

Technologie digitálních knihoven



Miroslav Bartošek
ÚVT MU Brno

bartosek@ics.muni.cz



Úvod

- **Computers in Libraries** (Washington, 2006)
 - Millennials
 - Informace k dispozici **OKAMŽITĚ, KDEKOLIV, KDYKOLIV** a nejlépe přes **MOBILNÍ ZAŘÍZENÍ**
 - knihovny **musí** inovovat
- **GO DIGITAL !**
 - EIZ
 - **digitální knihovny** (DL)

Perceptions of Libraries and Information Resources (OCLC, 2005)



Letem světem ...

1. 15 let digitálních knihoven
2. Budování sbírek
3. Metadata
4. Identifikace
5. Interoperabilita
6. Další technologie
7. Google-izace digitálních knihoven

1. 15 let digitálních knihoven



1.1 Dospívání DL

- **počátek 90.let – C3**

- computing
- connectivity
- content



- **15 let DL = dospívání (puberta?)**

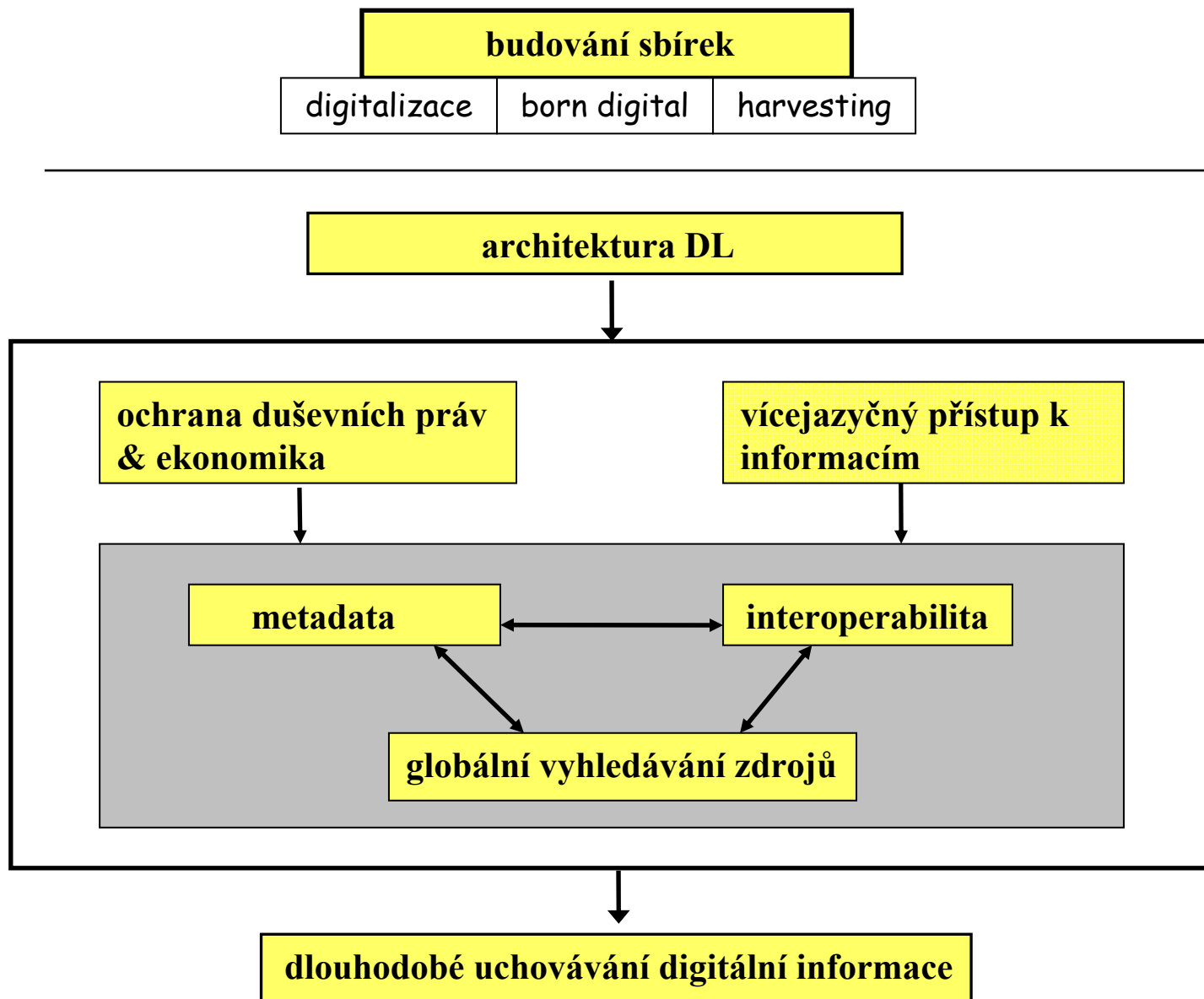
- první úspěchy
Google, Dublin Core, OAI, METS, DSpace
- první zklamání
identifikace, s-interoperabilita, bezpečnost



1.2 Infrastruktura DL

Trend výzkumu:

- **od budování (izolovaných) DL
k vytváření globální infrastruktury DL**
 - snazší implementace DL
 - spolupráce mezi DL
- ještě dlouhá cesta ...
 - ... ale již první **stavební kameny** !
- jejich stručné představení knihovníkům



2. Budování sbírek





2.1 Budování sbírek – digitalizace

- dobře zvládnuté (běžné materiály)
- Best Practices, formáty (DjVu)
- speciální dokumenty (DML-CZ)
- masové digitalizační projekty
 - 1.gen: American Memory, JSTOR, ...
 - 2.gen: Google Book Search
Open Content Alliance
Elsevier, Springer, ...
 - zdokonalení technologií + pád cen



2.2 Budování sbírek – born digital

- vědecká komunikace, open access
- pre-print archívy – **ArXiv.org**
- repozitáře
 - oborové (fyzika, knihovnictví, ...)
 - institucionální (MIT, CERN, ...)
 - OA-časopisy
 - ...
- **EPrints, DSpace, Greenstone, Fedora ...**



2.3 Budování sbírek – harvesting

- kulturní dědictví
 - technologie sklizení-archivace-zpřístupnění
 - obrovské objemy dat
 - **Internet Archive** (1996, 2PB = $2 \cdot 10^{15}$)
 - národní knihovny
 - ČR: [WebArchiv](#)
- specializované automatizované DL
 - úzce specializovaná data
 - **CiteSeer**

3. Metadata





3.1 Klasická metadata

- Klasická knihovnická metadata (MARC)
 - složitá
 - jen pro profesionály
 - nevhodná pro digitální prostředí
- Automaticky generovaná (Google)
 - primitivní
 - úzce specializovaná
- Internet – džungle specializovaných MD



3.2 Metadata pro DL

○ **Dublin Core**

- jednoduchá popisná metadata (15 prvků)
- velká očekávání
- střízlivá realita

○ **MODS**

- popisná metadata - nový „svatý grál“ ? (LoC)
- MARC + XML + flexibilita

○ **METS**

- XML kontejner pro digitální objekty
 - metadata (popisná, administrativní, technická)
 - digitální obsah (soubory)
 - struktura + propojení

4. Identifikace



4.1 Požadavky na identifikátory

- spolehlivé identifikátory – klíčový prvek !

požadavky:

- globální jednoznačnost
- persistence
- nezávislost na lokaci
- směrovatelnost

URL

- ✓
- **ne**
- **ne**
- ✓

- URL x URN (??)



4.2 Identifikátory pro DL

○ **PURL**

- URL s nepřímou adresací (OCLC)
- jednoduché řešení

○ **Handle Systém**

- Corporation for National Research
- propracované, nezávislé na URL, ne TCP/IP
- vlastní směrovací infrastruktura

○ **DOI**

- Asociace amerických nakladatelů
- aplikace nad Handle Systém
- samovýdělečný ekonomický model
- velmi úspěšný - [CrossRef](#)

5. Interoperabilita





5.1 Interoperabilní protokoly

○ **Z39.50**

- složitý, velmi rozsáhlý, search/retrieval
- úzká oblast využívání (drahé knihovní syst.)
- ne-webová technologie, zastaralý?

○ **SRU/SRW**

- Search Retrieve via URL / WebServices
- odlehčená a webová verze Z39.50
- metasearching

○ **OAI-PMH**

- jednoduchost !
- oddělení rolí (poskytovatelé dat x poskytovatelé služeb)
- sklízení metadat PMH + DC + XML

5.2 Linking

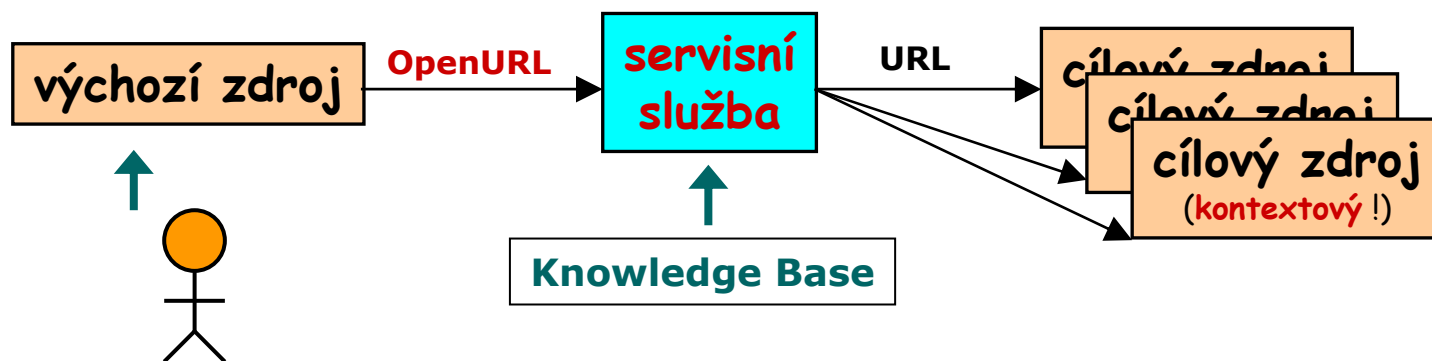
- kontextově citlivé vazby (appropriate copy)

- **OpenURL**

- zakódování metadat o cíli vazby do URL

**`http://xxx.com/demo?sid=ebsco:medline&aulast=Moll
&aunit=JR&date=2001103&stitle=J%20Biol%20Chem&
volume=275&issue=44&spage=34826`**

- servisní služba (SFX)



6. Další technologie





6.1 Další technologie

- IPR - **Creative Commons**
- přístup - **Open Access – BOAI**
- archivace - **OAIS model**
- distrib.systém - **Web Services**
- autentifikace - **Shibboleth**
- architektura - **peer-to-peer**
- aj.

7. Google-izace DL





7.1 Google-izace DL

- Google – skvělé služby, expanze
- Vše vyřeší za „nás“, stačí si počkat ... ??
- Nové vymezení DL
 - nejen
 - **vyhledávání** (jak to najít?)
 - **přístup** (jak to získat?)
 - ale i
 - kvalitní **výběr** materiálu
 - **kolaborativní** (spolupráce uživatelů)
 - **propojení** všeho se vším

7.2 Nové vymezení DL



- DL – mnohem víc než jen vyhledávací portál
- podobně jako klasické knihovny – musí poskytovat kvalitní *výběr zdrojů* a *služby* napomáhající cílové komunitě uživatelů efektivnímu využívání zdrojů
- protože nejsou omezovány prostorem, časem, médiem, měly by být mnohem *přizpůsobivější* a *vstřícnější*
- měly by být *kolaborativní* – umožnit uživatelům přispívat svými znalostmi (aktivně, pasivně)
- měly by být *kontextové* – pavučina vztahů a znalostních vrstev mezi primárními zdroji DL
- snoubit profesionální výběr s moudrostí davu.

What is Digital Library Anomore, Anyway?
C.Lagoze et al, DLib, November 2005

Děkuji za pozornost

Dotazy ?



Logo: www.NCSTRL.org