

Národní program mikrofilmování a digitálního zpřístupňování dokumentů ohrožených degradací kyselého papíru KRAMERIUS, 6. rok v knihovně Národního archivu – Digitalizace deníku Československá republika. Digitální knihovna Národního archivu

/ Dagmar Spáčilová – Martin Šuraba /

Národní program ochranného reformátování a digitálního zpřístupňování bohemikálních dokumentů, ohrožených totálním rozpadem kyselého dřevitého papíru¹ plní své cíle od roku 1992. V počátcích byly aktivity zaměřené vedle vlastního mikrosnímkování a digitalizace ohrožených dokumentů především na vybudování moderních mikrografických pracovišť a v neposlední řadě stanovení všech nutných standardů komplexního reformátování pro dosažení maximálně možné kvality všech výstupů na úrovni rozsáhlého národního projektu, který dnes svým významem a rozsahem přesahuje hranice České republiky.²

V posledních letech probíhá přechod jak mezi dvěma vývojovými etapami standardů a technologií ochranného reformátování ohrožených dokumentů³ tak mezi dvěma verzemi systému (resp. aplikace) pro zpřístupňování digitalizovaných dokumentů. Postupně je zaváděna verze 4 digitální

1 Veřejné informační služby knihoven. Podprogram č. 7 (VISK 7) – Národní program mikrofilmování a digitálního zpřístupňování dokumentů ohrožených degradací kyselého papíru KRAMERIUS.

2 Shrnující informace viz:

Polišenský, Jiří. Současné trendy v digitalizaci novodobých dokumentů. In: Knihovny současnosti 2008. Brno 2008, s. 51–60; týž: Souhrnná zpráva o realizaci projektů programu VISK 7 v r. 2011.

Dostupné z: <http://visk.nkp.cz/2011/VISK7/VISK7realizace.pdf>.

3 Podrobný popis nových standardů digitalizace dokumentů od roku 2012 pro Národní digitální knihovnu je dostupný na: <http://www.ndk.cz/digitalizace/nove-standardy-digitalizace-od-roku-2011>.

knihovny Kramerius, která bude nadále pracovat se stejnými standardy jako projekt Národní digitální knihovny⁴.

Hybridní technologie ochranného mikrofilmování jako pojistka dlouhodobého (trvalého) „přežití“ ohrožených dokumentů zatím zůstává součástí komplexního reformátování a digitalizace.

Vedle ochranného reformátování ohrožených bohemikálních dokumentů projekt Kramerius podporuje další aktivity na poli ochrany tištěného kulturního dědictví:

- instalace Systému Kramerius (verze 4) v knihovnách
- zhotovení ochranných obalů z lepenky archivních kvalit pro již reformátované svazky, jejichž původní vazba je silně poškozená a převazba, vzhledem k pokročilé degradaci kyselého dřevitého papíru, nepřichází do úvahy
- odkyselení svazků postižených degradací kyselého dřevitého papíru

Zhodnocení projektu „Úřední a oficiální seriálové dokumenty republiky Československé“, třetí etapa (pokračování) – VISK7 – Kramerius 2011 knihovny Národního archivu schváleného Ministerstvem kultury České republiky (Rozhodnutí č.j. 1610/2011 SOLK, o poskytnutí dotace ze státního rozpočtu ČR na rok 2011).

Rok 2011 ještě patřil k období přechodu od digitalizace podle dosavadních standardů a formátů⁵ k digitalizaci podle nových standardů a formátů používaných Národní digitální knihovnou⁶. Knihovna Národního archivu po konzultaci s Národní knihovnou v roce 2011 pokračovala v digitalizaci deníku Československá republika (Pražské noviny) ještě podle stávajících standardů.

4 <http://www.ndk.cz/narodni-dk>

5 <http://www.digit.nkp.cz/>

http://digit.nkp.cz/mw_frame_cz.htm;

<http://digit.nkp.cz/Kramerius/AdminMetaData/1.0/AdminMetaData.xsd>

<http://digit.nkp.cz/Kramerius/AdminMetaData/1.0/ImportAdminMetadatKramerius.doc>

6 http://www.ndk.cz/digitalizace/nove-standardy-digitalizace-od-roku-2011/metadata_periodika_v1.2-1 (bez názvové konvence složek a souborů – tabulka na str. 8)
<http://kramerius-info.nkp.cz/visk/zadavaci-dokumentace-visk-7-pro-rok-2011/obrazove-formaty-JPEG2000/view>

Digitalizací „vládního“ deníku „Československá republika (Pražské noviny), který vycházel v letech 1920–1938⁷ a který se i ve své hlavičce uváděným ročníkem CCXLI [241] – 259 po celou dobu svého vycházení hlásil k „Pražským novinám“⁸ se knihovna NA zapojila do procesu digitálního zpřístupnění jednoho z nejdéle vycházejícího titulu novin v dějinách české (československé) žurnalistiky. Zároveň tím naplňuje své cíle v oblasti péče o dokumenty specifické pro fondy knihoven archivů ale i jiných úřadů – tj. v oblasti péče o dokumenty vládní, úřední a/nebo oficiální speciálně nebo všeobecně publikované. Projekt digitalizace „Úřední a oficiální seriálové dokumenty republiky Československé“ v tomto ohledu plynule a úzce navázal na již uzavřený projekt „Sibiřské tisky“⁹.

Od počátku digitalizace dokumentů knihovny NA v rámci národního programu VISK – podprogramu č. 7 – Kramerius od roku 2006 až do roku 2010 (2011) byl projekt ochranného reformátování dokumentů zároveň koncipován i jako projekt pořízení moderního a výkonného technického zařízení – PC sestavy s přídatným externím HW – pro provoz digitální knihovny NA. Realizace tohoto záměru se však podařila až v roce 2011 díky schválení žádosti o rozšíření projektu a přidělení 40 000 Kč výhradně na pořízení moderní počítačové sestavy pro správu, dlouhodobé uložení a kvalitní zpřístupnění digitalizovaných dokumentů.

Přidělené finanční prostředky (včetně finanční spoluúčasti Národního archivu) byly v rámci stanovených podmínek pro nakládání s dotací daných zadávací dokumentací projektu VISK 7 2011 plně využity pro realizaci projektu knihovny Národního archivu:

7 Československá republika (Pražské noviny) / vydavatel a zodpovědný redaktor Josef St. Hevera. – Roč. CCXLI (č. 1, ve čtvrtek 1. ledna 1920) – 259 (č. 301, v sobotu 31. prosince 1938). – Praha : Státní tiskárna v Praze, 1920–1938. Přílohy: Nedělní čtení.

8 přidělená ISSN souvisejícím titulům:

ISSN1802-775X	Pražské nowiny (1825–1848)
ISSN 1802-7768	Rozličnosti Pražských nowin (1826–1833)
ISSN 1802-7776	Konstituční Pražské nowiny (1848)
ISSN 1802-7781	Pražské nowiny (1848–1864)
ISSN 1803-4837	Pražský denník (1866–1895)
ISSN 1803-1845	Pražské noviny (1895–1919)
ISSN 1803-4853	Československá republika (Pražské noviny) (1920–1932)
ISSN 1803-4861	Pražské noviny (1932) – Československá republika (1932–1938)

9 rovněž realizován v rámci projektu VISK 7 – v letech 2006–2008; i zde šlo vedle unikátnosti dokumentů o oficiální publicistiku zaštitěnou ve většině případů činností Československé národní rady

- třetí etapou¹⁰ pokračovala digitalizace deníku „Československá republika (Pražské noviny)“, byla digitalizována část 1923 (říjen–prosinec)–1925 (leden–červen), celkem 7 svazků, tj. 6924 všech stran včetně stran s technickými, administrativními a bibliografickými metadaty, z toho celkem 6720 stran originálních předloh (dokumentů).
- bylo pořízeno technické zařízení – PC sestava s příslušným HW a SW vybavením pro kvalitní zpřístupňování digitalizovaných dokumentů a zároveň i jako základní moderní vybavení pro správu a provoz digitální knihovny Kramerius knihovny NA
- a byla provedena instalace systému Kramerius, verze 4 na tuto PC sestavu včetně zavedení digitalizovaných dokumentů knihovny Národního archivu, dosud digitalizovaných v letech 2006–2011 v rámci národního projektu „Veřejné informační služby knihoven. Podprogram č. 7 – Národní program mikrofilmování a digitálního zpřístupnění dokumentů ohrožených degradací kyselého papíru KRAMERIUS“.

V roce 2011 byl položen dobrý základ moderního využívání fondu „Sbírka novin NA“¹¹, rozsáhlého sbírkového fondu, který obsahuje dokumenty (noviny všech i historických typů a forem) počínaje rokem vydání 1722 do současnosti. Fond vznikl postupně v průběhu mnoha desítek let ať už na základě odběru novin vlastním archivem (počínaje Archivem země České a dále pak Ústředním archivem ministerstva vnitra a Ústředním zemědělsko-lesnickým archivem) nebo získaných v přejímkách knihoven zanikajících úřadů a institucí v 19. a 20. století. Velmi cenné části fondu jsou rozsáhlé novinové archivy politických stran a stranických vydavatelství, které knihovna NA převzala do svých fondů v letech 1990–do současnosti (Vydavatelství Melantrich a stranický tiskový archiv ČSNS, Archiv Vydavatelství Rudé právo a Redakční archiv Lidové demokracie). Stejně cenný je celek zahraničních a českých (československých) novin převzatý z Ministerstva zahraničních věcí.

Můžeme tedy začít v plné míře zohledňovat ten fakt, že noviny jako dokument patří k velmi využívaným dokumentům v historickém bádání a tedy i ve výpůjčních a informačních službách knihovny, což na jedné straně plně naplňuje oprávněnost jeho existence a péče o něj v NA, na druhé

10 Projekt digitalizace úředních a / nebo oficiálních dokumentů republiky Československé byl zahájen v rámci třetí (závěrečné) etapy digitalizace tzv. „sibiřských tisků“ v roce 2008.

11 Je katalogově veřejnosti zpřístupněn aktualizovanou katalogovou pomůckou „Sbírka novin 2010 říjen“ v Intranetu knihovny NA / část Katalogy knihovny NA – Vyhledávání

straně ale patří k nejhorším faktorům jeho ničení. I knihovna NA byla a je nucena postupně „uzavírat“ některé nejvíce frekventované tituly a vylučovat je z výpůjčních služeb. V roce 2009 tak byl např. uzavřen deník Rudé právo (období let 1938 a 1948–1953). Venkov je vyloučen z výpůjčních služeb více jak 5 let.

I pro knihovnu NA je tedy stále více nutné zpřístupňování tohoto typu dokumentů formou jeho digitálních kopií. Vzhledem k tomu, že ve fondu Sbírka novin NA knihovna vlastní velké množství titulů (rozsah sbírky je cca 1200 b.m. a obsahuje cca okolo 1000 titulů bohemikálních a cizích novin všech typů a forem) může na jedné straně sama nebo formou výpůjček svých dokumentů jiným institucím pro dokompletování titulu přispět do systému Kramerius digitalizací svých dokumentů, na druhé straně může podle standardů a pravidel projektu „Veřejné informační služby knihoven (VISK)“, podprogram 7 Kramerius zpřístupňovat své originální dokumenty formou digitálních kopií pořízených na základě replikačních smluv.

Pociťujeme tedy velké uspokojení nad vlastními výsledky dlouhodobého národního digitalizačního projektu VISK 7, etapy roku 2011. Uspokojení pociťujeme rovněž nad úspěšně a smysluplně se rozvíjející spoluprací mezi knihovnou NA a Knihovnou Vojenského historického ústavu v rámci plnění záměrů národního projektu VISK 7 – Kramerius. V roce 2011 jsme zapůjčením odborného (profesního) časopisu „Vládní vojsko“¹² knihovně Vojenského historického ústavu navázali na roky 2006–2007 (2008), kdy naopak knihovně Národního archivu byly zapůjčeny dokumenty z Knihovny Vojenského historického ústavu pro projekt „Sibiřské tisky“.

Spolupráce knihoven obou ústavů v oblasti ochrany knihovních fondů a ochranného reformátování dokumentů prostřednictvím výpůjčních a informačních služeb obou knihoven pro projekt VISK 7 – Kramerius dává velkou šanci naplnit beze zbytku cíle tohoto národního projektu. Digitální knihovna Kramerius zpřístupňuje vzácné a ohrožené bohemikální tisky, starší noviny, časopisy a další vzácné tituly tištěné na kyselém papíře. Již dnes její význam např. jen pro dějiny českého knihtisku, dějiny české žurnalistiky a pro jiná historická bádání překračuje hranice České republiky. Je jen potěšitelné, že zpřístupňuje i dokumenty dochované v knihovních fondech knihovny NA a Knihovny Vojenského historického ústavu.

12 Časopis byl zapůjčen z dílčího historického knihovního fondu Knihovna Ministerstva vnitra knihovny NA, 6 svazků z let 1939–1945 (kompletní titul).

Shrnutí výsledků třetí etapy (pokračování) projektu „Úřední a oficiální seriálové dokumenty republiky Československé“ v roce 2011:

Formou komplexní digitalizace (mikrofilm 2 generací + digitalizace dokumentů z matriční negativní kopie archivního mikrofilmu) byl digitalizován 1 titul: Československá republika (Pražské noviny), část 1923 (říjen–prosinec)–1925 (leden–červen):

- seriálu bylo přiděleno ISSN již v roce 2008 (při zahájení digitalizace titulu v rámci třetí, závěrečné etapy projektu „Sibiřské tisky“). V případě titulu „Československá republika (Pražské noviny)“ byl pro potřeby přidělení ISSN proveden podrobný a časově náročný bibliograficko-dokumentační a bibliograficko-lokační průzkum
- knihovně NA byly předány mikrofilmy dvou generací (archivní negativ a matriční negativ) a komplexní digitalizované dokumenty na externím HDD 1,5 TB
- kopie komplexně digitalizovaných dokumentů v požadovaných formátech byly na výměnném HDD předány firmou Elsys Engineering do Národní knihovny k archivaci a pro zavedení do národní databáze „Kramerius“. Mikrofilmy dvou generací (tj. archivní negativ a matriční negativní kopie archivního negativu) budou uloženy ve specializovaném depotu pro foto-fono- kino dokumenty Národního archivu
- na mikrofilmech je celkem nasnímáno (dle faktury č. 201100021, vystavené firmou Věra Fialová, MICRONA) 3.462 „hrubých“ polí včetně snímků technický předloh / 3.380 „čistých“ polí stran dokumentů, celkem bylo digitalizováno (dle faktury č. 11010185 a dodatkové faktury 11010226, vystavené firmou Ing. Ivan Sabo, Elsys Engineering) 7.072 všech stran (včetně stran s technickými, administrativními a bibliografickými metadaty), z toho 6826 stran dokumentů
- bylo pořízeno technické zařízení (PC sestava s příslušným HW a SW vybavením) v celkové hodnotě Kč 41.298,- (Kč 40.000,- bylo hrazeno z dotace). PC byl zprovozněn, firmou UNIDATA byl instalován systém Kramerius v. 4, byla natažena data dokumentů digitalizovaných v letech 2006–2010 (2011) a bylo provedeno zaškolení administrátorů – v celkové hodnotě Kč 26.000,-

Celkové náklady projektu činily **Kč 158.333,-** (čtetně DPH), z toho bylo Kč 110.000,- hrazeno z dotace projektu (tj. cca 69,47 % celkových nákladů

a zároveň 100 % dotace) a Kč 48.333,- z finanční spoluúčasti Národního archivu (tj. cca 30,53 % celkových nákladů projektu).

Na závěr shrnutí všech výsledků můžeme konstatovat, že knihovna NA je plnohodnotně zapojena do národního projektu VISK 7 – Kramérius – projektu zachrany seriálových i monografických bohemikálních dokumentů 19. a 20. století. V rámci tohoto projektu digitalizuje unikátní a jedinečné dokumenty z vlastních fondů a zároveň spolupracuje s dalšími řešiteli v rámci kompletace digitalizovaných dokumentů pro dosažení maximálně možné kompletnosti dokumentů a dosažení maximálně možné kvality digitalizovaných dokumentů¹³.

Od roku 2009 jsou aktivity knihovny Národního archivu na poli digitalizace tištěných dokumentů zviditelněné kromě stávajících databází¹⁴ rovněž v Registru digitalizace České republiky¹⁵, projektu evidence digitalizovaných dokumentů a sledování procesu zpracování. Registr k 5. 1. 2012 eviduje skutečnosti do roku 2011.

Od roku 2011 (jesp. 2012) má knihovna NA k dispozici moderní počítačovou sestavu (v současnosti s vnitřní architekturou jako server / uživatelská stanice systému Kramérius) a stojí na prahu zprovoznění své digitální knihovny a zapojení do rozrůstající se rodiny digitální knihoven České republiky i digitálních knihoven zahraničních¹⁶.

13 Pro potřeby dokončování deníků a časopisů projektu „Sibiřské tisky“ knihovna NA v letech 2006–2008 úzce spolupracovala s knihovnou Vojenského historického ústavu, která knihovně NA zapůjčila své téměř 100% kompletní dokumenty; v roce 2008 naopak knihovna NA zapůjčila scházející čísla a měsíce titulu „Zlínská Pravda“ 1945 – dále jako „Naše Pravda“ 1945–1958, digitalizovaného Krajskou knihovnou Františka Bartoše Zlín; v roce 2011 knihovna NA zapůjčila knihovně Vojenského historického ústavu profesní časopis „Vládní vojsko“ 1939–1945.

14 Systém a databáze Manuscriptorium (digitalizace rukopisů, starých a vzácných tisků), systém a databáze Kramérius (digitalizace bohemikálních tisků – monografií a seriálů – 19. a 20. století ohrožených degradací kyselého papíru) a také bibliograficko-lokační databáze – řešitelé jsou povinni předat záznam (popřípadě podklady pro vytvoření záznamu) na každý reformátovaný titul pro centrální evidenci reformátovaných dokumentů (CZROMM – Czech Register of Microform Masters), při dokončení projektu pak záznam ve formátu MARC 21 (příp. UNIMARC) pro evropskou databázi EROMM (European Register of Microform Masters).

15 Společný projekt Knihovny AV ČR, Národní knihovny a firmy Incad (systém Relief, v. III – ve stejném systému je vyprojektovaná struktura i relační katalogové databáze a dokumentačních databází knihovny NA); projekt Registr digitalizace – on-line na <http://www.registrdigitalizace.cz>

16 Blíže viz Portál Knihovny.cz (<http://www.knihovny.cz/>) – kapitola České digitální knihovny (<http://www.knihovny.cz/digitalni-knihovny>) / kapitola Zahraniční digitální knihovny (<http://www.knihovny.cz/digitalni-knihovny/zahranicni-digitalni-knihovny>)

Poznámkový aparát k digitálnej knižnici Kramerius

Už od najstarších čias platí, že ten, kto nemal informácie, bol ochudobnený. Typickým príkladom môže byť aj zrada vo vojnovom konflikte. Nemáme na mysli teraz ľudské hľadisko zrady, ale hovoríme o získaní informácií. Bez informácií by sme nemali vedomosti a nedokázali by sme sa vyvíjať ďalej. V každej jednej oblasti sa informácie využívali, ale nás zaujímajú predovšetkým inštitúcie, ktoré informácie predávajú druhým. Hovoríme najmä o pamäťových inštitúciách, ktorých sme súčasťou.

Samotné slovo informácia si prešlo niekoľkými revolučnými zmenami. Pre nás sú najvýznamnejšie dve: vznik kníhtlače a web¹⁷. Získavanie a sprístupňovanie informácií prešlo obrovským historickým vývojom. Učili sme sa zaznamenávať informácie na steny jaskýň, hlinené tabuľky, papier, prišiel vek novín a neskôr sme zaznamenali obrovský vplyv informačných technológií. Poznali sme diskety, CD, USB kľúče a prišla doba, kedy sa informácie nachádzajú na sieti.

Už dlhšiu dobu sa začíname oboznamovať so slovami a slovnými spojeniami ako sú digitalizácia, digitálne dokumenty, digitálne zbierky, digitálne knižnice a mnoho iných slov. Samotné slová sa zdajú byť na prvý pohľad synonymami, ale cítiť v nich rozdiely.

Prídavné meno digitálny by sme si mohli dať do súvislosti s počítačom. Počítač je matematický stroj, ktorý počíta v dvojrovej sústave. Digitálny by sme mohli preložiť ako číselný. Niektoré jazyky slovo digitálny prekladajú do slov, ktorým rozumieme „číselný“. Príkladom Francúzsko (Bibliothèque numérique) a Poľsko (Biblioteka cyfrowa). Počítač pracuje s dátami, ktoré sú v digitálnej podobe. Komunikácia prebieha prostredníctvom výpočtov v dvojrovej sústave, prostredníctvom matematických operácií jednotiek a núl.

Digitálny dokument by sme mohli vnímať ako dokument v digitálnej podobe. Potrebujeme zabezpečiť dokument, aby s ním dokázal pracovať počítač. Základný dôvod digitalizácie nie je dať dáta do digitálnej podoby. Je potrebné zachovať kultúrne a archívne dedičstvo a hlavne digitálne dáta sprístupniť užívateľom.

Je potrebné vytvoriť vhodné užívateľské prostredie, ktoré bude pre každého intuitívne. Užívateľské prostredie je komunikačný nástroj, ktorý komunikuje s užívateľom a počítačom. Touto tematikou sa zaoberá odbor

17 Internet je sieť, ktorá má mnoho služieb, napríklad emaily, obrovský prenos súborov a predovšetkým najznámejšia služba, web.

HCI, dobré užívateľské prostredie musí mať jednoduchú navigáciu a užívateľ sa musí v ňom dokázať orientovať bez manuálu alebo inej pomôcky. Môžeme to porovnať s knižnicami. Ak sú knihy utriedené podľa nejakého systému, ktorý čitateľ pochopí, bude sa do knižnice vracaať.

Posledné slovné spojenie, ktoré by sme si mali vysvetliť, je digitálna knižnica. Najjednoduchší slovný výklad by bol, že ide o knižnicu, ktorá je v digitálnej podobe. Svojím spôsobom je to aj pravda, ale tomu výkladu málokto rozumie. Samotný pojem digitálna knižnica je tzv. multivýkladový. Čím viac sa objavujú digitálne knižnice, tým viac sa láme výklad toho pojmu. Mnoho ľudí vníma digitálnu knižnicu ako svoje oskenované knihy, iní zasa zdigitalizované knižničné katalógy, ďalší vnímajú pojem digitálna knižnica ako akúsi službu, ktorá im poskytne elektronické dokumenty¹⁸.

Všetky spomínané a iné výklady majú obrovský kus pravdy. Je potrebné uvedomiť si, že knižnice už od najstarších čias okrem iných cieľov majú za cieľ svoje dokumenty sprístupňovať verejnosti. Teda, že hocikto sa dostane do knižnice a môže využívať knižničné služby. (O tom ako má využívať služby je napísané v knižničnom poriadku.) Cieľom digitálnych knižníc v oblasti prístupnosti je aj to, aby sa k nim mohol dostať ktokoľvek bez návštevy kamennej budovy. Pre tento účel slúži internet. Ľudia si prezerajú digitálnu kópiu dokumentov na webových stránkach. Nie je to jednoduchá záležitosť z hľadiska finančného, technického a právneho.

Samotný proces digitalizácie je veľmi drahý. Je potrebné dokumenty dať do digitálnej podoby (skenery, fotoaparáty, digitalizačné steny, apod.), taktiež sa musia zaplatiť ľudia, potom nejaký server, kde bude samotný proces bežať a potom tie najtechnickejšie detaily ako je miestnosť, elektrina, atď.

Zväzuje nám ruky aj právny problém. Autorské práva sú mnohokrát problémom k sprístupňovaniu digitálnej kópii pre verejnosť. Digitálnu knižnicu by sme mali vnímať ako službu, ktorá je na sieti. Čo sa týka siete, je potrebné rozlišovať celosvetovú (internet) a vnútornú sieť organizácie (intranet). Máme mnohokrát autorsko-právne problémy, ktoré nám zväzujú ruky a preto nemôžeme dáta sprístupňovať po internete. Vtedy volíme alternatívne riešenie, spustiť dáta cez intranet.

Technické hľadisko je takisto potrebné vyriešiť. Ešte predtým ako si vysvetlíme technické hľadisko digitálnej knižnice Kramerius, pozrieme sa

18 Pre tento článok by bolo vhodné vysvetliť, že za elektronické dokumenty budeme považovať tie, ktoré nemajú papierovú podobu (tzv. born digital). Digitálne dokumenty vnímame oba druhy, teda aj tie, ktoré sú len born digital, ale aj zdigitalizované papierové dokumenty.

na architektúru digitálnych knižníc. Samotné slovo architektúra by sme zrejme hľadali v stavbárskom priemysle.

Poznáme však aj počítačovú architektúru, kde hovoríme o tom ako sú poskladané počítače. Vieme rozprávať aj o informačnej architektúre. Informačná architektúra stavia informačný systém tak, aby bol pre užívateľa príjemný. Architektúra digitálnych knižníc stavia digitálnu knižnicu. Pri popisovaní architektúry digitálnej knižnice využijeme architektúru Kahn – Willensky.

Základné pojmy, ktoré spomínaná architektúra rozoberá, sú digitálny objekt a digitálny materiál. Digitálny objekt je datová štruktúra, ktorej obsahom je digitálny materiál. Potom tu máme jedinečný identifikátor (handle¹⁹.) Jedinečný identifikátor by sme mohli analogicky porovnať k prírastkovému číslu. Príde kniha do knižnice, dostane prírastkové číslo, aby sme knihu mohli lepšie nájsť. Podobne je to aj s digitálnymi dokumentmi. Digitálny dokument dostane identifikátor, ktorý identifikuje dokument a vďaka ktorému sa odlišuje od ostatných.

Pre architektúru je dôležitý aj samotný pojem repozitár. Ide o priestor, kde sa ukladajú digitálne objekty. A potom tu máme vyhľadávač, ktorý na základe indexácie dokumentov dokáže vyhľadávať daný dokument v digitálnom repozitári. Spomínaná architektúra má 4 základné komponenty.

1. Užívateľské prostredie
2. Vyhľadávací systém
3. Repozitár
4. Identifikátor

Funguje to tak, že užívateľ v užívateľskom prostredí napíše svoju požiadavku do vyhľadávacieho systému. Vyhľadávač prehľadá repozitár a vyberie objekt, ktorý používateľ zadá. Keďže každý digitálny objekt má identifikátor, užívateľ získa predovšetkým handle identifikátor, s ktorým naďalej pracuje. (napríklad link adresy.)

Ako je to v digitálnej knižnici Kramerius? Technické hľadisko knižnice Kramerius by sme si mohli rozdeliť na hardvér a software. Hardvérovo ide o počítač, ktorý má obrovské komponenty. Softwareovo má taktiež Kramerius obrovské požiadavky.

19 Robert Khan a Robert Willensky sú americkí vedci a Američania za svoj identifikátor považujú Handle.

Potrebujeme:

1. Stabilný operačný systém
2. Programovací jazyk
3. Databázový server
4. Webový server
5. Digitálny repozitár
6. Vyhľadávač

Operačný systém je základný prvok každého počítača. Operačný systém komunikuje s počítačom a človekom. Prekladá prirodzený jazyk užívateľa do umelého jazyka, v ktorom vie pracovať počítač. Napríklad ak si chcem nahrať fotky do počítača, operačný systém komunikuje s počítačom ako dlho bude trvať kopírovanie, ako budú fotky vyzerajú, apod. Existuje mnoho operačných systémov. Niektoré sú komerčné (Microsoft Windows, Apple Macintosh), iné sú open source²⁰. Software Kramerius v NA beží na linuxovej distribúcii Debian.

Programovací jazyk potrebuje každý software, v ňom je napísaný a robí funkcie. Program je postupnosť príkazov. Kramerius je napísaný v programovacím jazyku Java.

Databázový server je nástroj, kde sú uložené dáta. Dáta musia mať nejakú štruktúru, aby sme s nimi dokázali pracovať. Databáze ukládajú dáta pre ďalšiu prácu s nimi. Na to nám slúžia dotazovacie jazyky. Databázový server, ktorý využíva Kramerius, sa volá PostgreSQL.

Webový server slúži na komunikáciu programu a webového prostredia. Kramerius využíva webový server Apache a jeho nadstavbu Tomcat.

Digitálny repozitár sme si vysvetlili; je potrebné povedať, že Kramerius využíva digitálny repozitár Fedora Commons.

Vyhľadávač sme si takisto vysvetlili. Kramerius využíva open source vyhľadávač Solr.

20 Open source je software s otvoreným kódom. To znamená, že software si môžeme upravovať na základe svojich potrieb. Ak sa bavíme o open source pri operačných systémoch, rozprávame sa o ceste linuxu. Linux má mnoho distribúcií, ktoré sú voľne šíriteľné a otvorené.