



<XML/>

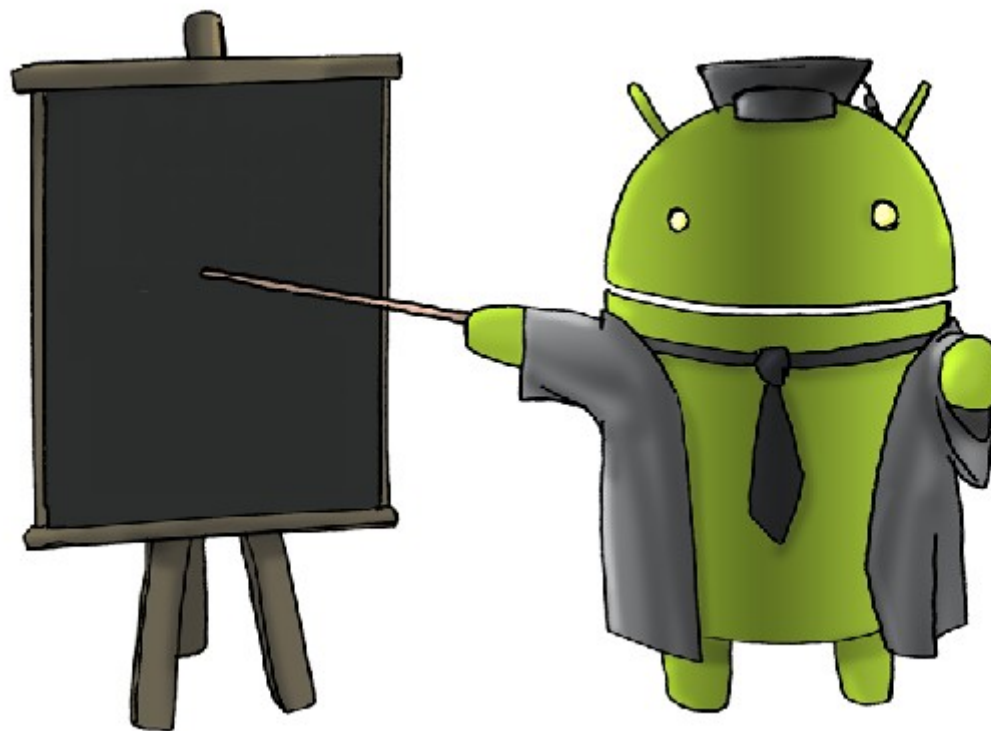
PAPLAŠINĀMĀS IEZĪMĒŠANAS VALODA

eXtensible Markup Language

Vladimirs Kotovs



Mentor @ New Horizons
Java Developer @ Citadele
vladimir.kotov@gmail.com



2 dienas

- Lekcijas
- Praktiskie uzdevumi

Plāns



- Elektroniskās iezīmēšanas vēsture un iezīmēšanas valodas
- XML paplašināmās iezīmēšanas valoda
- XML funkcijas
- XML motivācija un labumi
- Pamatrīki XML dokumentu apstrādei

Plāns



- XML dokumenta komponenti
- XML elements
- XML dokumenta parsēšana
- XML dokumenta atribūti
- XML apstrādes instrukcijas
- XML deklarācija
- XML atsauces
- XML validācija
- XML kritika

1. Iezīmēšana


- Dokumentam pievienota informācija, kas to uzlabo un paplašina
- Nosaka dokumenta daļas un kā tās ir saistītas
- Markup (angl.)



2. Elektroniskā iezīmēšana

```
1 <!DOCTYPE HTML PUBLIC "-//W3C//D
2 <html xmlns="http://www.w3.org/1
3 <head>
4   <meta http-equiv="cache-cont
5   <meta http-equiv="Pragma" co
6   <meta http-equiv="expires" o
7   <meta http-equiv="Content-Ty
8     <meta name="keywords" co
9     <meta name="description"
10 New Horizons - progresīvākās ang
11 <title>Mācību centrs New Hor
12 <link rel="stylesheet" href=
13 <script language="JavaScript
14 <script language="JavaScript
15 <script type="text/javascrip
16 </head>
17
18 <body>
```

www.nh.lv




New Horizons®

viens no pasaulē lielākajiem mācību centru tīkļiem. Latvijā mēs strādājam IT un angļu valodas apmācību jomā.

Mūsu misija – ar efektīvu apmācību palīdzēt cilvēkiem un organizācijām gūt izcilus panākumus.


Izvēlieties, kas Jūs interesē:



Svešvalodu kursi

Visi apmācību līmeņi, sagatavošana IELTS eksāmenam.


Izmēģini angļu valodas kursus nedēļu bez maksas!



IT apmācība

Lietotājiem: Microsoft Office, Adobe, dizains, HTML u.c.

IT Akadēmija - IT karjeras uzsākšanai!



Svešvalodu apmācība internetā

Angļu, vācu, franču, ķīniešu, spāņu, itāļu, zviedru, turku un 20 citas valodas

3. Iezīmēšanas valodas

- Definē noteikumus un konstrukcijas, kas palīdz pievienot jēgu dokumenta saturam un struktūrai
- Vārdu un simbolu kopa, lai aprakstītu dokumenta daļas (piem., “paragrāfs”, “virsraksts”)
- Stilistiskās, struktūras, semantiskās



4. Elektroniskās iezīmēšanas vēsture (1)

- **TeX & Troff (1960s)**

- Drukāto dokumentu atspoguļošana un formatēšana

```
\institute{Division of Applied System  
Latvia \\  
\mailsa}  
\maketitle  
\begin{abstract}  
The need in systematic approach and ef  
methods that facilitate collaboration  
to address recurring problems success  
important for e-business. The synergy  
methods for software solutions may cre  
opportunities This study describes th
```

- **GML (1969)**

- Generalized Markup Language (angl.)
- Dokumentu iezīmēšana izmantošanai ar vairākām informācijas sistēmām

```
:h1.Chapter 1: Introduction  
:p.GML supported hierarchical contain  
:ol  
:li.Ordered lists (like this one),  
:li.Unordered lists, and  
:li.Definition lists
```

4. Elektroniskās iezīmēšanas vēsture (2)

- **SGML (1986)**

- Standard Generalized Markup Language (angl.)
- Uzlabota GML valoda
- Moderno elektronisko iezīmēšanas valodu sencis
- Ļauj autoriem definēt savus tagus, specifiskus dokumenta saturam
- Prasa atsevišķu failu valodas noteikumu definēšanai



4. Elektroniskās iezīmēšanas vēsture (3)

- **HTML (1995)**
 - Hypertext markup language (angl.)
 - Vienkārša hiperteksta dokumentu iezīmēšanas valoda
 - Kompakta un efektīva



4. Elektroniskās iezīmēšanas vēsture (4)

- **HTML**

- Atvērts standarts (nav cieši saistītas ar specifisku tehnoloģiju)
- Specializēti stilistiskie tagi (komandas)
- Ierobežota tagu kopa
- Grūti paplašināms



4. Elektroniskās iezīmēšanas vēsture (5)

- **HTML minusi**

- Atspoguļošana bieži atrauta no satura
- Raksturīgs iezīmēšanas runīgums
- Nav standartizēta starptautiskā atbalsta
- HTML iezīmēšana grūti atkārtoti lietojama



4. Elektroniskās iezīmēšanas vēsture (6)

- HTML minusi

```
<body>
  <div id="all">
    <a name="top"></a>
<!-- START: visa kreisaa mala -->
<div id="left">
  <div id="logo"><a href="/">
  <!-- START: navigacijas bloks -->
  <div id="left1">

  <div>
```


4. Elektroniskās iezīmēšanas vēsture (7)

- **Jauna standarta tapšana**

- World Wide Web Consortium (W3C)
- SGML elastība + HTML vienkāršība
- Formai jāseko funkcija
 - Vairākas iezīmēšanās valodas specifiskas dokumenta saturam universālas valodas vietā



4. Elektroniskās iezīmēšanas vēsture (8)

- **Jauna standarta tapšana**
 - Jāatdala iezīmēšanu no atspoguļošanas
 - stilistiski iezīmēti dokumenti grūti konvertējami jaunās formās
 - stilistiskā informācija jāuzglabā atsevišķi
 - Maksimālā kļūdu pārbaude
 - Standartizēts starptautisks atbalsts
 - KIS (angl. keep it simple)

5. XML - paplašināmās iezīmēšanas valoda

- 1998 - W3C oficiālā XML 1.0 versija
- “Vienkāršota” SGML versija
- Standarts vispārēja nolūka datu transportēšanai un uzglabāšanai
- Fokusējās uz datu semantiku un dokumenta saturu
- XML dokuments ir rakstzīmju virkne



6. XML funkcijas (1)



- Datu glabāšana un apmaiņa
 - no platformas neatkarīgs un valodas ziņā neitrāls standarts
- Datu integritātes nodrošināšana
 - garantē minimālo uzticību datiem (nav bojāti, saīsināti vai nepilnīgi)

6. XML funkcijas (2)



- Dokumentu formatēšana
 - strukturēto datu atspoguļošanas dažādos skatos ar stila lapu palīdzību
- Mašīnlasāmais un cilvēka lasāms formāts
- Vairāku valodu atbalsts (Unicode)

6. XML funkcijas (3)

- Nodrošina paplašināmās iezīmēšanas iespējas
 - nav ierobežota formāta
 - ļauj definēt savu specifisku iezīmēšanu
 - meta-valoda, lai aprakstītu iezīmēšanas valodas
 - “sev aprakstošie dati”
 - definē datu vārdnīcu



7. XML dokumenta piemērs (1)

```
<?xml version="1.0"?>
<note>
  <to>Tove</to>
  <from>Jani</from>
  <heading>Reminder</heading>
  <body>Don't forget me this weekend!</body>
</note>
```



7. XML dokumenta piemērs (2)

```
<?xml version="1.0" encoding="utf-8"?>
<CBSCF:CB.SCF.BlkCdtTrf xmlns:CBSCF="urn:CBSCF:xsd:CB.SCF.BlkCdtTrf"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="urn:CBSCF:xsd:CB.SCF.BlkCdtTrf CB.SCF.BlkCdtTrf.xsd">
  <CBSCF:SndgInst>PARXLV22</CBSCF:SndgInst>
  <CBSCF:RcvgInst>PARXLV22</CBSCF:RcvgInst>
  <CBSCF:SrvId>SCT</CBSCF:SrvId>
  <CBSCF:TstCode>P</CBSCF:TstCode>
  <CBSCF:FType>SCF</CBSCF:FType>
  <CBSCF:FileRef>2009071000000001</CBSCF:FileRef>
  <CBSCF:FileBusDt>2010-06-30</CBSCF:FileBusDt>
  <CBSCF:FIToFICstmrCdtTrf xmlns="urn:iso:std:iso:20022:tech:xsd:sct:pacs.008.
    <GrpHdr>
      <MsgId>Payments001</MsgId>
      <CreDtTm>2010-06-30T08:31:00</CreDtTm>
      <NbOfTx>1</NbOfTx>
      <TtlIntrBkSttlmAmt Ccy="EUR">1.01</TtlIntrBkSttlmAmt>
      <IntrBkSttlmDt>2010-06-30</IntrBkSttlmDt>
      <SttlmInf>
        <SttlmMtd>CLRG</SttlmMtd>
        <ClrSys>
          <Prtry>EKS</Prtry>
        </ClrSys>
      </SttlmInf>
    .....
```

SAMPLE

7. XML dokumenta piemērs (3)

- XML Laika apstākļu serviss

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

```
  <credit>NOAA's National Weather Service</credit>  
  <credit_URL>http://weather.gov/</credit_URL>
```

```
  <image>  
    <url>http://weather.gov/images/xml_logo.gif</url>  
    <title>NOAA's National Weather Service</title>  
    <link>http://weather.gov</link>  
  </image>
```

```
  <location>New York/John F. Kennedy Intl Airport, NY</location>  
  <station_id>KJFK</station_id>  
  <latitude>40.66</latitude>  
  <longitude>-73.78</longitude>  
  <observation_time_rfc822>Mon, 11 Feb 2008 06:51:00 -0500 EST  
</observation_time_rfc822>
```

```
  <weather>A Few Clouds</weather>
```



8. XML motivācija

- Datu apmaiņa ir būtiska mūsdienu tīklā visā pasaulē
 - Finansi: līdzekļu pārskaitījums, pasūtījumu apstrāde
 - Zinātniskie dati:
 - ChemML, BSML (Bio-Sequence Markup Language)
- Papīra dokumentu plūsma starp organizācijām tiek aizstāta ar elektroniskas informācijas plūsmu
- Katrā priekšmetiskā jomā ir savs standartu kopums informāciju glabāšanai



9. XML labumi biznesam



- Informācijas apmaiņa
 - Ļauj uzņēmumiem definēt savus datu formātus XML
 - Nodrošina rīkus, XML lasīšanai, rakstīšanai un datu pārveidošanai no XML uz citu formātu
- XML vienu lietojuma ietvaros
 - Jaudīga, elastīga un paplašināma valoda
- Satura piegāde
 - Atbalsts dažādiem lietotājiem un kanāliem (digitālo TV, telefons, interneta un multimediju kioski)

10. XML Tehnoloģiskie labumi (1)

- Datu atdalīšanas no prezentācijas
 - dinamiskiem HTML datiem HTML dokumentā var būt nepieciešams daudz darba dokumenta rediģēšanai
- Semantiskā informācija
- Atkārtotā datu izmantošana
- Paplašināšanas iespējas



10. XML Tehnoloģiskie labumi (2)

- XML vienkāršo datu koplietošana
 - reālajā pasaulē, datorsistēmas un datu bāzes ietver datus nesaderīgā formātā.
 - XML dati tiek saglabāti vienkārša teksta formātā, kas nodrošina programmatūras-un aparatūras neatkarīgo datu uzglabāšanas veidu.
- XML vienkāršo datu transportu
 - datu apmaiņa starp nesaderīgām sistēmām Internetā



10. XML Tehnoloģiskie labumi (3)

- XML Vienkāršo platformas izmaiņas
 - jaunas sistēmas (aparatūras vai programmatūras platformas) ieviešana vienmēr ir laukietilpīga. XML dati tiek saglabāti teksta formātā, kas atvieglo parēju
- XML veicina jūsu datu pieejamību
 - Dažādas programmas var piekļūt jūsu datiem, ne tikai HTML lapās, bet arī no XML datu avotiem.
 - Ar XML, jūsu dati var būt pieejami visiem "lasīšanas ierīces" veidiem

10. XML Tehnoloģiskie labumi (4)

- XML veicina jaunas interneta Valodas
 - XHTML
 - WSDL, lai tīmekļa servisi
 - WAP un WML, iezīmēšanas valodās portatīvajām ierīcēm
 - RSS, valodas ziņu plūsmām
 - RDF un OWL, resursu aprakstīšanas un ontoloģija



11. Kāpēc XML ir noderīgs informācijas sistēmās?

- Paredzams un viegli validējams dokumentu saturs
- Datu pārbaudes implementētas dokumenta shēmas un sintaktiska analizatora līmenī
- Viegli manipulēt programmas kodā
- Multi-valodu (Unicode) atbalsts
- Dokuments pārsējams līdz kokveida datu struktūras, mezgliem

12. Starpība starp XML un HTML (1)

- XML nav paredzēts aizvietot HTML
- XML un HTML tika izstrādātas atšķirīgiem mērķiem:
 - XML paredzēts transportēt un uzglabāt datus, koncentrējoties uz to, kādi dati ir
 - HTML izstrādāts, lai atspoguļotu datus, koncentrējoties uz to, kā dati izskatās
 - XML ļauj definēt savus tagus
 - HTML tagu kopa ir ierobežota



12. Starpība starp XML un HTML (2)

```
<UL>
  <LI> TOM CRUISE
  <LI> CLIENT ID : 100
  <LI> COMPANY : XYZ Corp.
  <LI> Email : tom@usa.net
  <LI> Phone : 3336767
  <LI> Street Address: 25t
  <LI> City : Toronto
  <LI> State : Toronto
  <LI> Zip : 20056
</UL>
</UL>
```

```
<Details>
  <CONTACT>
    <PERSON_NAME>TOM CRUISE </PERSON_NAME>
    <ID>      100      </ID>
    <Company> XYZ Corp. </Company>
    <Email>   tom@usa.net </Email>
    <Phone>   3336767   </Phone>
    <Street>  25th St.   </Street>
    <City>    Toronto    </City>
    <State>   Toronto    </State>
    <ZIP>     20056      </ZIP>
  </CONTACT>
</Details>
```

13. Pamatrīki XML dokumentu apstrādei

- **XML redaktors**

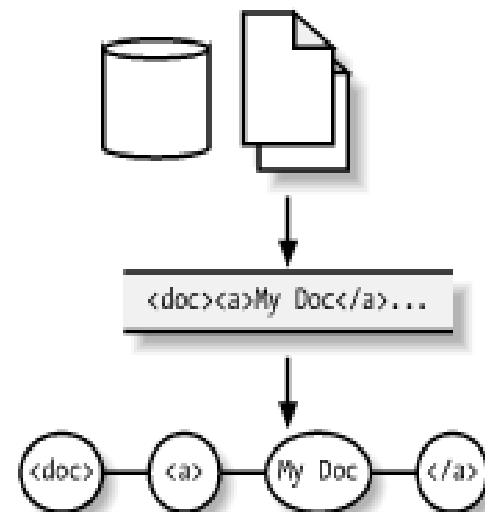
- Notepad, vim vai jebkurš cits teksta redaktors, kas atbalsta dokumenta rakstzīmju kopu
- XML Spy, Stylus, Eclipse IDE

- **XML skatītājs**

- Tīmekļa pārlūkprogramma (e.g. IE, Firefox, Chrome)

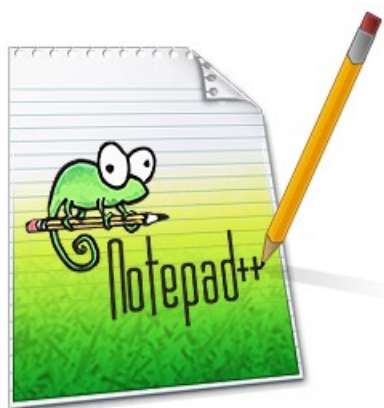
13. Pamatrīki XML dokumentu apstrādei

- **XML sintaktiskais analizators (parsētājs)**
 - Parsēšana ir process, kurā XML sintakse ir pārbaudīta un teksts tiek sadalīts atsevišķās, pārvaldāmās daļās.
- **XML transformētājs**
 - Dokumenta pārveidošana no viena veida uz citu (drukas formatēšana, XML pārveidošanu HTML)



14. XML redaktori (1)

- Populārākie bezmaksas XML redaktori



14. XML redaktori (2)

- Populārākie maksas XML redaktori



14. XML redaktori (3)

- XML redaktora svarīgākās iezīmes
 - XML labošana
 - Sintakses izcelšana
 - Validēšana
 - XML teksta auto pabeigšana
 - Grafiskā XML shēmas atspoguļošana
 - XSLT procesors un atklūdotājs
 - XPATH analizators
- <http://www.xml.com/pub/pt/3>
- http://en.wikipedia.org/wiki/Comparison_of_XML_editors

Praktiskais uzdevums



15. XML Dokumenta pamati

- XML dokumentus parasti saglabā “.xml” failos
- Konkrētus XML dokumentus bieži sauc par XML “instancēm”
- XML teksts ir jutīgs reģistram
- Dokuments sākās ar neobligāto XML deklarāciju



<xml version = "1.0" encoding = "UTF-8"?>

16. XML Dokumenta komponenti

- XML deklarācija
- Elementi
- Atribūti
- Komentāri
- Atsauces (angl. entity)



17. XML tags

- Tagi veido dokumenta iezīmēšanu
- Tags ir teksta konstrukcija, kas sākas ar “<” un beidzas ar “>”
- Sākuma tags “<name>”
- Noslēguma tags “</name>”
- Tagi ir jūtīgi reģistram
 - <CourseName> nav vienāds <courseName>
 - <Bad>Nekorekts XML</bad>



18. XML elements

- Loģiskais XML dokumenta komponents, ko izmanto datu iezīmēšanai
- XML elements sastāv no:
 - Sākuma tags
 - Noslēguma tags
 - Elementa saturs (teksts starp sākuma un beigu tagiem)
 - Iekļauts XML elements vai XML atribūts
- Piemērs: **<first>John</ first>**

19. Noteikumi XML elementu nosaukumiem

- Nosaukumā var būt burti, cipari un citas rakstzīmes
- Nosaukumu nevar sākt ar numuru vai pieturzīmi
- Nosaukumā nevar izmantot atstarpes
- Labas nosaukšanas prakse
 - Aprakstošs nosaukums, piemēram: <first_name>, <last_name>
 - Īss un vienkāršs, piemēram: <book_title> nevis <the_title_of_the_book>
 - Izvairīties no “-”, “.” un “:” zīmēm nosaukumā

20. Noteikumi XML elementiem (1)

- Noslēguma tags ir obligāts
 - Tas ir slikts XML `<p>`
 - Tas arī ir slikts `
`
 - `<p>` Tas ir OK `</p>`
 - Tas ir saīsinājums tukšam tagam `
`



20. Noteikumi XML elementiem (2)

- Tukšie XML elementi
 - Tukšam XML elementam nevar būt saturs
 - Piemēri:
 - `<product prodid="1345" />`
 - `<product></product>`
 - `<product/>`



20. Noteikumi XML elementiem (2)

- Elementu hierarhija

- Slikta elementu hierarhija

<tag1><tag2>Looks bad!</tag1></tag2>

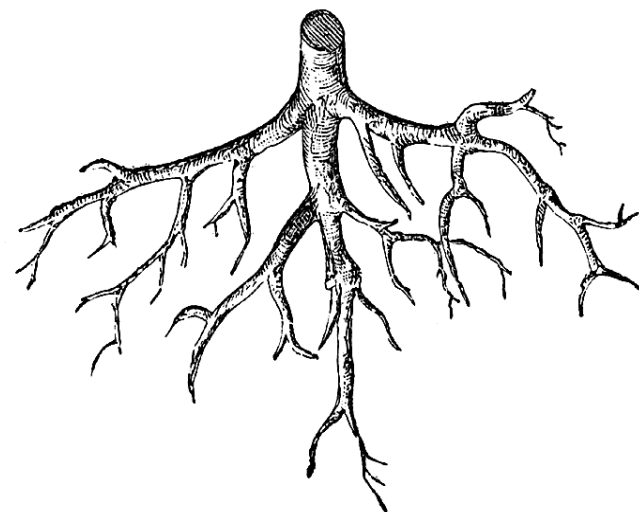
- Korekta elementu hierarhija

<tag1><tag2>Looks good!</tag2></tag1>



21. XML dokumenta saknes elements

- Katra XML dokumenta instance var saturēt tikai **vienu** saknes elementu
- Visi citi elementi jābūt hierarhiski iekļauti saknes elementā
- Piemēram, jebkuram XHTML tipa dokumentam varbūt tikai viens “<html>” tags

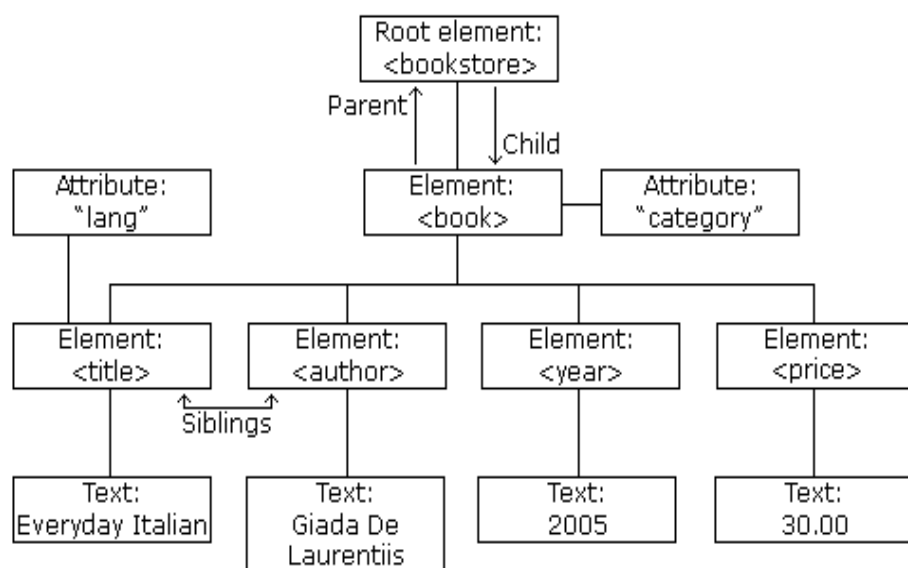


22. XML dokumenta koks

- XML dokumentam ir hierarhiskā struktūra, kas sākas "saknes" elementā un atzarojās "lapās"
- Visiem XML elementiem var būt apakš elementi
- Informācija ir strukturēta koka veidā ar parent-child saistībām

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

22. XML dokumenta koks



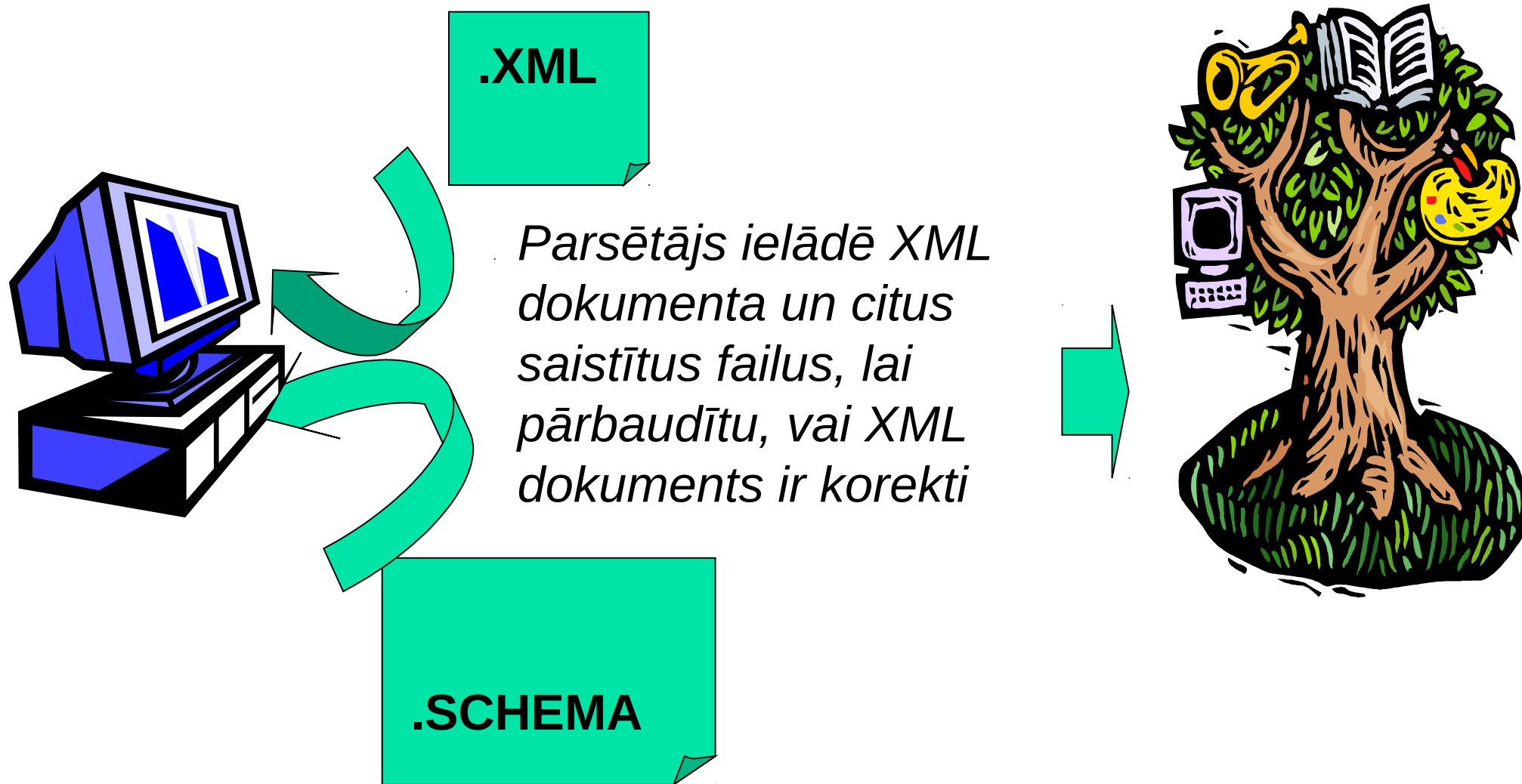
```
<bookstore>
  <book category="COOKING">
    <title lang="en">Everyday Italian</title>
    <author>Giada De Laurentiis</author>
    <year>2005</year>
    <price>30.00</price>
  </book>
  <book category="CHILDREN">
    <title lang="en">Harry Potter</title>
    <author>J K. Rowling</author>
    <year>2005</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>
```

23. XML dokumenta parsēšana (1)

- Katra programmēšanas valoda, un operētājsistēma atbalsta XML parsēšanu
- Lielākā daļa tīmekļa pārlūkprogrammās ir XML parsētāji
- Divu līmeņu XML parsēšana
 - "vāja pārbaude" - korekti veidots XML
 - "stingra pārbaude" – ir ievēroti noteikumi, noteikti konkrētā XML dokumenta shēmā



23. XML dokumenta parsēšana (2)



24. XML dokumenta atribūti (1)

- “Nosaukums - Vērtība” pāris saistīts ar elementu
- Sniedz papildus informāciju par elementu
- Pievienojami tikai elementa sākuma tagam
- Piemērs:

<name firstname="Vladimir"></name>

24. XML dokumenta atribūti (2)

- Atribūtam jābūt kādai vērtībai (vai "" virknei)
 - Atribūta nosaukums ir unikāls elementa ietvaros
 - Atribūta vērtībai var izmantot saskaņotus " vai '
 - Nav svarīga kārtība kādā atribūti ir iekļauti elementā
 - XML parsētājs negarantē secību elementa atribūtiem
- <name first="John" middle="Fitzgerald Johansen" last="Doe"></ name>**

24. XML dokumenta atribūti (3)

- Atribūti izmanto mazāk vietas par elementiem
- Elementi var būt sarežģītāki organizācijā
- Atribūti nevar saturēt vairākas vērtības - elementi var
- Atribūti nevar saturēt hierarhisko struktūru - elementi var
- Atribūtu kārtību nevar piespiest – elementu kārtību var



24. XML dokumenta atribūti (4)

```
<note date="10/01/2008">  
  <to>Tove</to>  
  <from>Jani</from>  
  <heading>Reminder</heading>  
  <body>Don't forget me this weekend!</body>  
</note>
```



```
<note>  
  <date>  
    <day>10</day>  
    <month>01</month>  
    <year>2008</year>  
  </date>  
  <to>Tove</to>  
  <from>Jani</from>  
  <heading>Reminder</heading>  
  <body>Don't forget me this weekend!</body>  
</note>
```

24. XML dokumenta atribūti (5)

- XML atribūti meta-datiem par elementu

```
<messages>
  <note id="501">
    <to>Tove</to>
    <from>Jani</from>
    <heading>Reminder</heading>
    <body>Don't forget me this weekend!</body>
  </note>
  <note id="502">
    <to>Jani</to>
    <from>Tove</from>
    <heading>Re: Reminder</heading>
    <body>I will not</body>
  </note>
</messages>
```

25. Komentāri

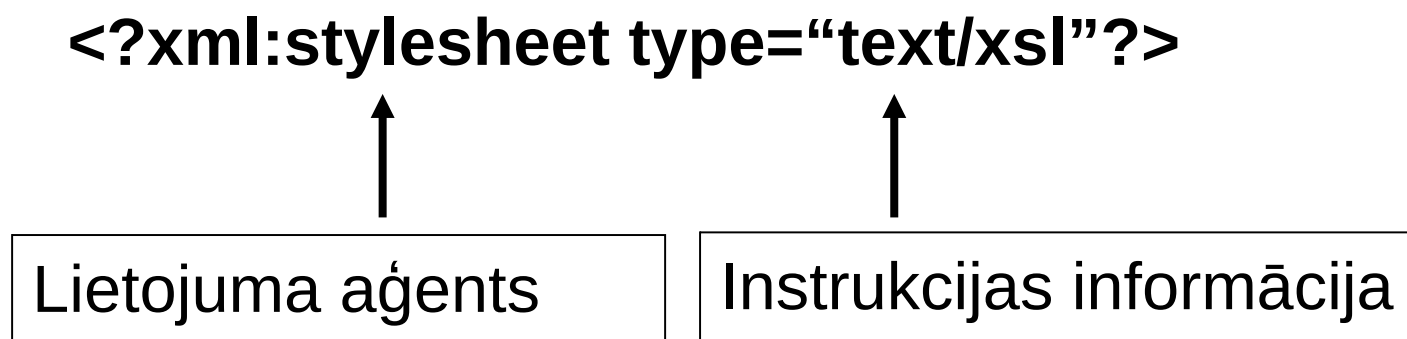
- XML komentārus ignorē lietojumi, kas apstrādā XML dokumentu.
- XML komentāri ir noderīgi
 - Dokumentācijai
 - Papildus informācijas dokumenta skatītājiem
- Sintakse



<!-- Komentārs -->

26. XML apstrādes instrukcijas

- Papildus informācija lietojumam, kurš izmanto XML dokumentu
- Norādījumus parsētājs tieši nodod lietojumam
- XML deklarācijā ir arī definēti apstrādes aģenti



27. XML deklarācija

- Apraksta detaļas XML parsētāja sagatavošanai darbam ar dokumentu
- XML deklarācija jābūt faila sākumā
- Sintakse:

**<?xml version='1.0'
encoding='UTF-8' standalone='yes' ?>**

- Tikai versijas atribūts ir obligāts
 - versija ir 1.0 vai 1.1

28. Standalone dokumenta atribūts

- Standalone = {yes or no}
 - **yes**: specificē, ka dokumenti pastāv neatkarīgi no citiem failiem
 - **no**: norāda, ka dokuments ir atkarīgs no ārējiem dokumentiem (piemēram, XML shēmas definīcijas).



29. XML rakstzīmju kodēšana (1)

- XML dokumenti var saturēt Unicode rakstzīmes
 - piemēram, Norvēģijas æ o, vai franču ê è é



<?xml version='1.0' encoding='UTF-16' ?>

<?xml version='1.0' encoding='UTF-8' ?>

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

29. XML rakstzīmju kodēšana (2)

- Rakstzīmju kodēšana ir metode, ko izmanto, lai pārstāvētu rakstzīmju koda numurus digitāli
 - cik baitu jāizmanto katram rakstzīmes numuram
- ASCII: apzīmē jebkuru rakstzīmi ar baitu
- UTF-16: izmanto divus baitus katrai rakstzīmei



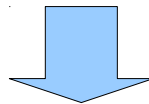
29. XML rakstzīmju kodēšana

- UTF-8 izmanto vienu baitu burtiem no ASCII
 - citas rakstzīmes var pārstāvēt ar diviem vai vairākiem baitiem
- UTF-8 v. UTF-16:
 - UTF-8 radīs mazāku failu angļu valodā
 - teksts citās valodās ar UTF-16 var būt mazāks (jo UTF-8 var prasīt trīs vai vairāk baitu atsevišķām rakstzīmēm)

30. XML dokumenta teksta dati (1)

- Teksta dati tiek definēti starp sākuma un noslēguma elementa tagiem
- Viss XML dokumenta teksts tiek parsēts
- Rakstzīmes "<" un "&" nav atļautas XML elementu tekstā
- Kad tiek parsēts XML elements, tiek arī parsēts teksts starp XML tagiem, jo XML element var saturēt citus elementus

```
<name><first>Bill</first><last>Gates</last></name>
```



```
<name>  
  <first>Bill</first>  
  <last>Gates</last>  
</name>
```

30. XML dokumenta teksta dati (2)

- PCDATA
 - (angl.) Parsed Character Data
 - XML parsētāji parasti izanalizē visu tekstu XML dokumentā
- CDATA
 - (angl.) Character Data
 - Teksts CDATA sadaļā netiek parsēts
 - CDATA sadaļa sākas ar "<[CDATA" ["!" un beidzas ar "]]>"

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

31. XML atsauces (1)

- Entity References (angl.)
- "<" simbols XML elementa iekšā radīs kļūdas, jo parsētājs to interpretē kā cita elementa sākumu
- Lai novērstu kļūdu, nomainiet "<" raksturu uz atsauci

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

31. XML atsauces (2)

- 5 iepriekš noteiktas atsauces XML
- Var definēt jaunas atsauces, lai izvairītos no teksta fragmentu atkārtotības dokumentā

<	<
>	>
&	&
'	'
"	"

```
<?xml version="1.0"?>
<!DOCTYPE CUSTOMERS [
<!ENTITY client "15 Downing St Floor 2">
]>
<message>
    <opening>Dear &client;</opening>
</message>
```


32. XML validācija (1)

- XML dokuments ar korektu sintaksi ir “Korekti noformēts”
- XML dokuments kurš atbilst XML shēmas noteikumiem ir “Valīds”



32. XML validācija (2)

- XML dokumenta sintakses pārbaude
 - dokumenta saknes elements
 - datu korekta hierarhiju
 - atribūtu sintakses korektums

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

33. Noderīgi XML Standarti

- XMLNS (XML Namespaces)
 - ļauj vienā dokumentā izmantot XML elementus un atribūtus no vairākām vārdnīcām bez jebkādas nosaukšanas kolīzijām
- XSLT (XML Transformations)
 - izmantota, lai transformētu XML dokumentus citā XML dokumentā, HTML, unc
- XPATH, XQUERY
 - definē sintaksi XPath izteiksmēm, kas identificē vienu vai vairākas iekšējas dokumenta daļas (elementi, atribūti, unc)
- XML Signature
 - definē sintaksi un noteikumus digitālam parakstam uz XML

34. XML kritika

- XML tiek regulāri kritizēts par runīgumu un sarežģītību
- XML koka kartēšana uz programmēšanas valodas tipu sistēmām varētu būt sarežģīta
- Par alternatīvām bieži tiek piedāvāti
 - JSON (angl. JavaScript Object Notation)
 - YAML (angl. YAML Ain't Markup Language)



```
{
  "firstName": "John",
  "lastName": "Smith",
  "age": 25,
  "address": {
    "streetAddress": "21 2nd Street",
    "city": "New York",
    "state": "NY",
    "postalCode": 10021
  },
  "phoneNumbers": [
    {
      "type": "home",
      "number": "212 555-1234"
    },
    {
      "type": "fax",
      "number": "646 555-4567"
    }
  ]
}
```

35. Kopsavilkums (1)



Iezīmēšanas valoda ir vārdu un simbolu kopa, lai aprakstītu dokumenta daļas

XML ir paplašināmās iezīmēšanas valoda, “vienkāršota” SGML versija

XML ir standarts vispārēja nolūka datu transportēšanai un uzglabāšanai

35. Kopsavilkums (2)



- XML nav paredzēts aizvietot HTML
- XML paredzēts transportēt un uzglabāt datus, koncentrējoties uz to, kādi dati ir
- Parsēšana ir process, kurā XML sintakse ir pārbaudīta un teksts tiek sadalīts atsevišķās, pārvaldāmās daļās

35. Kopsavilkums (3)



- XML dokuments sastāv no deklarācijas, elementiem, atribūtiem, komentāriem un atsaucēm
- XML tagi veido dokumenta iezīmēšanu
- XML dokumentam ir hierarhiskā struktūra, kas sākas “saknes” elementā un atzarojās “lapās”
- XML dokuments ar korektu sintaksi ir “Korekti noformēts”
- Korēkti noformēts XML dokuments validēts pēc XML shēmas ir “Valīds”

Praktiskais uzdevums

