



Česká digitální knihovna – agregace digitálního obsahu českých knihoven

Martin Lhoták

Knihovna AV ČR, v. v. i.

Národní agregátor ve světě eCulture, Praha, 14. 7. 2015

- Výzkumný projekt financovaný z programu MK ČR NAKI
- 2012-2015
- 20 mil. Kč
- Spolupráce Knihovny AV ČR, Národní knihovny ČR a dalších partnerů
- Hlavní cíl – **Česká digitální knihovna**
 - Agregace a zpřístupnění obsahu digitálních knihoven z území České republiky

<http://www.czechdigitallibrary.cz>



| Kramerius

- systém pro zpřístupnění digitálních dokumentů



| RDflow

- systém pro sledování digitalizačního workflow



| ProArc

- produkční a archivační systém

<http://www.czechdigitallibrary.cz>

- Zastřešení českých digitálních knihoven
- Sklizení metadat
- Jedno vyhledávání nad všemi digitalizovanými daty
- Jedno rozhraní pro doručení plných textů
- Dynamické nahrávání plných textů
- Nejedná se pouze o index
- Agregátor pro projekt Europeana
- Výstupy projektu volně dostupné pod licencí GNU GPL

Kramérius

- Open source řešení pro digitální knihovnu
- Vývojový tým – KNAV, NK, MZK, NLK, NTK + Incad
- Výchozí systém pro Českou digitální knihovnu
- Podporuje standardy NDK
- Více než 25 knihoven provozuje Kramérius 4/5
- Celkem více než 30 miliónů stran
- Vývoj od roku 2003 (verze K4 od r. 2009, K5 od 2014)
- Využívá Fedora Commons repository

Kramerius

- **Nová verze Kramerius 5**
- Uživatelské rozhraní je odděleno od administrátorského
- Nové API, které umožní snadné připojení dalších nadstaveb (MZK – aplikace pro Android)
- Orientace na využití v mobilní zařízení
- Jednoduché a intuitivní uživatelské rozhraní, které umožní snadné nalezení a čtení
- Maximální využití obrazovky pro digitální dokument
- SOLR4, HTML5, responsive design
- Uživatelské rozhraní ČDK bude vycházet z K5

K5 Procházet Casová osa Virtuální sbírky

zeptej se mne

Knihovna Václava Matěje Krameria

Digitální knihovna Václava Matěje Krameria obsahuje dokumenty, které byly digitalizovány v rámci různých projektů - VISK6, VISK7, EuropeanaTravel a další.

Václav Matěj Kramerius (9. února 1753 Klatovy – 22. března 1808 Praha) byl český spisovatel, nakladatel a novinář. Pokřtěn byl jako Matěj Valentin, jméno Václav přijal později jako vlastenecké. Sám své příjmení psal většinou bratrským pravopisem jako Kraméryus.

Krameriem založená Česká expedice byla prvním ryze českým nakladatelstvím v Českých zemích. Vyšla v něm většina českých knih své doby, a lze proto tvrdit, že založení tohoto nakladatelství bylo jedním z podstatných kroků pro úspěšné zavedení českého knižního trhu v letech 1800–1850.

vybrané
nejnovější
nejžádanější

<http://cdk-test.lib.cas.cz/client/index.vm>

Výsledky vyhledávání

K5 Procházet Časová osa Virtuální sbírky

440054 dokumentů

listy

Použité filtry Českokrumlovské Jihočeské listy 1994-2006 Periodikum

Klíčová slova

- numismatika (110)
- Boskovice (Česko) (86)
- regionální týdeníky inzerce (52)
- Literatura (17)
- Československo (14)
- Umění (10)

Dostupnost

- public (257697)
- private (182355)
- (2)

Virtuální sbírky

- Národní listy 1861-1941 Periodikum (261717)
- MZK-ostry (126437)
- NDK-Test (46665)
- KNAV (12506)
- MZK Demo (4472)
- NTK

Typ dokumentu

- Stránka (434189)
- Výtisk (5737)
- Monografie (5)

Entomologické listy 1938-1951 Periodikum

Listy Strakonicka Listy Strakonicka: Jihočeské listy 1993-2006 Periodikum

Listy Písecka Listy Písecka: Jihočeské listy 1994-2006 Periodikum

Listy Pelhřimovská Listy Pelhřimovská: Jihočeské listy 1994-2002 Periodikum

Chemické listy Chemické listy 2011 Periodikum

Listy filologické 1874- Periodikum

Zoologické listy 1956-1976 Periodikum

Hudební listy 1884-1888 Periodikum

Numismatické listy Česká numismatická společnost 1945- Periodikum

Numismatické listy 1945- Periodikum

Literární listy 1888 - 1904 Periodikum

Humoristické listy Periodikum

Ženské listy 1873 Periodikum

Listy chemické 1877 - 1881 Periodikum

Numismatické listy Česká numismatická společnost ko2002113080 1945- Periodikum

Italské listy Čapek, Karel 1947 Monografie

Listy hetér 1970 Monografie

Listy Nikodémovy Dobraczyfski, Jan 1989 Monografie

Listy filologické a paedagogické Periodikum

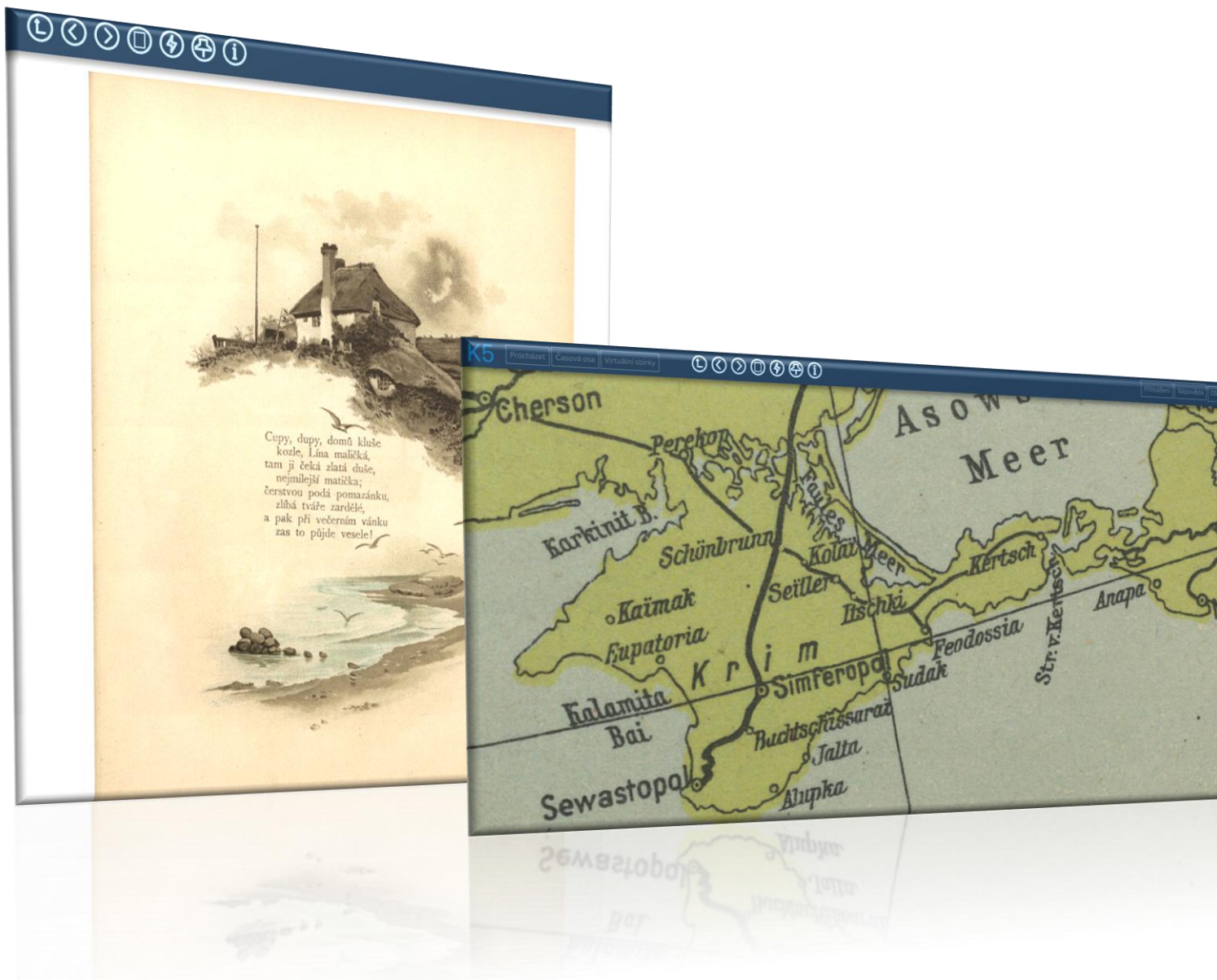
Náhled na dokument



Struktura dokumentu



Plné zobrazení dokumentu



Sbírky

The screenshot displays the CDK search interface. At the top left, the 'Časová osa' (Time axis) section shows search criteria: 'Od: 01.01.1390' and 'Do: 31.12.2014'. Below this is a vertical timeline with decade intervals from 1880-1890 to 2010-2020. The right side of the interface features a search bar with 'Pokročilé vyhledávání' and a navigation menu with options like 'Vybrané', 'Nejžadanější', 'Nejnovější', 'Navigace', 'Autoři', 'Tituly', 'Informace', and 'Zdroje'. The main content area is divided into two sections: 'Sources' and 'Collections'. The 'Sources' section lists: 'Národní digitální knihovna', 'Národní technická knihovna', 'Moravská zemská knihovna', and 'Knihovna Akademie věd ČR'. The 'Collections' section lists: 'První světová válka', 'Sborníky ELAG', 'Knihy Academia', 'Mapy', 'Historické dokumenty', 'Rajhrad', 'Rok smíření', 'První světová válka', 'Regionální muzeum v Mikulově', 'Řemesla a dovednosti', 'Vinařství', 'Komeniana', 'Muzeum Brněnska', 'Povinná četba', 'Zednářská literatura', 'EOD', 'Academia', 'ANL', and 'Atlasy'.

I Kramerius

- **Větší zapojení uživatelů**
- Zpětná vazba od knihoven
- Spolupráce s pracovníky ve službách
- UX – user experience
- Inspirace v dalších projektech (např. polona.pl)

- Nové možnosti dané úpravou legislativy – díla nedostupná na trhu, osiřelá díla, kolektivní smlouvy, e-povinný výtisk

- **System pro sledování digitalizačního workflow**
- Vychází z **Registru digitalizace CZ**
- Propojení s produkčním systémem **ProArc** umožní automatizované zaznamenávání jednotlivých digitalizačních kroků
- Automatická aktualizace dat v Registru digitalizace
- Provedeny analýzy a návrhy obrazovek
- Funkční verze plánována na podzim 2015

- **Produkční a archivační systém**
 - **Produkce digitálních dokumentů**
 - Dodržování standardů NK ČR - kompletní **PSP balíček NDK** (METS, MODS, MIX, PREMIS, ALTO, Dublin Core)
 - **Přebírání dat** z externích systémů (Aleph, Registrdigitalizace.cz)
 - Kompatibilita se systémem **Kramerius**
 - Podpora identifikátorů **URN:NBN, UUID**
 - **Dlouhodobá archivace digitálních dokumentů – LTP** (ve vývoji)
 - Dodržování **standardů pro LTP** - OAIS, ISO
 - Propojení s **Archivematicou**
- Dlouhodobě využíván v KNAV a SVKHK + ???
- Open source založený na Fedora Commons repository

Vytváření popisných metadat

Menu ▾ Česky ▾ nezbedova ProArc

Import - Příprava metadat pro načtení do úložiště

| Název souboru | Číslo | Typ |
|---------------------------------|-------|---------------|
| ABA007121085022009000100001.tif | [1] | Normální s... |
| ABA007121085022009000100002.tif | [2] | Prázdná st... |
| ABA007121085022009000100003.tif | 3 | Normální s... |
| Hřbet | 4 | Prázdná st... |
| Mapa | 5 | Normální s... |
| Normální strana | 6 | Normální s... |
| Obálka | 7 | Normální s... |
| Obsah | 8 | Normální s... |
| Prázdná strana | 9 | Normální s... |
| Přebal | 10 | Normální s... |
| Přední desky | 11 | Normální s... |
| Přední přidešti | 12 | Normální s... |
| Rejstřík | 13 | Normální s... |
| Reklama | | |
| Reprezentativní obálka | | |
| Seznam ilustrací | | |
| Seznam map | | |
| Seznam tabulek | | |
| Tabulka | | |
| Titulní strana | | |
| Volný list | | |
| Zadní desky | | |
| Normální strana | | |


Index strany :

Číslo strany :


Identifikátor

Typ : Hodnota : 188bbee3-b225-40b


Poznámka :




Normální strana: [1]
Index: 1




Prázdná strana: [2]
Index: 2




Normální strana: 3
Index: 3




Prázdná strana: 4
Index: 4




Normální strana: 5
Index: 5




Normální strana: 6
Index: 6



Normální strana: 7
Index: 7

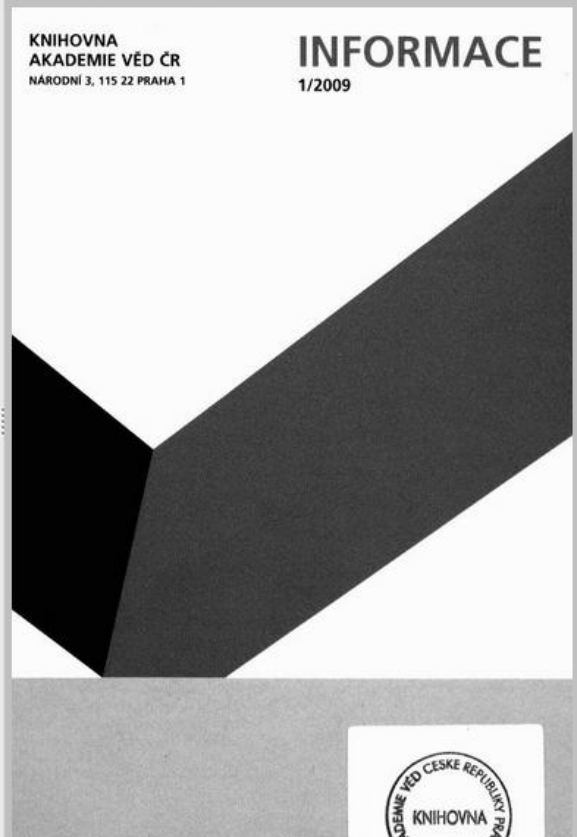


Normální strana: 8
Index: 8



Normální strana: 9
Index: 9

Celý



Možnosti exportu

Menu ▾ Český ▾ nezbedova ProArc

Správa digitálních objektů

Naposledy vytvořené
 Naposledy editované
 Dotaz
 Rozšířený dotaz

Model: ▾

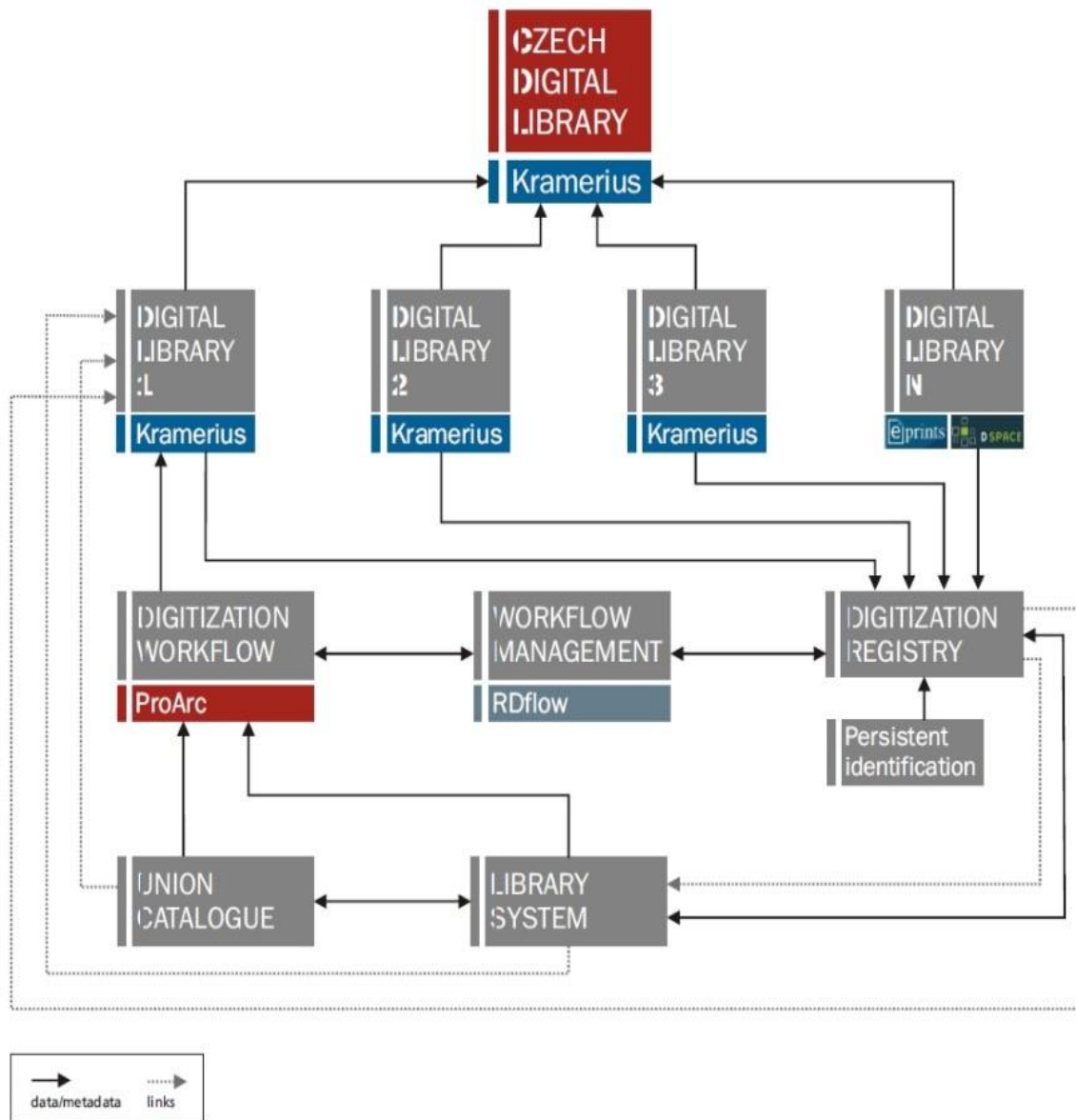
Filtr
 Obnovit
 Metadata
 Komentář
 Nadřazený objekt
 Vazby
 ATM
 FOXML
 Exporty ▾
 Smazat
 URN:NBN

| Popis | Model | PID | Vytvořen | Změněn | Vlastník | Export |
|---|-----------------------|---------------------|------------------|------------------|-----------|--------|
| Sociální svět středověkého města | NDK Svazek monografie | uuid:ddc4f5d3-77... | 24.10.2014 13:09 | 24.10.2014 13:11 | nezbedova | Ne |
| Archiv český Čili staré písemné památky české i ... | NDK Svazek monografie | uuid:f3616cb4-53... | 14.10.2014 16:21 | 14.10.2014 16:42 | nezbedova | Ano |
| Regionální elity 2004 | NDK Svazek monografie | uuid:bf4e3476-c1... | 02.10.2014 15:10 | 02.10.2014 15:12 | nezbedova | Ne |
| Etnografie národního obrození, 2 | NDK Svazek monografie | uuid:7fd0e017-1a... | 09.09.2014 12:33 | 09.09.2014 12:48 | kolarova | Ano |
| Etnografie národního obrození, 1 | NDK Svazek monografie | uuid:5aec4373-f7... | 09.09.2014 10:30 | 09.09.2014 10:33 | kolarova | Ne |

Obnovit
 Metadata
 Komentář
 Nadřazený objekt
 Vazby
 ATM
 FOXML
 Exporty ▾
 Smazat
 URN:NBN

| Popis | Model | PID | Vytvořen | Změněn | Vlastník | Export |
|----------------------------------|-----------------------|---------------------|------------------|------------------|-----------|--------|
| Sociální svět středověkého města | NDK Svazek monografie | uuid:ddc4f5d3-77... | 24.10.2014 13:09 | 24.10.2014 13:11 | nezbedova | Ne |

- Export pro Kramerius 4
- NDK PSP Export
- Export skenů
- Export původních skenů



Co je třeba řešit?

- **Deduplikace** sklízených dat
- Pravidelná **aktualizace**
- **Ochrana OCR plných textů**
- **Kompletní replikace** dokumentů
- Zajištění **stabilního prolinkování** na primární dokumenty
- Propojení ČDK na **Centrální portál knihoven**
- Sklizení **z jiných než K4/K5** dig. knihoven (min. standard = set pro Europeanu?)
- Sdílení dat s mezinárodními projekty – rozlišení **volně dostupného obsahu** pro sdílení přes OAI-PMH

- **Testovací provoz ČDK v Knihovně AV ČR**
 - 25 mil. stran
 - KNAV, NK, MZK, NTK
- **Přesun do Národní knihovny a zahájení ostrého provozu**
 - Nutno zajistit finanční zdroje pro stabilní provoz (institucionální + účelové)
 - Zajištění oficiálního statusu sektorového agregátora
 - Smlouvy s knihovnami o spolupráci
 - Poskytnutí metadat, náhledů a OCR textů
 - Zajištění trvalých odkazů
 - Předávání dat do TEL a Europeany
 - Předávání dat národnímu agregátorovi



Děkuji za pozornost!

Ing. Martin Lhoták

Lhotak@knav.cz

<http://www.knav.cz>