XML Vārdu Telpas un XML Schema

Plāns



- XML vārdu telpas
- XML dokumentu shēma un validācija
- Ieskats DTD dokumentu tipu definīcijā
- XML Schema
- XML Schema funkcijas un labumi
- XML Schema dokuments
- Vienkārša <element> deklarēšana
- Sarežģīta <element> deklarēšana
- Piemēri

1. XML vārdu telpas

- XML Namespaces (angl.)
- W3C rekomendācija

http://www.w3.org/TR/REC-xml-names/

- XML vārdu telpas ļauj izvairīties no dokumenta elementu vārdu konfliktiem
- Tiek izmantotas, lai atšķirtu elementus un atribūtus no dažādiem XML dokumentu veidiem (vārdnīcām), to apvienojot vienā dokumentā



2. Vārdu konflikts

- XML elementu vārdus definē izstrādātājs
- Vārdu konflikts var rasties apvienojot dažādus XML dokumentus
- Piemērā XML parsētājam nav starpības starp <title> dokumenta elementiem

```
<person>
   <name>
       <title>Sir</title>
       <first>John</first>
       <middle>Fitzgerald Johansen</middle>
       <last>Doe</last>
   </name>
   <position>Vice President of Marketing</position>
   <résumé>
       <html>
            <head>
               <title>Resume of John Doe</title>
            </head>
            <body>
                <h1>John Doe</h1>
                John's a great guy, you know?
           </body>
       </html>
   </résumé>
</person>
```

3. XML elementu prefikss (1)

XML elementu vārdu konflikta atrisināšanai katram elementam jābūt unikāli identificējamam

Jāizmanto prefiksu elementa vārdā

prefix:elementName/>



3. XML elementu prefikss (2)

```
<pers:person>
    <pers:name>
        <pers:title>Sir</pers:title>
        <pers:first>John</pers:first>
        <pers:middle>Fitzgerald Johansen</pers:middle>
        <pers:last>Doe</pers:last>
    </pers:name>
    <pers:position>Vice President of Marketing</pers:position>
    <pers:résumé>
        <xhtml:html>
            <xhtml:head>
                <xhtml:title>Resume of John Doe</xhtml:title>
            </xhtml:head>
            <xhtml:body>
                <xhtml:h1>John Doe</xhtml:h1>
                <xhtml:p>John's a great guy, you know?</xhtml:p>
            </xhtml:bodv>
        </xhtml:html>
    </pers:résumé>
</pers:person>
```

4. XML vārdu telpas (1)



- Lai izmantotu XML prefiksu, tam jādefinē vārdu telpa
- XML Namespaces rekomendācijas no W3C nosaka standartu vārda telpu deklarēšanai un elementa vai atribūta vārda telpas identificēšanai
- XML vārdu telpu izmantošanai dokumentā elementiem jāizmanto kvalificētos vārdus (QName)
- QName sastāv no divām daļām
 - Elementa lokāla daļa (<prefix:<u>elementName</u>>)
 - Vārda telpas prefikss (<<u>prefix</u>:elementName>)

4. XML vārdu telpas (2)

- XML vārdu telpa ir definēta izmantojot xmlns atribūtu elementa sākuma tagā
- XML vārdu telpas definēšanas sintakse

xmlns:prefix="URI"

 Lai deklarētu vārdu telpu "http://www.wiley.com/pers" un saistītu <person> elementu ar to, ir jādefinē:

<pers:person</pre>

xmlns:pers="http://www.wiley.com/pers"/>

4. XML vārdu telpas (3)

 Prefiksu var izmantot jebkurš no <pers:person> elementa pectečiem, lai indicētu to piederību http://www.wiley.com/pers vārdu telpai

5. XML vārdu telpa pēc noklusēšanas (1)

 Vārdu telpā pēc noklusēšanas nav jānorāda prefiksu visiem elementiem, kuri to izmanto



5. XML vārdu telpa pēc noklusēšanas (2)

 Elementam var definēt vairāk par vienu vārdu telpām, bet tikai viena varbūt pēc noklusēšanas

5. XML vārdu telpa pēc noklusēšanas (3)

- XML vārdu telpu pēc noklusēšanas var definēt elementam un tā pēctečiem
- Piemērā vārdu telpa pēc noklusēšanas ir pārdefinēta elementā uz http://www.w3.org/1999/xhtml.

```
<person xmlns="http://www.wiley.com/pers">
     <name />

          This is XHTML

</person>
```

6. Atšķirīgas notācijas – vienāds rezultāts

```
<pers:person xmlns:pers="http://www.wiley.com/pers"</pre>
xmlns:xhtml="http://www.w3.org/1999/xhtml">
<pers:name/>
<xhtml:p>This is XHTML</xhtml:p>
</pers:person>
<person xmlns="http://www.wiley.com/pers"</pre>
xmlns:xhtml="http://www.w3.org/1999/xhtml">
<name/>
<xhtml:p>This is XHTML</xhtml:p>
</person>
    <person xmlns="http://www.wiley.com/pers">
    <name/>
    This is XHTML
    </person>
```

7. XML vārdu telpu URI (1)

- Uniform Resource Identifier (angl.)
- Rakstzīmju virkne resursu identificēša

```
• URL (angl. Uniform Resource Locator) version="1.0" encoding="UTF-8" standalone="yes"?>
<Response xmlns="http://tempuri/XMLSchemas/v1"</pre>
          xmlns:ns2="http://ivis.eps.gov.lv/XMLSchemas/100017/fidavista/v1-2">
    <AccountStatementResponse>
        <ns2:FIDAVISTA>
                                                                           4
            <ns2:Header>
                <ns2:Timestamp>20130312214257000/ns2:Timestamp>
                <ns2:From>PARXLV22</ns2:From>
            </ns2:Header>
            <ns2:Statement>
                <ns2:Period>
                     <ns2:StartDate>2013-01-01/ns2:StartDate>
                     <ns2:EndDate>2013-02-01
                    <ns2:PrepDate>2013-09-27</ns2:PrepDate>
                </ns2:Period>
                <ns2:AccountSet>
```

<ns2:IBAN>LV11PARX000123456789/ns2:IBAN>

7. XML vārdu telpu URI (2)

URN (angl., Universal Resource Name).

7. XML vārdu telpu URI (3)

- W3C rekomendācija izmanto vārdu telpas nosaukuma terminu
- Bieži vārdu telpas URI definēšanai tiek izmantots URL HTTP shēmā (piem., http://www.w3.org/1999/xhtml), neskatoties uz XSD formālas saistības trūkumu ar HTTP protokolu
 - Izvēlēta URI apraksta resursu, kas atrodas XML vārdnīcas autora kontrolē. Piemēram, autora tīmekļa servera URL

Piemērs: XML vārdu telpas XSLT

```
<xsl:stylesheet version="1.0"</pre>
  xmlns:xsl="http://www.w3.org/1999/XSL/Transform">
  <xsl:template match="/">
     <html>
        <body>
           <h2>My CD Collection</h2>
           Title
                 Artist
              <xsl:for-each select="catalog/cd">
                 >
                       <xsl:value-of select="title" />
                    >
                       <xsl:value-of select="artist" />
                    </xsl:for-each>
           </body>
     </html>
  </xsl:template>
```

Piemērs: ISO20022 maksājumu iniciēšana

```
<?xml version="1.0" encoding="UTF-8"?>
<Document xmlns="urn:iso:std:iso:20022:tech:xsd:pain.001.001.03"</pre>
          xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
          xsi:schemaLocation="urn:iso:std:iso:20022:tech:xsd:pain.001.001.03
    <CstmrCdtTrfInitn>
        <GrpHdr>
            <MsqId>87fbf20111125/1</MsqId>
            <CreDtTm>2011-11-25T11:16:58.696</CreDtTm>
            <Nb0fTxs>3</Nb0fTxs>
            <CtrlSum>2500</CtrlSum>
            <InitgPty>
                <Nm>AS XML</Nm>
            </InitgPty>
        </GrpHdr>
        <PmtInf>
            <PmtInfId>PMTID001/PmtInfId>
            <PmtMtd>TRF</PmtMtd>
            <BtchBookg>true</BtchBookg>
            <Nb0fTxs>3</Nb0fTxs>
            <PmtTpInf>
                <SvcLvl>
                    <Cd>SEPA</Cd>
                </SvcLvl>
```

Piemērs: XHTML 1.1 + MathML 2.0 + SVG 1.1

```
<?xml version="1.0"?>
<html xmlns="http://www.w3.org/1999/xhtml"</pre>
     xmlns:svg="http://www.w3.org/2000/svg">
    <body>
       <h2>MathML</h2>
        >
           <math xmlns="http://www.w3.org/1998/Math/MathML">
               <mfrac>
                   <mi>a</mi>
                   <mi>b</mi>
               </mfrac>
           <h2>SVG</h2>
        >
           <svg:svg width="50px" height="50px">
               <svg:circle cx="25px" cy="25px" r="20px" fill="green" />
           </svg:svg>
        </body>
</html>
```

Piemērs: XHTML 1.1 + MathML 2.0 + SVG 1.1



MathML

a b

SVG



Praktiskais uzdevums



XML Dokumentu Shēma

DTD XML Schema



1. XML shēma

- XML dokumenta kvalitātes kontroles rīks
- Apraksta XML dokumenta struktūru
- Ļauj pārliecināties, ka XML dokumenta instance atbilst sintakses un struktūras noteikumu kopai
- XML shēmu var uztvert par programmu parsētājam kā lasīt dokumentu

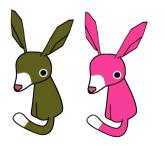
2. XML shēmas valodas

- DTD (angl. Document Type Definition)
 - vecākā XML dokumenta shēmas valoda
 - speciālā ne-XML sintakse



- W3C XML Schema
 - Pilnvērtīgs XML dokuments
- RELAX NG
- Schematron





3. XML dokumenta shēmas validācija

 Derīgs XML dokuments ir korekti noformēts XML dokuments, kurš arī atbilst XML dokumenta shēmas (DTD, XML Schema) noteikumiem



Validācijas process

- 1. Parsētājs ielāde shēmas noteikumus un deklarācijas
- 2. Tiek konstruēts speciāla tipa validācijas parsētājs
- 3. Parsētājs pieņem XML dokumenta instanci un ģenerē tas validācijas atskaiti

4. leskats DTD dokumentu tipu definīcijā

(angl.) Document Type Definition

- Nosaka
 - atlautus XML dokumenta blokus
 - dokumenta struktūru, izmantojot sarakstu ar atļautiem dokumenta elementiem un atribūtiem

 DTD var specificēt XML dokumenta iekšā (internal) vai atsevišķajā failā (external)

4. leskats DTD dokumentu tipu definīcijā

- lekšēja DTD piemērs
- <!DOCTYPE root-element[element-declaration]>

5. XML Schema (1)

- "XML Schema" ir W3C tehnoloģija XML dokumenta shēmas aprakstam
- Viens no svarīgiem XML tehnoloģiju pamatiem
- Līdzīgi DTD XML Schema ļauj aprakstīt XML dokumenta struktūru
- XML Schema ir DTD alternatīva un pēctecis balstīts uz XML

5. XML Schema (2)

- Nodrošina datu atbilstību specifikācijai
- XML Schema valodu parasti sauc arī par XSD (angl. XML Schema Definition)

- XSD ir XML valoda
 - izmanto XML vārdu telpas
- XSD ir specifiskāka par DTD
 - izmanto datu tipus



6. XML Schema pamata funkcijas

- Definē kādi elementi var parādīties dokumentā
- Definē kādi elementu atribūti var parādīties dokumentā
- Definē kādu elementus var iekļaut citos elementos
- Definē elementu kārtību
- Nosaka iespējamo elementu skaitu
- Definē vai elements ir tukšs vai var iekļaut tekstu
- Definē datu tipus elementiem un atribūtiem
- Definē vērtības pēc noklusēšanas elementiem un atribūtiem

7. XML Schema labumi (1)

- XML Schema tiek izveidota izmantojot XML pamata sintaksi (atšķirībā no DTD speciālās sintakses)
- XML Schema pilnīgi atbalsta W3C Namespace rekomendācijas
- Atbalsta teksta elementu satura validāciju izmantojot iebūvētus un lietotāja definētus datu tipus

7. XML Schema labumi (2)

- Atļauj vienkārša veidā veidot sarežģītus un atkārtoti lietojamus satura modeļus (piem., ISO20022 ziņojumu standarts)
- Atbalsta programmēšanas valodu koncepciju modelēšanu (piem. mantošanas saistības no objekt-orientētās programmēšanas)



7. XML Schema labumi (3)

ISO20022 ziņojumu formāta atkārtotā lietošana

8. XML Schema vs. DTD (1)

- XML Schema ir paplašināma nākotnē
- XML Schema mehānism ir jaudīgāks
 - var atkārtoti lietot shēmas
 - var definēt speciālus datu tipus, atvasinot no iebūvētiem
 - var norādīt vairākas shēmas vienā dokumentā.



8. XML Schema vs. DTD (2)

- XML Schema atbalsta datu tipus
 - Vieglāk aprakstīt atļauto dokumentu saturu
 - Vieglāk pārbaudīt datu pareizumu
 - Vieglāk apstrādāt datus no datu bāzēs
 - Vieglāk aprakstīt datu ierobežojumus un šablonus
 - Vieglāk konvertēt datus
- XML Schema tiek rakstīta XML valodā
 - Nav nepieciešamības mācīties jauno valodu
 - Var izmantot XML redaktoru un parsētāju
 - Var transformēt dokumenta shēmas aprakstu citā formātā

9. XML Schema standarts

- XML Schema standarts ir sadalīts trīs daļās:
 - XML Schema pamata koncepcijas
 - www.w3.org/TR/xmlschema-0/

- Struktūras, izmantotas XML Schemas
 - www.w3.org/TR/xmlschema-1/

- XML Schema data tipu apraksts
 - www.w3.org/TR/xmlschema-2/

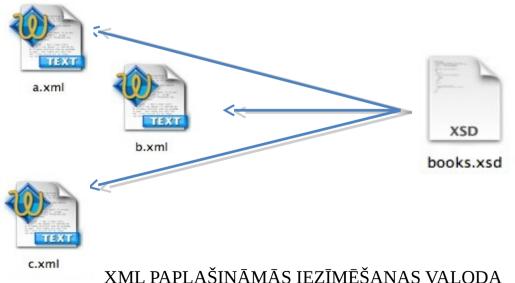
10. XML Schema dokuments (1)

- XML Schema tiek saglabāta atsevišķajā XML dokumentā
- XML dokuments kurš ir uzbūvēts atbilstoši noteiktai XML Schema shēmas vārdnīcai tiek saukts par XML Schema instances dokumentu
- XML Schema faili parasti izmanto .xsd paplašinājumu
- "schema" sāknes elements

10. XML Schema dokuments (2)

 XML dokumenti satur norādi uz XML Schema dokumentu, kas nosaka to vārdnīcu

<note xmlns="http://www.w3schools.com" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLocation="http://www.w3schools.com note.xsd">



10. XML Schema dokuments (3)

```
<?xml version="1.0" encoding="UTF-8"?>
<schema xmlns="http://www.w3.org/2001/XMLSchema"</pre>
    targetNamespace="http://www.example.org/NewXMLSchema"
    xmlns:tns="http://www.example.org/NewXMLSchema"
                                                                               (nameType)
                                                             e name
    elementFormDefault="qualified">
                                                                               (a) title
                                                                                      string
    <element name="name">
                                                                               e first
                                                                                      string
                                                                               e middle
                                                                                      string
        <complexType>
                                                                               e last
                                                                                      string
             <sequence>
                 <element name="first" type="string" />
                 <element name="middle" type="string" />
                 <element name="last" type="string" />
             </sequence>
             <attribute name="title" type="string" />
        </complexType>
    </element>
</schema>
```

11. <schema> elements (1)

 <schema> elements ir XML Schema dokumenta sāknes elements

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

<xs:schema

xmlns:xs="http://www.w3.org/2001/XMLSchema">

</xs:schema>

11. <schema> elements (2)

<schema> elementā definējami papildus atribūti

```
<?xml version="1.0"?>
<xs:schema
xmlns:xs= "http://www.w3.org/2001/XMLSchema"
targetNamespace= "http://www.w3schools.com"
xmlns= "http://www.w3schools.com"
elementFormDefault= "qualified">
</r></re>/xs:schema>
```

12. XML Schema vārdu telpa (1)

Var izmantot jebkuru no variantiem

```
<schema
xmIns="http://www.w3.org/2001/XMLSchema">

<xs:schema
xmIns:xs="http://www.w3.org/2001/XMLSchema">

<xsd:schema
xmIns:xsd="http://www.w3.org/2001/XMLSchema">
```

12. XML Schema vārdu telpa (2)

 Nosaka, ka elementi un datu tipi izmantoti XML shēmas definīcijā nāk no XMLSchema vārdnīcas (http://www.w3.org/2001/XMLSchema)

 Nosaka, ka elementi un datu tipi no "http://www.w3.org/2001/XMLSchema" jāraksta ar "xs:" vai "xsd:" prefiksu

13. Mērķa vārdu telpa (targetNamespace)

- Pirmatnēja XML Schema nozīme ir XML vārdnīcu deklarēšana vārdu telpai, specificētai "targetNamespace" atribūtā
- Deklarējot mērķa vārdu telpu ir svarīgi iekļaut atbilstošas deklarācijas

<xs:schema

xmlns:xs="http://www.w3.org/2001/XMLSchema"

targetNamespace="http://www.example.org/NewXMLSchema"

xmlns:tns="http://www.example.org/NewXMLSchema">

14. XML Schema datu tipi

Atļauj vairāk specifikas par DTD

 Var norādīt datumus, numurus, diapazonus, unc

Var iedalīt divās kategorijās:

Vienkāršie - nodarbojas ar pamatvērtībām

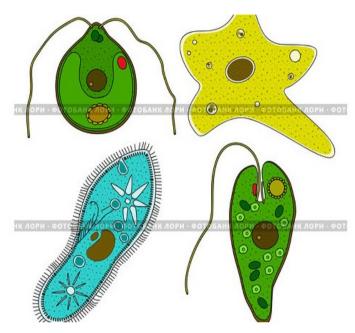
Sarežģītie - apraksta sarežģītākas konstrukcijas

Complicated

Simple

15. Vienkāršie datu tipi (1)

- Apraksta tekstu, ciparus, datumus
- Bieži sauktas par "primitīvajiem"
- Vienkāršo datu tipu elementiem nevar būt atribūtu
- lebūvētie datu tipi ir XML
 Schema vārdu telpā



15. Vienkāršie datu tipi (2)

String Date/time

Boolean Time

Numbers TimeInstant

Integer Duration

Decimal

Float

Double

Custom (simpleType)

16. Vienkāršā <element> deklarēšana (1)

<xs:element name="NAME" type="TYPE"/>

- NAME elementa nosaukums,
 jāseko XML noteikumiem elementu nosaukumiem
- TYPE elementa datu tips
- XML Schema definē vairākus iebūvētus datu tipus
 - xs:string
 - xs:decimal
 - xs:integer
 - xs:boolean
 - xs:date
 - xs:time

```
<lastname>Jones</lastname>
<age>36</age>
<dateborn>1970-03-27</dateborn>
```

```
<xs:elementname = lastname " type="xs:string" />
<xs:elementname = age " type="xs:integer" />
<xs:elementname = dateborn " type="xs:date" />
```

16. Vienkāršā <element> deklarēšana (2)

```
<xsd:element name="firstname"</pre>
               type="xsd:string" />
<xsd:element name="ableToSwim"</pre>
               type="xsd:boolean" />
<xsd:element name="date"</pre>
               type="xsd:date" />
```

16. Vienkāršā <element> deklarēšana (3)

 Kardinalitāte (angl. cardinality) nosaka cik reizes specifisks elements var parādīties satura modelī

- x : minimālais elementa parādīšanas skaits
- y : maksimālais elementa parādīšanas skaits

16. Vienkāršā <element> deklarēšana (4)

- Vienkāršiem elementiem XML shēmā var definēt vērtības pēc noklusēšanas vai nemainīgas vērtības
 - Vērtība pēc noklusēšanas tiek automatiski piešķirta elementam, kad netiek specificēta cita vērtība

<xs:element name="color" type="xs:string" default="red"/>

Nemainīgas vērtības tiek automātiski piešķirtas elementam.
 Elementa vērtību nevar izmainīt uz citu.

<xs:elementname="color" type="xs:string" fixed="red"/>

17. Vienkāršais vs. sarežģītais

- Vienkāršais
 - <grade>7</grade>
 - nav citu iekļauto elementu



- Sarežģītais
 - <students><student>Jack</student></students>
 - ir iekļautie elementi

18. Sarežģīta elementa deklarēšana (1)

- Sarežģīts XML elements (angl. complex element) satur citus elementus un/vai atribūtus
- Ir četri sarežģīto elementu tipi:
 - Tukšie elementi
 - Elementi, kuri ietver tikai elementus
 - Elementi tikai ar tekstu
 - Elementi, kuri ietver gan tekstu gan citus elementus
- NB! Katrs no tiem elementu tipiem var ietvert arī atribūtus



18. Sarežģīta element deklarēšana (2)

- Tukšie elementi
 - Sarežģīts XML elements, "product", kurš ir tukšs
 product pid="1345"/>

 Sarežģīts XML elements "employee", kurš satur tikai citus elementus

```
<employee>
    <firstname>John</firstname>
    <lastname>Smith</lastname>
</employee>
```

18. Sarežģīta element deklarēšana (3)

 Sarežģīts XML elements "description", kurš satur gan tekstu gan citus elementus

<description> It happened on
 <date lang="norwegian">03.03.99</date>
</description>

19. <complexType> deklarēšana (1)

"employee" elementu tipa deklarēšana

19. <complexType> deklarēšana (2)

- "employee" elementu tipa deklarēšana
 - Tikai "employee" elements var izmantot deklarēto sarežģītu tipu
 - Elementi "firstname" un "lastname" apvienoti zem viena <sequence> indikatora, kas nozīmē, ka dokumentā tiem jāparādās tādā pati kārtībā, kā deklarēti shēmā

20. <complexType> deklarēšana (3)

```
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema"</pre>
            targetNamespace="http://www.example.org/employee"
            xmlns:tns="http://www.example.org/employee"
            elementFormDefault="qualified">
            <xs:element name="employee" type="tns:EmployeeType" />
            <xs:complexType name="EmployeeType">
                <xs:sequence>
                    <xs:element name="firstname" type="xs:string" />
                    <xs:element name="lastname" type="xs:string" />
                </xs:sequence>
            </xs:complexType>
            <xs:element name="department">
                <xs:complexType>
                    <xs:sequence>
                         <xs:element name="employee"</pre>
                         type="tns:EmployeeType" maxOccurs="unbounded"/>
                    </xs:sequence>
                </xs:complexType>
            </xs:element>
Vladimir Ko
```

21. <attribute> deklarēšana

XML atribūta definēšanas sintakse:

<xs:attributename="ANAME"
type="ATYPE"/>

- ANAME atribūta nosaukums
- ATYPE atribūta datu tips

Vienkāršajiem alementiem nevar būt atribūti

22. Kopsavilkums

- XML Schema ir W3C XML tehnoloģijas XML dokumentu shēmu definēšanai
- Apraksta XML dokumenta struktūru
- Apraksta sintakses un struktūras noteikumu kopu

- lebūvēto bāzes tipu pārveidošana
 - Vērtību ierobežojumi
 - Diapazoni
 - Saraksti

```
<xsd:simpleType="Degrees">
  <xsd:restriction base="xsd:string">
  <xsd:enumeration value="AA" />
  <xsd:enumeration value="AS" />
  </xsd:restriction>
</xsd:simpleType>
```

Saraksts (angl. enumeration)

```
<xsd:simpleType name="car">
    <xsd:restriction base="xsd:string">
        <xsd:enumeration value="Audi"/>
        <xsd:enumeration value="Golf"/>
        <xsd:enumeration value="BMW"/>
        </xsd:restriction>
</xsd:simpleType>
```

Šablons ar REGEX

```
<xsd:simpleType>
     <xsd:restriction base="xsd:string">
          <xsd:pattern value="[a-z]"/>
          </xsd:restriction>
</xsd:simpleType>
```

REGEX Examples

```
<xs:pattern value="[A-Z][A-Z][A-Z]"/>
<xs:pattern value="[a-zA-Z][a-zA-Z][a-zA-Z]"/>
<xs:pattern value="[xyz]"/>
<xs:pattern value="[0-9][0-9][0-9][0-9][0-9]"/>
<xs:pattern value="([a-z])*"/>
<xs:pattern value="male|female"/>
<xs:pattern value="[a-zA-Z0-9]{8}"/>
```

Elementu kārtībā

- Sequence: elements parādās tajā pašā kārtībā, kā ir definēts XML Schema
- All: elementi parādās jebkura kārdība

Tukšs elements atribūtu

- XML
 - <student id="A1" />

Schema

Praktiskais uzdevums

