

Věc: **Zpráva z pracovní cesty**
Pracovník: **Pavčina Rouhová**
Akce: **Seminář krajské digitalizace**
Datum konání: **6. - 7. 4. 2016**
Místo konání: **Moravská zemská knihovna, Brno**

Stav digitalizace pro Krajské knihovny (KK) k dubnu 2016

- viz. tabulka - <http://goo.gl/ZhGr4p>
- **Jihočeský kraj:**
 - provozují K5 verze 5.1.0
 - vloženo cca. 1,8 milionu stran
 - pokračuje krajsky projekt digitalizace regionálních periodik
- **Plzeňský kraj:**
 - zpracování dat z hybridního snímkování je v rámci krajského projektu IOP ve fázi udržitelnosti
 - K5 s daty z předchozích externích digitalizací (VISK 7, Norské fondy) v rozsahu 366 300 nepřepočtených stran A4-A2 před spuštěním pro veřejnost
- **Karlovarský kraj:**
 - nikdo nedorazil, neví se nic nového
- **Středočeský kraj:**
 - od října 2015 vlastní digitalizační pracoviště
 - digitalizace na vlastní náklady výběrově z fondu, pro instituce v kraji, kompletace regionálních periodik
 - obsluha skeneru necelé 2 úvazky
 - dokončeno nahrávání předchozí externí digitalizace do Krameria (nyní řešení chyb, potíže s indexací)
 - od ledna 2016 přechod na verzi 5.3.1
 - přípravy pro ukládání do krajského úložiště
- **Ústecký kraj:**
 - 31.7.2015 byla úspěšně dokončena druhá fáze naplnění indikátorů projektu „Rozvoj služeb eGovernmentu v Ústeckém kraji“ a bylo zpracováno cca 42 tis. stran monografií
 - dohromady bylo zpracováno v rámci projektu cca 107 000 stran,
 - do teď zpracovávali jen monografie
 - nyní probíhá příprava na zpracovávání periodik
- **Liberecký kraj:**
 - Z KVLI opět nikdo nedorazil
- **Královéhradecký kraj:**
 - dál běží udržitelnost projektu, zajištěná provozem KDJ
 - proběhly reklamace a opravy
 - do Digitální knihovny SVK HK byly naimportovány balíčky z digitalizace službou a starší produkce KDJ
 - novou produkci KDJ zpřístupňují už průběžně
 - nyní v K5 přes 1 mil. stran
 - zajišťují zpřístupnění i dalším knihovnám, jež digitalizovaly v rámci krajského projektu
 - v únoru 2016 se začalo se sklizením DK SVK HK do ČDK

- v současné době mají na oddělení digitalizace 3 celé úvazky (ale nejen na práci na KDJ, na tu je jen jeden)
- digitalizují především fond SVK HK, v případě zájmu i fondy dalších knihoven kraje
- vedle vlastní digitalizace loni začali mj. s ručním propojováním článkových záznamů v Alephu se zdigitalizovanými čísly periodik v DK, zatím propojeno 6 818 článků z celkem 7 titulů
- **Pardubický kraj:**
 - nikdo nedorazil, neví se nic nového
- **Kraj Vysočina:**
 - pracoviště KDJ je v majetku kraje, provozem je pověřena Krajská knihovna Vysočiny v Havlíčkově Brodě
 - od ledna 2016 mají 3 pracovní úvazky (1 kmenový zaměstnanec knihovny + 4 půl úvazky DPČ - rozděleno mezi 6 lidí)
 - pracoviště spolupracuje s institucemi podle jejich zájmu (archivy, knihovny, muzea, obce, NPÚ)
 - digitalizují se monografie (v menším množství), periodika (v současné době se zaměřujeme na kompletaci a doplňování), obecní, školní apod. kroniky (cca 400 ks.)
 - v rámci digitalizace službou zpracovávají negativy, projednávají mapy s VÚGTK
 - v rámci jiného projektu vznikl ve spolupráci s Odborem informatiky z Krajského úřadu a s Muzeem Vysočiny Pelhřimov Portál digitálního kulturního dědictví Kraje Vysočina (www.digivysočina.cz), který má především pro laickou veřejnost zpřístupnit všechny digitalizované objekty z KDJ, z 3D digitalizace a dalších digipracovišť
 - nadstavbou portálu jsou geolokace, kterými každý objekt obohacují
- **Olomoucký kraj:**
 - dokončena replikace z K3, celkový obsah téměř 2,8 mil. stran
 - přidány i mapy z K3 (MEditor)
 - vytvořeny virtuální sbírky (např. Rukopisy)
 - dále probíhá vlastní digitalizace (monografie, periodika)
 - verze Krameria 5.3.2, zprovozněna také klientská verze (<http://kramerus.kr-olomoucky.cz/client/index.vm>)
 - pravděpodobně získají 2 nové úvazky, takže předpokládají vyšší využití digitalizační linky pro digitalizaci fondů VKOL
- **Jihomoravský kraj (MZK):**
 - přijali data některých institucí, mají v Krameriu
 - institucím zpřístupňují dle IP
 - několik dalších institucí řeší zpřístupnění vyrobených dat ve svém lokálním Krameriu
- **Moravskoslezský kraj:**
 - digitalizace vlastními silami na pracovišti v knihovně (necelé dva úvazky)
 - digitalizace vlastního fondu a fondu krajských paměťových institucí (muzea, archivy, knihovny)
 - využíván také program VISK 7
 - na začátku roku 2016 přechod na Krameria 5
 - v současnosti přes 550 000 zdigitalizovaných stran
- **Zlínský kraj:**
 - od konce projektu Digitalizace a ukládání nevznikají nová data
 - všechna data KKFBZ jsou importována do Krameria knihovny (<http://dlib.kfbz.cz>), tč. verze
- **Praha**
 - konec Pralib - naskenováno 250 000 s. - do NKP

Metadatové standardy a RDA / Iveta Lodrová (NKP)

Změny standardů z října 2014:

- nové verze Definic metadatových formátů (DMF) pro digitalizaci periodik a monografií:
<http://www.ndk.cz/standardy-digitalizace/metadata>
- zapracování RDA dodatků do metadatových standardů NDK
- revize pravidel
- plán implementace specifik (DMF) do digitalizační linky NDK do 16. 4. 2016
- zahájena činnost Formátového výboru NDK, což je poradní orgán NDK pro tvorbu metadat, 1. setkání bude na přelomu dubna/května

IDENT: prezentace nástroje / R. Vančo, J. Houserek (VKOL)

- představení IDENTU, nástroje pro vyhledávání, identifikaci, třídění a validaci PSP balíčků
- stačí vybrat složku s balíčky a program ukáže, co v kterém balíčku je (název periodika, datum vydání, ve kterém je to balíčku) - data zjišťuje z xml souborů, které jsou ve složkách
- funkce validace je v této chvíli nedostupná, ale je plánovaná validace podle standardů
- zatím jako freeware -> objeví se na Githubu

Mobilní aplikace K5 a vývoj Krameria / L. Damborská (MZK)

- <https://play.google.com/store/apps/details?id=cz.mzk.kramerus.app>
- připomínky, kód a další info na Github: <https://github.com/moravianlibrary/kramerus-for-android/issues>
- díky existenci API je možné vyvíjet další aplikace, které využívají data Krameria
- aplikace pro Android vyvíjena na půdě MZK
- v roce 2016 jsou v plánu re-designové práce, údržba aplikace, filtry pro výběr dokumentů a přidávání nových knihoven
- pro zapojení nové knihovny je potřeba Kramerus ve verzi min. 5.0.2 (ideálně 5.2.0)
- Kramerus v MZK na tabletech - speciální wifi síť pro autorsky chráněné dokumenty, nelze z ní přistupovat na internet
- tablety v knihovně MZK - 10 ks. - s mobilní aplikací K5 - není to moc využíváno - potřeba nastavit wifi, tak aby nevypadávala a více propagovat, výhody pro uživatele tam jsou
- zkušenosti s tablety v knihovnách ... v jedné knihovně je mají přišroubované ke stolům, používají je jako mediatéku
- v NTK je normálně půjčují
- MZK - v tabletech není připojení k internetu kvůli ochraně autorských práv
- verze mobilní aplikace pro IOS - zatím první veřejná beta verze, více info zde:
<https://itunes.apple.com/cz/app/kramerus/id1065771974?mt=8>
- plán na rok 2016: úprava vyhledávání, designová úprava dle Apple, lepší přehrávání zvukových nahrávek, uživatelské testování atp.
- vývoj Krameria:
 - projekt ČDK byl ukončen, projekt NAKI RightLib nebyl uznán
 - momentálně slíbená podpora z VISK 1
 - podpora vývoje a správa Krameria, vývoj nových funkcí - dvouvrstvé PDF, aktualizace a optimalizace, SOLR, implementace IIIF protokolu, Shibboleth, proof of concept přechodu na Fedoru 4

Alternativní webový klient pro Krameria / P. Rychtářová (MZK)

- www.digitalniknihovna.cz
- oficiální klient v rámci ČDK <http://cdk.lib.cas.cz/client/index.vm>
- je dodáván, požadavky na opravy/rozvoj na Githubu
- je to jednotné uživatelské rozhraní Krameria
- 23 Krameriů na jedné adrese
- je to určeno pro uživatele, ne pro knihovny
- je to jen rozhraní, a jakožto rozhraní Krameria využívá api Krameria, tj. je třeba mít instalovaného klasického Krameria nad kterým rozhraní pracuje
- uživatelské rozhraní je uživatelsky přívětivější než stávající Kramerius - nemá rozhraní pro administrátory
- rozšířená funkcionality oproti K4/K5
- nezávislé na nových verzích Krameria
- dynamický vývoj - otevřeno pro připomínky
- není závislé na penězích z projektu, protože jde o nezávislého vývojáře
- responzivní design
- přepínání mezi jednotlivými knihovnami - jazykové mutace - slovenština, angličtina, němčina

Správa Krameria: bezpečnost Krameria, K4 tools, Meditor, dotazy / M. Rumánek (MZK)

- je důležité si pravidelně aktualizovat Krameria - kvůli bezpečnostním updatům
- je potřeba si hlídat uživatelská práva - aby uživatelé, kteří už v knihovně nepracují, neměli moc práv (zapomenuté účty, uživatelé s vyšším oprávněním, administrátorská práva)
- je to vhodné používat protokol https://, nyní se dá certifikát pořídit i zadarmo
- zálohovat imageserver, nTriples a Fedora3 databáze, celý adresář SOLR, databáze uživatelů Krameria - nejlepší je udělat snapshoty celých virtuálních strojů
- stará pohyblivá zeď je nefunkční, nová pohyblivá zeď se teď testuje
- skripty na hromadné úpravy dat pro MZK - MZK může dát k dispozici:
 - martin.rumanek@mzk.cz
 - Java nebo Groovy, Python se neosvědčil
- neoficiální diskuze ke Krameriovi je na:
 - <https://gitter.im/moravianlibrary/kramerius>
- MEditor - metadatový editor (MZK):
 - <https://github.com/moravianlibrary/MEditor>
 - nové funkce importu
 - plná podpora RDA
- MEditor umí pracovat i s audiodokumenty, to ProArc neumí
- v MKP a VKOL používají oba dva, MEditor pro metadata, ProArc pro PS balíčky

Digital preservation (DP) / L. Cubr (NKP)

- <http://www.digitalpreservation.cz>
- <http://itlib.cvtisr.sk>
- anglické termíny:
 - digital preservativ (DP)
 - long term preservation (LTP)



- české termíny:
 - dlouhodobá ochrana digitálních dokumentů
 - dlouhodobé uchovávání digitálních dokumentů/dat/objektů
- základní normy DP, vyšly i v českém překladu:
 - ISO 14 721:2012 (OAIS)
 - ISO 16 363:2012 (TRAC)
- digital repository – univerzitní repozitáře jsou spíše pro krátkodobé uchovávání
- dlouhodobé uchovávání neznamená navěky
- uchováváme pro cílovou komunitu
- srozumitelnost pro techniku i pro uživatele - porozumění informacím, možnost technologické reprodukce (pro software)
- uchovává se informační obsah + metadata – datový objekt s obsahem (co reprezentuje informační obsah)
- bitová ochrana - ochrana konkrétních datových objektů (dat) s obsahem:
 - zastarávání úložných technologií
 - zálohy, migrace na nové nosiče, hashovací funkce (kontrolní součty)
- logická ochrana - ochrana samotné informace (informačního obsahu):
 - použití vhodných formátů (používat dobře zdokumentované a neproprietární formáty)
 - dostatečně bohatá metadata (technická, popisující historii operací s dokumentem)
 - organizace zajišťující dlouhodobé uchovávání:
 - dostatek zaměstnanců, trvalé finanční zajištění, znalosti (pokračující vzdělávání), technologické zajištění,
- v knihovnách jde hlavně o digitalizáty (obrazové, zvukové dokumenty).
 - informační obsah - vizuálně prohledatelná reprezentace tištěných dokumentů (to, co chceme chránit - řádky, odstavce, barva ...)
 - data = reprezentace informačního obsahu:
 - obrazová data – zobrazitelnost
 - textová data (ocr) – prohledatelnost
- příklad bitové a logické ochrany v NK ČR:
 - data uložena na mgf pásky, pouze zápis, zálohy, on-line i offline kopie (3x kopie, 2x offline), vzdálenost - kopie uloženy daleko od sebe
 - kontrolní součty (MD5)
 - sledování technologií
 - standardizace
 - vhodné formáty (JPEG2000, bezztrátový formát - jeden ze dvou hlavních formátů DP)
 - ALTO XML - textová data (výstup OCR)
- Metadata pro uchovávání:
 - záznam o původu (technická metadata o skenování a formátech – o digitalizaci)
 - popis formátů pro budoucí migrace
 - záznam historie v repozitáři (migrace)
 - využívání mezinárodních (mezinárodně uznávaných) metadatových standardů – jsme vázaní na mezinárodní komunitu - interoperabilita
- formátová migrace – změna původně dodaných souborů
- zachovat signifikantní vlastnosti - stejné rozlišení a barevnost textu
- nezávislost dat na technologiích:
 - samopopisnost - AIP balíček (archivní balíček)
 - globální trvalý identifikátor
 - autenticita

Virtuální národní fonotéka / Filip Šír (NM)

- <https://www.narodnifonoteka.cz/>
- <projekt.narodnifonoteka.cz>
- proč VNF nyní spadá pod NM? ... prý se nemáme ptát :-)
- cca 500 000 záznamů
- problémem je nepracovanost fondu v knihovnách i dalších institucích
- projekt VISK 7 na rekatalogizaci zvukových dokumentů v knihovnách
- standard pro zvukové dokumenty – měl by být do 4 měsíců, je to podmínka VISK 7
- v plánu je žádost do VISK na záchrannou digitalizaci audio válečků, šelakových desek a další ohrožených nosičů
- vznik pracovní skupiny “Pracovní skupina pro ochranu, digitalizaci a zpřístupňování zvukových dokumentů v ČR”
- MEditor - umí zvukové dokumenty
- zvukové dokumenty je potřeba dodat uživatelům v hezkém balení
- zvyšuje se návštěvnost webu VNF
- VNF nemá vlastního ajťáka - při konferencích v Paříži a v Kapském městě jim VNF vypadla :-). Jen doufají, že letos v Kongresové knihovně se jim to už nestane

LTP-pilot + Arclib a jak to s nimi vypadá / Zdeněk Hruška (MZK)

- projekt LTP-Pilot:
 - <http://ltp-portal.mzk.cz/ltp-pilot>
 - byl ukončen v říjnu 2015, testování opensource systému pro dlouhodobou ochranu digitálních dat Archivematica (www.archivematica.org)
 - zmíněno několik základních vlastností Archivematicy (např. je zdarma, ale náklady na hardware, administrátora, atp.)
 - textové výstupy projektu: Archivematica - instalace, systémová správy, výkonnostní testy, napojení na Cesnet a DSpace
 - zmapován stav LTP v ČR, úvod do OAIS ...
- projekt ARCLib:
 - <http://ltp-portal.mzk.cz/home/Novinky/arclib-ceskyopen-sourceltpprojekhttps://www.muni.cz/research/projects/33644>
 - cíle: vybudovat komplexní opensource LTP systém, metodika pro bit-level ochranu dat, metodika pro logickou ochranu dat aj.projekt poběží v letech 2016-2020, podílí se KNAV ČR, NK ČR, MZK a ÚVT MU
- otázka: Jak zajistit neměnnost (kompatibilitu) obsahu, když lze měnit utility? Jedná se o certifikovaný systém?
- odpověď: Zřejmě metadata, systém je teprve v začátcích ...

Strategie digitalizace novodobých dokumentů / A. Vandasová (NKP)

- cíle digitalizace:
 - destruktivní digitalizace - dary (vyřazované knihy) od knihoven, firem, organizací či soukromých osob, rezervní fondy, každá 6. kopie

- vzácné dokumenty - pouze jeden skleněný velký manuální skener, jehož kapacita je beznadějně naplněna na půl roku dopředu
- je plánována koupě dalšího ručního skeneru
- omezení fyzických výpůjček - frekventované dokumenty, u kterých je vysoká míra opotřebení
- digitalizace nových titulů
- kompletace neúplných periodik a vícesvazkových monografií (snaha zdigitalizovat je celé)
- v průběhu digitalizace dochází ke změnám v záznamech dokumentů
- zlepšení služeb čtenářům:
 - nedostačující stav technického vybavení v budovách NK
 - problém s terminály - nevyhovující uživatelské prostředí
 - bude nutné realizovat průzkum uživatelských potřeb
 - cíl – donutit uživatele využívat elektronické dokumenty
- do budoucna vytvoření obecnější strategie

Strategie digitalizace a její vývoj v MZK, NK ČR, obecně / M. Indrák (MZK)

- v MZK strategie na 2013-2020
- strategie digitalizace musí zohlednit:
 - ochranná digitalizace dokumentů, kterým hrozí fyzický zánik
 - unikátní a vzácné dokumenty
 - kritéria - dle zvyklostí zahraničních institucí - ochranná digitalizace, unikátní dokumenty, požadavky institucí a uživatelů, na konci destruktivní digitalizace
- otázky:
 - jaké knihy digitalizovat, jak je vybírat, kdo to má rozhodovat a dle čeho? A co knihy pod autorskou ochranou?
 - stojí destrukce za to? (velké množství dokumentů, jejichž informační i kulturní hodnota není příliš velká)
 - kdyby bylo zajištěno financování do budoucna, mohlo by se digitalizovat vše. Ale to není, tak se vybírají prioritní dokumenty
 - bylo by třeba mít kurátora, který bude rozhodovat o tom, co se bude skenovat (s ohledem na legislativu a jiné podmínky)
- odpovědi: Tomáš Foltýn (NKP)
 - to je sporné, bohemikální produkce je bohemikální produkce, kdo posuzuje hodnotu knížek? My jako knihovny máme uchovávat a digitalizovat bohemikální fond
 - NKP doporučuje nedigitalizovat knihy pod autorskou ochranou
- omezení digitalizace - rozložení skenerů na lince (na každý skener se hodí jiná kniha, co snesou skenery, co mohou skenovat a co snesou knihy), spolupráce v budově, stavění fondu
- stále více se volá po odborných kurátorech – zefektivnění digitalizace
- diskuze:
 - Ivan Ljubka (NKP) - spolupráce s KK je téměř nulová, dochází k duplicitní digitalizaci (předcházení duplicitám – implementace strategie – spolupráce knihoven)
 - T. Foltýn - není to zas tak špatné, duplicitní digitalizace je podle něj jen tak 5 % - na to, jaká masová digitalizace probíhá v ČR, je to ok
- skenery – variabilní ceny (NDK nemá peníze) - výrobce cenu jen tak neřeknou, odvíjí se od známostí a konkurenční nabídky

- kvalita skeneru - záleží, co chce knihovna zpracovávat, např. velké mapy se skenují po částech a jdou těžko spojovat (musí být dobře nastavená geometrická kalibrace, v MZK se části spojují ve Photoshopu)
- NKP – zastaralá technika, málo peněz, málo lidí (v NKP na rukopisech PC z roku 19 bambus ☺)
- V NKP mají problém s digitalizací starších (cenných) dokumentů – správci fondů chrání svůj fond a nechťejí nechat dokumenty z 19. století otočit (dotýkat se jich) 50x
- jak v ostatních zemích řeší autorská práva – někde pod licenci, něco volně dostupné dle veřejného zájmu, někde konsorcia – u nás je to problém (zákony, legislativa obecně)
- černý trh s digitálními daty (např. kolejni servery)
- tento systém u nás je dlouhodobě neudržitelný, uvědomuje si to Dilia? Její pracovníci tlačí na vládu – knihovny budou muset vyvinout větší tlak – ale možná větší vyvine sama veřejnost

NDK: současný stav a probíhající fáze udržitelnosti / M. Indrák (MZK)

- projekt NK ČR a MZK
- udržitelnost NDK není nijak kvalitativně ani kvantitativně stanovena
- fáze udržitelnosti od roku 2015
- v roce 2015 byly výkony v MZK různé a odvíjely se od funkčnosti linky/skenerů a dalších faktorů
- NDK je postavena na výkonnosti, nyní se snaží dostat ze stavu výkonnostního do stavu výběrového
- NDK naskenováno letos 900 000, do konce roku by měli mít 3 000 000.
- problém s technickou podporou během udržitelnosti - moc nebyla, opravy skenerů si knihovny loni dělaly víceméně svépomocí, v MZK nemají peníze ani čas na opravu od dodavatele (p. Popelka, zaměstnanec MZK je prý velmi šikovný☺)
- někdy, než nechat opravit skener, je levnější koupit nový

Problematika duplicitního zpracování / I. Ljubka (NKP)

- důležité je, aby záznamy přicházející do digitalizace měly ČČNB - na digitalizaci jdou dva záznamy trochu odlišného názvu, které nemají ČČNB (nebo jeden z nich), poté se to opraví a zjistí se, že se jedná o stejný dokument, a je tu duplicita
- problém s replikací mezi různými visky, knihovnám v podstatě nezbývá, než dokument digitalizovat znovu (L. Jirků - poznámka o titulech (např. regionální literatura), které nemohou replikovat kvůli financování z jiného projektu)
- bylo by vhodné upravit zadávací dokumentaci programu VISK tak, aby data digitalizací v rámci VISK7 nešla povinně jen do NK, ale i do MZK
- NKP a MZK v tom počtu, co produkují (digitalizují), mají větší počet duplicit, nežli jiné knihovny
- registr digitalizace – kontrolovat záznamy
- problémy se starými verzemi Krameria – při replikaci
- NKP – potřebuje, aby dokumenty, které se mají digitalizovat, měly čárový kód (např. knihy z 19. století jej nemají), když dokument nemá čárový kód, nedigitalizují jej
- Elsys – dokumentům, které nemají čárový kód, vygenerují vlastní



Diskuze (otázky a odpovědi)

- diskuze s Mgr. Tomášem Foltýnem, NKP, odborným garantem pro Visk 7:
 - Jak jsou na tom s digitalizací ostatní kraje?
 - Kolik lidí mají na digitalizaci?
 - Budou kraje podávat projekty?
 - Přispívání do Europeany:
 - Europeana (potažmo Evropská knihovna) zvýšila finanční nároky za jednotlivé přispěvatele (od NDK žádali cca 3700 €, od Krajských knihoven cca 1600-1700 € ...)
 - nově se vyjednává o lepších podmínkách, ale výhled je poměrně skeptický (z minulých zkušeností)
 - vznikla pracovní skupina ostatních evropských států (ČR, Polsko ...), která se to snaží řešit ... agregátorem za ČR je NDK
- v rámci IROP je digitalizace coby součást nějakého konkrétního projektu ...
- Sdružení IT: <http://sdruk.mlp.cz/sdruk/odborne-sekce/sekce-pro-informacni-technologie/clanek/sekce-pro-informacni-technologie/>
- projekt ČDK skončil, projekt NAKI "RightLib - elektronické zpřístupnění chráněných publikací" nebyl uznán
- Visk 1 - bude druhé kolo?