

<h2>Zpráva ze služební cesty</h2> <h3>Projekt „Vytvoření Národní digitální knihovny“</h3> <p>CZ 1.06/1.1.00/07.06386</p>
--

Jméno a příjmení účastníka cesty	Ing. Petr Žabička
Pracoviště – dle organizační struktury	Oddělení 1.7. IOP – Moravská zemská knihovna (knihovnické činnosti):
Pracoviště – zařazení	Zástupce a Koordinátor za partnera projektu – MZK
Důvod cesty	Účast na konferenci 3rd LIBER-EBLIDA Workshop on Digitization of Library Material in Europe a návštěva digitalizačního centra KB
Místo - město	Den Haag
Místo – země	Nizozemí
Datum (od – do)	4.10. - 7.10.2011
Podrobný časový harmonogram	viz http://indico.eblida.org/conferenceOtherViews.py?view=nicecompact&confId=22
Spolucestující z NK	Pavla Švástová, Tomáš Foltýn
Finanční zajištění	IOP „Vytvoření Národní digitální knihovny“
Vztah k projektu	<ul style="list-style-type: none"> ▲ získání informací o procesech digitalizace v digitalizačním centru KB ▲ zkoumání potencionálního využití systémů prezentovaných na této konferenci (např. Goobi, Digitalkoot apod.) ▲ převzetí zkušeností s různými metodikami vývoje softwaru pro využití v knihovnách ▲ diskuze kolem standardů užívaných v digitálních knihovnách a problematice open-metadata a open-data ▲ možnosti evaluace a user studies digitálních knihoven a archivů ▲ nastavení workflow digitalizace, optimalizace procesů a monitoring workflow

Cíle cesty	Zjistit aktuální informace o světových trendech v oblasti digitalizace, diskuze s účastníky s fungujícími linkami masové digitalizace o možných problémech, diskuze o vývoji metadatových standardů a jejich využívání pro potřeby digitalizace. Návštěva digitalizačního centra KB a načerpání získaných zkušeností.
Plnění cílů cesty	Na konferenci byly představeny poslední verze systémů, které by byly potencionálně využitelné v projektu NDK, možná šíře jejich využití, silné a slabé stránky. Významná část konference věnovaná evaluaci a user studies se bude hodit v pozdějších fázích projektu při zveřejňování výsledků v digitální knihovně. Zkušenosti z digitalizačního centra KB budou promítnuty do možných vylepšení digitalizačního procesu v rámci NDK.
Další podrobnější informace	viz strana 3
Podpora publicity projektu	-

Související materiály	
Materiál	Místo uložení

Datum předložení zprávy	20.10.2011
Podpis předkladatele zprávy	

Datum	Podpis
--------------	---------------

Podpis nadřízeného

Vloženo na intranet

Přijato v mezinárodním oddělení

Podrobnější informace o služební cestě:

K cestě do Prahy 4.10. využito místa ve voze p. Neumanna, v 10 hodin jsem se zúčastnil jednání řídicího výboru NDK (viz. Samostatný zápis). Odjezd z Prahy ve 13 hod. společně s P. Švástovou vozem Knihovny Akademie věd (řídil M. Lhoták, který také cestoval na konferenci do Haagu). Přespání v Osnabrucku (zaplatil jsem společný účet za ubytování i za P. Švástovou), dopoledne dojezd do Haagu (ubytování v Haagu rezervoval M.Lhoták a protože mu byla cena za ubytování stržena ze služební platební karty, bylo NK přefakturováno). Odjezd z konference v pátek 7.10., příjezd do Prahy 19:30. Mou jízdenku na autobus do Brna uhradila prostřednictvím kreditové jízdenky P. Švástová.

K programu konference

K nejzajímavějším přednáškám patřily zejména:

Crowdsourcing as a Way to Make Pictures Visible for the Research

Majlis Bremer-Laamanen (The National Library of Finland)

Příspěvek popisoval projekt Finské národní knihovny a firmy Microtask. Cílem projektu je indexovat digitální archivy tak, aby byly vyhledatelné na internetu. Jedním z produktů projektu je počítačová hra, která umožňuje kolaborativní opravy OCR ze starých finských novin. Informace i hra samotná jsou k nalezení zde: <http://www.digitalkoot.fi/en/>

Experience is the teacher of fools: Project management lessons learned

Frederick Zarndt (Global Connexions)

Příspěvek byl zaměřen na řízení vývoje softwarového projektu a projektové řízení obecně. Důraz byl kladen na komunikaci – to, že zadavatel specifikuje požadavek ještě neznamená, že jej dodavatel přesně pochopí a nevyloží si požadavky po svém. To je důležité zejména u psané komunikace – je lepší zadání zopakovat jinými slovy a druhá strana by měla slyšené zopakovat, aby bylo zřejmé, že zadání pochopila. Příspěvek se dále zabýval metodikami řízení softwarového vývoje a srovnával zejména klasické lineární metodiky typu vodopádu a moderní metodiky typu agilního programování. U lineárních metodik vývoj prochází dlouhotrvajícími cykly analýzy požadavků, návrhu systému, implementace, kontroly kvality a nasazení do produkce, což je finančně i časově náročné a může se stát, že systém není nikdy dokončen. Naopak agilní metody jsou levnější díky rozdělení aplikace na menší části, umožňující rychlé úpravy a nasazení drobnějších změn. Agilní metody jsou náročnější na komunikaci zadavatele s dodavatelem, ale přinášejí rychleji výsledky vývoje do produkce. Cyklus agilního vývoje se může opakovat tak dlouho, dokud zadavateli stačí finanční prostředky, je ale menší riziko, že musí být projekt ukončen bez použitelného výsledku.

IMPACT: Centre of Competence in Text Digitisation

Hildelies Balk (Koninklijke Bibliotheek - IMPACT)

Zajímavá byla informace o výsledcích projektu IMPACT, nejsem ale přesvědčen o ekonomické udržitelnosti výsledků projektu navrhovaným způsobem (vytvoření Centre of Competence s nutností platit relativně vysoké členské příspěvky).

Promoting Access to in Copyright Newspapers through Collaboration with the Finnish Copyright Society Kopiaisto

Maria Sorjonen (National Library of Finland) Majlis Bremer-Laamanen (The National Library of Finland)

V příspěvku bylo uvedeno, že finské knihovny mají při zpřístupňování periodik a fotografií (ne knih) možnost využít „rozšířenou kolektivní licenci“, kdy nemusejí jednat s jednotlivými autory, ale

s kolektivním správcem. Toto řešení dělá problém zpřístupnění zvládnutelný. Vývoj práva EU prochází v této oblasti stále vývojem, finské řešení by mohlo být inspirací i pro nás.

JISC Guide - Open bibliographic Data

Owen Stephens

Přednáška zdůraznila význam volné dostupnosti metadat a potenciál, který se otevírá jejich zpřístupněním prostřednictvím standardizovaných mechanismů.

Why We Need to Adopt APIs for Digitised Content

Dunning Alastair (JISC)

Přednáška na příkladech zdůraznila nutnost nejen data a služby zpřístupňovat na webu, ale nabízet k nim API rozhraní, pomocí nichž je možné propojovat různé systémy – v ideálním případě bez omezení přístupu třetích stran.

Upscaling Digitisation at the Wellcome Library, Showcasing the Goobi workflow system

Christy Henshaw (Wellcome Library)

Poměrně zásadní přednáška, popisující přechod od ručního zpracování a malých roztržitých digitalizačních projektů v rámci jedné instituce k systematické masové digitalizaci. Wellcome Library používá pro digital preservation systém SDB firmy Tessela, který si vybrala jako v podstatě první velká knihovna. Je proto zajímavé slyšet že si jako workflow systém zvolili právě Goobi.

Integrating Digitized Material Into an Institutional Repository: The Case Of "SOMNI" and EUROPEANA REGIA at the University of Valencia

Elisa Millás Mascarós (University of Valencia (Spain)) José Manuel Barrueco Cruz (University of Valencia (Spain))

V rámci projektu využívají image server IIPImage s podporou jpeg2000, vyvinutý v MZK, spojený s prohlížečkou METS souborů.

Workshop zaměřený na systémy pro podporu digitalizačního workflow

Různé knihovny používají jako workflow systém cokoli od Excel přes Goobi po DocWorks. Důležité je, o jak velký projekt se jedná. Zatímco několik desítek rukopisů lze ještě řešit ručně, stovky tisíc knih již vyžadují kvalitní nástroj (rakouská NK zjistila, že pro management Google digitalizace nestačí kombinace Excelu a Alephu a hledá lepší nástroj). Byl zmíněn český Registr digitalizace, který je poměrně unikátní.

Workflow systém je nutný zejména pokud na projektu pracuje více lidí. Ideální workflow systém by mělo být možné opakovaně využít ve více projektech různých velikostí. Musí sledovat výkon, podávat statistiky o každém kroku, sledování knih a činností na více lokalitách. Musí umožňovat psát poznámky/tagy a podrobnosti ke každé položce (např. Fyzický stav dokumentu, absence metadat, chyby). Měl by spolupracovat s knihovním systémem.

Workflow systémy by měly navzájem spolupracovat - například přenášet informace centrálním registrům, které by poskytnou celkový pohled na to, co bylo uděláno.

Po skončení digitalizace automaticky vložit dokument do digitální knihovny/LTP systému a odkaz na dokument vložit do knihovního systému.

System by měl být škálovatelný ... a pokud možno volně dostupný.

Návštěva digitalizačního pracoviště Královské knihovny (KB)

V průběhu služební cesty jsem navštívil také digitalizační centrum KB, kde se digitalizují tisky vydané před rokem 1700 na území Nizozemí nebo v holandštině). Digitalizaci provádí firma Diadeis, která má na místě 2-3 pracovníky provádějící vlastní digitalizaci. Mluvil jsem s p. Thierry Puissegur-Ripet (thierry.puissegur@diageis.com), ředitelem projektů.

Před zahájením digitalizace snímají digitálním fotoaparátem hřbet a ořízku knihy – mj. pro zachycení stavu svazku před zdigitalizováním. Vlastní digitalizace probíhá na strojích ATIZ BookDrive Mini – formát knihy do A4 a Zeuschel – formát knihy do A3. Cílové rozlišení je 400 dpi a plná barva. Před započítáním digitalizace se rozhoduje který stroj se pro snímání dané knihy použije. K tomu slouží maketa kolíčky definující úhel a rozměry snímané plochy scanneru ATIZ a plochá deska definující snímací plochu scanneru Zeuschel. Tyto makety se osvědčily v praxi, protože pak nedochází k chybám. Pro naskenování větších dokumentů se používají větší stroje, které jsou na místo přivezeny jen jednou nebo dvakrát za dobu trvání projektu. Provoz je jednosměrný a kopíruje tak otevírací dobu knihovny.

Při skenování bylo zajímavě vyřešeno zachování měřítka při fotografování digitálním fotoaparátem: zároveň se snímanou předlohou se vyfotografuje i pruh bílého papíru se dvěma černými kruhy o průměru cca 1 cm, umístěnými asi 10 cm od sebe. Tento papír je nalepen buď na přitlačném skle, nebo, při snímání hřbetu a vazby na černým sametem potažené krabici sloužící jako podložka knihy. Značky jsou tak vždy umístěny v rovině se snímaným dokumentem. Při dalším automatickém zpracování jsou tyto značky použity pro výpočet dpi a výsledek je vložen do metadat obrázku. Snímání probíhá do TIFFu (scanner) nebo do RAW (u fotoaparátů) a na PC je pak přes DLL Canonu konvertováno do TIFFu a dělá se jen hrubý ořez cca 2 cm zvenčí. Další zpracování pak probíhá v Indickém oceánu (Madagaskar, Mauritius, Reunion), kde má firma cca 500 pracovníků.

Zkušenosti se stojem ATIZ: krycí plexisklo bylo značně poškrábané, mění ho cca každý měsíc. Nekupují originální, ale kupují si čiré plexisklo s antireflexní úpravou (neexistuje antireflexní plexisklo odolné proti poškrábání). Při digitalizaci ale škrábance tolik nevadí, patrné jsou spíše v černém okolí knihy. ATIZ různým způsobem vylepšovali, například vše co šlo uvnitř potáhli černým sametem aby eliminovali odlesky. ATIZ menšího formátu zvolili proto, že do format A4 lze s nepříliš drahými fotoaparáty dosahovat rozlišení 400 dpi a vyšší, u větších formátů už ne. Navíc osvětlení u menšího modelu je kvalitnější (rovnoměrnější) než u modelu Pro.

