



Digitální knihovny

Daniel Jakubík
Fakulta Informatiky, MU

Co je to digitální knihovna?

Pohled pracovníka IT

Digitální knihovna je spravovaná sbírka informací spolu s odpovídajícími službami, přičemž informace jsou uloženy v digitální podobě a jsou dostupné prostřednictvím sítě.

- *W.Y.Arms, 2000*

Pohled knihovníka

Digitální knihovny jsou organizace, které poskytují zdroje (včetně specializovaného personálu) umožňující provádět výběr, strukturování a zpřístupnění sbírek digitálních prací, tyto práce dále distribuovat, udržovat jejich integritu a dlouhodobě uchovávat – a to vše s ohledem na snadné a ekonomické využití určitou komunitou nebo množinou komunit uživatelů.

- *Digital Library Federation, 1997*

Budování sbírek

- **digitalizace**
 - Digitalizace informačních zdrojů existujících v analogové podobě.
- **born-digital**
 - Začleňování zdrojů vzniklých přímo v digitální podobě.
- **harvesting**
 - Sklizení informačních zdrojů z webu.

Metadata

- Metadata členíme do tří kategorií:
 - metadata popisná
 - metadata strukturální
 - metadata administrativní
- V praxi nemusí být toto rozdělení striktně dodržováno a jednotlivé typy se často překrývají.

MARC

- MARC (Machine Readable Cataloging)

```
01042nam a2200325 a 4500
001 cpk20031253496
003 CZ PrNK
005 20080118124007.0
007 tu
008 030610s2003 xr a e 001 0 cze
015 $a cnb001253496
020 $a 80-7302-050-5 (váz.)
035 $a (0CoLC)56868300
040 $a ABA001 $b cze
072 7 $a 004.4/.6 $x Programové vybavení. Programové prostředky $2 Konspekt $9 23
080 $a 004.451.9Linux $2 MRF
080 $a (035) $2 MRF
100 1 $a Brandejs, Michal $7 jx20070925011 $4 aut
245 10 $a Linux : $b praktický průvodce / $c Michal Brandejs
250 $a 2. vyd., (V Konvoji 1.)
260 $a Brno : $b Konvoj, $c 2003
300 $a 304 s. : $b il. ; $c 24 cm
504 $a Obsahuje bibliografii a rejstřík
650 07 $a Linux $2 czenas
650 07 $a operační systémy $2 czenas
650 09 $a Linux (operating system) $2 eczenas
655 7 $a příručky $2 czenas
655 9 $a handbooks, manuals, etc. $2 eczenas
901 $b 9878073020505 $f 2. vyd., 1. vyd. v nakl. Konvoj $o 20040506
998 $a 001253496
```

Dublin Core

- nekvalifikovaný Dublin Core
 - 15 základních metadatových prvků
 - obsah (název, předmět, popis, pokrytí, typ, zdroj, vztah)
 - Intelektuální vlastnictví (tvůrce, přispěvatel, vydavatel, práva)
 - instance síťového zdroje (identifikátor, datum, jazyk, formát)
- kvalifikovaný Dublin Core
 - kvalifikátor prvku
 - kvalifikátor hodnoty
 - kvalifikátory musí splňovat princip "dumb-down"

Metadata Kongresové knihovny

- MODS (Metadata Object Description Schema)
 - navrženo odborníky z Kongresové knihovny (Library of Congress) a Marc Standard Office
 - založen na jazyce XML (titleInfo, name, typeOfResource, genre, originInfo, language, physicalDescription, abstract, tableOfContents, targetAudience, note, subject, classification, relatedItem, identifier, location, accessCondition, part, extension, recordInfo)
- obsahuje 20 základních prvků
 - dělí se na podprvky
 - mohou obsahovat atributy

METS

- METS – Metadata Encoding and Transmission Standard
 - Kontejnerový formát založený na jazyce XML
 - může obsahovat data v různých formátech (např. popisná metadata ve formátech MARC21, MARC XML, Dublin Core, MODS aj.)
- obsahuje 20 základních prvků
 - dělí se na podprvky
 - mohou obsahovat atributy

Identifikace

- koncept URN
 - globální rozsah, perzistence, škálovatelnost, legacy support, rozšiřitelnost
- klasické bibliografické identifikátory
 - ISBN, ISSN
 - ISADN, ISTC, ISAN, ISWC, SICI, BICI, ...
- PURL
- handle-system
- DOI

PURL

- URL poskytující nepřímou adresaci
- <http://purl.oclc.org/catalog/item1>
- na odkazované adrese je uloženo skutečné URL
- přesměrování pomocí http redirect
- libovolný počet PURL serverů

Handle System

- kompatibilní s konceptem URN
- hdl:cnri.dlib/magazine
 - pojmenovávající autorita
 - Jedinečný řetězec znaků v rámci autority
- dvojúrovňová architektura
 - globální registr
 - libovolný počet lokálních serverů
- distribuovaný systém s decentralizovanou administrací
- nutnost instalovat plugin do prohlížeče, jinak neúplná činnost

DOI

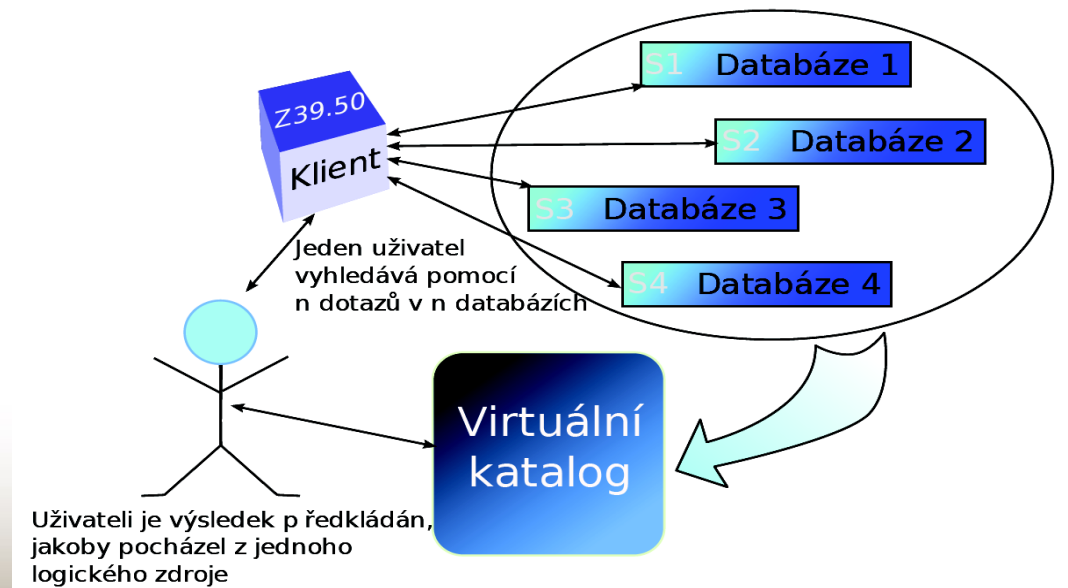
- směrovací mechanismus na základě DOI
- <http://dx.doi.org/10.1000/182>
 - prefix: 10. následovaný číslem registrující organizace
 - identifikátor digitálního objektu - jedinečný v rámci autority
- silně centralizovaný
 - povinná registrace u jediné registrační autority
- není bezplatný

Interoperabilita

- Schopnost vzájemně si rozumět, spolupracovat a dosáhnout součinnosti mezi nezávislými, technicky různorodými systémy.
 - Z39.50
 - OAI-PMH
 - Open URL

Z39.50

- architektura klient/server
- stavový protokol
- komunikaci inicializuje klient, provede sekvenci interakcí a spojení uzavře



Open Archives Initiative (OAI)

- OAI-PMH (OAI Protocol for Metadata Harvesting)
- protokol pro hromadné sklízení metadat
- založen na základních otevřených standardech a protokolech (HTML, XML, HTTP)
- 6 příkazů kódovaných do URL
 - Identify, ListSets, ListIdentifiers, ListMetadataFormats, GetRecord, ListRecords
- příklad dotazu:
 - http://memory.loc.gov/cgi-bin/oai2_0?verb=ListIdentifiers&set=gmd&metadataPrefix=oai_dc

OpenURL a SFX

- Open URL
 - standard pro kódování metadat o zdroji do jeho URL (ANSI norma Z39.88)
- SFX
 - aplikační rámec nad OpenURL

Konec

Děkuji za pozornost