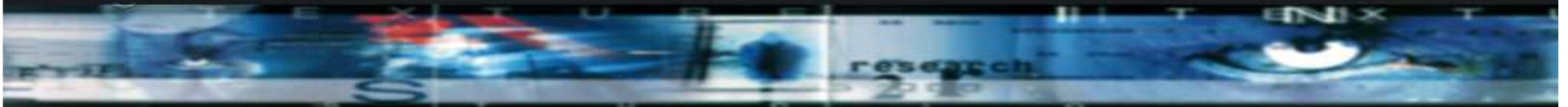


# XML



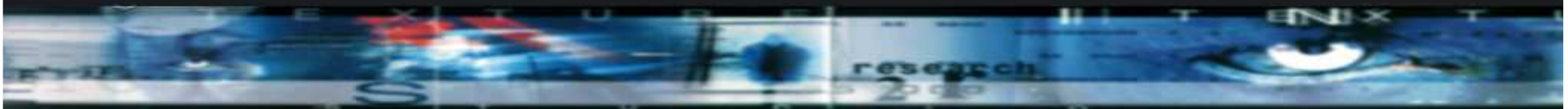
# Sadržaj

- **Osnovna pravila XML-a**
- **Šta znači kad je XML dokument dobro formiran**
- **Komponente iz kojih se sastoji XML dokument**



# Šta je XML?

- **XML je napravljen u cilju čuvanja, prenosa i razmene podatka.**
- **XML tagovi nisu predefinisani. Autor sam definiše tagove koji će se koristiti.**
- **Može biti smešten (prikazan):**
  - **Kao normalan dokument smešten na disku**
  - **Poruka koja se šalje preko HTTP-a**
  - **Znakovni niz u programskom jeziku**
  - **Svaki drugi način korišćenja tekstualnih podataka**



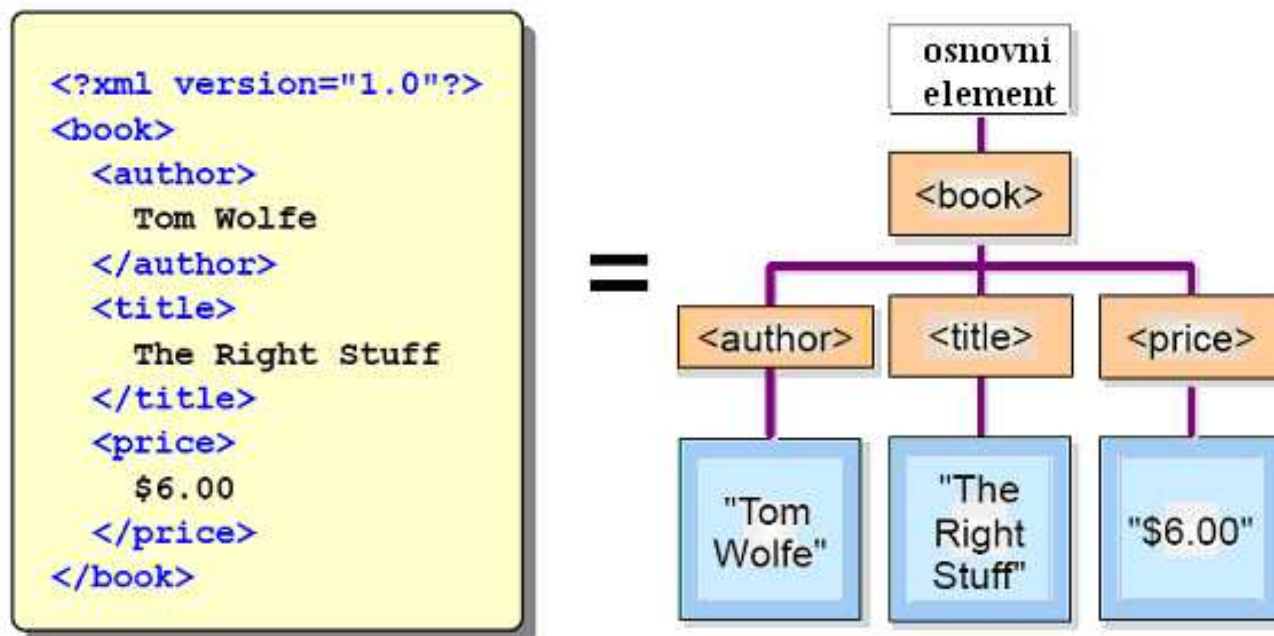
# Šta je XML?

- Korišćenjem XMLa podaci se smeštaju izvan HTML dokumenta
- XML omogućava razmenu podataka između nekompatibilnih sistema
- XML podaci se čuvaju u *plain text* formatu što omogućava hardver i softver nezavistan način razmene podataka
- XML se koristi za čuvanje podataka u fajlovima ili bazama podataka
- XML se koristi i za kreiranje novih jezika, npr. WML (Wireless Markup Language)



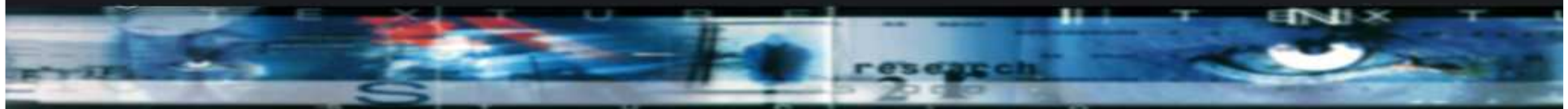
# XML kao stablo– primer

- **Možemo smatrati da XML dokument ima strukturu hijerarhijskog stabla.**



# Osnovna struktura XML dokumenta

<code>&lt;?xml version="1.0"?&gt;</code>	"Opciona" prva linija; obavezna ukoliko kodiranje nije UTF-8 ili UTF-16
<code>&lt;book&gt;</code>	Početni kontrolni kod osnovnog elementa
<code>&lt;title&gt;</code> Alphabet from A to Z <code>&lt;/title&gt;</code>	Prvo dete element sa tekstom
<code>&lt;isbn number="1112-23-4356" /&gt;</code>	Prazan element (bez teksta)
<code>&lt;author&gt;</code>	Početni kontrolni kod elementa
<code>&lt;firstName&gt;Boreng&lt;/firstName&gt;</code> <code>&lt;lastName&gt;Riter&lt;/lastName&gt;</code>	Ugnježdena deca elementi
<code>&lt;/author&gt;</code>	Završni kontrolni kod elementa
<code>&lt;chapter title="Letter A"&gt;</code> The letter A is the first in the alphabet. It is also the first of five vowels . <code>&lt;/chapter&gt;</code>	Element koji sadrži atribut i parsirane znakovne podatke (PCDATA)
<code>&lt;!-- The rest of the letter chapters are missing --&gt;</code>	Komentar
<code>&lt;chapter title="Letter Z"&gt;</code> The letter Z is the last letter in the alphabet. <code>&lt;/chapter&gt;</code>	Poslednji element u dokumentu
<code>&lt;/book&gt;</code>	Završni kontrolni kod osnovnog elementa



# Sintaksa

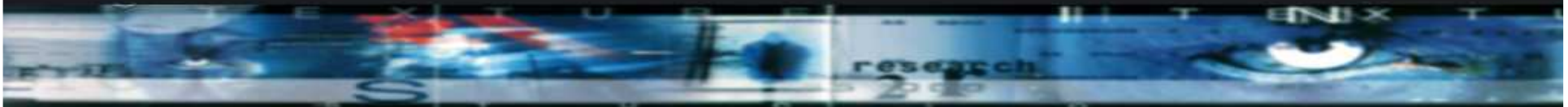
- Elementi moraju biti valjano ugnježdjeni
  - XML tagovi su *case sensitive*
  - Svi elementi moraju biti zatvoreni
  - Svaki dokument mora imati osnovni element
  - Vrednosti atributa se navode između navodnika
  - Za razliku od HTMLa XML čuva razmak
- 
- **<?xml version="1.0" encoding="ISO-8859-1"?>**
  - **ISO-8859-1**
  - **ISO-8859-2**
  - **ISO-8859-5**



# Jedinstven osnovni element

- **Svaki XML dokument mora da ima samo jedan osnovni element**

Ispravno:	Neispravno:
<pre>&lt;?xml version="1.0"?&gt; &lt;colors&gt;   &lt;color&gt;red&lt;/color&gt;   &lt;color&gt;green&lt;/color&gt; &lt;/colors&gt;</pre>	<pre>&lt;?xml version="1.0"?&gt;   &lt;color&gt;red&lt;/color&gt;   &lt;color&gt;green&lt;/color&gt;</pre>
Colors je ovde osnovni element.	Color predstavlja dva osnovna elementa.





# Kontrolni kodovi

- Elementi se sastoje od početnog i završnog kontrolnog koda.

Primer: `<color>red</color>`

- Elementi mogu sadržati atribute u okviru početnog kontrolnog koda.

Primer: `<book isbn="34323"></book>`

- Prazni elementi ne sadrže decu elemente niti tekst

Primer: `<record key="123"></record>`

`<record key="123" />`

`<record />`



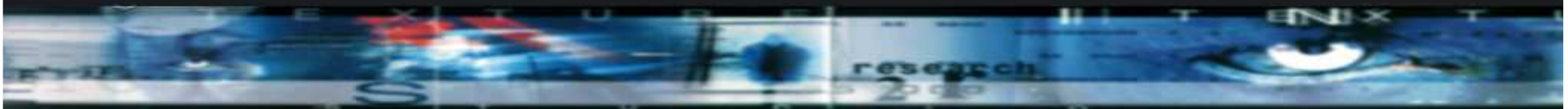
# Ugnježdavanje elemenata

- **Elementi moraju biti propisno ugnježdjeni.**  
**Završni kontrolni kodovi unutrašnjih elemenata moraju da se pojave pre završnih kontrolnih kodova spoljašnjih elemenata.**
- **Ne postoji pravilo koje određuje broj dece elemenata ili podataka koji mogu biti ugnježdjeni unutar početnog i završnog koda elementa**



# Konvencija o imenima

- **Prvi znak mora biti A-Z, a-z, ili \_ (donja crta)**
- **Posle prvog znaka može doći bilo koji broj slova, crtica, razmaka, dvotački i donjih crta**
- **XML razlikuje mala i velika slova**
- **Imena ne smeju da sadrže razmake**
- **Imena ne smeju imati prefiks xml, jer to imaju rezervisane reči u XML-u**
- **Nije neophodno pisati kratka imena, koristiti opisna imena za elemente i attribute**



# Sadržaj elemenata

- **Prazan**
- **Jednostavan**
- **Sadrži druge elemente**
- **Mešovit sadržaj - sadrži tekst i/ili druge elemente**

**<title><ref>XML</ref> Example</title>**

**<chapter>**

**Chapter information**

**<para>What is XML</para>**

**<para>What is HTML</para>**

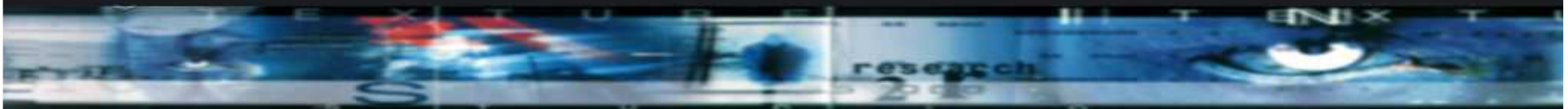
**More chapter information**

**</chapter>**



# Atributi

- Dodatne informacije o elementu
- Vrednosti atributa se navode između znakova navoda " " ili ' '
- `<osoba pol="muški">` ili `<osoba pol='muški'>`
- Ako atribut sadrži duple znake navoda onda može da se navede između jednostrukih znakova navoda.  
`<osoba ime='Pera "Pekar" Petrović'>`
- Element može da ima nekoliko atributa , svaki sa jedinstvenim imenom



# Kada koristi elemente a kada attribute?

```
<?xml version="1.0" encoding="ISO-8859-1"?>
```

```
<poruka datum="25/05/2005" >
```

```
<za>Peru</za>
```

```
<od>Mike</od>
```

```
<naslov>Pozdrav</naslov>
```

```
<sadrzaj>Vidimo se sutra!</sadrzaj>
```

```
</poruka>
```

```
<?xml version="1.0" encoding="ISO-8859-1"?>
```

```
<poruka>
```

```
<datum>25/05/2005</datum>
```

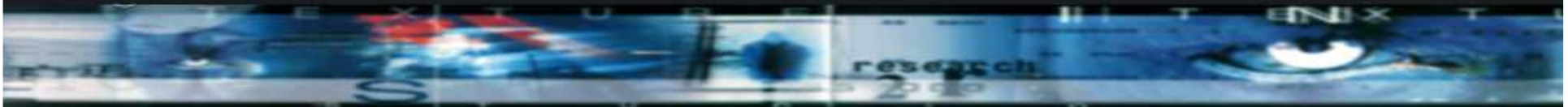
```
<za>Peru</za>
```

```
<od>Mike</od>
```

```
<naslov>Pozdrav</naslov>
```

```
<sadrzaj>Vidimo se sutra!</sadrzaj>
```

```
</poruka>
```



# Kada koristi elemente a kada attribute?

## ▪ Prošireni datum element

```
<?xml version="1.0" encoding="ISO-8859-1"?>
```

```
<poruka>
```

```
<datum>
```

```
<dan>25</dan>
```

```
<mesec>5</mesec>
```

```
<godina>2005</godina>
```

```
</datum>
```

```
<za>Peru</za>
```

```
<od>Mike</od>
```

```
<naslov>Pozdrav</naslov>
```

```
<sadrzaj>Vidimo se sutra!</sadrzaj>
```

```
</poruka>
```



## Kada koristi elemente a kada attribute?

- **Nedostatci atributa:**

**Ne mogu sadržati višestruke vrednosti (elementi mogu)**

**Ne mogu se lako proširiti**

**Teži su za korišćenje**

**Teško se testiraju korišćenjem Document Type Definition (DTD), koji se koristi za definisanje legalnih elemenata XML dokumenta**





# Kada koristi elemente a kada attribute?

<?xml version="1.0" encoding="ISO-8859-1"?>

<poruka id="153">

<za>Peru</za>

<od>Mike</od>

<naslov>Pozdrav</naslov>

<sadrzaj>Vidimo se sutra!</sadrzaj>

</poruka>



# XML validacija

- **Tekst može da bude**
- **Dobro formirani (well formed) XML dokumenti su oni XML dokumenti koji imaju korektnu sintaksu**
- **Validni XML dokumenti su oni koji odgovaraju DTD pravilima**
- **DTD definiše strukturu XML dokumenta sa listom njegovih validnih elemenata**
- **Greške u XML dokumentu će zaustaviti program. Svi XML dokumenti treba da budu kompatibilni, a XML softver jednostavan**



## XML i *browser-i*

- Većina *browser-a* podržava XML, ali svi *browser-i* nemaju celokupnu podršku za XML
- XML fajlovi se mogu prikazati u Mozilla, Firefox, Opera, Internet Explorer i Netscape 6+ browser-ima
- XML document se prikazuje sa obojenim elementima. Plus (+) ili minus znak (-) sa leve strane elementa se može klikuniti u cilju proširenja ili smanjenja strukture elementa. Za prikazivanje XML *source-a* (bez + i - znaka) treba selektovati "View Page Source" ili "View Source" iz browser menija.



# XML parser i CDATA

- **XML parser omogućava čitanje, kreiranje i obradu XML dokumenta**
- **Microsoft's XML parser je COM komponenta i sadrži sve potrebne funkcije koje omogućavaju kreiranje čvorova stabla, pristup njihovim atributima, brisanje, dodavanje čvorova i konvertovanje čvorova stabla nazad u XML dokument**
- **Ceo tekst XML dokumenta se parsira osim onog dela koji se nalazi u okviru CDATA sekcije**
- **Kada se parsira XML element parsira se i tekst koji sadrži , jer XML elementi mogu sadržati i druge elemente**



# XML parser i CDATA

`<poruka>Ovaj tekst se takodje parsira</poruka>`

`<osoba>`

`<ime>Pera</ime>`

`<prezime>Peric</prezime>`

`</osoba>`

`<message>if salary < 1000 then</message>`

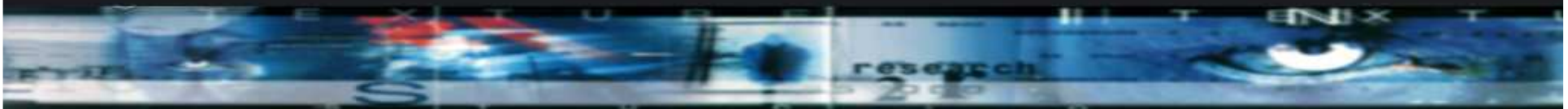
`<message>if salary &lt; 1000 then</message>`

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



# CDATA

- **Sintaksa:** `<![CDATA[ ...Ovde može bilo šta... ]]>`  
**Primedba:** Sve osim znakovnog niza `"]]>"`; To ubacujemo sa `"]]>"`
- **CDATA nije parsiran pa se kao takav i posmatra**
- **Korisno za ubacivanje drugih jezika u okviru XML-a.**
  - HTML dokumenata.
  - XML dokumenata.
  - JavaScript izvornog koda.
  - Ili nekog drugog teksta sa dosta specijalnih znakova



# CDATA - primer

```
<script><![CDATA[  
function matchwo(a,b) {  
  if (a < b && a < 0)  
  then  
    { return 1 }  
  else  
    { return 0 }  
}  
]]></script>
```

```
<script><![CDATA[  
function matchwo(a,b) {  
  if (a < b &#38;&#38; a < 0)  
  then  
    { return 1 }  
  else  
    { return 0 }  
}  
]]></script>
```

```
<nameXML>  
  <![CDATA[  
    <name common="freddy" breed="springer-spaniel">  
      Sir Frederick of Ledyard's End  
    </name>  
  ]]>  
</nameXML>
```





# XML - namespaces

- **S obzirom na to da elementi u XMLu nisu predefinisani može doći do konflikta imena.**
- **Informacije u tabeli:**

```
<table>  
  <tr>  
    <td>Apples</td>  
    <td>Bananas</td>  
  </tr>  
</table>
```

- **Informacije o tabeli:**

```
<table>  
  <name>African Coffee Table</name>  
  <width>80</width>  
  <length>120</length>  
</table>
```





# XML - namespaces

- **Ovaj problem se rešava korišćenjem prefiksa:**

```
<h:table>
```

```
  <h:tr>
```

```
    <h:td>Apples</h:td>
```

```
    <h:td>Bananas</h:td>
```

```
  </h:tr>
```

```
</h:table>
```

```
<f:table>
```

```
  <f:name>African Coffee Table</f:name>
```

```
  <f:width>80</f:width>
```

```
  <f:length>120</f:length>
```

```
</f:table>
```



# XML - namespaces

## ▪ Namespaces:

```
<h:table xmlns:h="http://www.w3.org/TR/html4/">  
  <h:tr>  
    <h:td>Apples</h:td>  
    <h:td>Bananas</h:td>  
  </h:tr>  
</h:table>
```

```
<f:table xmlns:f="http://www.w3schools.com/furniture">  
  <f:name>African Coffee Table</f:name>  
  <f:width>80</f:width>  
  <f:length>120</f:length>  
</f:table>
```



# Pitanja?

Elektrotehnički fakultet  
Univerzitet u Beogradu

