

XML

Miroslav Trajanović

Šta je XML

- XML je skraćenica za EXtensible Markup Language
- XML je proširivi jezik označavanja (markup language)
- XML je projektovan da strukturira, čuva i prenosi podatke između aplikacija
- Za razliku od HTML, XML tagovi nisu standardizovani. Korisnici sami definišu tagove
- XML je projektovan da opisuje samog sebe
- XML je preporuka W3C iz 1998.

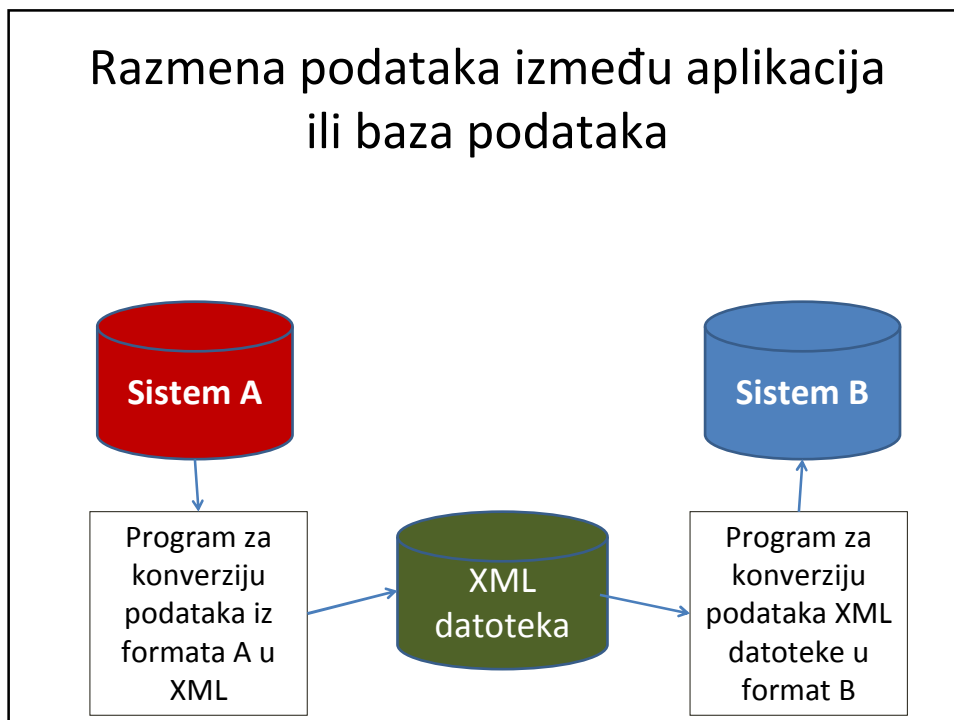
Primer poruke

- Cola Cola podseća Delta Holding da ističe rok plaćanja
- `<poruka>`
`<primalac>Delta Holding</primalac>`
`<posiljaoc>Coca Cola D.O.</posiljaoc>`
`<zaglavlje>Podsecanje</zaglavlje>`
`<telo>Sutra vam ističe rok za plaćanje računa</telo>`
`</poruka>`
- Poruka je samo opisujuća
- XML dokument ne radi ništa
- Potreban je softver za slanje, prijem i prikaz poruka

Korišćenje XML

- Korisnici sami definišu strukturu dokumenata i tagove koji opisuju informacije
- Aplikacije mogu da interpretiraju tagove po potrebi
- XML pomaže da se podaci odvoje od HTML koda. Na taj način je moguć dinamički prikaz strane. HTML se koristi za definisanje izgleda, a pomoću XML se podaci mogu menjati
- Pomoću umetnutog JavaScript-a mogu se pročitati podaci iz neke XML datoteke i prikazati unutar HTML strane

Razmena podataka između aplikacija ili baza podataka



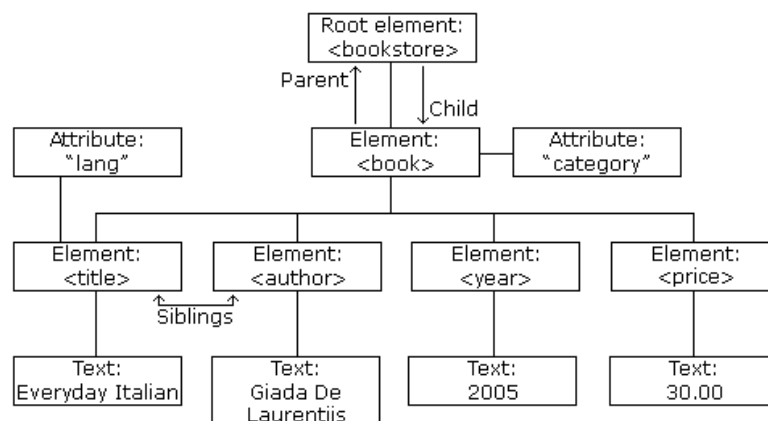
XML dokument ima strukturu stabla

- ```
<?xml version="1.0" encoding="ISO-8859-1"?>
<poruka>
 <primalac>Delta Holding</primalac>
 <posiljaoc>Coca Cola D.O.O</posiljaoc>
 <zaglavlje>Podsecanje</zaglavlje>
 <telo>Sutra vam ističe rok za plaćanje računa</telo>
</poruka>
```

## XML dokument ima strukturu stabla

- `<root>`
  - `<child>`
    - `<subchild>.....</subchild>`
  - `</child>`
- `</root>`

## Primer



## Primer

```
• <bookstore>
 <book category="COOKING">
 <title lang="en">Everyday Italian</title>
 <author>Giada De Laurentiis</author>
 <year>2005</year>
 <price>30.00</price>
 </book>
 <book category="CHILDREN">
 <title lang="en">Harry Potter</title>
 <author>J K. Rowling</author>
 <year>2005</year>
 <price>29.99</price>
 </book>
 <book category="WEB">
 <title lang="en">Learning XML</title>
 <author>Erik T. Ray</author>
 <year>2003</year>
 <price>39.95</price>
 </book>
</bookstore>
```

## XML - Sintaksna pravila

- Mora da postoji završni tag  
`<price>39.95</price>`
- XML pravi razliku između velikih i malih slova  
Ova dva taga nisu ista `<Cena>` i `<cena>`
- Ugnježdavanje mora biti ispravno  
`<b><i>This text is bold and italic</b></i>` neispravno  
`<b><i>This text is bold and italic</i></b>` ispravno

## XML - Sintaksna pravila

- XML mora da ima koreni element koji je roditelj svim podelementima (child)
- Svi atributi moraju biti unutar navodnika

Netačno

```
<note datum=12/11/2007>
 <to>Tove</to>
 <from>Jani</from>
</note>
```

Tačno

```
<note datum="12/11/2007">
 <to>Tove</to>
 <from>Jani</from>
</note>
```

## Entitetske reference

&lt;	<	less than
&gt;	>	greater than
&amp;	&	ampersand
&apos;	'	apostrophe
&quot;	"	quotation mark

## Komentari i blanko karakter

- <!-- Ovo je komentar-->
- Blanko karakteri unutar teksta se ne odbacuju  
Ovde ima    nekoliko blanko karaktera i oni se ne  
  odbacuju

## XML element

- XML element je bilo šta od startnog taga do završnog taga
- Jedan element može da sadrži druge elemente, običan tekst ili oba
- Elementi mogu da imaju attribute

## Primer

```
• <bookstore>
 <book category="CHILDREN">
 <title>Harry Potter</title>
 <author>J K. Rowling</author>
 <year>2005</year>
 <price>29.99</price>
 </book>
 <book category="WEB">
 <title>Learning XML</title>
 <author>Erik T. Ray</author>
 <year>2003</year>
 <price>39.95</price>
 </book>
</bookstore>
```

Atribut

Tekstualni  
sadržaj

Sadržaj u  
obliku  
ugnježenih  
elemenata

## Pravila za davanje imena

- Imena mogu da se sastoje od slova, brojeva i drugih karaktera
- Ime ne može da počne brojem ili tačkom
- Ime ne može da započne stringom xml
- Ime ne može da sadrži blanko karakter
- Bilo koje ime može da se koristi
- Nema rezervisanih reči



## Dobra praksa za davanje imena

- Imena treba da budu opisna. <datum\_prijema>, <datum\_slanja>.
- Imena treba da budu kratka: <broj\_racuna> a ne: <ovo\_je\_broj\_racuna>.
- Izbegavati znak "-" jer neki programi to mogu doživeti kao oduzimanje
- Izbegavati znak "." jer neki programi mogu drugi deo imena doživeti kao osobinu prvog dela. Primer: <broj.racuna>
- Izbegavati znak ":" . Dvotačka je rezervisana za namespaces
- Ne engleska slova šžèòá su sasvim legalna u XML-u, ali se može dogoditi da ih neki softveri ne podržavaju

## Moguće je proširiti XML elemente

- Nekom XML dokumentu je moguće dodati nove elemente, ali će aplikacije sa kojima su radili stari dokumenti i dalje moći da funkcionišu
- <?xml version="1.0" encoding="ISO-8859-1"?>  
<poruka>  
    <primalac>Delta Holding</primalac>  
    <posiljaoc>Coca Cola D.O.O</posiljaoc>  
    <zaglavlje>Podsecanje</zaglavlje>  
    <telo>Sutra vam ističe rok za plaćanje računa</telo>  
    <iznos>578300</iznos>  
</poruka>



Novi element

## XML atributi

- Obezbeđuju dodatne informacije o elementu
- Atributi se pišu unutar startnog taga
- Atributi se koriste da bi se obezbedile informacije koji nisu deo podataka
- Izbegavati korišćenje atributa. Bolje preko elemenata
- Attribute koristiti samo kao metadata  
`<file type="gif">computer.gif</file>`

## Validacija XML-a

- XML dokumenti sa korektnom sintaksom su "Well Formed" XML dokumenti.
- XML koji je proveren na osnovu DTD je "Valid" XML dokument.
- DTD - Document Type Definition

## Well Formed XML Dokumenti

- "Well Formed" XML document ima korektnu XML sintaksu.
- XML dokument mora da ima koreni element
- XML element mora da ima završni element
- XML tagovi razlikuju velika i mala slova (case sensitive)
- XML elementi treba da budu ispravno ugnježđeni
- XML atributi treba da budu unutar navodnika

## Validni XML Dokumenti

- Validni XML dokument je "Well Formed" XML dokument koji takođe odgovara pravilima Document Type Definition (DTD):

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<!DOCTYPE note SYSTEM "Note.dtd">
<note>
 <to>Tove</to>
 <from>Jani</from>
 <heading>Reminder</heading>
 <body>Don't forget me this weekend!</body>
</note>
```

## Note.dtd

- DTD definiše strukturu XML dokumenta preko liste legalnih elemenata
- <!DOCTYPE note  
[  
<!ELEMENT note (to,from,heading,body)>  
<!ELEMENT to (#PCDATA)>  
<!ELEMENT from (#PCDATA)>  
<!ELEMENT heading (#PCDATA)>  
<!ELEMENT body (#PCDATA)>  
]>
- PCDATA - parsed character data

## XML Schema alternativa za DTD

- <xs:element name="note">  
  
<xs:complexType>  
  <xs:sequence>  
    <xs:element name="to" type="xs:string"/>  
    <xs:element name="from" type="xs:string"/>  
    <xs:element name="heading" type="xs:string"/>  
    <xs:element name="body" type="xs:string"/>  
  </xs:sequence>  
</xs:complexType>  
  
</xs:element>

## XML Validator

- XML dokument koji nije validan će zaustaviti izvršenje aplikacije
- Svaki XML dokument treba validirati pre upotrebe posebnim programima za validaciju

## Prikazivanje podataka

- XML dokumenti ne sadrže podatke o tome kako će se podaci prikazati, nego samo definišu značenje podataka
- Za prikazivanje podataka se koriste tehnologije:
  - CSS (Cascading Style Sheets)
  - XSLT (eXtensible Stylesheet Language Transformations)
- W3C konzorcijum preporučuje XSLT

## CSS

- Poseban CSS dokument se kreira sa informacijama o načinu prikazivanja podataka u veb čitaču

## Primer XML sa CSS definicijom

- ```
<?xml version="1.0" encoding="ISO-8859-1"?>
<?xml-stylesheet type="text/css" href="cd_catalog.css"?>
<CATALOG>
  <CD>
    <TITLE>Empire Burlesque</TITLE>
    <ARTIST>Bob Dylan</ARTIST>
    <COUNTRY>USA</COUNTRY>
    <COMPANY>Columbia</COMPANY>
    <PRICE>10.90</PRICE>
    <YEAR>1985</YEAR>
  </CD>
  <CD>
    <TITLE>Hide your heart</TITLE>
    <ARTIST>Bonnie Tyler</ARTIST>
    <COUNTRY>UK</COUNTRY>
    <COMPANY>CBS Records</COMPANY>
    <PRICE>9.90</PRICE>
    <YEAR>1988</YEAR>
  </CD>
  .
</CATALOG>
```

Primer CSS definicije

```
CATALOG
{
background-color: #ffffff;
width: 100%;
}
```

```
CD
{
display: block;
margin-bottom: 30pt;
margin-left: 0;
}
```

```
TITLE
{
color: #FF0000;
font-size: 20pt;
}
```

```
ARTIST
{
color: #0000FF;
font-size: 20pt;
}
```

```
COUNTRY,PRICE,YEAR,COMPANY
{
display: block;
color: #000000;
margin-left: 20pt;
}
```

XSLT - eXtensible Stylesheet Language Transformations

- Pomoću XSLT-a se XML dokument prevodi u HTML dokument pre nego se on prikaže u veb čitaču
- Transformacija se može obaviti u samom veb čitaču ili na serveru