

DSpace 1.4

Vlastimil Krejčíř, krejcir@ics.muni.cz

Ústav výpočetní techniky, Masarykova univerzita, Brno

PV070 Digitální knihovny, podzim 2006

Obsah přednášky

- Co je to **DSpace**?
- Technické detaily a datový model.
- Co to všechno umí. . .
- . . . a jak to vypadá.
- Co může dělat uživatel. . .
- . . . a co administrátor.
- Kdy a kde nasadit systém **DSpace**.
- Praktická ukázka OAI-PMH.



„The DSpace **digital repository** system captures, stores, indexes, preserves, and distributes **digital research material**.“

– dspace.org



Co je to DSpace?

Jasnější definice?

DSpace je softwarový systém, který je v podstatě *digitální knihovnou*. Kromě samotného datového úložiště poskytuje připravené uživatelské rozhraní pro přístup k uloženým datům.

- => Menší flexibilita :-),
- => ale také méně práce(?) :-).

Technické detaily

DSpace 1.4

- vývojová platforma **Java** (servlety + JSP)
- databázový server **PostgreSQL** (možno i **Oracle**)
- WWW server doporučen **Jakarta Tomcat**
- operační systém – UN*X (Linux, Solaris, HP/UX, ...) nebo MS Windows

Solidní dokumentace (*JavaDocs*) + dobrá architektura (tradiční třívrstvá s jednotlivými komponentami).

Datový model I

Item, Bundle, Bitstream

Základní jednotkou pro ukládání informací je **Item**, což je *digitální objekt* (DO) systému DSpace. Obsahuje:

- metadata a
- soubory **Bitstreams** uložené v
- „přihrádkách“ tzv. **Bundles**.

Bundles slouží zejména k logickému rozdělení dat, která DO (**Item**) obsahuje (jiný **Bundle** je užit pro samotná data, jiný **Bundle** je užit pro uložení licence DO, jiný **Bundle** pro náhledy obrázků apod.).

Datový model II

Collection, Community

Items jsou ukládány v kolekcích (**Collection**) - např. kolekce esejí do předmětu PV070 :-).

Každá kolekce musí patřit alespoň do jedné komunity (**Community**).

Každá komunita může mít nadřazenou komunitu.

Datový model III

Příklad struktury

- Fakulta informatiky
 - Publikace
 - Výuka
 - PV070: Digitální knihovny
 - Studentské eseje 2004 [15]
 - Studentské eseje 2005 [16]
 - Studentské eseje 2006 [0]

Komunity jsou tučně, kolekce normálním fontem

Co všechno **DSpace** umí...

Základní vlastnosti

- popisná metadata v **Dublin Core** (možná i jiná schémata, ale nutno nakonfigurovat)
- **CNRI Handle** systém (lokální registrovaný *handle* server a přidělený **DOI** prefix => globální resoluce přes `hdl:<DOI-prefix>.<DOI-suffix>`)
- podporuje **Storage Resource Broker (SRB)**
- ukládání **HTML** dokumentů (upravuje linky v dokumentu)

Co všechno **DSpace** umí...

Základní vlastnosti II

- má vlastní **OAI** server (podpora protokolu **OAI-PMH**)
- podporuje **OpenURL** pro **SFX** (minimální jednoduchá podpora)
- podpora **Creative Commons** licencí
- *process accounting* dle **Harmony/ABC** (pomocí **RDF**)

Co všechno **DSpace** umí...

Základní vlastnosti III

- import a export (export do **METS**)
- propracovaný proces vkládání (*ingest* a *workflow*)
- checksum checker
- statistiky přístupů (analýzou *logů*)
- ...



Search DSpace

 [Advanced Search](#)
[Subject Search](#)[Home](#)

Browse

- [Communities & Collections](#)
- [Titles](#)
- [Authors](#)
- [Subjects](#)
- [By Date](#)
- [By ID](#)

Sign on to:

- [Receive_email updates](#)
- [My DSpace authorized users](#)
- [Edit Profile](#)
- [Administrator](#)
- [Help](#)
- [About DSpace](#)

DSpace at MU >

DSpace is Live

Welcome to digital repository of **Masaryk University!**

The repository presents collections of communities at MU.

Search

Enter some text in the box below to search DSpace.

Communities in DSpace

Choose a community to browse its collections.

- [Archiv MU](#) [41]
- [Fakulta filozofická](#) [0]
- [Fakulta informatiky](#) [31]
- [Ústav výpočetní techniky](#) [2]

MU-DSpace

Digital collections of MU research communities

News

Sept 2006 -
DSpace 1.4
instalation



Kdy a kde nasadit DSpace

DSpace není specializovaný systém => mnoho věcí z principu neumí (a umět nebude).

Úpravy nejsou vždy jednoduché – je potřeba zvážit, kdy se vyplatí.

DSpace je zadarmo(?)

Před nasazením je dobré si rozmyslet, co od takového systému očekávám...

Protokol OAI-PMH prakticky

Co potrebují znát

Server OAI v DSpace na MU (dále \$URL):

`http://dspace.muni.cz/dspace-oai/`

Specifikace protokolu **OAI-PMH** verze 2.0:

`http://www.openarchives.org/OAI/
openarchivesprotocol.html`

Protokol OAI-PMH prakticky II

Dotazy na server

Dotaz na OAI server probíhá přes protokol HTTP (pro účely prezentace metodou GET). Jméno skriptu pro předání parametrů je request.

Dotazy budeme předávat v obvyklém tvaru:

`$URL/request?parametr1=hodnota1¶metr2=hodnota2...`

Speciální znaky v URL musíme správně zakódovat (například „/“ = %2F, ...)!

Protokol OAI-PMH prakticky II

Identifikace serveru

Požádáme server o **identifikaci**:

`$URL/request?verb=Identify`

`verb=Identify` – *verb* určuje, že hodnota parametru je příkaz, *Identify* je pak samotný příkaz, který umí provést OAI server

Protokol OAI-PMH prakticky III

Identifikátory všech objektů

Chceme získat **identifikátory všech objektů** v repozitáři:

```
$URL/request?verb=ListIdentifiers&  
metadataPrefix=oai_dc
```

`ListIdentifiers` – vypíše identifikátory všech objektů
v repozitáři

`metadataPrefix=oai_dc` – stanovuje, že mají být vráceny
jen identifikátory těch DOs, které mají metadata
v Dublin Core

Protokol OAI-PMH prakticky IV

Metadata objektu

Samotné sklizení **metadat objektů** je po zjištění všech identifikátorů snadné:

```
$URL/request?verb=GetRecord&  
identifier=oai:dspace.muni.cz:ics_muni_cz%2F47&  
metadataPrefix=oai_dc
```

`GetRecord` – příkaz pro získání metadat záznamu

`identifier` – identifikátor DO

`metadataPrefix=oai_dc` – zde určuje v jaké formátu mají být metadata vrácena

Protokol OAI-PMH prakticky V

Další příkazy

- GetRecord
- Identify
- ListIdentifiers
- ListMetadataFormats
- ListRecords
- ListSets

Odkazy

Storage Resource Broker

http://www.sdsc.edu/srb/index.php/Main_Page

Creative Commons

<http://creativecommons.org/>

Harmony/ABC

<http://www.metadata.net/harmony/>