

# MODS A DALŠÍ METADATOVÁ SCHÉMATA V OBLASTI DIGITALIZACE DOKUMENTŮ

*Pavla Švástová, Veterinární a farmaceutická univerzita Brno*

## Úvod

Se současnou digitalizací dokumentů a jejich seskupováním do širších objektů souvisí mnoho otázek, na které je potřeba odpovědět dříve, než dojde k samotné realizaci projektu. Kromě zajištění technické stránky, identifikátorů atd. je jedním z aspektů výběr vhodného schématu popisných metadat. Ve svém příspěvku stručně popíšu jedno ze zástupců popisných metadat, dnes ve světě již celkem běžně využívané schéma MODS.

## MODS

Metadata Object Description Schema (MODS) je metadatové schéma, které obsahuje soubor bibliografických prvků navržených přímo pro potřeby knihoven, ale použitelných i v dalších případech. Tímto souborem prvků lze popsat prakticky jakékoliv médium od monografií přes CD-ROM či mapy, hudební partitury, obrazy, video i 3D objekty až k webovým stránkám. MODS je založen na jazyce XML, a díky tomu lze pomocí šablon lehce převádět data z MARC či Dublin Core.

Myšlenka, která umožnila vznik schématu MODS, vzešla od provozovatelů digitálních knihoven, kteří požadovali projekt, který by zahrnul možnost bohatého popisu jakéhokoliv dokumentu v XML, spravování složitých digitálních objektů a sjednocení metadat z knihovnických databází založených na MARC s těmi, které využívají jiná schémata. Původně byl zamýšlen jako podmnožina MARC s jinými názvy prvků, ale nakonec vznikl jako samostatné schéma. První verze schématu MODS (verze 1.2) byla představena v polovině června roku 2002. Postupně byly přidávány nové atributy a prvky a byla zlepšována schopnost převoditelnosti dat z a do jiných formátů, aby se snížila ztrátovost a zpětně převedená data se co nejvíce blížila originálu. Systém byl zdokonalován až k současné verzi 3.3 z června 2008.

## Katalogizační pravidla

Samotný obsah prvků může být instruován jakýmkoliv katalogizačními pravidly, ale doporučuje se aplikovat AACR2 (Anglo-American Cataloging Rules). Interpunkce ISBD je vygenerována automaticky na základě šablony

XSLT, která vkládá požadovanou interpunkci podle jednotlivých tagů. V případě záznamů převedených z MARC 21, do kterého se musí interpunkce vkládat ručně, se může stát, že ta původní není úplně odstraněna a v převedeném záznamu se objeví dvakrát.

## Struktura

MODS ve verzi 3.3 obsahuje 20 základních prvků, kterými můžeme dokument popsat. Ty se dále dělí na podprvky (díličí prvky) a k jednotlivým prvkům i podprvkům lze přiřadit atributy. Všechny prvky i jejich atributy jsou nepovinné. Prvky a podprvky jsou opakovatelné. Zatímco podprvky se musí vyskytovat v přesně daném pořadí, atributy nemají pevné pořadí a jsou neopakovatelné. Celý obsah je uzavřen v kořenovém prvku <mods>, který určuje použitou verzi schématu, popřípadě <modsCollection>, který značí, že se nejedná pouze o jediný záznam, ale o celou kolekci záznamů (každý zvláštní záznam je uzavřen v elementu <mods>).

Následuje seznam hlavních prvků MODS, podrobný popis, seznam všech jejich podprvků a atributů naleznete na stránkách:

<http://www.loc.gov/standards/mods/v3/mods-userguide-elements.html>

[http://www.diglib.org/aquifer/dlffmodsimplesimplementationguidelines\\_finalnov2006.pdf](http://www.diglib.org/aquifer/dlffmodsimplesimplementationguidelines_finalnov2006.pdf)

1. titleInfo – názvové informace
2. name – údaje o autorovi (osoba, korporace, konference)
3. typeOfResource – druh zdroje (text, mapa, notový záznam, zvuk, obraz, ...)
4. genre – žánr díla
5. originInfo – nakladatelské údaje
6. language – jazyk díla
7. physicalDescription – fyzický popis dokumentu (materiál, rozsah díla, formát – pdf, jpeg, Huml apod.)
8. abstract – abstrakt
9. tableOfContents – obsah díla (seznam kapitol, skladeb na CD apod.)
10. targetAudience – cílová skupina uživatelů (děti, dospívající apod.)
11. note – všeobecná poznámka
12. subject – předmět, téma díla
13. classification – signatura (MDT, DDT, ...)
14. relatedItem – dílo, které se vztahuje k popisovanému dílu (odkaz na originál, jiné vydání, lze využít všech prvků MODS jako podmnožinu pod relatedItem)
15. identifier – identifikátor

16. location – lokace díla (zeměpisně – kde dílo vzniklo nebo kde se nachází, elektronicky – URL apod.)
17. accessCondition – informace o omezení přístupu k dílu
18. part – část díla
19. extension – informace, které nebylo možno zahrnout pod jiný prvek, např. údaje z národních polí
20. recordInfo – informace o záznamu MODS

Příklad popisu webové stránky:

```

<mods version="3.0">
  <titleInfo>
    <title>FranUlmer.com -- Home Page</title>
  </titleInfo>
  <titleInfo type="alternative">
    <title>Fran Ulmer, Democratic candidate for Governor, Alaska,
2002</title>
  </titleInfo>
  <name type="personal">
    <namePart>Ulmer, Fran</namePart>
  </name>
  <genre>Web site</genre>
  <originInfo>
    <dateCaptured point="start" encoding="iso8601">20020702
    </dateCaptured>
  </originInfo>
  <language>
    <languageTerm authority="iso639-2b">eng</languageTerm>
  </language>
  <physicalDescription>
    <internetMediaType>text/html</internetMediaType>
    <internetMediaType>image/jpg</internetMediaType>
  </physicalDescription>
  <subject>
    <topic>Elections</topic>
    <geographic>Alaska</geographic>
  </subject>
  <relatedItem type="host">
    <titleInfo>
      <title>Election 2002 Web Archive</title>
    </titleInfo>
    <location>
      <url>http://www.loc.gov/minerva/collect/elec2002</url>
    </location>
  </relatedItem>

```

```
<location>
  <url displayLabel="Active site (if available)">
http://www.franulmer.com/</url>
</location>
</mods>
```

## Porovnání metadatových schémat MARC 21, Dublin Core a MODS

Pokud chceme porovnávat tři u nás nejdiskutovanější schémata – MARC21, Dublin Core a MODS, musíme brát ohled na to, že každý formát byl vytvořen k jinému účelu. Soubor prvků MARC je velice rozsáhlý, výhodou je dokonalý popis objektu, nevýhodou náročnost popisu, potřeba zkušeného katalogizátora a též se již ne zcela hodí k popisu elektronických zdrojů. Oproti tomu Dublin Core (DC) byl vytvořen jako požadavek minimálního popisu zdroje, pro běžného uživatele, který si potřebuje popsat webové stránky, nebo v projektech chudých na popisná metadata bohatě postačí, avšak myslím, že k podrobnějšímu popisu zdroje se nehodí. MODS lze využít jak pro rozsáhlý, tak pro stručný popis zdroje. Stejně jako DC využívá tzv. language-based tags – tagy založené na přirozeném jazyce, ze kterých lze ve většině příkladů na první pohled pochopit, o jaké informace se jedná.

## Konverze záznamů mezi MARC 21, DC a MODS

Záznamy z MODS lze jednoduše pomocí konverzních šablon XSLT převést jak na MARCXML (MARC 21 v XML), tak na DC záznamy (i opačným směrem). Knihovny tedy nemusí mít strach, že by záznamy vytvořené v MODS byly „nekompatibilní“ se záznamy jiných knihoven. Šablony založené na technologii XML fungují zjednodušeně takto: šablona vezme z původního XML souboru prvek, obsah zachová, ale formu a umístění změní dle zadání. Vygenerovaný soubor může být jak ve formátu XML, tak např. HTML nebo PDF, možností je spousta.

Příklad:

### ***Původní záznam v MARCXML:***

```
<datafield tag="245" ind1="1" ind2="0">
  <subfield code="a">Dějiny, rasa a kultura :</subfield>
  <subfield code="b">sborník příspěvků z interdisciplinárního symposia
o problematice ras, Plzeň, 12.5.2005 </subfield>
  <subfield code="c">editoři Ivo T. Budil, Vladimír Blažek, Vladimír
Sládek</subfield>
</datafield>
```

**Část šablony:**

```
... <xsl:for-each select="marc:datafield[@tag='245']">
<titleInfo> ...
```

**Záznam převedený do MODS:**

```
<titleInfo>
  <title>Dějiny, rasa a kultura</title>
  <subTitle> sborník příspěvků z interdisciplinárního symposia
o problematice ras, Plzeň, 12.5.2005
  </subTitle>
</titleInfo>
```

**Využití**

Metadatové schéma MODS je v současnosti používáno v mnoha projektech. Ve všech případech se jedná o aplikaci v digitálním světě, např. webový archiv Minerva nebo hudební archiv MusicAustralia (seznam projektů lze získat na stránkách <http://www.loc.gov/standards/mods/registry.php>). Nejčastěji bývá použití MODS spojeno s nasazením kontejnerového formátu METS.

**METS (Metadata Encoding Transmission Schema)**

METS je standard pro ukládání strukturálních, popisných a administrativních metadat v digitální knihovně. Obsahuje tyto části:

- dmdSec – popisná metadata, buď jako součást záznamu METS nebo odkaz na samostatný soubor metadat (využití již vytvořených záznamů), doporučeno je použít MARCXML, DC nebo právě MODS
- admSec – administrativní metadata (technické údaje – formát, velikost souboru, údaje o skenu, legislativní metadata – autorská práva, informace o původci objektů apod.)
- fileSec – soubory, skupiny souborů a jejich fyzické umístění (odkaz na digitální úložiště)
- structMap – strukturální mapa, nejdůležitější část, která propojuje jednotlivé části (soubory mezi sebou, popisná a administrativní metadata, která se vážou k jednotlivým souborům) díky identifikátorům
- behaviourSec – chování podle obsahu

METS je implementován již v řadě světových projektů (Fedora, DSpace). Zdá se být robustním formátem, ale díky tomu, že lze využít jen některé jeho části (např. propojit pouze jednotlivé soubory mezi sebou s návazností na popisná metadata), lze ho bez obav nasadit i v menších projektech.

## Závěr

Při vybírání metadatových schémat je i u menších projektů potřeba myslet do budoucna a připustit, že by se digitalizovaná sbírka milostných dopisů Františka Nováka z Horní Dolní mohla později stát součástí většího projektu, který bude využívat osvědčené standardy, kterými METS v kombinaci s MODS dnes jistě jsou. Jejich výhoda spočívá v tom, že jsou uzpůsobeny elektronickým zdrojům, založeny na moderních technologiích, standardizovány a navíc je obsah lehce převoditelný do jiných formátů.

## Použitá literatura

GUENTHER, Rebecca S. Using the Metadata Object Description Schema (MODS) for resource description: guidelines and applications. Library Hi Tech. Vol. 22, No 1, January 2004, p. 89–98. MCB University Press. ISSN 0737-8831. Dostupný z WWW: <<http://www.ingentaconnect.com/content/mcb/238/2004/00000022/00000001/art00010>>

MODS: Metadata Object Description Schema [online]. Dostupný z WWW: <<http://www.loc.gov/standards/mods/>>.

MODS User Guidelines : Introduction and Implementation [online]. Dostupný z WWW: <<http://www.loc.gov/standards/mods/v3/mods-userguide-intro.html>>.

ŠVÁSTOVÁ, Pavla. Konverze bibliografických formátů MARCXML, MODS a MADS. Brno: Masarykova univerzita, Filozofická fakulta, Ústav české literatury a knihovnictví, Kabinet informačních studií a knihovnictví, 2008. 157 s. Vedoucí diplomové práce Ing. Zdeněk Kadlec, Dr.

ŠVÁSTOVÁ, Pavla. Metadata object description schema. Brno: Masarykova univerzita, Filozofická fakulta, Ústav české literatury a knihovnictví, kabinet knihovnictví, 2006. 63 s. Vedoucí diplomové práce Ing. Zdeněk Kadlec, Dr.

VOJNAR, Martin. Stručný úvod do METS [online]. 2007. Dostupné z: <[www.sdruk.cz/it/vojnar\\_mets.ppt](http://www.sdruk.cz/it/vojnar_mets.ppt)>