

MASARYKOVA UNIVERZITA

FAKULTA INFORMATIKY



## **DSpace XMLUI: Manakin Project**

<http://di.tamu.edu/projects/xmlui>

Esej z predmetu PV070 Digitálne knižnice

**Ľuboš Kohút**

1. ročník N-IN INS

Brno, 7.12.2009

## Úvod

DSpace Manakin je užívateľské rozhranie pre systém DSpace ako alternatíva k pôvodnému rozhraniu založenému na platforme JSP<sup>1</sup> (*Java Server Pages*). Systém DSpace je softvér pre budovanie digitálnych repozitárov, primárne digitálnych knižníc. Manakin vylepšuje jeho webovú prezentáciu, dodáva jej väčšiu mieru flexibility a zabezpečuje separáciu prezentačnej vrstvy od aplikačnej. DSpace Manakin je postavený na technológiách XML<sup>2</sup>, XSL<sup>3</sup>, SAX<sup>4</sup> a Apache Cocoon framework<sup>5</sup>.

---

<sup>1</sup> [http://en.wikipedia.org/wiki/JavaServer\\_Pages](http://en.wikipedia.org/wiki/JavaServer_Pages)

<sup>2</sup> Extensible Markup Language: <http://www.w3.org/XML>

<sup>3</sup> Extensible Stylesheet Language: <http://www.w3.org/TR/xslt>

<sup>4</sup> Simple API for XML: [http://en.wikipedia.org/wiki/Simple\\_API\\_for\\_XML](http://en.wikipedia.org/wiki/Simple_API_for_XML)

<sup>5</sup> <http://cocoon.apache.org>

# 1. História a ciele projektu, popis systému DSpace

## 1.1. Vznik projektu a história

Existencia digitálnych repozitárov si prirodzene vynucuje ich vhodné prezentovanie koncovým užívateľom. Tieto repozitáre majú však často individuálne požiadavky zobrazovania vzhľadom na ich rozličný obsah. Užívateľské rozhranie musí byť schopné adaptácie na špecifické a často sa meniace požiadavky užívateľov. Na základe týchto potrieb vznikol Manakin ako prezentačná nadstavba systému DSpace. Celá idea rozhrania je vo vytváraní XML dokumentu tzv. DRI dokumentu [1], ktorý je následne transformovaný XSL šablónami na zobrazované XHTML dokumenty.

DSpace Manakin je projekt organizácie **Digital Initiatives** v rámci **Texas A&M University** (TAMU). Táto organizácia vznikla za účelom podpory a propagovania vedeckých a štipendijných programov na TAMU v oblasti výskumu a vývoja digitálnych nástrojov a zdrojov. [2]

Prvá vývojárska príručka Manakinu bola zverejnená v októbri 2005 a prvá pracovná verzia DRI schémy vo februári 2006. Manakin je však až druhou verziou užívateľského rozhrania pre DSpace založeného na XML (*DSpace XML UI*). Prvou verziou bol systém **Moa**. Ten využíval k spracovaniu XML aplikačné rozhranie DOM<sup>1</sup>, na rozdiel od Manakinu, kde je použité SAX. Moa prerazil cestu pre Manakin, ktorý čerpá zo skúseností svojho predchodcu. [3]

## 1.2. Stručný popis systému DSpace

Systém DSpace je open source softvér pre budovanie digitálnych repozitárov navrhnutých pre ukladanie, archivovanie, indexovanie, zálohovanie a prezentovanie dát v digitálnej podobe v rôznych formátoch. Pôvodne vyvíjaný **MIT** (Massachusetts Institute of Technology) **Libraries** a firmou **Hewlett-Packard**. Prvá verzia bola zverejnená v roku 2002. [4, 5]

Základnou organizačnou jednotkou je **komunita** (*Community*). Komunity môžeme do seba ľubovoľne vnárať a môžu obsahovať **kolekcie** (*Collection*). Každá kolekcia musí byť súčasťou práve jednej komunity. Kolekcia nemôže obsahovať iné kolekcie. Základnou archivačnou jednotkou je **položka** (*Item*), ktorá predstavuje digitálny objekt v DSpace. Položky obsahujú metadáta a sú rozdelené na **zväzky** (*Bundles*), ktoré obsahujú konkrétne **bitové prúdy** (*BitStream*). Bitové prúdy reprezentujú samotné súbory ako PDF dokumenty, obrázky a pod.. Bitové prúdy sú vždy konkrétneho **formátu** (*Bitstream format*). Položka musí byť umiestnená v nejakej kolekcii, ktorá ju vlastní, ale zároveň môže byť súčasťou iných kolekcii. [5, 6]

<sup>1</sup> Document Object Model: [http://en.wikipedia.org/wiki/Document\\_Object\\_Model](http://en.wikipedia.org/wiki/Document_Object_Model)

### 1.3. Ciele projektu

Medzi hlavné ciele projektu DSpace Manakin patrí [7, 8, 9]:

- ľahko zmeniť vzhľad a vnímanie (*look-and-feel*<sup>1</sup>) repozitáru DSpace pomocou motívov (*themes*), čím sa reflektujú požiadavky užívateľov.
- umožniť, aby jednotlivé stránky komunit, kolekcí a položiek mohli mať svoj vlastný motív zobrazenia a zároveň umožniť dedenie motívov prostredníctvom organizačnej štruktúry repozitáru.
- oddeliť prezentačnú vrstvu (*presentation tier*) od aplikačnej vrstvy (*business tier*).
- umožniť internacionalizáciu a lokalizáciu obsahu.
- poskytnúť alternatívne rozhranie k rozhraniu JSP (JSP UI), ktoré neoddeľovalo dôsledne prezentačnú vrstvu a bolo zložité pre rozširovanie a údržbu.
- zakryť nepotrebnú funkcionálnosť repozitáru, v niektorých prípadoch umožnenie odtienenie platformy repozitáru a zároveň modulárne rozširovať systém o nové funkcie.
- uplatniť princíp modularity užívateľského rozhrania pre paralelný vývoj.
- zlepšenie vizualizácie netypických metadát ako priestorových (*geospatial*) a časových údajov.

### 1.4. Aktuálny stav

V súčasnosti je Manakin integrovaný v DSpace 1.5.2 (už od verzie 1.5) ako jedna z dvoch možností užívateľských rozhraní. Tou druhou alternatívou je pôvodné rozhranie založené na JSP (predtým bolo nutné Manakin doinštalovať ako samostatný produkt napr. v spojení s DSpace 1.4.2).

Cieľom vývojárov je začlenenie celého Manakinu (XML UI), samozrejme, ako aj pôvodného JSP UI ako modulárnu súčasť hlavných rozhraní (*core service interfaces*) v DSpace 2.0. [10, 11]

Týmto sa vlastne Manakin plynule stal plnohodnotnou súčasťou DSpace a jeho vývoj pokračuje v rámci celého DSpace.

Dlhodobou víziou vývojárov je z Manakinu vytvoriť aplikáciu použiteľnú aj pre iné repozitáre ako DSpace. [7]

<sup>1</sup> [http://en.wikipedia.org/wiki/Look\\_and\\_feel](http://en.wikipedia.org/wiki/Look_and_feel)

## 2. Popis, architektúra a použitie systému

Manakin je systém, v ktorom je možné budovať užívateľské rozhranie, momentálne pre digitálne repozitáre pod systémom DSpace. Predstavujú ho tri unikátne koncepty: **DRI schéma** (*DRI schema*), **aspekty** (*aspects*) a **motívy** (*themes*). Manakin je postavený na Apache Cocoon framework v jazyku Java. [7]

### 2.1 DRI dokument

DRI (*Digital Repository Interface*) je XML dokument, ktorý vytvára abstraktnú reprezentáciu zobrazovanej stránky v repozitári. Uchováva informácie ako o štruktúre stránky, tak aj metadáta zobrazených objektov. Zakódovanie štruktúry vychádza z TEI<sup>1</sup> a na zachytenie metadát sa využíva schéma METS<sup>2</sup>. Popisné metadáta obsiahnuté v METS môžu byť v týchto podporovaných formátoch: DIM<sup>3</sup>, MODS<sup>4</sup> a Dublin Core<sup>5</sup> (či už kvalifikovaný alebo nekvalifikovaný). [7]

Podrobný popis DRI schémy je dobre spracovaný či už na stránkach projektu DSpace Manakin [1], alebo v bakalárskej práci Tomáša Hofmana [5].

### 2.2 Aspekty

Aspekty sú komponenty, ktoré modifikujú existujúce funkcie a vlastnosti digitálneho repozitáru, alebo môžu poskytovať novú funkcionálnosť. Každý aspekt očakáva na vstupe nejaký DRI dokument, ktorý modifikuje a na jeho výstupe je potom ním upravený DRI dokument. Takto jednotlivé aspekty pridávajú do výsledného DRI dokumentu svoj vlastný obsah (obr. 2.1). [7]

Celý proces je riadený prostredníctvom konfiguračných XML súborov sitemap.xmap. Ich štruktúra je definovaná schémou Cocoon Sitemap<sup>6</sup>. Podstata spočíva v tom, že v konfiguračných súboroch sa do tzv. *pipeline* reťazia jednotlivé aspekty a ich aplikovaním vzniká spomínaný výsledný DRI dokument.

<sup>1</sup> Text Encoding Initiative: <http://www.tei-c.org/release/doc/tei-p4-doc/html/ST.html>

<sup>2</sup> Metadata Encoding & Transmission Standard: <http://www.loc.gov/standards/mets>

<sup>3</sup> DSpace Intermediate Metadata Format:

[http://www.dspace.org/1\\_5\\_2Documentation/DSpace-Manual.pdf](http://www.dspace.org/1_5_2Documentation/DSpace-Manual.pdf), strana 99

<sup>4</sup> Metadata Object Description Schema: <http://www.loc.gov/standards/mods>

<sup>5</sup> <http://dublincore.org>

<sup>6</sup> <http://cocoon.apache.org/2.1/userdocs/concepts/sitemap.html>

V Manakine sú štyri základné aspekty (*core aspects*) [5, 7]:

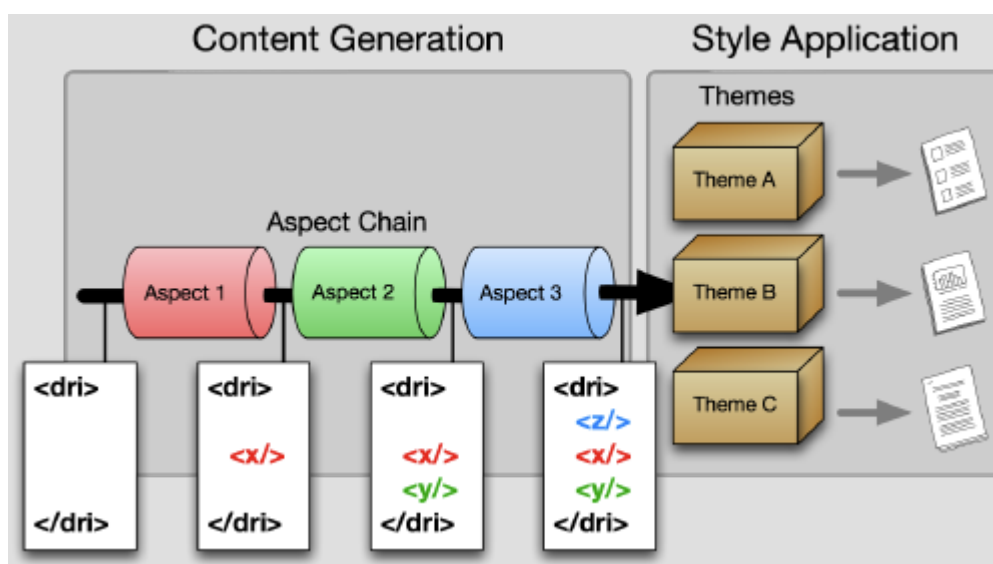
- Artifact Browser – je zodpovedný za zobrazovanie a vyhľadávanie jednotlivých objektov v repozitári.
- E-Person – zabezpečuje prihlasovanie, odhlasovanie a registrovanie užívateľov a je zodpovedný za manažment hesiel a profilov.
- Submission – zaisťuje pridávanie nových položiek do repozitáru a celý schvaľovací proces, ktorý s pridávaním súvisí.
- Administrative – slúži k administratívnym úkonom, ako je autorizácia; vytváranie, modifikovanie a odstraňovanie komunit, kolekcí, užívateľov a skupín.

Okrem preddefinovaných aspektov si užívatelia definujú vlastné aspekty, ktoré realizujú ich vlastné požiadavky a rozšírenia.

## 2.3 Motívy

Motívy zabezpečujú výsledný vzhľad celého repozitáru alebo môžu vytvárať samostatný vzhľad pre jeho konkrétne časti – jednotlivé komunity a kolekcie. Manakin obsahuje už preddefinované motívy, ktoré sa dajú upravovať alebo si užívateľ vytvorí svoje vlastné. Dobrou zásadou je však vhodne upravovať alebo importovať predefinované motívy k svojim vlastným, čo môže ušetriť prácu.

Implementačne ide o množiny XSL šablón, CSS štýlov a iných súborov dotvárajúcich vzhľad stránky (obrázky, JavaScript,...). XSL šablóny sa aplikujú na DRI dokument, ktorý pretransformujú na XHTML dokument a ten je následne prezentovaný užívateľovi (obr. 2.1).



Obr. 2.1: Vizualizácia architektúry Manakinu [7]

## 2.4 Kto používa Manakin?

Medzi uvádzané inštitúcie, ktoré používajú systém DSpace s Manakinom sú napríklad Texaská digitálna knižnica<sup>1</sup> (*Texas Digital Library*) a Geologický atlas USA<sup>2</sup> (*Geologic Atlas of the United States*). V českom prostredí je to napríklad Česká digitálna matematická knižnica<sup>3</sup> (*Czech Digital Mathematics Library*), ktorá je vyvíjaná v Knihovnícko-informačnom centre Ústavu výpočetní techniky MU.

---

<sup>1</sup> <http://www.tdl.org>

<sup>2</sup> <http://txspace.tamu.edu/handle/1969.1/2490>

<sup>3</sup> <http://dml.cz>

## Záver, vlastné zhodnotenie projektu

Projekt DSpace Manakin by som zhodnotil ako úspešný. Zo snahy stať sa alternatívou pre pôvodné DSpace rozhranie sa z neho stalo rozhranie štandardné. Zaujímavý je unikátny koncept aspektov a motívov založených na XSL šablónach. Nespornou výhodou je fakt, že Manakin ako i DSpace sú vyvíjané v perspektívnom programovacom jazyku Java a sú ľahko prenositeľné na rôzne platformy operačných systémov.

Manakin som si osobne vyskúšal v rámci práce v tíme, ktorý pracuje na projekte DML-CZ (Czech Digital Mathematics Library). Moja vlastná skúsenosť s prispôbovaním Manakinu je dobrá. Zároveň by bolo možné vhodné vytvoriť ďalšiu jednoduchšiu úroveň abstrakcie, ktorá by spojila koncepty Manakinu, pretože úpravy, vytváranie vlastných aspektov a XSL šablón nie je vždy najjednoduchšia úloha a vyžaduje skúsenosti programátora. Na druhej strane je dobré, že tvorba rozhrania sa dá rozdeliť medzi skupinu ľudí, ktorí pracujú každý na svojej úrovni.



## Literatúra

- [1] Digital Initiatives. DRI Schema Reference. Dostupné na adrese: <http://di.tamu.edu/projects/xmlui/schemaReference> (december 2009).
- [2] Digital Initiatives. About Digital Initiatives. Dostupné na adrese: <http://di.tamu.edu/about> (december 2009).
- [3] Digital Initiatives. DSpace Manakin. Dostupné na adrese: <http://di.tamu.edu/projects/xmlui> (december 2009).
- [4] DSpace. About DSpace. Dostupné na adrese: <http://www.dspace.org/about-dspace/introducing> (december 2009).
- [5] T. Hofman. *Šablonovací systém DSpace Manakin*. Bakalárska práca, Brno, 2007. Dostupné na adrese: [http://is.muni.cz/th/143156/fi\\_b/bakalarska\\_prace.pdf](http://is.muni.cz/th/143156/fi_b/bakalarska_prace.pdf) (december 2009).
- [6] DSpace. DSpace 1.5.2 Manual. Data Model. Dostupné na adrese: [http://www.dspace.org/1\\_5\\_2Documentation/DSpace-Manual.pdf](http://www.dspace.org/1_5_2Documentation/DSpace-Manual.pdf) (december 2009).
- [7] S. Phillips. *Manakin A New Face for DSpace*. D-Lib Magazine, 2007, volume 13, number 11/12. Dostupné na adrese: <http://www.dlib.org/dlib/november07/phillips/11phillips.html> (december 2009).
- [8] Texas Digital Library. Publications: Manakin/DSpace. Introducing Manakin: Overview and Architecture. 04-2009. Dostupné na adrese: <http://www.tdl.org/wp-content/uploads/2009/04/Introducing%20Manakin.pdf> (december 2009).
- [9] Texas Digital Library. Publications: Manakin/DSpace. Learning to use Manakin. 04-2009. Dostupné na adrese: <http://www.tdl.org/wp-content/uploads/2009/04/LearningToUseManakin.pdf> (december 2009).
- [10] DSpace. Newsletters – NewSpace, January 2008. Dostupné na adrese: <http://www.dspace.org/newsletter-newspace/NewSpace-January-2008.html> (december 2009).
- [11] DSpace. About Newsletters – NewSpace, November 2008, vol. 1, issue 10. Dostupné na adrese: <http://www.dspace.org/about-dspace/introducing> (december 2009).

## Metadáta

```
<link rel="schema.DC" href="http://purl.org/dc/elements/1.1/" />
<meta name="DC.Title" content="DSpace XMLUI: Manakin Project" />
<meta name="DC.Creator" content="Ľuboš Kohút" />
<meta name="DC.Subject" content="DSpace Manakin" />
<meta name="DC.Subject" content="Manakin" />
<meta name="DC.Subject" content="XML UI " />
<meta name="DC.Description.abstract" content="DSpace Manakin je užívateľské
rozhranie pre systém DSpace ako alternatíva k pôvodnému rozhraniu založenému na
platforme JSP (Java Server Pages). Systém DSpace je softvér pre budovanie
digitálnych repozitárov, primárne digitálnych knižníc. Manakin vylepšuje jeho webovú
prezentáciu, dodáva jej väčšiu mieru flexibility a zabezpečuje separáciu prezentačnej
vrstvy od aplikačnej. DSpace Manakin je postavený na technológiách XML, XSL,
SAX a Apache Cocoon framework" />
<meta name="DC.Date.created" content="2009-12-07" />
<meta name="DC.Type" content="Text" />
<meta name="DC.Format" content="application/pdf" />
<meta name="DC.Format.medium" content="computerFile" />
<meta name="DC.Language" content="sk" />
<meta name="DC.Source" content="http://di.tamu.edu/projects/xmlui"/>
<meta name="DC.Source" content="http://wiki.dspace.org/index.php/Manakin"/>
<meta name="DC.Source"
content="http://www.dlib.org/dlib/november07/phillips/11phillips.html"/>
<meta name="DC.Source"
content="http://is.muni.cz/th/143156/fi_b/bakalarska_prace.pdf"/>
```