

MASARYKOVA UNIVERZITA  
FAKULTA INFORMATIKY



## Česká digitální knihovna

PV070

Mário Vereš – 422767

3. ročník, obor Umělá inteligence a zpracování přirozeného jazyka

4.12.2016

## Obsah

1. Úvod.....	3
2. O projekte .....	3
3. Digitálna knižnica KRAMERIUS.....	6
4. Register digitalizácie .....	6
5. Proarc – produkčno archivačný systém .....	7
6. RDflow .....	8
7. Štandardy digitalizácie .....	8
8. Záver.....	9
9. Zdroje.....	10

Oficiálna stránka Českej digitálnej knižnice: <http://www.czechdigitallibrary.cz/>

# 1. Úvod

Česká digitální knižnice (ČDK) je národným agregátom digitálnych knižníc prevádzkovaných v Českej republike. Umožňuje vyhľadávanie v dokumentoch, obsiahnutých v digitálnych knižniciach prevádzkovaných na území Českej republiky. Hlavným cieľom ČDK je zaistenie prístupu k digitálnym dokumentom v knižniciach z jedného miesta. Ďalším cieľom je poskytnutie informácií o digitálnych dokumentoch do ďalších zastrešujúcich projektov, hlavne do Centrálného portálu knižníc KNIHOVNY.CZ a do projektu Europeana.

Projekt ČDK bol financovaný a podporovaný Ministerstvom kultúry Českej republiky v rámci Programu aplikovaného výskumu a vývoja národnej a kultúrnej identity (NAKI) v rokoch 2012-2015. Administrátorom ČDK je Knižnica Akadémie vied v Českej republike. Na riešení projektu ČDK sa podieľali Knižnica AV ČR, v. v. i, Národná knižnica ČR a spoločnosť INCAD, s. r. o. Na niektorých ďalších častiach projektu (napríklad vývoj systému Kramerius) spolupracovali aj zamestnanci ďalších knižníc: Moravská zemská knižnica v Brne, Národná technická knižnica a Národná lekárska knižnica.

## 2. O projekte

Hlavným cieľom projektu Českej digitálnej knižnice (ČDK) bolo vytvoriť jednotné rozhranie pre koncového užívateľa ale aj ako primárny poskytovateľ dát pre zastrešujúce národné ale aj medzinárodné projekty. ČDK je významným zdrojom digitálnych dát a je hlavným pilierom pre zaistenie centralizovaných služieb v Českej republike definovanej v koncepcii rozvoja knižníc ČR na roky 2011 – 2014.

Východiskovým softwarovým riešením pre ČDK mal byť Kramerius 4, jeho jadrom je repozitár Fedora. Momentálne to je Kramerius s označením K5, čo je nová generácia open source systému pre sprístupnenie digitálnych dokumentov, ktorá vychádza z verzie Kramerius 4.

Ako prepojujúci prvok a zdroj informácií o digitálnych dokumentoch, pôvodnom tlačovom zdroji, vlastníkovi, identifikátoroch a ďalších súvisiacich údajoch slúži Register digitalizácie. V minulosti Register digitalizácie bol založený na komerčných nástrojoch a bolo ho problematické nasadiť v lokálnych inštitúciách tak, aby mohli bez akýchkoľvek závislostí na iných inštitúciách sledovať a evidovať vlastný digitalizovaný workflow. Nové riešenie je založené na voľne dostupných softwarových komponentoch, umožňuje prevádzku samostatných inštancií pre evidenciu čiastkových krokov digitalizovaného workflowu priamo na jednotlivých digitalizovaných pracoviskách.

Pri riešení projektu bolo prepojených viacero knižných informačných systémov (knižné katalógy, digitálne knižnice, register) a bol braný veľký dôraz na využitie perzistentných identifikátorov. Sú to hlavne identifikátory: URN:NBN – identifikátory digitálnych dokumentov pridelené Národnou knižnicou ČR, ČNB – identifikátor Českej národnej bibliografie, RDCZ – identifikátor Registrov digitalizácie a UUID – identifikátor jednotlivých častí digitálnych objektov.

Komunikácia užívateľa vrátane zobrazenia vybraného digitálneho dokumentu prebieha cez jednotné rozhranie ČDK. Existujú dve možnosti uloženia primárnych dokumentov, vo väčšine prípadov zostávajú primárne digitálne dokumenty v pôvodných digitálnych knižniciach a v ČDK sú zobrazené na vyžiadanie stiahnutím pomocou odkazu, druhou možnosťou je prenesenie digitálnych dokumentov na lokálne úložisko ČDK. Druhý spôsob sa využíva len v prípadoch, kde nie je možné zaistiť trvalé odkazy na pôvodný zdroj a pokiaľ to je v súlade s dodržaním práv k dokumentom.

Pre potreby digitalizovaných pracovísk bol vyvinutý voľne dostupný systém pre dlhodobú archiváciu založený na medzinárodne uznávaných štandardoch (Proarc). Obdobné riešenie v českom prostredí úplne chýbalo. Riešenie tiež slúži zároveň ako systém pre vytváranie a editáciu metadát a správy digitálnych dokumentov.

V súčasnej dobe sú v programe ČDK k dispozícii dáta z Národnej digitálnej knižnice, Moravsko zemskej knižnice v Brne, Knižnice Akadémie vied ČR, Národnej technickej knižnice a Študijná a vedecká knižnica v Hradci Králové. Ďalšie knižnice sú postupne pridávané a začleňované.

### Česká digitální knižnica - webový portál:

ČESKÁ DIGITÁLNÍ KNIHOVNA

O ČDK Pro knihovny CZ

Vstoupit do knihovny

ČDK – národní agregátor digitálních knihoven

Vyhledávání

Jen veřejně dostupné | Nápověda

Česká digitální knihovna umožňuje vyhledávání v dokumentech obsažených v digitálních knihovnách provozovaných jednotlivými knihovnami v České republice. Cílem je zajištění přístupu k digitálním dokumentům v knihovnách z jednoho místa.

MONOGRAFIE (107 270) PERIODIKUM (1 664) RUKOPIS (425) MAPA (1 058) HUDEBNINA (10) GRAFIKA (15)

Bremen, Johann Gottfried  
Die symbolische Weisheit der

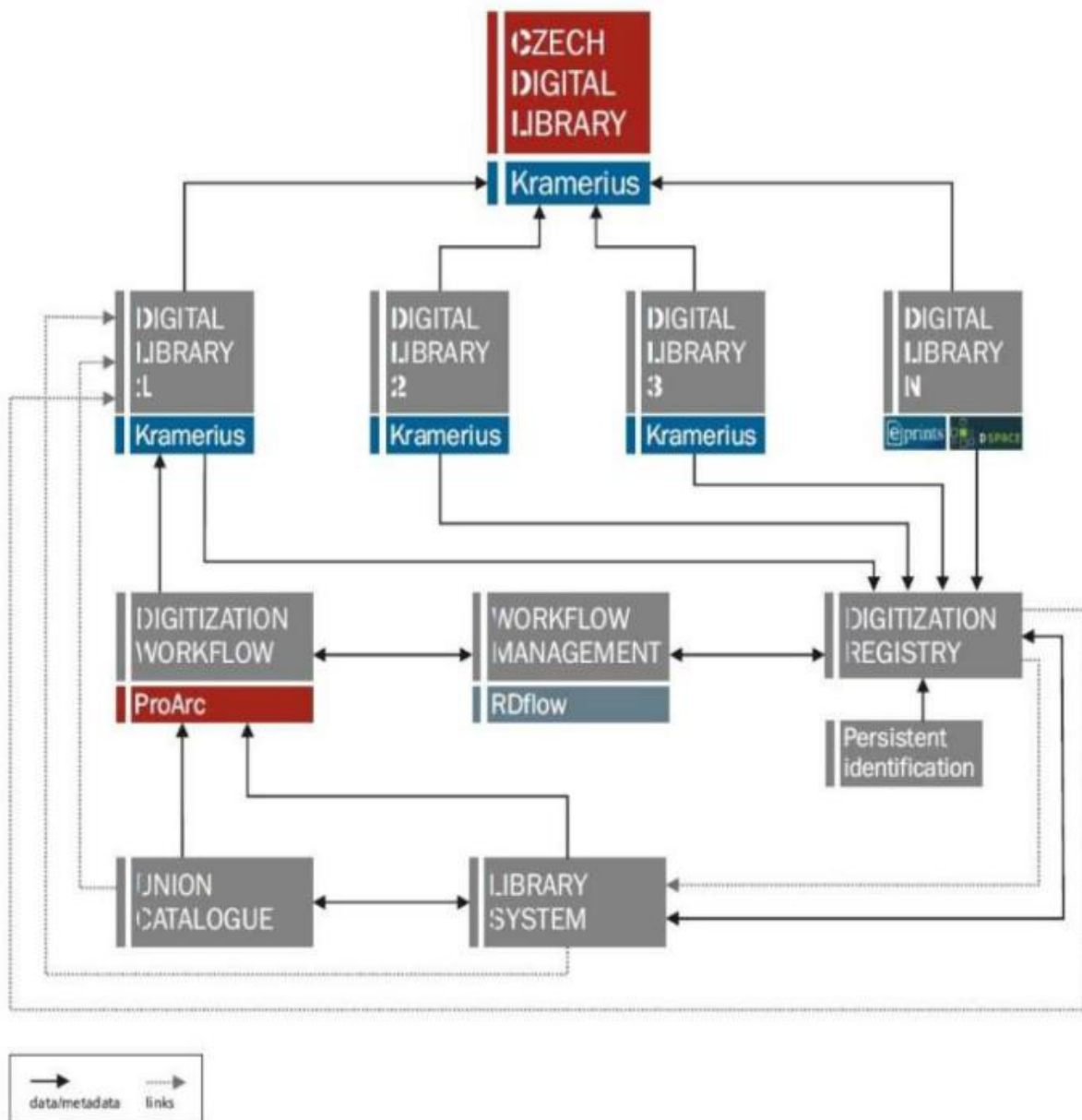
Mahen, Jiří,  
Mrtvé moře: hra o čtyřech

Kosmák, Václav,  
Potkalo ji štěstí: obraz ze

Každý uživatel má možnost registrácie a po zaregistrovaní má možnosť si ukladať do svojho konta záznamy obľúbených dokumentov. Registráciou užívateľ nezíska prístup

k neverejným dokumentom. Na ČDK sa nachádzajú dva druhy dokumentov: verejné a neverejné. Verejné dokumenty sa zobrazujú v celom rozsahu, pretože už nie sú chránené autorským zákonom, čo znamená že od smrti autora prešlo viac ako 70 rokov, alebo prešlo 50 rokov od vydania. Neverejné dokumenty sú chránené autorským zákonom, nie je možné ich verejné sprístupnenie. Pre zobrazenie dokumentov označených ako neverejné, je potrebné navštíviť budovu knižnice, ktorá požadovaný dokument vlastní.

**Schéma toku dát :**



### 3. Digitálna knižnica KRAMERIUS

System Kramerius je voľne dostupné softvérové riešenie, ktoré je v Českej republike využívané. Služi ako hlavný prístupový bod k digitálnym dokumentom pochádzajúcich z digitalizácie. Je určený predovšetkým pre sprístupnenie digitalizovaných knižných fondov, najmä monografií a periodík. Možné je aj sprístupnenie ďalších typov dokumentov, napr. mapy, hudobné médiá a staré tlače, alebo časti dokumentov, ako sú články a kapitoly. Tento systém je vhodný aj pre tzv. "Digital born" dokumenty, teda dokumenty, ktoré vznikli v elektronickej podobe. Kramerius je priebežne vyvíjaný tak, aby zodpovedal metadátovým štandardom Národnej knižnice Českej republiky. Systém poskytuje rozhranie pre koncových užívateľov, ktoré umožňuje vyhľadávanie v metadátach aj v plných textoch, generovanie viacstránkových PDF dokumentov z vybraných stránok, vytváranie virtuálnych zbierok a ďalších operácií nad uloženými zbierkami digitálnych dokumentov. Nadväzuje na funkčnosť predchádzajúcej verzie systému Kramerius, ktorá končila označením 3.3.1. Od verzie 4 je ako úložisko využívaný open source repozitár Fedora. Ďalej sú využívané voľne dostupné technológie tretích strán - napr. Apache, Apache Solr a Postgres SQL. Systém je založený na programovaní jazyku Java a môže byť spustený ako webová aplikácia. Viac ako 30 knižníc v Českej republike (väčšina z nich patrí medzi najväčšie v ČR) používa systém Kramerius na prevádzkovanie svojej vlastnej digitálnej knižnice. Všetky inštalácie dohromady obsahujú v súčasnosti viac ako 90 miliónov strán.

Open source systém Kramerius je východiskovým softvérovým riešením pre Českú digitálnu knižnicu. Zaisťuje interoperabilitu s najvýznamnejšími digitálnymi knižnicami v Českej republike. Česká digitálna knižnica bude slúžiť aj ako národný poskytovateľ dát pre Europeanu a pre Centrálny portál českých knižníc - KNIHOVNY.CZ.

### 4. Register digitalizácie

Register digitalizácie bol založený ako projekt Knižnice Akadémie vied a Národnej knižnice Českej republiky. Z hľadiska Českej digitálnej knižnice sa jedná o spolupracujúci prepojovací systém a relevantný zdroj informácií. Obsahuje veľké množstvo informácií o digitalizovaných dokumentoch v Českej republike, vrátane identifikácie pôvodných tlačených dokumentov, vlastníka a podľa pôvodnej digitálnej knižnice, kde je k dispozícii digitálny dokument. K dôležitým údajom patrí tiež stála identifikácia napr. číslo Českej národnej bibliografie a ďalšie relevantné záznamy. Hlavným cieľom národného registra digitalizovaných dokumentov je zabrániť nežiadúcim duplicitám pri digitalizácii a umožniť zdieľanie výsledkov digitalizácie po celej Českej republike, čo vedie k lepšiemu využitiu finančných prostriedkov. V súčasnej dobe Register digitalizácie obsahuje takmer 352 000 titulov kníh, novín, zborníkov a ďalších dokumentov. Register digitalizácie poskytuje zároveň riešenie, ktoré monitoruje digitalizovaný workflow. Spolupracuje pritom s knižnými katalógmi a digitálnymi knižnicami a repozitármi.

V rámci interoperability Register komunikuje s knižnými informačnými systémami a umožňuje dávkové importy vo formáte MARCXML. Register je schopný zbierať dáta z digitálnych knižníc prostredníctvom OAI-PMH a importovať popisné metadáty digitalizovaných dokumentov. Na konci celého procesu odošle Register informáciu o digitalizovanom dokumente do verejného knižného OPACu spoločne s odkazom na tento dokument v digitálnej knižnici. Informácie sú odoslané tiež do Súborného katalógu Českej republiky.

**RegistrDigitalizace.CZ**  
evidence digitalizovaných dokumentů a sledování procesu zpracování

průvodce po království  hledat jiné zdroje: google books BSB

Počet nalezených položek: 3 titulů 1-5 |

**Nastavení filtrů**

- průvodce po království
- Vlastník (3)
  - Městská knihovna v Praze (2)
  - Národní knihovna České republiky (2)
  - Moravská zemská knihovna, Brno (1)
- Hledání podle digitální knihovny (3)
  - Digitální knihovna Městské knihovny v Praze (2)
  - Digitální knihovna NK ČR (2)
  - Digitální knihovna Moravské zemské knihovny (1)
- Finanční program (2)
  - Norské fondy (2)
  - EOO (1)

Sloučeno podle: [Bez sloučení](#) | [ISSN/ISBN](#) | [ČNB](#) | [SIGLA+ID](#) Seřazeno podle: [relevance](#) | [abecedy](#)

**Náhled není k dispozici** Řivnáčův průvodce po království Českém. II. Mapy a plány (1 předloha)  
Rok(y) vydání: 1882 ČNB: cnb001227342

**Náhled není k dispozici** Řivnáčův průvodce po království Českém. I. Část popsaná (1 předloha)  
Rok(y) vydání: 1882 ČNB: cnb001227341

Kafičův ilustrovaný průvodce po království českém. Sv. 13. Železná a Žďárské hory (1 předloha)  
Hlavní autor: Kafka, Josef, 1858-1929  
Rok(y) vydání: 1908

**Náhled není k dispozici** Královské Hradčany (1 předloha)  
Hlavní autor: Tomek, Vratislav Václav  
Místo vydání: Praha  
Nakladatel/vydavatel: Körber  
Rok(y) vydání: 1905  
Část/ročník: Körbnův ilustrovaný průvodce po památných a zajímavých místech Království Českého Seš. 5  
Počet stran: 59

## 5. Proarc – produkčno archivačný systém

ProArc je voľne dostupný open source systém. ProArc je produkčno- archivačný systém, čiže nástroj pre výrobu a úpravu popisných, administratívnych a technických metadát a zároveň riešenie pre dlhodobé uchovávanie digitálnych dokumentov. V systéme ProArc je taktiež možné konvertovať obrazové súbory do textových dokumentov pomocou softwaru ABBY Recognition Server 3.0. Ďalej je možná obrazová konverzia do formátu Jpeg2000 pomocou softwaru KAKADU. ProArc je založený na repozitári Fedora Commons. Využitie rovnakého výrobného a archivačného nástroja posilňuje interoperabilitu a zdieľanie dát medzi jednotlivými digitalizovanými projektmi a pracovníkmi. ProArc umožňuje rýchle čiastočné automatizované vytváranie štandardných metadát. Skladajú sa zo štruktúrálnej, deskriptívnej a archívnej časti a OCR. Z hľadiska archivačnej súčasti riešenia sú zohľadnené štandardy pre dlhodobú archiváciu, najmä ISO normy referenčný model OAIS. ProArc exportuje tzv. SIP - submission information package, ktorý je možné za účelom archivácie importovať do systému Archivematica.

Produkční a archivační systém - Editor nežbedova Český English

Import

- Nová dávka
- Historie
- Edtovat
- Nový objekt
- Hledat
- Konzole

Import - Připravit metadata pro import

Obnovit Vybírat vše Editovat FOXML Smazat

Název souboru	Číslo	Typ
ABA00726000202504200000002	[1a]	Titulní strana
ABA00726000202504200000003	[1b]	Prázdná strana
ABA00726000202504200000004	[1]	Běžná strana
ABA00726000202504200000005	[2]	Běžná strana
ABA00726000202504200000006	[3]	Běžná strana
ABA00726000202504200000007	[4]	Běžná strana
ABA00726000202504200000008	[5]	Běžná strana
ABA00726000202504200000009	[6]	Prázdná strana
ABA00726000202504200000010	[7]	Běžná strana
ABA00726000202504200000011	8	Běžná strana
ABA00726000202504200000012	9	Běžná strana
ABA00726000202504200000013	10	Běžná strana
ABA00726000202504200000014	11	Běžná strana
ABA00726000202504200000015	12	Prázdná strana

MODS Poznámka OCR DC

Typ strany: Titulní strana  
 Index stránky: 1  
 Číslo stránky: [1a]

Uložit

Celý Originál Pozadí Přizpůsobit panelu

Titulní strana: [1a] Index: 1  
 Prázdná strana: [1b] Index: 2  
 Běžná strana: [1] Index: 3  
 Běžná strana: [2] Index: 4  
 Běžná strana: [3] Index: 5  
 Běžná strana: [4] Index: 6  
 Běžná strana: [5] Index: 7  
 Prázdná strana: [6] Index: 8

ORIENTALIA BOHEMICA - OPERA MINORA  
 VOLUME I  
 VINCENC POŘÍZKA  
 OPERA MINORA  
 STUDIES IN THE BHAGAVADGĪTĀ  
 AND NEW INDO-ARYAN LANGUAGES  
 EDITED BY  
 JAROSLAV STERNAD  
 ORIENTAL INSTITUTE  
 ACADEMY OF SCIENCES OF THE CZECH REPUBLIC  
 PRAGUE  
 2000

Nacist další adresář Další >

## 6. RDflow

RDflow je súčasťou systému ProArc. Jeho využitie je voliteľné. Záleží na rozhodnutí každého digitalizovaného pracoviska, či má záujem tento modul používať. RDFlow je pracovné prostredie určené pre sledovanie priebehu digitalizácie. Umožňuje nielen sledovanie všetkých naplánovaných prác na jednotlivých predlohách určených na digitalizáciu, ale aj jednotlivých prebiehajúcich prác naprieč všetkými predlohami. Pre každú inštaláciu ProArc je možné nakonfigurovať workflow podľa špecifických potrieb jednotlivých digitalizovaných liniek pomocou XML.

## 7. Štandardy digitalizácie

Pre digitalizáciu kultúrneho dedičstva, jeho následnú správu a sprístupnenie je potreba stanoviť a dodržiavať štandardy vytvárania dát a metadát. V rámci projektu ČDK sú rešpektované štandardy definované pre Národnú knižnicu ČR.

Ďalšie informácie ohľadne štandardov digitalizácie je možné nájsť na stránkach <http://www.ndk.cz/standardy-digitalizace>



## 8. Záver

Česká digitálna knižnica je veľmi vydarený projekt v Českej republike. ČDK je základnou infraštruktúrou na národnej úrovni, ktorá zaisťuje prístup k digitálnym dokumentom a zároveň poskytuje dáta ďalším centrálnym riešeniam na národnej i medzinárodnej úrovni. Systémy a nástroje vyvinuté v rámci projektu umožňujú vzájomnú súčinnosť a spolupracujú aj s ďalšími už zavedenými knižnými informačnými systémami. Hlavným cieľom je umožniť vytváranie, zdieľanie a opätovné využívanie digitálneho obsahu, ľahko a efektívne poskytnúť prístup k všetkým digitalizovaným knižným dokumentom v Českej republike pod jednou strechou. Tento projekt má veľké uplatnenie a je veľmi nápomocný pre bežného užívateľa. Užívateľ dokáže všetky dokumenty nájsť na jednom mieste, a nestráca čas hľadaním na internete, čiže ČDK značne zjednodušuje hľadanie a šetrí čas.

## 9. Zdroje

<http://www.czechdigitallibrary.cz/>

<http://www.czechdigitallibrary.cz/archiv/>

<http://www.czechdigitallibrary.cz/archiv/popis-projektu/>

<http://www.czechdigitallibrary.cz/archiv/harmonogram-projektu/>

<https://www.lib.cas.cz/casopis-informace/aktualni-stav-produkcniho-a-archivacniho-systemu-proarc/>

<http://readgur.com/doc/1045185/zah%C3%A1jen%C3%AD-ostr%C3%A9ho-provozu-a-zapojen%C3%AD-dal%C5%A1%C3%ADch-knihoven>

<http://readgur.com/doc/1027078/%C4%8Desk%C3%A1-digit%C3%A1ln%C3%AD-knihovna>

<http://www.ndk.cz/standardy-digitalizace>

<http://ikaros.cz/registr-digitalizace>

<http://readgur.com/doc/1027078/%C4%8Desk%C3%A1-digit%C3%A1ln%C3%AD-knihovna>

<http://www.czechdigitallibrary.cz/softwareva-reseni/>

<http://www.registrdigitalizace.cz/rdcz/uzivatele/>