

XML

Literatura:

1. Michael J.Young: XML Step by Step
2. Charles Goldfarbs: XML Handbook 5th Edition
3. Mark Wilson: XML Programming with VB and ASP

internetne strani:

www.w3c.org

www.xml.com

različna orodja npr. xmlSpy

CILJ vaj: pridobiti osnovna znanja o XML tehnologiji in njen pomen v poslovnih rabi.

XML – EXtensible Markup Language

XML – je danes jezik, ki se uporablja za izmenjavo podatkov preko World Wide Weba.

- SGML – prednik XML, precej zahteven in sintaktično strog jezik.*
- XML je jezik namenjen opisu in prenosu podatkov.*
- oznake (tag-i) v XML niso predefinirani(kot v HTML-ju), lahko definiramo svoje.*
- XML ne nosi s sabo informacije, kako bodo podatki predstavljeni. V ta namen so bile razvite sheme XSLT in CSS.*

<http://edavki.durs.si/OpenPortal/Pages/Technicals/FormsXml.aspx>

Razlike med XML in HTML

XML je jezik namenjen opisu in prenosu podatkov.

XML ni nadomestilo za HTML.

XML in HTML sta bila razvita za različne namene.

XML je bil razvit z namenom opisa podatkov in pomena podatkov.

HTML je bil razvit za namene prikaza podatkov in se osredotoča na to kako podatki izgledajo.

HTML je namenjen prikazu informacij, XML pa opisu informacij.

XML je razvit tako, da je informacije lahko strukturirati, hraniti in pošiljati.

XML je razširljiv

V HTML-ju lahko uporabljamo samo v naprej definirane oznake(tags) in predpisano strukturo.

```
<HEAD> .... </HEAD>
```

V XML-u sami razvijamo elemente, od tu oznaka razširljiv (extensible).

```
<?xml version="1.0" encoding="win-1250"?>
<?xml-stylesheet type="text/xsl" href="tecajnica.xsl"?>
<!DOCTYPE Tecajnica SYSTEM "tecajnica.dtd">
<DevizniTecaji>
  <Datum>20050430</Datum>
  <Tecaj>
    <NSifra>978</NSifra>
    <TSifra>EUR</TSifra>
    <Srednji>0000239.669400</Srednji>
    <Enota>1</Enota>
  </Tecaj>
</DevizniTecaji>
```

XML je komplementaren HTML

XML ne nadomešča HTML-ja, ampak ga dopolnjuje, služi kot orodje za prenos podatkov, ki je neodvisno od software-a, hardware-a ali platforme.

Napovedi za prihodnost pravijo, da bo XML najbolj razširjeno orodje za upravljanje s podatki in prenose podatkov.

XML je lahko berljiv za človeka in temelji na javnih, odprtih standardih.

XML in B2B

XML se uporablja za izmenjavo poslovnih informacij preko interneta.

XML omogoča, da so podatki lahko na voljo uporabnikom, ki delujejo na različnih platformah.

XML se lahko uporablja za kreiranje novih jezikov

XML je osnova za WAP in WML (Wireless Markup Language) – jezik za internetne aplikacije, ki delujejo preko mobilnih telefonov.

Pravijo, da je XML mati WAP in WML. Slednji (Wireless Markup Language) uporabljajo za uporabo mobilnih naprav na Internetu. Ta jezik je bil na primer napisan v XML.

Zahteve po povezovanju različnih sistemov in izmenjavi podatkov med njimi so sprožile hitro rast uporabe XML-ja.

Struktura XML

XML ima ***drevesno strukturo***, elementi so vgnezdjeni znotraj korenskega elementa (root element)

```
<ZALOGA>  
<KNJIGA>  
  <NASLOV>Vesela šola</NASLOV>  
  <CENA>300 SIT</CENA>  
</KNJIGA>  
<KNJIGA>  
  <NASLOV>Vesela šola</NASLOV>  
  <CENA>300 SIT</CENA>  
</KNJIGA>  
</ZALOGA>
```

Pri poimenovanju elementov XML je priporočljivo, da izberemo kar najbolj opisne labele.

XML strogo zahteva striktno korektnost pri strukturi. Vsak element mora imeti zaključni element in mora biti pravilno vgnezden.

Za XML dokumente obstajata dve različni stopnji sintaktične pravilnosti:

1. “well-formed” dokument ustreza osnovnim sintaktičnim pravilom, ki zanj veljajo; npr. da so vsi tag-i zaključeni, da so elementi pravilno vgnezdeni
2. “valid” – dokument zadošča strožjim sintaktičnim zahtevam kot dokumenti v prejšnji točki .

Opomba: večina Microsoft Office XP programov uporablja XML kot standardni format dokumenta. Excel 2002 omogoča shranjevanje ali branje XML dokumentov. Microsoft Access 2002 omogoča kreiranje tabele podatkov v bazi s pomočjo XML dokumentov, prav tako omogoča izvoz podatkov bazne tabele v XML dokument.

Prikaz XML dokumenta

Obstajajo trije najosnovnejši načini prikaza elementov XML dokumenta:

- **Style Sheet povezava**: pri uporabi te možnosti povežemo XML dokument z izbranim style sheet dokumentom. Style sheet je poseben dokument v katerem so zbrani ukazi za formatiranje posameznih XML elementov. Uporabljamo lahko CSS (cascading style sheet), ki se uporabljajo tudi za HTML strani ali XSLT (extensible stylesheet language transformation), ki ima na voljo precej več možnosti kot CSS in je razvita posebej za XML.
- **Povezovanje podatkov (data binding)**: kreiramo HTML stran in jo povežemo z XML dokumentom; nato povežemo standardne HTML dokumente, npr. elemente tabele z XML dokumenti. HTML elementi avtomatično izpišejo informacije iz ustreznih XML elementov.
- **XML DOM skripte**: v tem načinu kreiramo HTML stran in ustrezen link na XML dokument, nato izpisujemo vrednosti XML elementov s pomočjo programske kode (JavaScript ali VB Script). Explorer pretvori XML dokument na XML Document Object Model (DOM).

SGML, HTML in XML

SGML – Structured Generalized Markup Language je bil prvi opisni (markup) jezik, iz katerega sta se razvila HTML in XML. Slednja dva sta že v osnovi namenjena različnim potrebam.

SGML definira osnovno sintakso, vendar omogoča kreiranje novih elementov – od tod oznaka “generalized”.

Za opis dokumenta v SGML-ju moramo najprej definirati osnovno množico elementov in strukturo dokumenta: KNJIGA, POGLAVJE, UVOD, ODSTAVEK...

Množico elementov, ki opisuje določen tip dokumenta se imenuje SGML aplikacija.

HTML je eden od primerov SGML aplikacije, ki je bil razvit leta 1991 za opis WEB strani.

XML je poenostavljena izvedenka SGML, optimirana za WEB.

Kje se uporablja XML :

- v strukturiranih dokumentih: idealen je za strukturirane dokumente kot so igre, novele,..primer: Jon Bosak, ki je vodil delovno skupino za XML, je vsa Shakespear-ova dela oblikoval v XML obliko. <http://www.ibiblio.org/bosak>;
- shranjevanje vektorske grafike (Vector Markup Language – VML);
- opis multimedijskih predstavitev;
- kreiranje zvočnih vmesnikov za internetne programe (VoxML ali Voice Markup Language);
- opis software-skih paketov: omogoča distribuiranje in popraviljanje programaskih paketov preko omrežij;
- komuniciranje med programi preko WEB-a, s pomočjo sporočil na osnovi XML-ja (SOAP, Simple Object Access Protocol).

Primeri:

[Vecer.xml](#)

[finance.xml](#)

Kreiranje XML

Ker so XML dokumenti napisani v obliki navadne tekstovne datoteke, lahko za njihovo kreiranje uporabljamo katerikoli tekst editor (Notepad,..) Visual Studio .NET – omogoča hitrejše kreiranje XML dokumentov, zaradi posebnih že vgrajenih funkcij za XML dokumente.

Struktura XML dokumenta

<?xml version="1.0" ?>

PROLOG – XML deklaracija

<!-- File Name: Book Instance.xml -->

... komentar

- če vključimo XML deklaracijo, mora vedno biti v prvi vrstici

<KNJIGA>korenski element
<NASLOV>**The Marble Faun**</NASLOV> elementi
<AVTOR>**Nathaniel Hawthorne**</AVTOR>
<VEZAVA>**trade paperback**</VEZAVA>
<STRANI>**473**</STRANI>
<CENA>**10.95**</CENA>
</KNJIGA>

** element v XML ima podobno vlogo kot BODY element v HTML dokumentu, edina razlika je v tem, da ga lahko poljubno poimenujemo.

Prolog XML dokumenta

```
<?xml version="1.0" ?>
```

PROLOG – XML deklaracija

... komentar

```
<!-- File Name: Book Instance.xml -->
```

- če vključimo XML deklaracijo, mora vedno biti v prvi vrstici
- v prologu lahko tudi definiramo DTD (document type declaration), ki definira podatkovne tipe in strukturo dokumenta;
- nekatere navodila za procesiranje dokumenta, style sheet,...

```
<?xml version="1.0" encoding="win-1250"?>
```

```
<?xml-stylesheet type="text/xsl" href="tecajnica.xsl"?>
```

```
<!DOCTYPE Tecajnica SYSTEM "tecajnica.dtd">
```

```
<DevizniTecaji>
```

Osnovna pravila za XML dokumente

- dokumenta mora imeti samo en korenski element (root element)
- elementi morajo biti pravilno vgnezdjeni
- vsak element mora imeti začetni in končni tag
- XML je **case-sensitive** za imena elementov

XML dokument, ki izpolnjuje ta osnovna pravila je dobro formuliran (well-formed), kar pomeni da ga lahko procesira brkljalnik ali drug XML program.

Internet Explorer omogoča prikaz XML dokumenta na enak način kot katerikoli HTML dokument. V kolikor XML ni povezan s posebnim Style Sheet dokumentov, IE izpiše celotno vsebino dokumenta. V nasprotnem primeru IE pobere iz XML dokumenta vrednosti posameznih elementov, izpiše pa jih v takšni obliki, ki je navedena v Style Sheet dokumentu.

Osnovna pravila za XML dokumente

Primer:

inventory.xml

Kontrola napak v IE:

preden IE izpiše dokument, XML parser preveri strukturo in vsebino dokumenta. V kolikor ugotovi, da obstajajo sintaktične napake, izpiše obvestilo o prvi napaki, ki jo je zaznal.

Naj hitrejši in učinkovit test za testiranje XML dokumenta, v kolikor je dobro definiran je, da ga odpremo v IE. Če ga odpiramo znotraj HTML dokumenta obvestila o napakah ne bomo prejeli !!

1. Naloga

Kreirajmo XML, katerega root element se imenuje 'potnilisti', njegovi elementi so 'listi':

št_lista

ime

priimek

naslov

EMSO

pošta

2. Naloga

Kreirajmo XML, katerega root element se imenuje 'bazaclankov', njegovi elementi so clanki:

clanek

 stclanka

 avtor (eden ali več)

 naslov

 jezik

 stbesed

V XML dodaj tri zapise in pravilnost preveri z IE.

Izpis XML dokumenta z uporabo CSS

Primer:

```
/* File Name: Inventory01.css */
```

```
BOOK  
{display:block;  
margin-top:12pt;  
font-size:10pt}
```

```
TITLE  
{font-style:italic}
```

```
AUTHOR  
{font-weight:bold}
```

```
<BOOK>  
  <TITLE>The Adventures of Huckleberry</TITLE>  
  <AUTHOR>Mark Twain</AUTHOR>  
  <BINDING>mass market paperback</BINDING>  
  <PAGES>298</PAGES>  
  <PRICE>$5.49</PRICE>  
</BOOK>
```

Ta CSS pove IE, da naj izpiše vsak element iz zbirke knjig z zgornjo mejo širine 12 točk, vsak element ločen s presledki od prejšnjega elementa, naslov se zapiše s velikostjo črk 10, v obliki Italic, avtor je izpisan s krepko pisavo.

The Adventures of Huckleberry Finn **Mark Twain** mass market paperback 298 \$5.49

Leaves of Grass **Walt Whitman** hardcover 462 \$7.75

The Legend of Sleepy Hollow **Washington Irving** mass market paperback 98 \$2.95

The Marble Faun **Nathaniel Hawthorne** trade paperback 473 \$10.95

Moby-Dick **Herman Melville** hardcover 724 \$9.95

The Portrait of a Lady **Henry James** mass market paperback 256 \$4.95

The Scarlet Letter **Nathaniel Hawthorne** trade paperback 253 \$4.25

The Turn of the Screw **Henry James** trade paperback 384 \$3.35

Primer: želimo imeti izpis v obliki:

The Scarlet Letter

Nathaniel Hawthorne

hardcopy

\$50.3

```
<BOOK>  
  <TITLE>The Scarlet Letter</TITLE>  
  <AUTHOR>Nathaniel Hawthorne</AUTHOR>  
  <BINDING>trade paperback</BINDING>  
  <PAGES>253</PAGES>  
  <PRICE>$4.25</PRICE>  
</BOOK>
```

Izpis XML dokumenta z uporabo CSS - 2

/* File Name: Inventory02.css */

BOOK
{display:block;
margin-top:12pt;
font-size:10pt}

TITLE
{display:block;
font-size:12pt;
font-weight:bold;
font-style:italic}

AUTHOR
{display:block;
margin-left:15pt;
font-weight:bold}

BINDING
{display:block;
margin-left:15pt}

PAGES
{display:none}

PRICE
{display:block;
margin-left:15pt}

The Scarlet Letter
Nathaniel Hawthorne
hardcopy
\$50.3

<?xml-stylesheet type="text/css"
href="Inventory01.css"?>.....

se ne izpiše

Izpis XML dokumenta z uporabo CSS - 2

```
<?xml version='1.0'?>  
<!-- File Name: Parts.xml --> .....komentar  
<?xml-stylesheet type="text/css" href="Inventory01.css"?>....navodila za izpis
```

```
<INVENTORY>  
  <BOOK>  
    <TITLE>The Adventures of Huckleberry Finn</TITLE>  
    <AUTHOR>Mark Twain</AUTHOR>  
    <BINDING>mass market paperback</BINDING>  
    <PAGES>298</PAGES>  
    <PRICE>$5.49</PRICE>  
  </BOOK>  
</INVENTORY>
```

Vpišemo ime datoteke
V kateri je navedeno
kako naj se XML izpiše.

Vaje XML in CSS

1. Za šolski telefonski imenik sestavi ustrezen XML dokument, ki bo vseboval podatke:

Profesor (ime, priimek)

predmet

kabinet

stevilka.

Vaje XML in CSS

2. V dokument vpiši 4 zapisov.
3. S pomočjo ustreznega CSS izpiši podatke v IE v naslednjih oblikah
 - a) **Ime Priimek** *predmet* *stevilka*
 - b) *Ime* *priimek* **stevilka**
4. Oblikuj podatkovno strukturo za XML dokument Prebivalci s podatki (Priimek, Ime, Datum-rojstva, zaposlen, poklic, izobrazba, st-otrok, roj-leto-otrok)

Vnesi v strukturo 3 različnih podatkov.

Vaje XML in CSS

Izpiši podatke v obliki:

Priimek Ime Datum-rojstva st-otrok

roj-leto-otrok
roj-leto-otrok
roj-leto-otrok

Priimek Ime Datum-rojstva st-otrok

roj-leto-otrok
roj-leto-otrok

Primer uporabe XML

V našem informacijskem sistemu imamo vse dokumente, ki so potrebni za pripravo pošiljke za določenega kupca. Ta dokument pretvorimo "izvozimo" iz informacijskega sistema, pretvorjenega v XML obliko, pošljemo preko izbranega komunikacijskega kanala do partnerja. Le-ta potrdi prejem dokumenta, pripravi pošiljko in pošlje nazaj podatke o pošiljki.

Primer XML naročilo dobavitelju

```
<?xml version="1.0" encoding="utf-8"?>
<eSupply>
  <Glava>
    <st_dok>5500000279</st_dok>
    <dobavitelj>PLASTIKA SKAZA D.O.O.</dobavitelj>
    <datum_spr>4.5.2007</datum_spr>
  </Glava>
  <Poz_zaost>
    <st_poz>000090</st_poz>
    <material>000000000000104208</material>
    <naziv>ŽLICA ZA LED</naziv>
    <zaostanek>500</zaostanek>
    <datum_zaost>4.5.2007</datum_zaost>
    <razkl_mesto>0314</razkl_mesto>
    <skladisce>HZA PL-KOOP.600</skladisce>
  </Poz_zaost>
```

```
<Pozicija>  
  <st_poz>003260</st_poz>  
  <material>000000000000114561</material>  
  <naziv>ROZETA GUMBA GE5-12 EN-N-S 040/7021</naziv>  
  <nar_kol>500</nar_kol>  
  <dob_kol>0</dob_kol>  
  <datum_potr>15.5.2007</datum_potr>  
  <razkl_mesto>0212</razkl_mesto>  
  <skladisce>KA PLINSKI MATER</skladisce>  
</Pozicija>  
</eSupply>
```

“Well-Formed” dokument

<?xml version='1.0'?>priporočen stavek	
<!-- File Name: Parts.xml -->komentar	PROLOG
.....prazna vrstica	
<?xml-stylesheet type="text/css" href="Inventory01.css"?>navodila za izpis	
.....prazna vrstica	
<INVENTORY>		
<BOOK>		
<TITLE>The Adventures of Huckleberry Finn</TITLE>		
<AUTHOR>Mark Twain</AUTHOR>		
<BINDING>mass market paperback</BINDING>		
<PAGES>298</PAGES>		
<PRICE>\$5.49</PRICE>		
</BOOK>		
</INVENTORY>		
.....prazna vrstica	
<!-- Comments, processing instructions, and white space can also appear after the document element. -->		vrstice za elementi dokumenta
<?MyApp Parm1="value 1" Parm2="value 2" ?>		...navodila za procesiranje

“Well-Formed” dokument – 2

- priporočen je stavek, ki označuje verzijo XML dokumenta
- komentar je lahko dodan v prologu ali za podatkovnimi dokumenti
- prazne vrstice v dokumentu so priporočljive boljše preglednosti dokumenta
- vsebuje samo en root element

Pravila za poimenovanje elementov:

- ime elementa se mora začeti s črko ali počrtajem (_), ostali zanki v imenu so opcijski
- XML je rezervirano ime, ki se ne sme pojavljati kot ime elementa
- ker je XML case sensitive, je potrebno biti pozoren, da je ime v prednjem tag-u enako imenu v zadnjem tag-u

“Well-Formed” dokument – 2

Primer:

Dobro definirana imena

_1stPlace

A

B-SECTION

Street.Address.1

Napačna imena:

1stPlaceime se ne sme začeti s številko

B Section.....presledek v imenu ni dovoljen

B/Section...../ ni dovoljen v imenu

:poglavje.....znak : ni dovoljen kot prvi znak imena

“Well-Formed” dokument – 3

```
<BOOK>
  <TITLE>The Adventures of Huckleberry Finn</TITLE>
  <AUTHOR>Mark Twain</AUTHOR>
  <BINDING>mass market paperback</BINDING>
  <PAGES>298</PAGES>
  <PRICE>$5.49</PRICE>
</BOOK>
```

vsebina elementa

s petimi
vgnezdenimi
elementi

```
<TITLE>The Adventures of Huckleberry Finn</TITLE>
```

.....vsebina elementa TITLE...

nedovoljeni znaki v vsebini so >> &

Element ima lahko poleg vrednosti tudi vgnezden element:

```
<TITLE>Huckleberry
```

```
  <SUBTITLE> povest o fantu </SUBTITLE> ... ne uporablaj posebnih  
znakov /
```

```
</TITLE>
```

“Well-Formed” dokument – 3

CDATA stavek omogoča vnos katerega koli teksta v katerem ne smejo biti znaki >] vsebina elementa

```
<TITLEPAGE>
```

```
<![CDATA[
```

```
Document Name: “Kako dodati < ali & “ ]]>
```

```
</TITLEPAGE>
```

s petimi
vgnezdenimi
elementi

CDATA odstavek nekateri poimenujejo tudi “Marked section” .

“Well-Formed” dokument – 3

Komentar v dokumentu:

`<! -- vstavljanje komentarjev -->`

v tekstu znotraj komentar se ne smeta skupaj pojaviti znaka “- -” .
Komentar se lahko piše kjer koli v prologu. Ne sme pa biti znotraj deklaracij, tagov ali drugih komentarjev.

Prazen element

`<HR> </HR>` ali `<HR/>`

Uporablja se na primer pri ukazu IMG, ki vsebuje samo atribut, ki procesorju pove kje najde sliko, ki jo mora izpisati.

2. Naloga

Odpravi napake v XML (imenik-napake.xml) ,
tako da se bodo izpisali v IE.

Atributi

Vsak element ima lahko tudi enega ali več atributov, ki opisujejo značilnosti ali lastnosti elementa. Atribut ima svoje ime in vrednost.

```
<KNJIGA Založnik="DZS" Leto="1990">
```

```
<OSEBA visina="170cm" teza="56"> Tom Sawyer </person>
```

Z atributi lahko dodajamo informacije v elemente. (več o tem pri DTD-jih)

Za imena atributov veljajo enaka pravila kot za imena elementov XML-ja. Isto ime atributa ne sme biti ponovljeno znotraj enega elementa, lahko pa se uporablja v različnih elementih.

Vrednosti atributa so navedene znotraj enojnih ali dvojnih narekovajev, vendar znotraj vrednosti tega znaka kasneje ne smemo uporabljati. Znotraj vrednosti atributov sta tudi prepovedana znaka > ali &.

```
<KANDIDAT ime="John.K. 'Jim' Clinton"> .... pravilna uporaba atributa
```

Atributi elementa

```
<?xml version="1.0"?>
```

```
<!-- File Name: Inventory03.xml -->
```

```
<?xml-stylesheet type="text/css" href="Inventory02.css"?>
```

```
<INVENTORY> <!-- Inventory of selected 19th Century  
American Literature -->
```

```
<BOOK>
```

```
<COVER_IMAGE Source="Huck.gif" /> .....predstavlja prazen element z atributom Source...
```

```
<TITLE>The Adventures of Huckleberry Finn</TITLE>
```

```
<AUTHOR>Mark Twain</AUTHOR>
```

```
<BINDING>mass market paperback</BINDING>
```

```
<PAGES>298</PAGES>
```

```
<PRICE>$5.49</PRICE>
```

```
</BOOK>
```

Atribute lahko uporabljamo za različne namene: lažjo selekcijo elementov, za podatke, ki jih ne bi želeli dati kot nov podatkovni element....

