
Katalog obchodu s gramodeskami

Dokumentace k semestrální práci XML 4IZ238

Jiří Čech <cecj02@vse.cz>

Table of Contents

Zadání	1
XML Schéma	1
Schematron	2
XSL transformace	2
XSLT - HTML	3
XSLT procesor	3
XSL FO - formátovací objekty	3
FO procesor	3
Dokumentace	3

Zadání

XML dokument eviduje gramofonové desky pro obchod.

U každého díla jsou uchovávány informace:

- název díla
- obrázek
- umělci
- vydavatel
- datum vydání
- typ díla
- formát nosiče
- žánry
- stav díla
- cena
- dostupnost
- obsah

XML Schéma

XML schéma popisuje strukturu, kterou musejí ke schématu provázané XML dokumenty dodržovat. XML schéma definuje entity a jejich atributy, které se mohou v dokumentech vyskytovat. Dále popisuje jejich vlastnosti, jako jejich pořadí, množství výskytů a datové typy. Následná validace XML dokumentu pomocí provázaného XML schématu nám potvrdí, zda XML dokument dodržuje stanovený formát a strukturu.

Schéma pro XML dokument `vinyl_shop.xml` nalezneme v souboru `vinyl_shop.xsd`. Při vytváření XML schématu byla použita technologie W3C XML Schema.

Pro obchod s gramofonovými deskami je vhodné se vyhnout duplicitních položek. Zde však vzniká problém s velmi velkým množstvím nejrůznějších verzí pro stejné hudební dílo. Integrita tedy nemohla být kontrolována pomocí názvu díla. Jediná vždy unikátní hodnota pro každou verzi díla je její kód, proto datový typ *kod* je hlavním elementem entitní integrity.

```
<xs:key name="unikatni_kod_dila">
  <xs:selector xpath="ns:dilo"/>
  <xs:field xpath="ns:kod"/>
</xs:key>
```

Některé informace (*země, den a měsíc vydání*) nemusí být uvedeny, proto jsou nepovinné. Datové typy string jsou omezeny maximálním počtem znaků.

Číselné datové typy (*den a měsíc*) jsou omezeny intervalovým rozpětím. V případě atributů *jednotka* a *mena* je využita enumerace jako například:

```
<xs:simpleType name="menaType">
  <xs:restriction base="xs:string">
    <xs:enumeration value="CZK"/>
    <xs:enumeration value="EUR"/>
    <xs:enumeration value="GBP"/>
  </xs:restriction>
</xs:simpleType>
```

Další enumerace byla využita pro *dostupnost díla (skladem/na obědnání)*, *typ díla (album/ep/single)* a *stav poškození nosiče (dle systému největší vinylové burzy na světě discogs.com)*:

```
<xs:simpleType name="stavType">
  <xs:restriction base="xs:string">
    <xs:enumeration value="M" />
    <xs:enumeration value="NM" />
    <xs:enumeration value="VG+" />
    <xs:enumeration value="VG" />
    <xs:enumeration value="G+" />
    <xs:enumeration value="G" />
    <xs:enumeration value="F" />
    <xs:enumeration value="P" />
  </xs:restriction>
</xs:simpleType>
```

Element *hudebni_dila* obsahuje elementy *dilo*, předpokládá alespoň 1 výskyt a maximální počet výskytů je neomezen pomocí hodnoty *"unbounded"*.

Schematron

V souboru *vinyl_shop.sch* je schematron, pomocí kterého validujeme XML dokumenty. Schematron kontroluje existenci povinných tagů a atributů.

XSL transformace

XSL nám umožňuje definovat více stylů pro jeden XML soubor a měnit tak vzhled XML dokumentů bez zásahu do originálního dokumentu.

XSLT - HTML

Soubor se skládá z jednotlivých šablon, které jsou pomocí dotazovacího jazyka XPath přiřazovány k jednotlivým elementům původního XML souboru a určují, jakým způsobem se dané elementy zpracují. Styl pro transformaci souboru `vinyl_shop.xml` do podoby html je v souboru `vinyl_shop_html.xsl`.

Výstupem této transformace je katalog gramodesek na hlavní stránce `index.html`. Dále jsou vytvořeny stránky s detailem každého hudebního díla a jsou provázané odkazy zpět na hlavní stránku. Pomocí funkce *generate-id* jsou vytvořené unikátní názvy stránek. v souboru `styles.css` jsou kaskádové styly, které upravují vzhled všech stránek.

V případě, kdy lze očekávat větší množství umělců, nebo hudebních žánrů je využita podmínka `if`, která upravuje výstup jednotlivých položek a odděluje je čárkou.

```
<xsl:template match="fd:umelci">
  <xsl:for-each select="fd:umelec">
    <xsl:if test="position() > 1">,< /xsl:if>
    <xsl:value-of select="." />
  </xsl:for-each>
</xsl:template>
```

XSLT procesor

Pro transformaci základního XML dokumentu `vinyl_shop.xml` do formátu HTML pomocí stylu `vinyl_shop_html.xsl` je zapotřebí XSLT procesor. Byl použit Saxon-EE 9.7.0.19 integrovaný v editoru oXygen.

XSL FO - formátovací objekty

XSL FO definuje styly pro převod dokumentu `vinyl_shop.xml` do formátovacích objektů, které následně umožňují dokument převést do formátu PDF a mnoha dalších. Je uložen v souboru `vinyl_shop_fo.xsl`. Opět se jedná o šablony, které se aplikují na jednotlivé elementy původního XML dokumentu a generují formátovací objekty.

Výstupem je PDF dokument s informacemi o nabídce gramofonových desek.

FO procesor

XML dokument se nejprve pomocí XSLT procesoru transformuje na formátovací objekty. Ty se následně zpracují FO procesorem a vytvoří tak výsledný dokument PDF. Byl použit FO procesor Apache FOP integrovaný v oXygenu.

Dokumentace

Tato dokumentace k semestální práci byla vytvořena ve formátu DocBook a vygenerována do výstupů HTML a PDF.