatočná odbornosť, ktorú môže projekt vyžadovať.

Projekt by sa nemal začať bez toho, aby sa uskutočnil prieskum iných projektov v rovnakej oblasti. Takýto výskum identifikuje problémy, ktorým sa treba venovať, bude stimulovať nové nápady a oblasti, ktoré ešte neboli zvažované a pridať hodnotu výstupu projektu.

Výskum taktiež pomôže určiť množstvo práce, ktorú treba naplánovať pre realizáciu projektu pomocou stretnutí alebo inej komunikácie s organizáciami, ktoré realizovali podobné projekty. Takáto interakcia pomôže stanoviť, či vaša organizácia má potrebný personál, vedomosti a technologickú infraštruktúru na realizáciu projektu, alebo či bude potrebné školenie a iná príprava.

Je výhodné investovať určitý čas do zistenia statusu copyrightu materiálu, ktorý sa ma digitalizovať. Neúspech pri získaní povolenia digitalizovať a zverejniť na webe môže spôsobiť neúspešnosť celého projektu, bez ohľadu na technickú odbornosť a skúsenosti.

Môže sa zväziť aj možnosť realizácie technického pilotného projektu na začiatku projektu, aby sa zabezpečilo to, že akékoľvek anomálie alebo problémy pri technickej realizácii sa odstránia pred spustením hlavného projektu

## 8.5 Zdôvodnenie projektu

### Definícia problému

Každý projekt digitalizácie má vlastné zdôvodnenie. Medzi dôvody často patrí **sprístupnenie** kultúrnych zbierok cez internet, ktoré by boli v opačnom prípade využívané len málo. Ďalším dôvodom je **ochrana** krehkých zbierok pred opotrebovaním pri bežnom používaní. V ďalších prípadoch sú projekty digitalizácie realizáciou spolupráce medzi organizáciami a zahŕňajú zriadenie portálov, sietí atď.

Dôvody projektu majú silný vplyv na kritéria výberu materiálu, ktorý sa má digitalizovať. Často majú vplyv aj na riadenie projektu, metaúdaje a on-line zverejnenie výstupu projektu (ak k nemu dôjde), kontrolu kvality atď. „**Prečo**“ je najdôležitejšou otázkou, ktorú si treba položiť pred začatím projektu digitalizácie.

### Pragmatické návrhy

- Projekt musí mať konkrétne, explicitné ciele, ktoré sa musia zdokumentovať.
- Ciele projektu by mali byť realistické v porovnaní s prostriedkami, ktoré sú k dispozícii.
- Všetky kroky projektu sa musia overiť podľa týchto cieľov, aby sa zaručilo, že práca vykonávaná v projekte prispieva k dosiahnutiu splnenia smerníc.
- Ciele projektu by mali dokumentovať hodnotu, akú projekt prinesie inštitúciám zapojeným do projektu. Ak sa má do projektu investovať čas a úsilie, musí byť jasné zaradenie projektu z inštitucionálneho pohľadu.

## 8.6 Ľudské zdroje

### Definícia problému

Pred tým, ako môže projekt začať, je dôležité, aby bol k dispozícii personál potrebný pre prácu na projekte. Mnohé kultúrne inštitúcie nemajú veľký počet zamestnancov s dostatkom voľného času na vykonávanie digitalizácie mimo ich zvyčajnej pracovnej náplne. Aj vedomosti potrebné pre projekt digitalizácie sa môžu líšiť od zručností potrebných na vykonávanie bežných úloh každodennej

prevádzky. Preto je potrebné identifikovať hardvérové a softvérové riešenia potrebné pre projekt digitalizácie.

### **Pragmatické návrhy**

- Zabezpečte dostatočný počet zamestnancov na realizáciu projektu
- Pridel'te zamestnanca ku každej úlohe alebo pracovnému balíku projektového plánu.
- Identifikujte požiadavky na školenie, vrátane IT školení a inštruktáži o zaobchádzaní s jemnými artefaktmi a dokumentmi.
- Ak je to možné, realizujte školenia pred začatím projektu pomocou tých istých hardvérových a softvérových riešení, ktoré sa budú používať počas projektu (firmy niekedy ponúkajú technické riešenia zdarma na školenie, na krátky čas sa dajú prenajať aj potrebné zariadenia)
- Zamerajte sa skôr na malé jadro šikovných zamestnancov nadšených projektom ako na veľkú skupinu "príležitostných" zamestnancov.

### **Poznámky/komentáre**

Aj keď je materiál prezentovaný v tomto odporúčaní spoločný pre všetky scenáre riadenia digitalizácie, je potrebné túto vec zopakovať: existuje reálne riziko, že môžu vzniknúť nenapraviteľné škody na nenahraditeľných artefaktoch a dokumentoch pri nesprávnom zaobchádzaní.

## **8.7 Výskum**

### **Definícia problému**

Bez ohľadu na rozsah konkrétneho projektu sa dá predpokladať, že v minulosti sa už realizovali podobné projekty. Je vysoká pravdepodobnosť, že informácie o takýchto projektoch sú dostupné na internete alebo publikované inde v časopisoch a pod.

Výskum v danej oblasti môže ako súčasť procesu plánovania projektu pomôcť identifikovať hardvérové a softvérové riešenia prichádzajúce do úvahy, naplánovať pracovný postup a proces a vyhnúť sa problémom a prekážkam, ktoré sa vyskytli v iných projektoch.

### **Pragmatické návrhy**

- Už vo fáze plánovania vykonajte prieskum všetkých podobných projektov, ktoré sa zaoberajú podobným záležitosťami ako plánovaný projekt. Na internete je ich mnoho.
- Výskum pomáha predchádzať chybám. Môže skontaktovať projektový tím s inými, ktoré uskutočnili podobné projekty a poskytnúť možnosť učiť sa z ich skúseností.
- Vlastný výskum pridáva uznanie a hodnotu výstupu akéhokoľvek projektu. Výsledok projektu zlepší aj štúdium diel iných, aby sa zaručilo, že projekt sa nebude uskutočňovať vo vákuu.

### **Poznámky/komentáre**

Mnohé kultúrne projekty digitalizácie sú financované z verejného rozpočtu a je na ne kladená požiadavka, aby ich zistenia a správy boli zverejnené. Publikovanie môže byť na internete, alebo cez iné vhodné médiá.

Projektové tímy sa často s radosťou delia o ich skúsenosti a výsledky – pridáva to hodnotu ich dielu.

## 8.8 Hrozby

### Definícia problému

Na začiatku akéhokoľvek projektu je potrebné naplánovať, ako sa zaručí úspešnosť výstupu. Cieľom však nie je odstrániť všetky hrozby ale pripraviť sa na ne tak, že sa vytvorí projektový rámec, v rámci ktorého bude možné efektívne reagovať na nepredvídané okolnosti. Cieľom je vytvorenie projektu s personálom a postupmi, ktoré sa môžu prispôbiť zmenám. Preto je do plánovania projektu potrebné zahrnúť analýzu rizík.

### Pragmatické návrhy

- Distribúcia digitalizovaných obrazov cez Internet je forma publikovania a preto sa na ňu vzťahujú autorské zákony (copyright) a právo duševného vlastníctva (IPR). Pri analýze rizík sa kladú nasledujúce otázky:
  - Aké by mohli byť následky použitia materiálu bez špecifického povolenia?
  - Vyvinulo sa úsilie na zistenie držiteľa práv?
  - V prípade porušenia copyrightu, aký by mohol byť dopad na projekt?
- Pri verejných informáciách je dôležitou otázkou právna hodnota informácie. Aké kroky sa podnikli na zaručenie toho, že digitalizovaný zdrojový materiál nie je narušený a že sa vyprodukoval oprávnenou inštitúciou?
- Musí sa zaručiť autenticita. Aké úkony sa podnikli na udržiavanie obrazových súborov a aké nástroje sa použili?
- Financovanie projektu by mohlo predstavovať problém a možné ohrozenie možností dosiahnuť ciele projektu.
- Kľúčovou je otázka úrovne zručností v projekte. Je možné najat' nové vysoko zručné a skúsené osoby? Ak nie, aký to bude mať vplyv na projekt?

## 8.9 Výber zdrojového materiálu na digitalizáciu

Výber materiálu na digitalizáciu je dôležité rozhodnutie. Ideálnou voľbou je zvyčajne digitalizácia celej zbierky alebo fondu, to je však realizovateľné len v ojedinelých prípadoch, preto je nutný určitý výber. Kritéria výberu sa budú meniť v závislosti od cieľov projektu digitalizácie, napríklad on-line zdroj pre školy sa bude riadiť učebnými osnovami, múzeá vyberú najznámejšie predmety na zvýšenie počtu návštevníkov alebo najvzácnejšie artefakty, aby sa znížil dopyt po „blízkom“ skúmaní. Toto samozrejme nie sú jediné problémy, ktorým sa treba venovať pri stanovení kritérií výberu – dôvody výberu určitého materiálu sa budú meniť od projektu k návrhu, rovnako aj dôvody, prečo nedigitalizovať. Medzi príklady ďalších dôvodov patria právne obmedzenia, inštitucionálna politika, technická náročnosť digitalizácie, existencia digitálnej kópie atď.

## 8.10 Stanovenie kritérií výberu

### Definícia problému

Pri plánovaní projektu digitalizácie je kriticky dôležitá voľba, aký materiál sa má digitalizovať. Kritéria výberu budú závisieť od cieľov projektu, rovnako aj od technických a finančných obmedzení, otázok copyrightu a práv duševného vlastníctva a od iných projektov v danej oblasti.

### **Pragmatické návrhy**

- Musia sa stanoviť kritériá na výber materiálu, ktorý sa má digitalizovať. Kritériá výberu musia odrážať celkové ciele projektu. Mali by sa zväžiť minimálne nasledujúce kritériá:
  - Prístup k materiálom by inakšie nebol možný, alebo by bol obmedzený.
  - Širší a ľahší prístup k veľmi obľúbenému materiálu
  - Stav originálov
  - Uchovávanie chýlostivých materiálov tým, že sa sprístupnia ich digitálne verzie ako alternatíva
  - Téma projektu
  - Copyright a práva duševného vlastníctva
  - Dostupnosť existujúcich digitálnych verzií
  - Náklady na digitalizáciu
  - Vhodnosť zdrojového materiálu na prezeranie v režime online
- Kritériá výberu musia byť explicitné a prediskutované všetkými zainteresovanými účastníkmi pred výberom alebo digitalizáciou.
- Kritériá výberu sa musia plne zdokumentovať, aby boli všetky rozhodnutia digitalizovať/nedigitalizovať jasné počas celého projektu.

### **Poznámky/komentáre**

Kultúrne organizácie majú bežne jadro vysoko hodnotného a atraktívneho materiálu, ktorý je bez diskusie zahrnutý do projektu digitalizácie a jeho úlohou je reprezentovať danú inštitúciu.

Veľká časť všetkých digitalizačných projektov majú za cieľ zverejnenie na webe. To znamená, že je potrebné zväžiť záležitosti copyrightu a práv duševného vlastníctva týkajúcich sa akéhokoľvek materiálu ešte pred výberom.

## **8.11 Výber podľa kritérií**

### **Definícia problému**

Po stanovení kritérií, aký materiál sa má digitalizovať sa môže začať vlastný proces výberu. Táto príručka navrhuje, ako sa má tento proces riadiť.

### **Pragmatické návrhy**

- Každý kandidát na digitalizáciu sa musí hodnotiť podľa kritérií výberu. Mali by sa poznamenať prípady, kedy sa kritérium výberu nespĺnilo. Ak budú v takýchto prípadoch odmietnuté dôležité alebo kritické objekty, môže sa stať, že bude potrebné znovu revidovať kritériá. Ak sa toto vyskytne, nové kritériá by sa mali zaznamenať.
- Po výbere objektu na digitalizáciu by sa mali podrobnosti o ňom zadať do bázy poznatkov riadenia digitalizácie (pozri kapitolu *Manažment projektu digitalizácie*)

### **Poznámky/komentáre**

V tomto štádiu sa projekt zaoberá každým dokumentom/objektom, ktorý sa má digitalizovať, po prvý krát. Toto je optimálna príležitosť pre projekt, aby vytvoril bázu poznatkov o všetkých dokumentoch/objektoch, ktoré sú v rozsahu projektu. Takáto báza poznatkov podporí manažment projektu a pomôže zaručiť, že na zaobchádzanie so vzácnymi artefaktmi boli nadobudnuté odborné znalosti a obsahuje aj také bežné informácie, ako je umiestnenie originálov.

## 8.12 Príprava na digitalizáciu

Pred tým, ako môže začať digitalizácia, musí sa pripraviť vhodné prostredie a systém hardvéru a softvéru. Medzi prvky takéhoto prostredia patrí hardvér pre samotný proces digitalizácie (skenery, digitálne fotoaparáty, stojany, iný hardvér), počítačová infraštruktúra, ku ktorej je pripojený hardvér, softvér na zachytávanie a spracovanie obrazu, rovnako aj softvér pre metaúdaje a kontrolu metaúdajov. Pracovné prostredie by malo vyhovovať digitalizovanému materiálu a zvláštnu pozornosť je potrebné venovať svetlu, vlhkosti, vibráciám, rušivým vplyvom, presunu originálov atď.

### 8.12.1 HARDVÉR

#### **Definícia problému**

Je potrebné zabezpečiť vhodné technické zariadenia na digitalizáciu, akými sú zariadenia na zachytenie digitálneho obrazu (digitálne fotoaparáty a kamery, skenery na knihy, dokumenty alebo mikrofilmy, audio a video hardvér) pripojené na vhodnú počítačovú platformu (počítač, operačný systém, sieť). Je možné rozlíšiť dve rôzne metódy digitalizácie: skenovanie a používanie digitálnych kamier/fotoaparátov.

#### **Pragmatické návrhy**

- Pred začatím digitalizácie sa musí nainštalovať vhodný hardvér a skontrolovať jeho kvalita a funkčnosť.
- Na účely ohodnotenia zariadení na zhotovenie digitálneho obrazu by sa mali použiť relevantné skúšobné materiály.
- Kým sa hardvérové prostredie úplne nepripraví a neodskúša na necitlivých materiáloch, nemali by byť prítomné žiadne zdrojové materiály.
- Pre materiál, ktorý sa nepoškodí pritlačením o tvrdý povrch, bude vhodný plošný skener (napríklad rozviazaný tlačенý materiál a rukopisy, fotografie)
- Na účely projektu by sa mal zaobstarať čo najväčší skener. Nemalo by sa pristupovať k mozaikovitému skenovaniu alebo k skladaniu materiálov. Je potrebné mať na pamäti, že preprava veľkého skenera (napr. A0) nie je triviálna záležitosť.
- Plošný skener by sa mal použiť len vtedy, ak je materiál plochý a nepoškodí sa pri pridržaní na plochom tvrdom povrchu. Pre mnohé viazané dokumenty bude vhodný skener s knižnou kolískou, až do príslušnej veľkosti. Mnohé materiály, ktoré nie sú ploché, ani vhodné pre knižnú kolísku, budú na zhotovenie obrazov vyžadovať digitálny fotoaparát/kameru.
- Pri použití skenera by malo mať toto zariadenie aspoň taký rozsah, ako je veľkosť dokumentu/objektu, ktorý sa má skenovať.

- Zhotovenie obrazu by sa malo realizovať pri čo najvyššom primeranom rozlíšení. Takto budú vznikať veľké súbory, z ktorých sa môžu extrahovať menšie verzie - napríklad na účely prezentácie cez Web. Z obrazu s nižšou kvalitou nie je za žiadnych okolností možné získať obraz s vyššou kvalitou.
- Definícia “primeraného” rozlíšenia závisí od povahy snímaného materiálu a od spôsobov využitia digitálnych obrazov. Napríklad ak sa skenované obrázky majú použiť len ako miniatúrne náhľady, skenovať sa môže pri nižšom rozlíšení. Vo všetkých prípadoch však musí rozlíšenie umožniť zachytenie najvýznamnejších detailov dokumentu alebo predmetu. Je ťažké odôvodniť použitie vyššieho rozlíšenia, ak skenovanie s vyšším rozlíšením neposkytuje viac informácií ako skenovanie pri nižšom rozlíšení.
- Zachytávaním obrazu by sa mal vytvárať formát súborov, ktorý je bezstratový, čiže nekomprimovaný. Typicky sa používa formát TIFF (Tagged Image File Format).
- Projekt by mal použiť čo najvýkonnejšiu a flexibilnú digitálnu kameru (fotoaparát), akú je možné zaobstarať. Žiadnym ďalším spracovaním nie je možné prekonať obmedzenia digitalizačného zariadenia. Je potrebné poznamenať, že digitálny “zoom” neposkytuje lepšiu kvalitu obrazu, len sa ním zobrazuje menej bodov (pixelov) na jednotku plochy zobrazenia. Na zachytenie detailov sú najdôležitejšie tieto tri parametre: počet bodov (pixelov) v obraze, bitová hĺbka a kvalita použitých optických šošoviek.
- Je dôležité mať k dispozícii vhodné stojany na uchytenie a pridržanie materiálov pri digitalizácii.
- Mala by sa použiť digitálna kamera/fotoaparát s účelovým stojanom. Kamera alebo fotoaparát by mali byť nainštalovane na trojnohom statíve a podľa potreby by mali mať doplnkové osvetlenie, filtre a pod.
- Fotografická rovina a rovina materiálu musia byť rovnobežné, aby sa obraz neskreslil.
- Súčasťou fotografických príprav musí byť vhodné osvetlenie. Je málo pravdepodobné, že bude postačovať len okolité svetlo. Svetelné podmienky musia byť stabilné.
- Na zníženie skreslenia farieb by sa mali použiť vhodné filtre.
- K zariadeniam musí byť pripojený počítač s veľkým úložným priestorom. Údaje na tomto počítači by sa mali v krátkych a pravidelných intervaloch zálohovať.
- Ak sa musí obraz skenovať po častiach, mal by sa ponechať priestor niekoľko centimetrov na prekrytie, aby sa zabránilo vzniku medzier medzi jednotlivými časťami. Pre všetky časti by sa mali použiť rovnaké nastavenia, aby nedošlo ku efektu „zlátaniny“.

### **Poznámky/komentár**

Použitý hardvér je hlavným obmedzením kvality konečného výstupu projektu digitalizácie. Pokiaľ projekt sa v rámci projektu nedigitalizujú len ploché materiály, ktoré sa dajú skenovať bez poškodenia väzby, rámov alebo samotného materiálu, bude potrebné použiť digitálnu kameru/fotoaparát. Môže sa použiť aj analógový fotoaparát a následne z neho skenovať diapozitívy, no digitálny fotoaparát má mnohé výhody z hľadiska času, úsilia a kvality.



Ak má projekt obmedzený životný cyklus, môže byť výhodný prenájom hardvéru. Ďalšou alternatívou je využitie externých agentúr na realizáciu digitalizácie v prospech kultúrnych subjektov zapojených do projektu.

### 8.12.2 SOFTVÉR

#### **Definícia problému**

Po vytvorení digitálnej verzie objektu je pravdepodobné, že výsledný súbor bude požadovať spracovanie pred tým, ako sa bude môcť použiť. Medzi úpravy patria korekcia farby, orezanie obrazu, kompresia na menší súbor (napr. náhľady, obraz na prezentáciu cez web).

#### **Pragmatické návrhy**

- Kalibračný proces má začať ihneď po zapnutí skeneru alebo digitálnej kamery.
- Na zužitkovanie matričných súborov bude potrebný vhodný softvér na spracovanie obrazov. Aj keď digitalizačný hardvér je často poskytovaný s nejakým softvérom, pre projekt digitalizácie nemusí byť dostatočne výkonný a flexibilný.
- Požiadavky na softvér závisia na cieľoch projektu. Je vhodné poznamenať, že ak sa matričné súbory akýmkoľvek spôsobom nezmenia, na ich spracovanie sa môžu použiť rôzne typy softvérov. Avšak náklady na čas a úsilie môžu byť významné a zvyčajne prevýšia náklady na výkonnejší softvérový balík.
- Projekt by mal získať ten najvhodnejší a najvýkonnejší softvér, ktorý si môže dovoliť.
- Absolútne minimum, ktoré musí softvér byť schopný vykonať je:
  - otvárať veľmi veľké obrazové súbory
  - meniť rozlíšenie a hĺbku farby
  - uchovávať viaceré rôzne verzie, s rôznymi veľkosťami súborov.
  - vyberať a kopírovať časť obrazu a uchovať ho ako iný súbor.
  - exportovať obrazy v rôznych formátoch súborov, vrátane webových štandardov JPEG a GIF.
- Túto úroveň funkčnosti poskytuje niekoľko voľne dostupných softvérových balíkov; avšak investícia do komerčného produktu sa oplatí z hľadiska ušetrenia času, úsilia, dokumentácie a technickej podpory.
- Ak má projekt aj zložku optického rozoznávania znakov (OCR), je dôležitá voľba softvéru pre OCR. Všetky úkony spojené s OCR vyžadujú určitú dávku ručných úprav a opráv, spôsob, akým tieto funkcie softvérový produkt podporuje, má významný vplyv na čas a úsilie vyžadované projektom. Lepšie balíky OCR umožňujú prehľad a úpravy na jednej obrazovke, navrhujú opravy zle prečítaných slov, podporujú rôzne spôsoby rozloženia textu a obrazov, obsahujú mnohé slovníky a pod. Oplatí sa zhodnotenie viacerých softvérových balíkov pre OCR, ak projekt zahŕňa viac ako jednu osobu za rok.

#### **Poznámky/komentár**

Správna voľba softvéru ušetrí projektu množstvo času a práce. Ak má projekt značné trvanie (napr. viac ako dve osoby počas viac ako pol roka), je vhodné ohodnotiť viac softvérových balíkov, aby sa vybral produkt, ktorý najlepšie vyhovuje požiadavkám projektu

## 8.13 Proces digitalizácie

Táto kapitola poskytuje praktické odporúčania pre samotný proces digitalizácie. Technické riešenia na digitálne snímanie sa môžu meniť. Detailnejšie je opísaná práca so skenermi, digitálnymi kamerami/fotoaparátmi alebo softvérovými aplikáciami na optické rozoznávanie znakov (OCR), nakoľko ide o oblasti relevantné pre väčšinu projektov. Neuvažuje sa tu o digitalizácii priehľadných originálov (napr. mikrofilm).

### 8.13.1 POUŽÍVANIE SKENEROV

**Skenery 4DigitalBooks**

- **Automatické knižné skenery**  
Digitizing Line  
Exkluzívna technológia založená na automatickom bezobslužnom obracaní strán (2-6 ks)
- **Manuálne knižné skenery**  
CopiBook & Scanbase  
IS spoločnosť je podielníkom a obchodným partnerom
- **Softvér na zlepšovanie obrazu**  
Page Improver  
Exkluzívna technológia automatického delenia a usporiadania strán

#### Definícia problému

Plošné skenery sú najbežnejším nástrojom na digitalizáciu. Najrozšírenejšie modely A4 a A3 sú relatívne lacné, nenáročné na použitie a ak sa raz zabehne pracovný tok, zvládnu relatívne veľké množstvo materiálu. Veľkoplošné skenery (až do A0) a skenery vybavené knižnými kolískami sú veľmi drahé a požadujú dlhodobé projekty/programy, veľkokapacitnú digitalizáciu alebo nadrozmerný zdrojový materiál.

#### Pragmatické návrhy

- Na plošnom skeneri snímajte len materiály, ktoré sa nepoškodia pri pritlačení o pevný povrch. V prípade pochybností sa poraďte s odborníkmi.
- Zabezpečte neustálu a úplnú čistotu sklenenej skenovacej plochy, čím zvýšite kvalitu obrazu a chránite materiál pred znečistením.
- Ak je to možné, skenujte len také dokumenty/predmety, ktoré sa celé zmestia na plošný skener alebo skener vybavený knižnou kolískou.
- Ak je potrebné skenovať dokument/predmet po častiach, zabezpečte, aby zostalo dostatočné prekrytie umožňujúce znovu zmontovať obraz pri budúcom spracovaní (pomocou mozaikovacieho softvéru).

- Otestujte skener a jeho výstup na necitlivom materiáli pred tým, ako začnete skenovať originálny zdrojový materiál. Používateľov tiež vyškoľujte pomocou toho istého necitlivého materiálu.
- Zavedte konvenciu pridelovania názvov súborom vytvorených skenerom - napríklad podľa existujúceho katalogizačného systému – a/alebo pridávania názvov s významom. Názov súboru by mal umožniť mapovanie so zdrojovým dokumentom/predmetom.
- Aby sa maximalizovala schopnosť prenosu súborov cez viaceré počítačové platformy, mal by sa používať taký názov súboru, ktorý má osem znakov a po ktorom nasleduje pozostávajúce z maximálne troch znakov.
- Pred stanovením toku práce alebo dávkového pracovného procesu vykonajte ucelený proces skenovania a spracovania obrazu, aby ste sa uistili, že výsledkom toku práce je to, čo sa očakáva.
- Rozlíšenie a bitová dĺžka by mali byť čo najvyššie, ako je prípustné podľa dôvodu projektu, obmedzení skenera, podmienok na ukladanie údajov a atribútov zdrojového materiálu.
- Skenujte pri maximálnej vhodnej hĺbke farby, s tými istými limitmi uvedenými vyššie.
- Denne zálohujte pevné disky, na ktorých sú uložené údaje.
- Ďalej je dôležitá kontrola kvality digitálneho obrazu a metaúdajov. V čase skenovania je najlepšie zaoberať sa otázkou kvality. Je potrebné mať na pamäti nasledovné body:
  - Pre jednotlivé skupiny dokumentov/predmetov a ich častí sa musí stanoviť minimálne rozlíšenie.
  - Preskúmajte zoskenovaný výstup na obrazovke, na papieri, premietacom a v akomkoľvek inom formáte, v ktorom sa môže používať (napríklad na mobilnom zariadení)
  - Uistite sa, že obrazovky (monitory) sú spoľahlivo vykalibrované. Zabráňte tomu, aby na obrazovke alebo v jej okolí bol iný materiál, ktorý môže mať vplyv na vnímanie dokumentu/predmetu.
  - Matricové obrazy sa musia vytvárať s viditeľnými mierkami, farebné obrazy alebo obrazy so stupnicou šedej farby musia obsahovať aj štandardizované referenčné hodnoty.

### **Poznámky/komentáre**

Samotné skenovanie je relatívne nenáročná operácia. Je však užitočné vytvoriť si pracovný tok na zvýšenie efektívnosti a zníženie výskytu chýb.

Skenovanie nadrozmerných dokumentov/objektov alebo skenovanie pri veľmi vysokej kvalite vyžaduje veľké investície času a úsilia na jeden dokument/objekt. Tie je možné znížiť používaním hardvéru, ktorý vyhovuje povahe dokumentu (napr. veľkoplošný skener, knižná kolíska). V prípade, že nie je k dispozícii veľa prostriedkov na hardvér, vyhradte na skenovanie dostatok času.. Nemalo by sa zanedbať školenie o práci s nadrozmernými a nepravidelne tvarovanými materiálmi.



### 8.13.2 POUŽÍVANIE DIGITÁLNYCH KAMIER/FOTOAPARÁTOV

#### **Definícia problému**

Používanie digitálnych kamier/fotoaparátov sa stáva v digitalizačných projektoch čoraz bežnejšie. To vypovedá o ich flexibilita, hlavne ohľadom schopnosti digitalizovať neploché objekty, akými sú zviazané knihy, poskladané alebo pokrčené rukopisy a trojrozmerné predmety. Normálne sa však pri digitalizácii zviazaných kníh a nadrozmerných materiálov ako mapy a nákresy uprednostňuje skener vybavený knižnou kolískou

#### **Pragmatické návrhy**

- Pri obmedzenom rozsahu projektu je potrebné zvážiť prenájom vysoko-kvalitnej kamery/fotoaparátu.
- Digitálnu kameru/fotoaparát nainštalujte na motorizovaný vozík s možnosťou vertikálneho posuvu a umiestnite dokumenty/predmety, ktoré sa majú digitalizovať, na pevný podklad pri špeciálne prispôsobenom osvetlení.
- Zorganizujte cvičenie špecialistom na digitálne fotografovanie – rozdiel medzi obrázkami zhotovenými amatérom a tými istými fotografiami zhotovenými profesionálom môže byť obrovský.
- Vytvorte také pozadie, aby dokument/predmet na ňom vynikal.
- Zabráňte zmenám svetelných podmienok medzi jednotlivými zábermi medzi obrazmi rôznych častí alebo strán dokumentu/objektu – toto môže viesť k mylným dojmom z variácie farieb.
- Používajte apochromatické šošovky a vhodné filtre, aby nedošlo k nesprávnej registrácii farieb a skresleniu obrazu.

#### **Poznámky/komentáre**

Stúpajúce používanie digitálnych kamier/fotoaparátov v digitalizácii odráža ich dostupnosť ako všeobecne rozšírený konzumentský produkt s klesajúcou cenou. Stále však zostáva veľký rozdiel v cene aj kvalite medzi profesionálnymi digitálnymi kamerami/fotoaparátmi a masovo vyrábanými konzumentskými produktmi.

### 8.13.3 SOFTVÉROVÉ APLIKÁCIE NA OPTICKÉ ROZoznÁVANIE ZNAKOV (OCR)

#### **Definícia problému**

Mnohé projekty digitalizácie zahŕňajú digitalizáciu tlačенých dokumentov, akými sú knihy a noviny. To sa odohráva v tandeme s využívaním skenerov, aj keď nie výlučne. Používanie softvéru na optické rozoznávanie znakov (OCR) je populárnym spôsobom extrakcie informácií zo naskenovaných informácií a otvorenia možností spracovania informácií. Softvér OCR rozoznáva písmená a číslice, ktoré vytvárajú naskenovaný obraz (obrazový súbor s mapovanými bitmi), a exportuje ich ako textové súbory ASCII namiesto obrazových súborov. To umožňuje vykonať vyhľadávanie, indexovanie, konverzie formát a iné operácie na spracovanie.

#### **Pragmatické návrhy**

- Pred výberom konkrétneho produktu ohodnoťte viaceré ponúkané riešenia na rozoznávanie znakov (OCR, optické rozoznávanie znakov). Aj keď je softvér na OCR zahrnutý v cene skenera, výkonnejší softvér sa typicky predáva osobitne.
- Hlavným prvkom projektu OCR je identifikácia a manuálna oprava chýb, dvojzmyslov a miest, na ktorých sa text nedal spracovať. Balík OCR, ktorý poskytuje rozhranie prívetivé k používateľovi vykonávajúcemu túto úlohu, ušetrí čas a úsilie.
- OCR pracuje najlepšie pri dokumentoch v dobrom stave – skladané, pokrčené a vyblednuté zdrojové materiály zvýšia počet chýb v procese OCR. Aby ku takýmto chybám nedochádzalo, zdrojový materiál by sa mal vopred ošetriť. Pri materiáloch, ktoré nie sú v perfektnom stave, by sa pred použitím softvéru OCR malo zvážiť aj použitie softvéru na spracovanie obrazu, aby sa zamedzilo strate farby a bolo možné zlepšiť kontrast.
- Pri výbere softvéru OCR je nutné zistiť, či v balíku OCR je súčasťou aj jeho schopnosť (alebo neschopnosť) pracovať so slovníkmi v jazyku zdrojového materiálu.

#### **Poznámky/komentáre**

Medzi relevantné produkty na trhu patria:

- OmniPage
- TextBridge
- Adobe Capture
- Abby FineReader – má vynikajúcu funkčnosť vykonávania úprav a rozpoznávania chýb
- Transkribus (rozpoznávanie rukopisných a historických textov)

### 8.14 Metaúdaje (metadáta)

Oblasť metaúdajov je jedna z najaktívnejšie skúmaných a najdynamickejších v celej oblasti digitalizácie, rovnako aj v oblastiach, akými je informačný prieskum, vyhľadávanie na Webe, výmena údajov, integrácia podnikových aplikácií. atď.

Vybraný model metaúdajov je obzvlášť dôležitý, nakoľko ovplyvňuje výber atribútov na opísanie objektu.

### 8.14.1 METAÚDAJE POUŽITÉ PRE OPIS PREDMETU (DESKRIPTÍVNE METADÁTA)

#### **Definícia problému**

Pred výberom modelu metaúdajov pre projekt digitalizácie by sa mal skontrolovať materiál opísaný prostredníctvom príslušných metaúdajov. Umožní sa tak identifikácia existujúcich metaúdajových modelov, rovnako aj určenie všetkých nepokrytých aspektov a nedostatkov pri porovnaní existujúceho metaúdajového modelu s metaúdajmi potrebnými pre váš projekt.

#### **Pragmatické návrhy**

- Použitie vhodných metaúdajov je veľmi dôležité pre to, aby bolo možné nájsť a získať materiál z digitálnych zbierok. Ide viac o prípad, kedy je celkovým zámerom vyhľadávanie vo viacerých zbierkach uložených na rôznych miestach (logické súborné katalógy, virtuálne kombinované múzeá atď.).
- Existuje niekoľko modelov metaúdajov. Preto si musí každý projekt vybrať model metaúdajov v súlade s jeho cieľmi. Odporúča sa, aby sa nevytváral nový model, len pokiaľ požiadavkám projektu nevyhovuje žiaden z existujúcich štandardov.
- Dobrou investíciou je čas strávený modelovaním dôležitých charakteristík digitalizovaného materiálu a identifikácia jeho kľúčových atribútov. Takýto model možno porovnať s rozsahom a vlastnosťami existujúcich modelov metaúdajov.
- Mali by sa identifikovať príslušné riadené slovníky (napr. na miestopis alebo popis umelca).
- Už existuje niekoľko takýchto slovníkov, ktoré môžu vo veľkej miere zvýšiť úspešnosť vyhľadávanií.

#### **Poznámky/komentáre**

Projekt *Making of America II* (Kongresová knižnica USA) použil tri kategórie metaúdajov:

- Deskriptívne – na účely opisu a identifikácie informácií
- Štrukturálne – na účely navigácie a prezentácie
- Administratívne – na účely riadenia a spracovania

Osobitný význam majú metaúdajové modely, ktoré sa majú vybrať pre projekt digitalizácie – ide rozhodnutie o tom, ktoré množiny atribútov sa použijú na charakterizáciu diel a dokumentov/objektov, ktoré sa majú digitalizovať, a výsledných obrazov, na opis spracovania, techník a technológie a na riadenie práv.

Národná knižnica Austrálie má na toto silný model. Existencia širokého spektra existujúcich modelov a konkurenčných štandardov pre metaúdaje spôsobila, že sa mnohé projekty zameriavali čisto len na preklad z jedného štandardu do druhého.

### 8.14.2 VYHOVUJÚCE ŠTANDARDY PRE METAÚDAJE

#### **Definícia problému**

Pre metaúdaje už existujú určité dôležité štandardy. V bibliografickej doméne (a v zvyšujúcej sa miere aj v kultúrnej neknihovnickej doméne) má veľký význam štandard Dublin Core (Dublinské jadro).

### **Pragmatické návrhy**

- Pred vytvorením vlastného modelu si preštudujte existujúce štandardy a modely pre metaúdaje
- Vyhnite sa vytváraniu nového modelu metaúdajov pre kultúrne zbierky.
- Je pravdepodobné, že pre projekt digitalizácie budú relevantné modely metaúdajov, ktoré sú výsledkom práce na podobných projektoch. Medzi projektmi v oblasti kultúry sa modely metaúdajov prenášajú dobre.
- Do modelu metaúdajov by sa mali zahrnúť polia Dublin Core, pokiaľ neexistuje rozumný dôvod neurobiť tak. Aj keď múzeá môžu zistiť, že ich zbierkam lepšie vyhovuje model CIMI, mal by sa sledovať cieľ existencie spoločnej jadrovej množiny atribútov, ktorá umožní vyhľadávanie vo viacerých zbierkach súčasne.
- Ak sa má použiť súkromne vytvorený model metaúdajov, mal by sa vyvinúť spôsob mapovania prvkov tohto modelu do Dublin Core.
- Aj keď sú schémy pridelovania názvov alebo národné konvencie pomenovania užitočné, úplný model metaúdajov je lepší, aj ohľadom množstva údajov o dokumente/predmete a na umožnenie výkonnejšieho vyhľadávania a spolupráce s inými projektmi a inými krajinami.

### **Poznámky/komentáre**

Existuje ohromný počet existujúcich štandardov, ktoré pokrývajú rôzne aspekty metaúdajov. Medzi jednotlivými štandardmi je však veľké prekrývanie a veľmi veľké množstvo modelov špecifických pre inštitúcie, pri tvorbe ktorých sa zanedbali rezortné alebo medzidoménové modely.

## **8.15 Zverejnenie**

V tejto fáze projektu sa vytvoril a záložoval/uložil originálny digitálny materiál. Bol identifikovaný vyhovujúci metaúdajový model a vytvorili sa metaúdaje súvisiace s každým artiklom.

Príprava na zverejnenie zahŕňa spracovanie novo vzniknutého materiálu ešte pred zverejnením. Zverejnenie typicky znamená vystavenie na internete a spracovanie znamená typicky redukciu kvality obrazu/zvuku/videozáznamu, kvôli tomu, aby sa prispôbili prevádzkovým podmienkam na internete.

## **8.16 Spracovanie obrazu**

### **Definícia problému**

Súbory TIFF vytvorené počas procesu digitalizácie sú typicky veľmi veľké (až niekoľko desiatok megabajtov). Takéto súbory nie sú vhodné na zverejnenie na internete kvôli dlhému času, ktorý je potrebný na prevzatie obrazu koncovým používateľom.

### **Pragmatické návrhy**

- Vytvorte doručiteľné verzie matričného materiálu. Musí existovať minimálne jedna doručiteľná verzia. Môže byť užitočné vytvoriť aj druhú

verziu 'miniatúru, náhľad', čo závisí od architektúry webového sídla, na ktorom sa má materiál verejne sprístupniť.

- Doručiteľné verzie sa vytvárajú otvorením matričného súboru TIFF v softvéri na spracovanie obrazu a jeho exportom vo formáte JPEG, PNG.
- Farebné rozlíšenie sa môže zredukovať na 256 farieb. Ak pri takomto rozlíšení dochádza k značnej strate kvality môže sa použiť vyššie rozlíšenie pre farbu.
- Výber správneho farebného rozlíšenia zvyčajne vyžaduje subjektívne rozhodnutie.
- Obraz vytvorený pri rozlíšení 72 DPI približne ukáže jeho pôvodnú veľkosť na mnohých počítačových monitoroch. Preto je rozlíšenie 72DPI rozumnou voľbou pre mnohé obrazy, ktoré sa majú prezerať cez obrazovku. Pre nižšie rozlíšenia bude potrebné subjektívne rozhodnutie ohľadom 'prijateľnej kvality'.
- Výber formátu súborov, farebného rozlíšenia a bodového rozlíšenia zahŕňa rozhodnutie o tom, čo je 'prijateľná' kvalita. Medzi kvalitou a veľkosťou súborov je potrebné nájsť rovnováhu.
- Vo všeobecnosti by všetky obrazové súbory na webstránke nemali presahovať veľkosť 100 kilobajtov. Môžu sa publikovať aj väčšie súbory; mali by však byť prístupné cez linku z webovskej stránky s upozornením, že ich získanie môže trvať dlhšie.
- Pokiaľ materiál nie je určený na spúšťanie cez sieť (*streaming*), video a audio materiál má typicky väčšie súbory, pričom súbor sa preberie pred jeho prezretím/vypočutím v režime offline.
- Čas na preberanie môže byť ovplyvnený zmenou počtu obrázkov (okienok) na sekundu pri video materiále, vzorkovacou rýchlosťou pri audio materiále atď.

### **Poznámky/komentáre**

Rozhodnutie ohľadom spracovania obrazu závisia vo veľkej miere na osobnom zvážení. Tu uvedené predpisy môžu byť pre niekoho príliš striktné, pre iného voľné, čo sa odvíja od projektu a koncových používateľov.

Softvér na spracovanie obrazu ako Paint alebo Paintshop je voľne dostupný na Internete. Ďalšie softvérové nástroje na spracovanie obrazu môžu ušetriť čas a úsilie a podľa toho je potrebné zvážiť výdavky na softvér.

Voľne dostupný je aj softvér na úpravu zvuku a videa. Audio a video softvér sa spravidla dodáva na spracovanie vytvorených údajov spolu s príslušným hardvérom.

## **8.17 Trojrozmerné materiály a virtuálna realita**

### **Definícia problému**

Predpisy v predchádzajúcej časti o zverejnení obrazov nie sú priamo aplikovateľné na digitálne prevedenie materiálu virtuálnej reality a v trojrozmerné materiály. V súvislosti s prezentáciou na internete je dôležitá rovnováha medzi kvalitou a veľkosťou súboru.

### **Pragmatické návrhy**



- Prehliadače trojrozmerných (3D) a VR materiálov nie sú ešte všeobecne rozšírené spolu so softvérom operačného systému. Ide o protiklad rozšírenia softvéru pre obraz, audio a video, ktorý je bežne zahrnutý v softvéri Windows.
- Uistite sa, že sú k dispozícii prehliadače 3D a VR materiálu. Sprístupnite prehliadač z toho istého sieťového miesta, na ktorom je materiál. Tým sa predíde situácii, že softvér nebude dostupný ani z iných zdrojov.
- Pred rozhodnutím sa pre nejaký prehliadač ohodnotte viaceré. Kompatibilita formátov súborov a prehliadačov nie je štandardizovaná do takej miery, ako pri statických obrazoch.
- Moderné PC s úpravou pre hry majú hardvérové akcelerátory a vyššiu grafickú pamäť, čo má významný vplyv na zážitok z pobytu vo virtuálnej realite.

### **Poznámky/komentáre**

VRML prehliadač, ktorý sa úspešne používa v jednom z referenčných projektoch (írsky projekt ACTIVATE) je prehliadač Blaxxun Contact.

## **8.18 On-line zverejnenie**

### **Definícia problému**

Mnohé projekty digitalizácie v oblasti kultúry vedú k vytvoreniu on-line kultúrnych zdrojov, zvyčajne ide o web sídlo s obrazmi, metaúdajmi, trojrozmernými artefaktmi atď. Môže ísť od sídel s jednoduchým obsahom až po zložité softvérom ovládané portály a nástroje na prezeranie. V tejto príručke je len niekoľko odporúčaní ohľadom tvorby web sídel, nakoľko vyžaduje rozsiahle vedomosti. Na stránke projektu Minerva je možné nájsť odkazy na príklady najlepších postupov nominovaných partnermi projektu Minerva.

### **Pragmatické návrhy**

- Vo webových sídlach by mala byť ľahká navigácia – odšadiaľ by mala viesť linka na úvodnú stránku alebo obsah.
- Náležitá pozornosť sa musí venovať univerzálnemu prístupu a využívaniu web stránok slabozrakými a inými hendikepovanými osobami.
- Webové stránky by mali byť dostatočne krátke, aby ich používateľ nemusel veľa rolovať.
- Obrázky by mali byť dostatočne malé na to, aby nenarušovali zážitky pri prezeraní.
- Väčšie obrázky by sa mali byť prepojené z webových stránok s poznámkou, že obraz je veľký a preberanie môže byť pomalé.
- K používaniu animácií, vysúvacích prvkov (*pop-ups*), *flashových* programov a podobných technológií by sa malo pristupovať s obozretnosťou. Musí byť možné obísť dlhé úvodné animované sekvencie.
- Webové sídla by mali byť v ideálnom prípade viacjazyčné, pričom by sa mal podporovať domáci jazyk a jeden alebo dva iné jazyky (spravidla anglický jazyk, ktorý je *de facto* jazykový štandard on-line zdrojov).
- Prepojenia na externé zdroje by sa mali overovať denne, aby sa minimalizoval výskyt mŕtvych liniek a s nimi súvisiace mrzutosti.

## Poznámky/komentáre

Samotný proces sprístupnenia materiálu na Webe je najširšie chápaný a zdokumentovaný. Táto príručka neposkytuje inštrukcie o tom, ako vytvárať webové sídla, programovať v HTML, budovať databázy prístupné cez Web a vykonávať ďalšie úlohy, ktoré sú potrebné na tvorbu a údržbu web prezentácií. Očakáva sa. Že mnohé z kultúrnych inštitúcií, ktoré využívajú tieto odporúčania, budú mať už k dispozícii nejaký web server, ktorý využijú pri ich projekte digitalizácie.

### 8.19 Práva duševného vlastníctva a copyright

#### 8.20 Právne aspekty

Právnu situáciu je potrebné overiť pri každom objekte alebo skupine (kategórii) objektov a práva na používanie digitálnych dokumentov by sa mali získať čo najskôr, podľa možnosti už vo fáze plánovania.

##### 8.20.1 AUTORSKÉ PRÁVA (COPYRIGHTS)

Podľa zákona o autorských právach napríklad na Slovensku sú diela chránené až 70 rokov po úmrtí autora, po ktorom sú voľne dostupné ako takzvané "verejné domény" (§ 32 *Trvanie majetkových práv*) "(1) Majetkové práva trvajú od okamihu vytvorenia diela počas autorovho života a 70 rokov po jeho smrti. Pri spoluautorskom diele majetkové práva trvajú počas života posledného zo spoluautorov a 70 rokov po jeho smrti. Ak je audiovizuálne dielo vytvorené ako spoluautorské dielo, majetkové práva trvajú počas života poslednej osoby spomedzi režiséra, autora scenára, autora dialógov a autora hudby, ktorá bola vytvorená osobitne pre toto dielo a 70 rokov po jeho smrti."

<http://www.zakonypreludi.sk/zz/2015-185#cast2>

Napríklad práca zverejnená v roku 1905, ktorého autor zomrel v roku 1955, je preto pod autorským právom do roku 2025.

##### 8.20.2 VYDAVATELSKÉ A LICENČNÉ ZMLUVY

Je potrebné overiť, či *vydavateľ* diela vlastní práva na jeho (dodatočné) zverejnenie na *internet*. (1990, t. j. koniec systému ARPANET a prvé nasadenie WWW v Cerne vo Švajčiarsku).

Znamená to, že do roku 1990 neexistovali vydavateľské zmluvy, v ktorých by si vydavatelia dohodli s autormi práva zverejnenia na internete. Podľa našich skúseností sa na Slovensku podmienky zverejňovania dohodnutého diela na internete v autorských zmluvách upravovali a upravujú len zriedka, ale prax v tomto ohľade sa postupne mení. V prípade starších publikácií (kategória staré a vzácne tlače a seriály) to možno vylúčiť hneď od začiatku, pretože internet neexistoval v čase dohody medzi autorom a vydavateľom.

##### 8.20.3 COPYRIGHT V KNIŽNICI PODĽA AMERICKÉHO A AUSTRÁLSKEHO PRÁVA

<https://www.nla.gov.au/key-exceptions-in-the-copyright-act>

#### 8.20.4 PRAKTICKÝ PŘÍKLAD

Okrem niekoľkých výnimiek sa staré a vzácne tlače, ktoré boli publikované pred rokom 1900 a na ktoré sa už nevzťahujú žiadne obmedzenia týkajúce sa autorských práv, *môžu voľne digitalizovať a sprístupňovať* organizáciami, ktoré ich vlastnia alebo spravujú vo svojich zbierkach.

Praktický príklad – seriálové publikácie, noviny, časopisy, periodické zborníky  
Predpokladom na zaradenie týchto dokumentov na digitalizáciu je súhlas držiteľov autorských práv. Výnimku tvoria časopisy až do 19. storočia, kedy už nie je možné zistiť držiteľov autorských práv.

Skúsenosti ukázali, že iba vydavatelia novších obchodných časopisov uzatvárajú zmluvy s ich autormi. Pritom *autori prenášajú plné práva na používanie daného článku* na vydavateľa (dohoda o autorských právach, licencia). Pokiaľ ide o historický obsah, takéto jasné dohody zvyčajne neexistujú a získanie spätného povolenia od autorov a ich potomkov prostredníctvom správnych právnych kanálov by znamenalo neúmerné úsilie.

#### 8.20.5 DISPENZÁCIA S FORMÁLNYM POVOLENÍM

Pre digitalizáciu novín a časopisov je možné prijať pragmatický postup (ako napr. vo Švajčiarsku): autorský súhlas nie je získaný, ale autor môže následne požadovať zablokovanie alebo vymazanie jeho práce z online služby.

Tento prístup predpokladá, že držiteľia autorských práv sú akademicky aktívni ľudia alebo autori na čiastočný úväzok, ktorí podporujú ďalšie šírenie svojej práce v záujme slobodnej vedy.

Pri on-line zverejnení akéhokoľvek materiálu musia byť vysporiadané práva duševného vlastníctva (IPR) súvisiace s materiálom.

Materiál, ktorý je vo verejnej doméne (konkrétne napríklad staré knihy alebo noviny, starý materiál umiestnený explicitne vo verejnej doméne) predstavuje relatívne nízku náročnosť. Mnohé kultúrne inštitúcie však z využívania obrazov artefaktov alebo obrazov z ich zbierok odvodzujú príjmy, preto obhajujú copyright.

Materiál, ktorého copyright držia tretie strany, sa môže zverejniť len so súhlasom takýchto tretích strán.

Našťastie existuje niekoľko technických možností na ochranu copyrightu materiálu umiestneného na Internet. Tu je ich prehľad.

#### 8.20.6 VYTVORENIE COPYRIGHTU

##### **Definícia problému**

Prvým krokom pri skúmaní situácie súvisiacej s copyrightom pre nejaký kultúrny dokument/predmet je stanovenie vlastníctva daného copyrightu.

##### **Pragmatické návrhy**

- Stanovte právnu situáciu s ohľadom na copyright a publikovanie v krajine, kde sa projekt v realizuje. Každá krajina má svoje zákony o copyrighte zvyčajne aspoň od 19. storočia. Takéto zákony sa zvyčajne vzťahujú na všetky formy publikácií, vrátane ich on-line zverejnenia.

Môžu, ale nemusia pokrývať zákon o digitalizácii, ktorá sa môže chápať ako archivačný proces, alebo ako kopírovanie (rozmnožovanie).

- K verejnému sprístupneniu v režime on-line by v žiadnom prípade nemalo dôjsť pred zistením copyrightu.
- Niektoré dokumenty/predmety, napr. staré noviny, majú jasné pravidlá copyrightu, ktorými sa riadia. Typicky je umožnené voľné kopírovanie až vtedy, keď noviny dosiahnu určitý vek. Dokumenty/predmety, ktoré spadajú do tejto kategórie, sa môžu voľne digitalizovať a zverejniť.
- Pre dokumenty/predmety, ktorých copyright prináleží inštitúcii realizujúcej projekt, bude požadované udeľovať interné povolenie na digitalizáciu a on-line zverejnenie.
- Pre dokumenty/predmety, ktorých copyright drží tretia strana – napríklad prenajímateľ alebo darca zbierky historických dokumentov/predmetov, sa musí získať písomné povolenie tejto strany. K zverejneniu môže dôjsť až po prijatí takéhoto rozhodnutia.
- Zaručenie povolenia digitalizovať a zverejniť môže zahŕňať aj potrebu platiť. Medzi výškou platby a hodnotou vyplývajúcej zo zahrnutia relevantných dokumentov/predmetov do on-line zdroja musí byť rovnováha

### **Poznámky/komentáre**

Situácia týkajúca sa copyrightu sa mení z krajiny na krajinu.

#### 8.20.7 OCHRANA COPYRIGHTU

### **Definícia problému**

Zverejnenie dokumentov/objektov na webe je vydanie ich na voľné kopírovanie. Nie je možné zamedziť kopírovaniu materiálu vystaveného na webe. Je však možné zvážiť niekoľko postupov, z ktorých má každý nejaký vplyv na ochranu copyrightu.

### **Pragmatické návrhy**

- Stanovte, či sa copyright musí ochraňovať, alebo nie.
- S držiteľmi copyrightu sa dohodnite na postupoch, ktoré sa majú použiť na ochranu copyrightu.
- Do úvahy prichádzajú nasledujúce postupy:
  - Pridanie viditeľného vodoznaku alebo známky o copyrighte na každý obraz.
  - Pridanie neviditeľného digitálneho vodoznaku na každý obraz. Takéto značky sa môžu použiť na dokázanie vlastníctva 'ukradnutého' obrazu, takisto aj na vystopovanie využívania obrazu na Internete.
  - Enkrypcia (šifrovanie) obrazov, pričom správny kľúč sa vydá len registrovaným používateľom. Tým sa samozrejme zníži hodnota on-line obrazu z pohľadu zvyšnej verejnosti.
  - Obmedzenie zverejnenia len na obrázky s nízkym rozlíšením, napr. 75 DPI na prezeranie na obrazovke. Tým sa zníži úroveň použitia obrazu v iných doménach, napr. tlač, odevy atď.
  - Obmedzenie zverejnenia len na malé časti obrazu.

- Zobrazovanie obrazov len registrovaným, oprávneným členom nejakej komunity.
- Preskúšajte výsledky procesu na ochranu copyrightu pomocou prvých niekoľkých dokumentov/predmetov, aby ste sa uistili, že proces nemá žiadne neočakávané alebo neželané následky.

### **Poznámky/komentáre**

Prístup, ktorý je najvhodnejší pre akýkoľvek projekt, bude závisieť vo veľkej miere od cieľov projektu a kultúrnej inštitúcie, rovnako aj od povahy materiálu. Najčastejším riešením galérií a múzeí. Relatívna jedinečnosť kultúrnych zbierok poskytuje v mnohých situáciách dôkaz o vlastníctve.

## **8.21 Manažment projektov digitalizácie**

Úspech každého projektu, vrátane projektov digitalizácie, je ovplyvnený vo veľkej miere riadením projektu. Táto sekcia poskytuje niekoľko odporúčaní špecifických pre manažment digitalizačných projektov.

### *8.21.1 RIADENIE PROCESU DIGITALIZÁCIE*

#### **Definícia problému**

V typickom projekte digitalizácie sú desiatky, stovky a dokonca aj tisíce dokumentov/objektov určených na digitalizáciu. Aby sa dosiahla efektívnosť projektu, je dôležité ustanoviť pracovný postup, ktorý maximalizuje výkonnosť digitalizačného tímu. Okrem toho bude mať veľký význam informačný zdroj, ktorým je báza znalostí projektu digitalizácie.

#### **Pragmatické návrhy**

- Stanovte a zdokumentujte každý krok, ktorým musí dokument/predmet prejsť v priebehu digitalizačného procesu. Týka sa to:
  - Získavania zo skladu / obvyklého umiestnenia
  - Čistenia a prípravy
  - Skenovania a fotografovania
  - Vrátenia do obvyklého miesta
  - Pridelovania názvov súborov
  - Ukladania súborov
  - Vytvárania verzií z veľkých matričných súborov na dodanie on-line
  - Zálohovania serverov / úložných médií
- Vytvorte bázu poznatkov projektu digitalizácie, ktorá sa môže použiť na vystopovanie objektu počas digitalizačného procesu a umožňuje kedykoľvek dohliadať na stav projektu. Táto báza poznatkov môže mať formu databázy (napr. v MS Access, Oracle, MySQL a pod.), môže byť aj tabuľkovým rozvrhom (*spreadsheet*) alebo zbierkou dokumentov. Dôležitým nie je formát bázy poznatkov, ale proces, ktorým sa zabezpečuje zaznamenávanie vykonávaných akcií.
- Do bázy poznatkov projektu digitalizácie by sa mali zadať názov, identifikátor a ďalšie relevantné informácie ku každému dokumentu/predmetu hneď po tom, ako sa vyberie na digitalizáciu.

Priebežne sa musí zaznamenávať aj status dokumentu/predmetu (jeho posledný ukončený krok).

- Bude potrebné prijať rozhodnutie ohľadom procedurálnych možností, či sa majú dokumenty/predmety zhromažďovať na digitalizačnom pracovisku na začiatku dňa, každého týždňa, alebo po jednom kuse.
- Dokumenty/predmety, ktoré vyžadujú podobné činnosti alebo nastavenia hardvéru by sa mali digitalizovať spolu. Ušetrí sa tak čas potrebný na nastavenie digitálnej kamery/fotoaparátu. Parametre nastavenia hardvéru by sa mali zdokumentovať, aby sa digitalizácia mohla opakovať v prípade straty súborov a pod.
- Miesto, telefónne číslo a pohotovostný personál dodávateľa kľúčovej služby (napríklad IT podpora) by sa mali zaznamenať na začiatku projektu a byť stále k dispozícii.

### **Poznámky/komentáre**

Čím je väčší projekt, tým je užitočnejšie ustanoviť postupnosť a pracovný tok. Efektívnosť, ktorú to prinesie, vykompenzuje čas strávený ich prípravou. Na stránke Minerva sú odkazy na projekty, ktoré sa zameriavajú čisto týmto aspektom digitalizácie.

## *8.21.2 VYTVORENIE TÍMU*

### **Definícia problému**

Projekty digitalizácie často umožňujú zamestnancom kultúrnych inštitúcií prvý kontakt s novými technológiami. Medzi takéto technológie patrí digitalizačný hardvér, zverejnenie na webe, spracovanie obrazu, pripisovanie metaúdajov, budovanie databázy atď.

### **Pragmatické návrhy**

- Podľa možností by mala v projektovom tíme byť aspoň jedna osoba s potrebnými zručnosťami práce s informačnými technológiami.
- Zhodnoťte stav poznatkov personálu, ktorý má pracovať na projekte, a identifikujte IT zručnosti, ktoré budú potrebovať. Identifikujte požiadavky na školenie a splňte ich pred začatím projektu.
- Nebudú potrebné len IT zručnosti. Môžu byť potrební aj špecialisti, ako je uvedené vyššie, na prácu s jemnými dokumentmi, artefaktmi a pod. Príslušné pokyny sa dajú získať od jednotlivcov, ktorí sú zodpovední za zdrojové materiály.

### **Poznámky/komentáre**

Je lepšie, aby na projekte pacovalo malé jadro zručných zamestnancov ako väčšia skupina príležitostných pracovníkov. Pre pracovníkov môže byť pri budovaní ich špecifických zručností lepšie, ak budú vystavení celému cyklu digitalizácie. Digitalizácia a pripisovanie metaúdajov nie je pre pracovníkov vďačná činnosť – zoznámenie sa s ďalšími prvkami projektu zvýši ich spokojnosť.

## **8.22 Školenie zamestnancov**

### **Definícia problému**

Pokiaľ ľudia pracujúci na projekte digitalizácie nemajú dostatočné skúsenosti z predchádzajúcich projektov, bude potrebné školenie zamestnancov. To zahŕňa dve rôzne oblasti – použité technológie a spracovanie zdrojového materiálu.

### **Pragmatické návrhy**

- Nedá sa predpokladať, že zamestnanci archívov, múzeí knižníc budú mať automaticky všetky relevantné odborné znalosti.
- Požiadavky na školenie sa musia identifikovať na začiatku projektu, čiže už vo fáze plánovania.
- Požiadavky na školenie by sa mali zahrnúť v báze poznatkov projektu digitalizácie a treba konať podľa nich pred tým, ako bude projekt vyžadovať školenie.
- Niektoré školenia, napríklad tie, čo sú zamerané na používanie digitálnych technológií, sa dajú realizovať “pri práci”; iné školenia – napríklad o zaobchádzaní so zdrojovým materiálom sa musia konať dopredu.
- Namiesto väčšej príležitostne vytvorenej skupiny s častou fluktuáciou členov sa uprednostňuje menšie jadro zamestnancov, ktorí sa počas celého projektu vzdelávajú a rozvíjajú svoje praktické skúsenosti.
- Technologické školenie sa môže pridružiť ku inému projektu v tej istej inštitúcii; školenie môže prípadne poskytnúť aj externá agentúra zaoberajúca sa digitalizáciou.
- Školenie kurátorov poskytujú najlepšie osoby zodpovedné za starostlivosť o originálny materiál.

### **Poznámky/komentáre**

Nedostatočne vyškolený personál môže spôsobiť nešťastné a nenapraviteľné nehody alebo škody už na začiatku projektu, to isté sa môže prihodiť, ak z projektu nejaký personál vystúpi a začne na ňom pracovať nový. Najžiadanejším aspektom takýchto projektov je malé, dobre vyškolené jadro.

Čas investovaný do školení na začiatku projektu sa vyplatí vo forme zvýšenej produktivity a menšieho počtu problémov počas celého projektu.

## **8.22.1 ZÍSKAVANIE TECHNICKÉHO PORADENSTVA OD TRETÍCH STRÁN**

### **Definícia problému**

Často je vhodné, aby projekt digitalizácie využíval služby jednej alebo viacerých tretích strán. Medzi najčastejšie poskytované služby patrí samotná digitalizácia, riadenie projektu, integrácia s inými systémami, vývoj softvéru a pod. To umožňuje, aby sa kultúrna inštitúcia zameriavala na svoju vlastnú odbornú činnosť bez potreby školenia a držania personálu s pokročilými znalosťami v oblasti informačných technológií atď.

### **Pragmatické návrhy**

- Vzťahy s technickými partnermi sa budú riadiť jasnými a striktnými zmluvami. Pred vykonaním akejkoľvek práce sa musí dohodnúť, zdokumentovať a podpísať špecifikácia produktov alebo služieb, ktoré sa majú poskytnúť.

- Je potrebné pravidelne vykonávať kontrolu plnenia vykonávanej práce, aby sa zaručilo, že výsledkom je skutočne to, čo projekt požaduje.
- Aj keď používanie tretích strán môže byť pohodlné, je potrebné mať na mysli, že odborné poznatky a praktické skúsenosti, ktoré sa mohli získať pri vykonávaní danej práce, budú pre kultúrnu inštitúciu na konci projektu navždy stratené. To isté platí aj v prípade dočasných zamestnancov pracujúcich len do ukončenia projektu. Lepšie je vyhradiť na projekt dlhodobého zamestnanca, ako ho nahrádzať iným na základe dohody o krátkodobej práci.

### **Poznámky/komentáre**

Určité veľké projekty, akým je francúzsky národný program digitalizácie, identifikovali uprednostneného dodávateľa, vzťah s ktorým sa preniesol aj na iné projekty a na niekoľko rokov. Po ustanovení pracovného vzťahu s dodávateľom bude možno potrebné zväziť hodnotu zmeny dodávateľa medzi projektmi.

## 8.22.2 SPOLUPRÁCA S PROJEKTOVÝMI PARTNERMI A VYTVORENIE SPOLOČNÉHO OBSAHU

### **Definícia problému**

Mnohé projekty digitalizácie sú kooperatívnou snahou a sú do nich zapojené viaceré kultúrne organizácie. Projekty a rámcové projekty financované EÚ majú takmer vždy mnohých partnerov z viacerých krajín. Odporúčania týkajúce sa spustenia a riadenia projektov s viacerými partnermi sú mimo rámec tohto dokumentu. Je tu však uvedených niekoľko usmernení.

### **Pragmatické návrhy**

- Ustanovte roly a zodpovednosť v rámci projektu a pravidelne ich aktualizovať.
- Zriadte spôsob komunikácie medzi partnermi, a zabezpečte príjem špecializovaných informácií. Na tento účel je ideálna elektronická pošta, pokiaľ ju partneri čítajú a odpovedajú na ňu.
- Dodávatelia sa musia riadiť striktnými obchodnými zmluvami s jasne a jednoznačne stanoveným spôsobom plnenia.
- Jasná dokumentácia práv duševného vlastníctva všetkých partnerov písomne potvrdených všetkými partnermi: pred začatím projektu by sa mala zostaviť partnerská dohoda, ktorá jasne stanovuje práva duševného vlastníctva vzťahujúce sa na materiál, ktorý vstupuje do projektu a materiál, ktorý sa vytvára v rámci projektu.
- Každý partner by mal mať v projekte jasnú rolu - ak rola partnera nie je jasná, skontrolujte, či je daný partner z pohľadu projektu potrebný.

### **Poznámky/komentáre**

Vyššie uvedené poznámky sú len malou časťou materiálu, ktorý je možné poskytnúť o spustení a riadení projektov s viacerými partnermi. Partneri a dodávatelia sú najčastejšou príčinou omeškaní a zmätku v projekte – jasné zmluvy a spoločná dohoda o úlohách a zodpovednosti všetkých partnerov môže pomôcť vyhnúť sa takýmto problémom.

## 8.23 Náklady

### **Definícia problému**



Budovanie digitálnej zbierky je za bežných okolností veľmi nákladné. Projekty musia brať do úvahy všetky počiatočné náklady a náklady na infraštruktúru a prevádzku projektu. To predstavuje náklady na úvodné plánovanie, pridanie opisných znakov, dokumentačné systémy, školenie personálu a tak ďalej, rovnako aj postupné náklady na digitalizáciu zdrojového materiálu.

### **Pragmatické návrhy**

Projekty digitalizácie by mali zväžiť nasledovné náklady spojené s návrhnutím, implementáciou a údržbou digitálnej zbierky:

- **Ľudské zdroje a ich rozvoj** - mzdy pre manažéra projektu, web programátora, technickú podporu, úradníkov pre oblasť vzdelávania, cestovné náklady a školenia
- **Zariadenia** – Náklady na zariadenia súvisia s otázkami požadovaných funkčností a potreby kompromisov. Buď sú dôležité náklady, alebo najvyšší možný štandard kvality snímania obrazu.
- **Prevádzkové náklady**
  - Čas manipuláciu so zdrojovým materiálom (od vybratia z regálu k snímaciemu zariadeniu a späť) ako percentuálny podiel celkových denných mzdových nákladov.
  - Príprava zdrojového materiálu (konzervácia, čistenie atď.).
  - Čas potrebný na snímanie (od prípravy až po pridelenie názvu a uloženie) vyjadrený ako percentuálny podiel denných mzdových nákladov na operátora.
  - Katalogizácia a práca s metaúdajmi ako percentuálny podiel celkových mzdových nákladov.
  - Náklady na hardvér a softvér na jeden digitalizovaný dokument/predmet (uprednostňuje sa skôr výpočet na základe amortizácie alebo nákladov za výmenu ako na základe nákladov na získanie)
  - Čas na zabezpečenie kvality ako percentuálny podiel mzdových nákladov.
  - Údržba hardvéru a softvéru.
  - Čas technickej podpory súvisiacej so snímaním.
  - Čas projektového manažmentu súvisiaci so snímaním.
  - Školenie súvisiace so snímaním.
  - Snímanie obrazu je najmenej nákladná časť projektu. V priemere je jedna tretina celkových nákladov spojená s digitalizáciou, menej ako tretina ide na tvorbu metaúdajov a o trochu viac ako tretina súvisí s administratívou a zabezpečením kvality. Zvyšok ide na dlhodobé poplatky za údržbu.
- **Náklady na ukladanie** – náklady na údržbu, počítajú sa ako náklady na jeden gigabajt.

### **Poznámky/komentáre**

Dlhodobá trvalá udržateľnosť sa často dostáva na zozname priorít na nižšie miesta kvôli bezprostrednejším a naliehavejším potrebám. Bez ohľadu na kvalitu a robustnosť, digitálne zdroje vytvorené projektom digitalizácie nevydržia dlho, ak projekt nezoberie do úvahy náklady na ich údržbu.

8.23.1 ORIENTAČNÉ CENY DIGITALIZÁCIE VIAZANÝCH KNÍH<sup>5</sup>

8.23.2 ĎALEJ UVEDENÉ INFORMÁCIE ILUSTRUJÚ FINANČNÚ NÁROČNOSŤ DIGITALIZÁCIE TLAČENÝCH VIAZANÝCH DOKUMENTOV NA ZÁKLADE INFORMÁCIÍ O SLUŽBÁCH KOMERČNEJ DIGITALIZÁCIE S POUŽITÍM ŠPIČKOVEJ TECHNOLOGIE.

Tabuľka 2 Orientačné ceny digitalizácie viazaných kníh

	Farba	Odtiene šedi	čiernobiela
Cena za prvých 100 strán	44,92 €	26,26 €	26,26 €
Každá ďalšia strana nad 100 strán	0,32 €	0,13 €	0,13 €
Poštovné	V cene	V cene	V cene

<b>Stanovenie ceny spolu</b>	
<b>Príplatok za rýchlosť EXPRES<sup>6</sup> (10%)</b>	0,00 €
<b>Podiel DPH 19% (DPH je v cene)</b>	0,00 €

Služba zahŕňa okrem digitalizácie dodávku skenovaného titulu zákazníkovi na DVD alebo on-line (FTP) v nekomprimovanom formáte TIFF a prezentačnom formáte JPG. Výstup je PDF s textom OCR na pozadí. Fulltextové vyhľadávanie je zabezpečené. Štandardná kvalita skenovania 300 dpi a viac, podľa dohody.

Na základe dostupných informácií sme dospeli k odhadu nákladov na masovú digitalizáciu všetkých slovacikálnych tlačených viazaných kníh, novín a časopisov v projekte Slovenská digitálna knižnica. Nezaoberali sme sa prepočtom nákladov na digitalizáciu archívnych materiálov, muzeálnych predmetov a audiovizuálnych diel. V prepočtoch uvažujeme o cenách podľa štúdie, ktorú pre Francúzsku národnú knižnicu<sup>7</sup>.

Podrobnejšie informácie o nákladoch na digitalizáciu obsahujú príručka európskeho projektu Minerva<sup>8</sup>

8.23.3 ODHAD NÁKLADOV NA DIGITALIZÁCIU SLOVACIKÁLNEHO FONDU KNÍH V SLOVENSKEJ DIGITÁLNEJ KNIŽNICI (SDK)

8.23.4 NÁKLADY NA DIGITALIZÁCIU TLAČENÝCH VIAZANÝCH DOKUMENTOV V SDK

Počet titulov kníh	Počet titulov seriálov	Spolu	Farba	Odtiene šedi alebo čiernobiele
500000	10000	510000 titulov	0,13 €/strana	0,10 €/strana

175000000 strán	25000000 strán	<b>200000000 strán</b>	<b>26000000 €</b>	20000000 €
		300000000 strán	39000000 €	

<sup>5</sup> Spracované podľa prieskumu komerčnej digitalizácie na požiadanie (business on demand)

<sup>6</sup> Normálna rýchlosť je 8 dní. Expresné skenovanie EXPRES do 3 dní.

<sup>7</sup> *Etude pour la mise en place d'une organisation de numérisation de masse de la Bibliothèque nationale de France*. 2006. Rapport de synthèse. V 1.0. Confidentiel BnF et IBM. Paris : BnF & IBM, 17. juillet 2006. 86 s. (Dostupné u autora. Použité s láskavým súhlasom riaditeľstva BnF).

<sup>8</sup> TANNER, Simon. 2006. *Handbook on Cost Reduction in Digitisation*. Working Draft – Version 1 - June 2006. King's College London. London : MINERVA Plus Project, 2006. 33. s. Dostupné na: [http://www.minervaeurope.org/publications/CostReductioninDigitisation\\_v1\\_0610.pdf](http://www.minervaeurope.org/publications/CostReductioninDigitisation_v1_0610.pdf)

		(15 % farba)	(15 % farba)	
--	--	--------------	--------------	--

Náklady na skenovanie ca 200 mil strán slovacikálnej literatúry odhadujeme na **23900000 €**

Pri tejto kalkulácii sa predpokladá, že 15 % strán z celkového objemu strán sa mal digitalizovať farebne v najvyššej kvalite.

Z prostriedkov OPIS bolo možné financovať aktivity na západnom, strednom a východnom Slovensku, pričom sa sleduje vyrovnávanie regionálnych disparít a zvýšenie zamestnanosti a inovatívnosti v regiónoch mimo hlavného mesta.

## 8.24 Prostredie

### Definícia problému

Mnohé vzácne alebo chýlostivé materiály vyžadujú špeciálne prostredie. Je kriticky dôležité pre akýkoľvek projekt, aby proces digitalizácie mal čo najmenší negatívny dopad na zdrojové materiály. Vyhovujúce prostredie pre digitalizáciu je dôležité pre mnohé projekty digitalizácie.

### Pragmatické návrhy

- Prostredie, v ktorom prebieha digitalizácia, je veľmi dôležité.
- Mali by sa získať poradenstvo od expertov, aby sa čo najlepšie zabezpečili všetky aspekty práce s originálnym materiálom. To sa týka aj prostredia, v ktorom sa digitalizuje.
- Priestor na digitalizáciu by mal byť vyhradený projektu digitalizácie počas celého projektu. Nadmerný pohyb a premiestňovanie pracovného priestoru môže viesť k poškodeniu, strate, alebo k iným negatívnym vplyvom na zdrojové materiály, takisto aj k strate času projektom.
- Ak majú zdrojové materiály zvláštne požiadavky týkajúce sa svetla, vlhkosti a podobne, tieto podmienky by sa mali čo najvernejšie vytvoriť v prostredí na digitalizáciu. Pri niektorých materiáloch, napríklad kožených dokumentoch môže krátkodobé zvýšenie vlhkosti napomôcť uvoľneniu materiálov pred vyrovnaním na účely odfotografovania alebo skenovania.
- Takmer vo všetkých prípadoch sa neodporúča priame vystavenie silnému svetlu (napr. slnečnému žiareniu) na dlhší čas. Fajčenie a konzumácia jedla a nápojov v blízkosti dokumentov/predmetov nie sú povolené – nevchádzajte do pracovného priestoru s kávou!

### Poznámky/komentár

Špecializované digitalizačné pracovisko nemusí byť v závislosti od veľkosti a projektu realizovateľné. Napriek tomu je potrebné mať na pamäti tu vytýčené ciele, aby sa minimalizoval pohyb, narušenie a manipulácia s materiálmi.

V prípade vzácných kultúrnych materiálov by diskusia s tými, ktorí sú zodpovední za starostlivosť o materiály, nemala byť nahradená žiadnymi referenciami.



### 9.1 Digitální knihovna a digitální archiv (DIKDA). Slovensko

V duchu Lundských zásad bol pripravený napríklad jeden z najväčších projektov digitalizácie v Európe. Ide o slovenský projekt Digitálna knižnica a digitálny archív (DIKDA)<sup>9</sup>.

V projekte s celkovým rozpočtom 49 miliónov eur boli tieto aktivity:

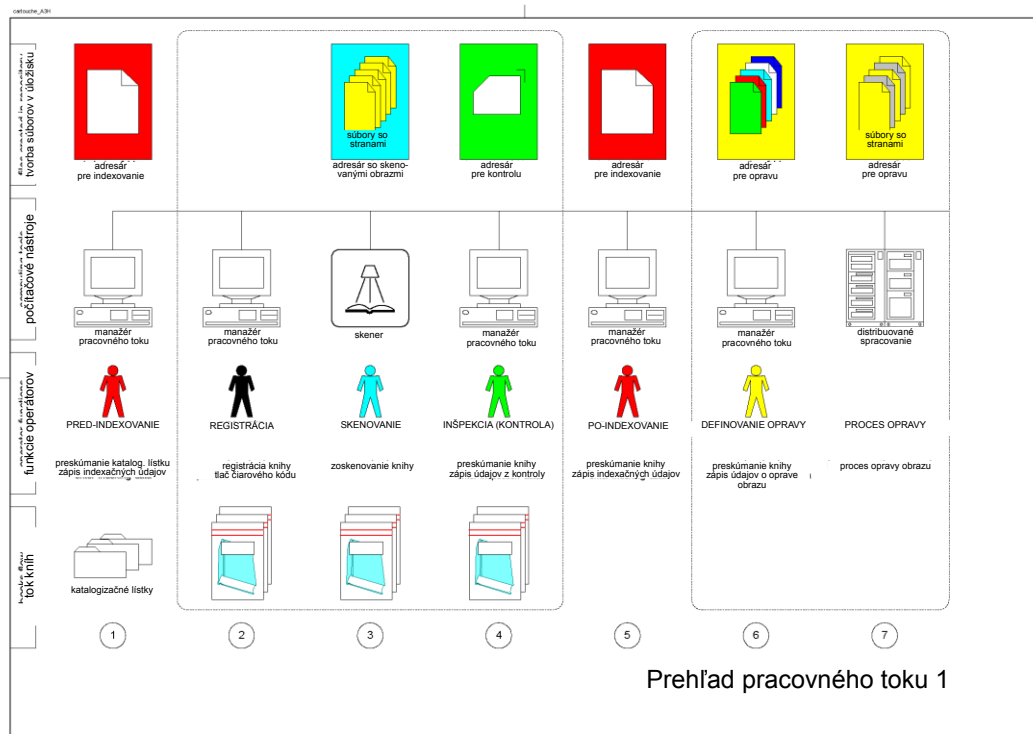
AKTIVITA PREHLAD AKTIVÍT PROJEKTU	
<b>PA1</b>	<b>Aktivita č.1 - Riadenie projektu (Podporná aktivita projektu)</b>
<b>PA2</b>	<b>Aktivita č.2 - Publicita a informovanosť (Podporná aktivita projektu)</b>
<b>A3</b>	<b>Aktivita č.3 – Príprava dokumentov na digitalizáciu a ich chemické ošetrenie</b>
A3.1	Podaktivita č.3.1 – Príprava a logistika dokumentov
A3.2	Podaktivita č.3.2 – Rekonštrukcia depozitov
A3.3	Podaktivita č.3.3 – Dobudovanie infraštruktúry DC
A3.4	Podaktivita č.3.4 - Sterilizácia
A3.5	Podaktivita č.3.5 - Deacidifikácia
A3.6	Podaktivita č.3.6 – Vstupno-výstupná kontrola kvality
<b>A4</b>	<b>Aktivita č.4 – Digitalizácia</b>
A4.1	Podaktivita č. 4.1 – Triedenie dokumentov do digitalizovaných tried
A4.2	Podaktivita č. 4.2 – Digitalizačné linky a ich obsluha
A4.3	Podaktivita č. 4.3 – Spracovanie digitalizovaných dokumentov a riadenie digitalizačných procesov
A4.4	Podaktivita č. 4.4 – Pracovné úložisko dát a technologická infraštruktúra
<b>A5</b>	<b>Aktivita č.5 – Archív (Digitalizácia a konzervovanie dokumentov), ak sú súčasťou archívne dokumenty.</b>

### 9.2 Proces digitalizácie prebieha:

1. *reťazcom skenerov*, ktoré sú generátormi digitálneho obsahu,
2. *dočasným pracovným úložiskom*, ktoré slúži na dočasné uloženie naskenovaných obrazov, vygenerovaných medziproduktov, kópie finálnych produktov určených na migráciu do centrálného archívu,
3. *pracovným priestorom*, nad ktorým sa vykonávajú všetky potrebné medzikroky digitalizácie a podpornou infraštruktúrou.

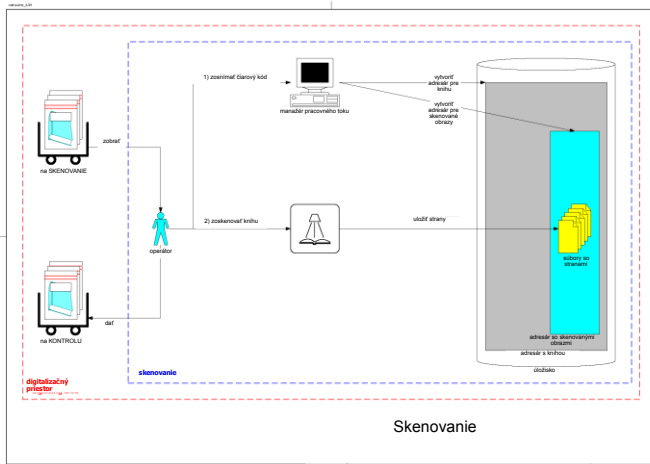
<sup>9</sup> Stránka DIKDA: Digitálna knižnica a digitálny archív (DIKDA). <http://dikda.eu/stav-projektu/>. Hlavným autorom projektu bol Dušan Katusčák, ktorý projekt pripravoval od roku 2005 do roku 2011. Zmluvu o poskytnutí nenávratného finančného príspevku z eurofondov podpísal v marci 2012. Následne ho minister kultúry nahradil nekompetentným manažmentom a z projektu sa realizovalo len malé torzo a finančné prostriedky sa využili extrémne neefektívne. Pôvodný Opis projektu (kolektívu autorov J. Dzivák, M. Čapkovič, O. Ardo, M. Hvolka, J. Gabrielová a i.), na základe ktorého bola uzatvorená zmluva: <https://dusan.katuscak.net/2013/01/20/dikda-digitalna-kniznica-a-digitalny-archiv-opis-projektu/>.

## 9.3 Digitalizačný workflow

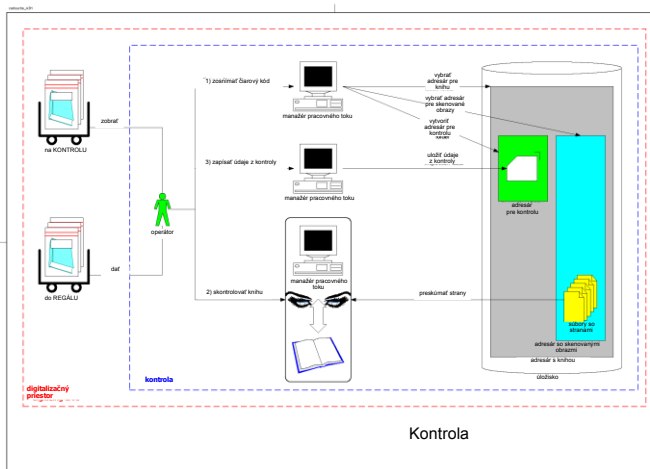


Dôverný!

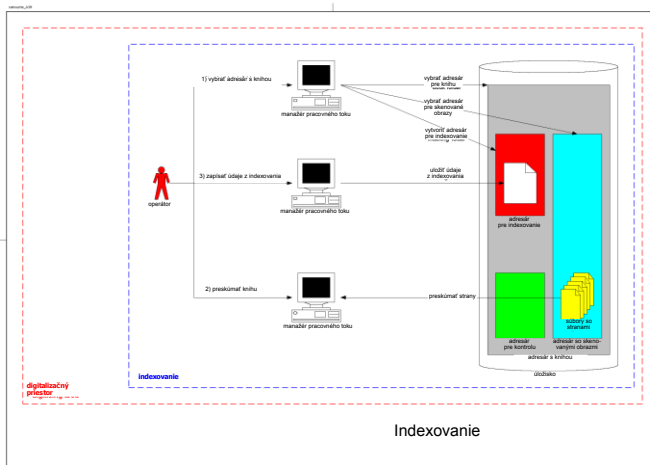




Dôverný!

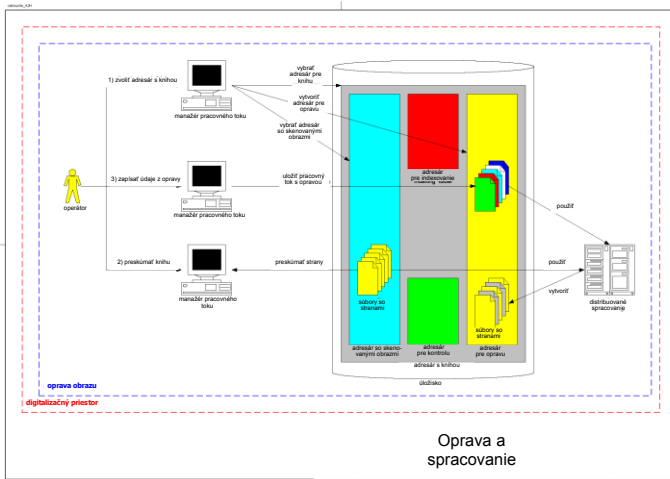


Dôverný!

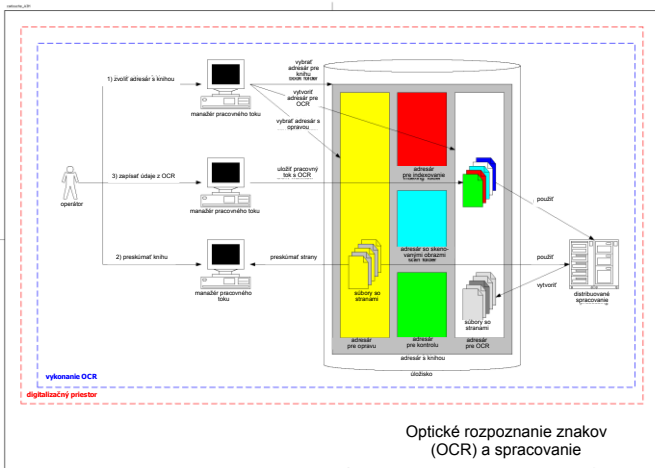


Dôverný!

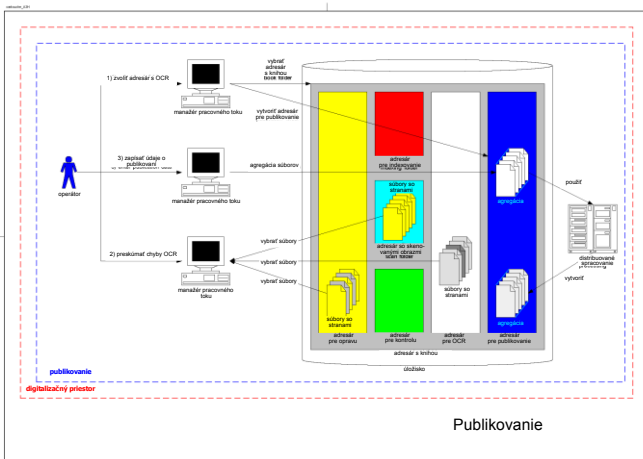




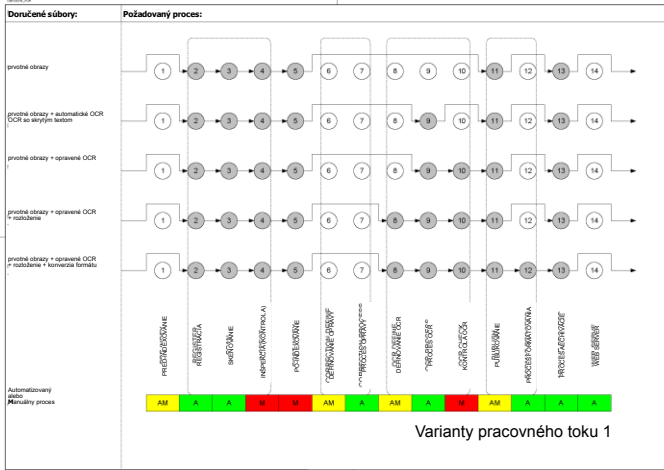
Dôverný!



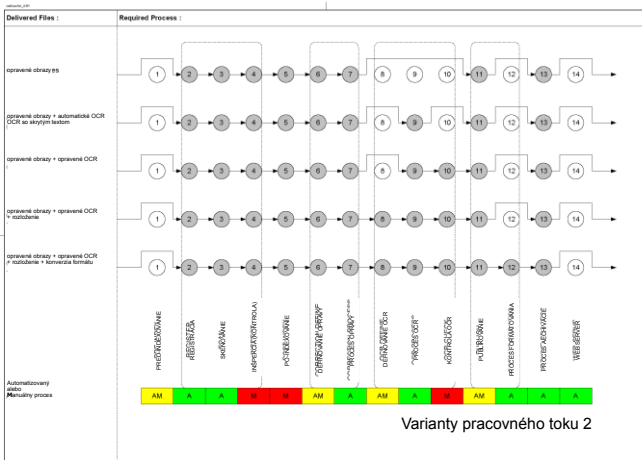
Dôverný!



Dôverný!



Dôverné!



Dôverné!

Samotný proces *všeobecného* postupu digitalizácie je vzhľadom na veľký dátový tok realizovaný paralelne *workflow enginom*, pričom *workflow* zabezpečuje a rozhoduje o presnom toku digitálneho objektu cez digitalizačnú linku v nadväznosti na technologické parametre daného objektu.

### **Generovanie primárnych údajov: Skenovanie**

Digitálny obsah sa bude ukladať na lokálne úložiska digitalizačných liniek a skenerov, ktoré budú súčasťou digitalizačných zariadení. Každý skener (digitalizačné zariadenie) má svoje lokálne úložisko. Pri skenovaní dokumentu sa vytvára digitálny objekt, ktorý je identifikovaný jednoznačným identifikátorom podľa fyzického dokumentu. Každý digitálny objekt je reprezentovaný identifikátorom skenovaného dokumentu. Krok tiež zahŕňa manuálnu a automatickú kontrolu kvality.

### **Prenos primárnych údajov zo skenerov do dočasného úložiska**

Skenery sú pritom z pohľadu toku dát pasívne zariadenia, digitálny obsah je distribuovaný zo skenerov priamo do pracovného archívu ďalej na spracovanie prostredníctvom NFS protokolu a riadené workflowov.

### **Prebranie údajov z dočasného úložiska do pracovného diskového priestoru**

Z pracovného archívu je do pracovného diskového priestoru prenesená pracovná kópia, ktorá je následne upravovaná špecializovaným softvérom na úpravu obrazu po skenovaní.

### **Postprocessing: Úprava obrazu – image treatment a kontrola kvality**

Rieši štandardizáciu obrazu zo skenou a prípravu na OCR a základnú kontrolu kvality skenovania.

### **Generovanie textu a odvodených údajov**

Generovanie plného textu (fulltext) OCR z obrazového materiálu a ďalej tzv. finálnej prezentácie objektu plus špecifické údaje, ako sú indexy a pozičné metadáta.

### **Generovanie metadát**

Generovanie odvodených údajov, Asociácia bibliografických metaúdajov, metaúdajov archívneho dokumentu, digitálneho obsahu, autorských práv. Generovanie Metaúdajov: Bibliografické / popisné metadáta, Technické metadáta

### **Kontrola kvality**

Kontrola kvality tvorby finálnych objektov metódou ISO štandardov.

### **Archivácia derivátov**

Finálne objekty sa uložia do dočasného archívu.

### **Migrácia do CA**

Migrácia objektov z dočasného archívu do CA. Technologické prostredie, ako status quo pozostáva z dvoch lokalít: Martin a Vrútky. V lokalite Vrútky sú umiestnené skenery, chemické technológie sterilizácie a deacidifikácie, neumiestniteľné z dôvodu kapacitných a hygienických v lokalite Martin a sieťovou infraštruktúrou LAN. Vrútky nemajú žiadne vhodné priestory na umiestnenie komplexnejšej IT infraštruktúry a keďže medzi oboma lokalitami je realizovaný optický prepoj. Ďalšie technologické časti IT sú umiestnené v lokalite Martin v rámci rekonštruovaného datového centra z dôvodu efektívnosti – technologickej aj energetickej (takže to je dôvod, prečo nie je možné umiestniť všetko na jednom mieste). Technologicky to sú servery, diskové polia, páskové knižnice, sieťová infraštruktúra a softvérové vybavenie, ktoré sumárne realizuje nasledovné technologické procesy:

- Proces logistiky toku fyzických dokumentov medzi jednotlivými technologickými *procesmi* a tiež medzi *lokalitami* Vrútky a Martin
- *Riadenie* samostatných technologických procesov: *spracovanie a distribúciu* vytvoreného digitálneho obsahu a spracovanie (vytváranie, úprava, dopĺňanie) *metaúdajov*
- *Uchovanie* digitálneho obsahu pre pracovné a procesné účely.

#### 9.4 Konzervácia dokumentov (konzervovanie)

Je to odborné ošetrovanie dokumentov s cieľom zachovať ich súčasný stav, resp. zachrániť poškodené dokumenty pred zánikom. Predstavuje priamu fyzickú ochranu zabezpečujúcu zlepšenie rizikových vlastností dokumentov (nosičov informácie) pri zachovaní ich originality a dlhodobej stability s minimálnou mierou rozpadu.

##### 9.4.1 REŠTAUROVANIE

Na rozdiel od konzervácie reštaurovanie predstavuje takú úpravu dokumentov, aby sa dosiahol ich pôvodný vzhľad. Odborné ošetrovanie konzerváciou sa týka najmä tradičných nosičov, t. j. písomných dokumentov (písomného dedičstva, tj. kníh a archívnych dokumentov, najmä starých a vzácných tlačí, historických rukopisov ap.).

#### 9.5 Integrácia konzervovania a digitalizácie

Konzervovanie a konzervačná veda sa však zaoberá všetkými objektmi kultúrneho dedičstva ako aj nosičmi digitálneho obsahu. Pri audiovizuálnych dokumentoch (filmových dielach) sa za najvhodnejší spôsob konzervácie z dlhodobého hľadiska pokladá pasívna konzervácia na filmových pásoch, hoci za spôsob konzervácie sa považuje aj *digitalizácia*.

Digitalizácia však fakticky nemôže byť konzervačným zásahom, pretože nezabraňuje objektívnej degradácii originálnych nosičov až po ich zánik. Digitalizácia môže nanajvýš spomaliť poškodzovanie dokumentov ich priamym používaním. Je ideálne, ak projekt digitalizácie zahŕňa aj konzervovanie, najmä čistenie, sterilizáciu a deacidifikáciu. Vyplýva to z jednoduchšej úvahy, že ak sa dokument raz vyberie zo skladu, treba s ním urobiť maximum operáciu v rámci jednej technologickej linky. To znamená, že konzervačné procesy je potrebné zahrnúť do komplexného procesu digitalizácie.

Digitalizácia sa spravidla realizuje po konzervácii tradičných dokumentov, v súčasnosti je však využívaná najmä pri ich sprístupnení verejnosti, čím sa zamedzuje ich opotrebeniu. Nosiče digitalizovaných dokumentov sú však spojené s rizikami, ktorými sú napr. zastarávanie dátových formátov, obmedzená životnosť pamäťových zariadení a i.

Konzervácia tradičných (papierových) dokumentov je odborné ošetrovanie dokumentov konzervačnými prostriedkami s cieľom zastaviť alebo spomaliť proces ich rozpadu. Pripúšťajú sa však iba zásahy, ktoré zabezpečujú existujúci stav dokumentu bez akýchkoľvek doplnkov (k používaným metódam patria deacidifikácia, sterilizácia, dezinfekcia, lyofilizácia a i.), v širšom chápaní sa pripúšťajú aj opatrenia, ktoré sú nevyhnutné na technické zabezpečenie fyzického stavu nosiča dokumentu ( laminovanie dokumentov).

#### 9.5.1 STERILIZÁCIA

Sterilizáciou a dezinfekciou dokumentov sa zabraňuje ich ďalšiemu poškodzovaniu mikroorganizmami a predchádza sa riziku poškodenia zdravia pri ich ďalšom spracovaní (alebo používaní).

#### 9.5.2 DEACIDIFIKÁCIA

Významnou súčasťou procesu konzervácie dokumentov je deacidifikácia (odkyslenie) papiera, ktorou sa upravuje jeho nevhodné nízke pH signalizujúce jeho kyslosť, čím sa papier stabilizuje. *Degradácia kyslého papiera je nevratný termodynamický dej, ktorý môže spomaliť deacidifikáciu.* Kyslosť papiera na báze drevnej celulózy je zapríčinená kyselinou sírovou, ktorá sa tvorí pri styku oxidov síry prítomných v atmosfére s vlhkým papierom, ako aj karboxylovými kyselinami, ktoré v ňom vznikajú pri oxidačných procesoch, ktorým časom podlieha. Výsledky testov potvrdzujú, že papier má pomerne dobrú stálosť, ak nemá nižšiu hodnotu pH ako 5,5, pričom všeobecne sa papier považuje za odkyslený, keď jeho hodnota pH vzrastie na 6,2. Na neutralizáciu sa používajú vodné roztoky a suspenzie hydroxidov, uhličitanov alebo hydrogenuhličitanov vápniaka, príp. horčíka, alebo sa dokument umiestni do plynnej atmosféry amoniaku (krátkodobá stabilizácia).

Cieľom deacidifikácie je jednak zníženie kyslosti papiera, jednak vytvorenie dostatočného alkalického pufru, ktorý zabráni reacidifikácii v priebehu času. Poškodený papier možno spevniť aj pomocou spevňujúcich zmesí, ktoré sa nanášajú štetcom alebo ponorením do roztoku. Používajú sa prírodné polyméry (pšeničný škrob, želatína), modifikované prírodné polyméry rozpustné vo vode (sodná soľ karboxymetylcelulózy a i.) alebo syntetické polyméry (polyvinyletanol plastifikovaný glycerínom a i.).

#### 9.5.3 LYOFILIZÁCIA

Lyofilizácia sa používa pri ošetrovaní a záchrane dokumentov postihnutých prírodnými katastrofami (búrky, povodne). — K najvýznamnejším konzervačným (a reštaurátorským) pracoviskám na Slovensku patrí pracovisko Ochrany fondov Slovenskej národnej knižnice v Martine, Odbor ochrany dokumentov Univerzitetnej knižnice v Bratislave, Oddelenie ochrany archívnych dokumentov Slovenského národného archívu v Bratislave a i. Na svete sú to najmä: Getty Conservation Institute a Rochester Institute of Technology.

## 10 PLÁNOVÁNÍ PROCESU DLOUHODOBÉHO UCHOVÁNÍ DIGITÁLNÍCH DĚL V KNIHOVNICTVÍ

### 10.1 Uchovávanie originálneho digitálneho materiálu

Dlhodobé archivovanie digitálneho obsahu je najväčší problém digitalizácie. Bežne sa generuje obrovské množstvo digitálnych dát, avšak problém ich uspokojivého dlhodobého archivovania nie je dostatočne vyriešený. V našom odbore dlhodobosť vymedzuje dĺžkou 100 rokov, čo dnešné efemérne médiá nespĺňajú.

Dôležitým dlhodobým cieľom akéhokoľvek projektu digitalizácie je chrániť a zachovať prístupnosť údajov, ktoré sa v rámci neho vytvorili. To zahŕňa zaobranie sa nevyhnutným zastarávaním formátov digitálnych súborov a rozličných typov počítačových úložných médií.

Uchovávanie originálneho digitálneho materiálu a korešpondujúcich metaúdajov pomáha tomu, by bolo možné vyhnúť sa opätovnej digitalizácii dokumentov/objektov, čím sa ochraňuje krehký zdrojový materiál a nie je potrebné opakovanie namáhavého procesu digitalizácie vrátane generovania metaúdajov.

### 10.2 Formáty súborov

#### Definícia problému

Digitálny výstup procesu digitalizácie je zvyčajne matričný súbor v nekomprimovanom formáte TIFF s určitými začlenenými metaúdajmi (pozri kapitolu Metaúdaje použité pre objekt). Formát súborov a kompresná technika budú mať veľký vplyv na použiteľnosť digitalizačného výstupu. V tomto bode je potrebné vziať do úvahy záležitosti ako formát súboru, štandardnú veľkosť súboru, čas prenosu cez sieť a rôzne typy výstupu (monitor, tlač).

#### Pragmatické návrhy

- Pre rozhodnutím sa pre nejaký formát súborov je potrebné vziať do úvahy súvisiace štandardy, existujúcu globálnu používateľskú základňu a stupeň, do akého sú formáty súborov podporované softvérom používaným organizáciou - a používateľské určenie. Veľkosť globálnej používateľskej základne a súčasná podpora sú dobrými indikátormi budúcej podpory pre konkrétny formát, ktoré vypovedajú aj o pravdepodobnosti trvalo udržateľných migračných ciest po zmene formátov súborov.
- Najbežnejším výstupom digitalizácie vo forme digitálnych obrazov je formát *Tagged Image File Format* (TIFF). Pokiaľ projekt nemá jasný a argumentmi podložený dôvod používať iný formát súborov, výstup digitalizácie, čiže matričné súbory, by mali používať tento formát.
- Výstupný súbor bude typicky veľmi veľký (rádovo aj stovky megabajtov a viac). Je bežné, že vznikajú takéto veľké matričné súbory, ktoré sa uchovávajú lokálne a nie sú určené na prenos cez Internet.
- Z matričného súboru je možné vytvoriť menšie verzie pomocou softvéru na spracovanie obrazu, buď vo formáte TIFF alebo zvyčajne vo formáte vhodnejšom na prenos, akým je JPEG 2000, PNG or GIF (pozri časť o štandardoch).

- Viac informácií o formátoch súborov je v prehľade štandardov ďalej v tomto dokumente.
- Bez ohľadu na to, aký je prítlačivý súkromný alebo národný formát z technického hľadiska, je dôležité mať na mysli, že nepoužívanie štandardných formátov a médií bude hlavnou prekážkou brániacej medzinárodnej výmene súborov s rastrovými obrazmi a náležitých včlenených metaúdajov, rovnako aj vytváraniu sieťových zdrojov.

### **Poznámky/komentáre**

Výber formátu súborov musí byť podriadený pravidlu vytvárať čo najkvalitnejší výstup digitalizácie a dostupnosťou migračných ciest pre budúce uchovávanie digitálneho originálu. Veľkú úlohu v tejto oblasti zohrávajú štandardy.

## 11 POSTUPY DLOUHODOBÉHO UCHOVÁNÍ DIGITÁLNÍCH DĚL

### 11.1 Výber médií

#### Definícia problému

Otázka výberu médií (nosičov) je dôležitá pri projektoch, ktoré chcú udržiavať ich digitálne zbierky počas obdobia niekoľkých rokov. Mnohé projekty (napríklad Domes Day vo Veľkej Británii) sa stratili vďaka zastarávaniu médií.

#### Pragmatické návrhy

- Výstup projektu digitalizácie sa bude uchovávať na serverových strojoch, vrátane tých, ktoré poskytujú digitálny obsah používateľom na Internete. Tieto stroje je potrebné zálohovať. Ak server nie je určený výhradne pre daný projekt digitalizácie, digitálny obsah by sa mal ukladať na odpojiteľné médiá oddelené od iných údajov na servri.
- Všetky matričné súbory (a príslušné metaúdaje) sa budú zálohovať na dvoch druhoch nosičov skladovaných oddelene od seba.
- V súčasnosti (2004) sa bežné zálohovacie médium CD-R nahrádza médiami DVD, ktoré ponúkajú oveľa väčší úložný priestor, pričom zariadenie na čítanie DVD je prítomné na takmer všetkých nových PC a notebookoch. Zariadenia na zapisovanie DVD sú stále dosť drahé, ale ich cena klesá, prípadne sú úplne nahradené inými nosičmi (USB kľúče, pamäťové karty SD a i.)
- Ešte pred 10 rokmi sa neočakávalo, že v blízkej budúcnosti nosiče DVD nahradia magnetické pásky, akou je napríklad páska Digital Linear Tape (DLT) ako úložné médiá použiteľné na zálohovanie počítačových diskov. Je potrebné seriózne zváženie oboch technológií na uchovávanie digitálneho obsahu. V skutočnosti sú pásky LTO5 a LTO6 vhodné na dlhodobé archivovanie. Bežne sa používajú aj pomerne lacné veľkokapacitné úložiská 10 a viac TB aj pre bežné domáce použitie na archivovanie digitálneho obsahu.
- Bez ohľadu na voľbu média je potrebné mať na pamäti, že dané médium bude v blízkej a stredne vzdialenej budúcnosti zastaralé. Je veľmi pravdepodobné, že v priebehu piatich rokov bude nevyhnutná migrácia na nové úložné médiá.

#### Poznámky/komentáre

Projekty digitalizácie v minulosti boli vo veľkej miere projekty poznačené zmenami médií, ktoré boli spôsoboval hlavne konzumentský elektronický priemysel.

Stúpajúci trend „ukladať údaje na internete“ na veľkých serveroch a na mobilných jednotkách s pevnými diskmi však uľahčuje migráciu údajov z jedného miesta na iné a z média na médium. Ak sa z času na čas servre zálohujú a migrujú na nové servery, očakáva sa, že bude klesať závislosť na oddeliteľných médiách ako jediných záznamoch z procesu digitalizácie.

Medzitým ostáva stále dôležitá otázka výberu médií. Nič nenasvedčuje tomu, že sa naráža na obmedzenia pri uchovávaní komprimovaného digitálneho obsahu.



## 11.2 Migračné stratégie

### Definícia problému

Pri výbere formátu súborov a úložného média je potrebné vziať do úvahy realizovateľnosť presunu dát do nového formátu a/alebo na iné úložné médium.

### Pragmatické návrhy

- Preskúmajte príslušné súvisiace štandardy ohľadom formátov súborov a úložných médií, ako bolo uvedené vyššie. Vyhovenie štandardom je rozumným indikátorom toho, že konkrétny formát alebo médium budú nejakým spôsobom podporované aj v budúcnosti.
- K používaniu súkromných formátov súborov a neštandardných formátov médií by sa malo pristupovať s obozretnosťou.
- Pri migrácii z jedného formátu na iný by sa malo zamerať na konverziu maticového digitálneho materiálu z bezstratového formátu (napr. TIFF v doméne obrazov) do stratového (napr. JPEG). Ak sa raz informácie stratia, nie je možné ich nahradiť.
- Je nutné mať na pamäti, že každý formát súborov a/alebo každé úložné médium bude v dohľadnej budúcnosti zastaralé (možno v priebehu piatich rokov, pravdepodobne do desiatich rokov).
- O tom, aká bude pravdepodobná realizovateľnosť migrácie z jedného úložného média na iné pri jeho zastarávaní, vypovedá veľkosť trhu s úložnými médiami.
- Po vytvorení digitálneho materiálu by sa mali úložné médiá (napr. CD-R, DVD) periodicky obnovovať (jedenkrát za dva až tri roky), aby sa zabránilo strate údajov. To sa týka aj kopírovania všetkých údajov na nové médiá.
- Status digitalizovaného materiálu, vrátane údajov o tom, kedy sa obnovoval nosič, by sa mali zaznamenávať v príslušnom protokole.

Ak je to možné, kópie digitalizovaného materiálu sa majú uchovávať na viacerých miestach, čím sa zníži riziko straty údajov pri nehodách ako požiar, záplavy a pod.

## 12 PRÁCE S PŮVODNÍMI DIGITÁLNÍMI DOKUMENTY

### 12.1 Zaobchádzanie a práca s originálmi

Táto časť sa zaoberá tým, ako by sa v rámci projektu malo zaobchádzať s materiálom, ktorý sa digitalizuje. Zdrojový materiál je v mnohých prípadoch vzácny alebo hodnotný, preto sa musí minimalizovať negatívny vplyv, ktorý naň môže mať digitalizácia.

V každom prípade je nutné zdôrazniť, že pre projektový tím budú cenné poznatky osôb, ktoré bežne zodpovedajú za zbierky.

### 12.2 Premiestňovanie a manipulácia s originálmi

#### Definícia problému

V mnohých prípadoch je materiál určený na digitalizáciu veľmi citlivý a krehký. Nahradenie bezprostredného prístupu on-line zverejnením je v prvom rade často dôležitým dôvodom projektu digitalizácie. Je kriticky dôležité, aby projekt prijal opatrenia, aby sa počas procesu digitalizácie materiál žiadnym spôsobom nepoškodil. Tieto opatrenia môžu mať podobu používania správneho hardvéru, zabezpečenia vyhovujúcej mikroklimy alebo presunu digitalizačného centra na miesto uloženia materiálu namiesto sťahovania dokumentov a predmetov.

#### Pragmatické návrhy

- Pred manipuláciou a prácou so zdrojovým materiálom sa poraďte s osobou, ktorá za materiál zodpovedá.
- Vyžaduje sa určitá flexibilita – nevyhovenie z pohľadu projektu digitalizácie sa dá prekonať, zatiaľ čo poškodenie jedinečného artefaktu môže byť nenapraviteľné.
- Ak to bude potrebné, digitalizačné zariadenie (digitálna kamera) sa môže preniesť k zdrojovému dokumentu alebo predmetu a nemusí sa prepravovať objekt určený na digitalizáciu.
- Vyhnite sa rozvázovaniu kníh a spisov. Namiesto plošného skenera použite skener s knižnou kolískou alebo digitálnu kameru.
- Vždy odstráňte skoby, spinky, svorky a iné spony; môžu poškodiť digitalizačné zariadenie i zdrojový materiál.
- Pred prácou s originálom sa treba poradiť s odborníkom (napr. kurátorom dokumentu/predmetu určeného na digitalizáciu).
- Konzultácie sa majú realizovať pred digitalizáciou, ideálne v čase výberu daného kusu na digitalizáciu. Pokyny by sa mali zaznamenať v báze poznatkov digitalizácie a mali by sa preštudovať pred premiestnením alebo digitalizáciou. Ak je to potrebné, schopnosti každého možného hardvérového riešenia by sa mali prekonzultovať s odborníkom.

#### Poznámky/komentár

Aj keď sú tieto odporúčania zrejmé, pri zaobchádzaní so zdrojovým materiálom je nutné a dôležité dodržiavať disciplínu.