

XML standardeja

- nimiavaruudet, namespaces
- XHTML
- XML Schema
- linkitys

Namespaces - ongelma

```
<?xml version="1.0"?>
<person>
  <name>
    <title>Mr. President</title>
    <first>Zaphod</first>
  <last>Beeblebrox</last>
</name>
<position>President of the Galaxy</position>
  <résumé>
    <html>
      <head><title>Resume of Zaphod Beeblebrox</title>
    </head>
      <body>
        <h1> Resume of Zaphod Beeblebrox</h1>
        <p>Zaphod's a great guy, you know?</p>
      </body>
    </html>
  </résumé>
</person>
```

Ratkaisu 1:

```
<?xml version="1.0" encoding="UTF-8"?>
<ps:person xmlns:ps="http://frogstar.mil/people"
           xmlns:html="http://www.w3.org/1999/xhtml">
  <ps:name>
    <ps:title>Mr. President</ps:title>
    <ps:first>Zaphod</ps:first>
    <ps:last>Beeblebrox</ps:last>
  </ps:name>
  <ps:position>President of the Galaxy</ps:position>
  <ps:r  sum  >
    <html:html>
      <html:head><html:title> Resume of Zaphod Beeblebrox
</html:title></html:head>
      <html:body>
        <html:h1>Zaphod Beeblebrox</html:h1>
        <html:p>Zaphod's a great guy, you know?</html:p>
      </html:body>
    </html:html>
  </ps:r  sum  >
</ps:person>
```

URL, URI ja URN

- URL Uniform Resource Locator, osoite, josta kyseinen resurssi löytyy
[Scheme]://[Domain]:[Port]/[Path]?[QueryString]#[FragmentID]
- URI Uniform Resource Identifier; yksiselitteinen merkkijono, joka ei välttämättä osoita mihinkään;
- sekä URL että URN ovat myös URIja
- URN Uniform Resource Name
urn:[namespace identifier]:[namespace specific string]
- esim: urn:isbn:87445536911

Namespaces, nimiavaruudet - teoriaa

- Nimiavaruus on puhtaasti käsitteellinen asia: sillä ryhmitellään joukko nimiä yhteenkuuluviksi URIn avulla
- millä taataan, ettei samaa namespace-nimeä käytetä tarkoittamaan erilaisia nimikäytäntöjä?
- ratkaisu: otetaan käyttöön jo olemassaoleva osoitteisto eli domain name -hallinto: URL Uniform Resource Locator
- nettiosoitteissa voi olla XML:ssä kiellettyjä merkkejä, joten tehdään ensin viittaus ja käytetään sitä

```
<person xmlns="http://frogstar.mil/people"  
        xmlns:html="http://www.w3.org/1999/xhtml">
```

Nimiavaruudet

- yhdistelmädokumentit
 - joskus olisi toivottavaa että voisimme liittää mukaan
 - eri kielillä kirjoitettua merkkausta
 - HTML-taulukon XML-dokumenttiin
 - ongelma: nimien yhteentörmäykset
- DTD omistaa nimiavaruuden, jossa
 - kaikki elementtinimet ovat uniikkeja
 - kaikki tietyn elementin attribuuttinimet ovat uniikkeja
 - siten kaikki viittaukset elementtiin ja tietyssä elementissä olevaan attribuuttiin ovat yksiselitteisiä
- dokumentti voi sisältää tietoa, joka on esitelty useammassa nimiavaruudessa

Nimiavaruuden tunnistaminen

- Useimmat standardit löytyvät nykyään verkosta
 - esim. <http://www.w3.org/TR/REC-html40>
- Nimiavaruus-standardi käyttää URL-osoitteita yksiselitteisinä tunnuksina
- Nimiavaruutta käyttävällä sovelluksella ei kuitenkaan tarvitse olla yhteyttä internetiin, URL on (vain) hyödyllinen uniikki merkkijono (jossa ei edes tarvitse olla mitään)
- Merkitään prefix:name
- Lisäksi annetaan nimiavaruudelle sopiva lyhenne
 - `<X:html xmlns:X="http://www.w3.org/TR/REC-html40">`
 - `<xsl:stylesheet xmlns:xsl="http://www.w3.org/1999/XSL/Transform" version="1.0">`

Namespaces, oletusnimiavaruudet

- Oletusnimiavaruus:

```
<person xmlns="http://frogstar.mil/people"
        xmlns:xhtml="http://www.w3.org/1999/xhtml">
  <name/>
  <xhtml:p>Tähän XHTML-avaruuden elementti</xhtml:p>
</person>
```

- esitellään nimiavaruus ilman prefix-osaa:
 <book xmlns="...">
 <para>A normal paragraph</para>
 </book>
- oletusnimiavaruus voidaan muuttaa missä tahansa elementissä
- myös jonkun nimiavaruuden attribuutteja voidaan käyttää toisen nimiavaruuden elementeissä

Ratkaisu 2:

```
<?xml version="1.0" encoding="UTF-8"?>
<person xmlns="http://frogstar.mil/people"
        xmlns:html="http://www.w3.org/1999/xhtml">
  <name>
    <title>Mr. President</title>
    <first>Zaphod</first>
    <last>Beeblebrox</last>
  </name>
  <position>President of the Galaxy</position>
</position>
  <résumé>
    <html:html>
      <html:head><html:title>Resume of Zaphod Beeblebrox
        </html:title></html:head>
      <html:body>
        <html:h1>Zaphod Beeblebrox</html:h1>
        <html:p>Zaphod's a great guy, you know?</html:p>
      </html:body>
    </html:html>
  </résumé>
</person>
```

19.1.2013

XHTML

- XHTML - W3C:n suositus
- XML-DTD HTML-kielelle
- Siirryttäessä HTML:stä XML:ään voi joutua tarkistamaan että "vanha" HTML-tiedosto (todellakin) on
 - hyvinmuodostettu
 - validi
 - XHTML-spesifikaation mukainen
- Selaintuki vielä epäjohdonmukainen

XHTML standardit

- XHTML - W3C recommendations 1.0 ja 1.1 perustuvat HTML 4.0:ään Working draft 2.0
- XHTML 1.1 suunnattu modulaariseen käyttöön
 - XHTML 1.1 näyttää samantaiselta kuin XHTML 1.0 Strict
 - Laajennetut XHTML Family document types
 - Modulaarinen rakenne mahdollistaa käytön yhdessä muiden merkintäkielien kanssa
 - XHTML 1.1 plus MathML 2.0 document type on yksi laajennettu XHTML Family dokumenttityyppi

XHTML - hyvinmuodostuneisuus

- HTML-tiedostossa voi joutua
 - lisäämään puuttuvat lopputunnisteet
 - muuttamaan tunnisteet niin että elementit ovat sisäkkäin eivätkä limittäin
 - yhtenäistämään alku- ja lopputunnisteiden nimiosat (esim. `<code>` ja `</code>`, tai `<CODE>` `</CODE>`)
 - ympäröimään attribuuttien arvot lainausmerkeillä
 - muuttamaan tyhjät tunnisteet esim. `<hr>` muotoon `<hr/>`
 - jne., jne., ...

XHTML - validius

- HTML-tiedostoon lisätään
 - DOCTYPE-esittely (3 mahdollisuutta)
 - peruselementit:
 - `<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">`
 - löysempi:
 - `<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">`
 - kehykset:
 - `<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Frameset//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-frameset.dtd">`

XHTML - muut vaatimukset

- HTML-tiedoston juuren pitää olla html
- Nimiavaruus pitää määritellä juuressa ja se on
 - <http://www.w3.org/1999/xhtml>
- (Mahdollisesti voi myös lisätä tyylitiedosto-prosessointikäskyn tiedostoon)

XHTML, esimerkki

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE html
  PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
  "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">

<html xmlns="http://www.w3.org/1999/xhtml" xml:lang="en"
  lang="en">
  <head>
    <title>Virtual Library</title>
  </head>
  <body>
    <p>Moved to
      <a href="http://vlib.org/">vlib.org</a>.</p>
  </body>
</html>
```

XML Schema

- XML Schema Recommendation sisältää:
 - elementtityypit,
 - hahmonsovituskielesioppi,
 - Alielementtien järjestämisen määrittely siten että dokumentin rakenne voidaan kontrolloida tiukasti,
 - Elementtien valinta eri tavoin niin, että sama skeema tuottaa erilaisia dokumentti-ilmentymiä

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

```
<xsd:schema
```

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

```
<xsd:element name="letter">
```

```
</xsd:element>
```

```
</xsd:schema>
```



```
<?xml version="1.0"?>
<xsd:schema
  xmlns:xsd="http://www.w3.org/2001/XMLSchema"
  xmlns:target="http://www.someone.net/name"
  targetNamespace="http://www.someone.net/name"
  elementFormDefault="qualified" >
  <xsd:element name="letter">
  </xsd:element>
</xsd:schema>
```

Schema namespaces

```
<?xml version="1.0"?>  
<schema xmlns = http://www.w3.org/2001/XMLSchema  
  xmlns:pers ="http://frogstar.mil/pers"  
  targetnamespace = "http://frogstar.mil/pers">
```

In the document:

```
<?xml version="1.0"?>  
<person xmlns ="http://frogstar.mil/pers"  
  xmlns:xsi = http://www.w3.org/2001/XMLSchema-instance  
  xsi:schemaLocation ="http://frogstar.mil/pers people2.xsd"  
  version = "1.0">
```

DTD vai Schema?

- XML Schema
 - noudattaa XML määrittelyä
 - tukee nimiavaruuksia
 - sallii monimutkaiset elementtimäärittelyt
 - olioiden periytyvyys
- DTD
 - voi sisältyä XML dokumenttiin
 - sallii entiteetit
 - suuri määrä olemassaolevia DTD-määrittelyjä

Schema: elementtityypit

```
<xsd:simpleType name="userType">  
  <xsd:restriction base="xsd:string">  
    <xsd:pattern value="U\w{2,6}\d{2}" />  
  </xsd:restriction>  
</xsd:simpleType>
```

```
<xsd:complexType>  
  sisältää alielementtejä  
<sequence> tai <choice> tai <all>
```

Globaalit elementit ovat schema-elementin lapsia
Lokaalit elementit ovat muun elementin lapsia

Elementtien määrittely

```
<?xml version="1.0"?>
<xsd:schema xmlns:xsd="http://www.w3.org/2001/XMLSchema">
  <group name="nimiryhmä">
    <sequence>
      <element name="etunimi" type="string"/ >
      <element name="sukunimi" type="string"/ >
    </sequence>
  </group>
  <complexType name="nimiryhmä">
    <group ref="target:nimiryhmä"/>
    <attribute name="arvo" type="string"/ >
  </complexType>
  <element name="Nimet" type="target:nimiryhmä"/>
</schema>
```

Schema: attribuutit

```
<xsd:complexType name="name">  
  <xsd:sequence>  
    <xsd:element name="title" type="xsd:string" maxOccurs="1"  
      default="Miss"/>  
    <xsd:element name="firstname" type="xsd:string" minOccurs  
      ="2"/>  
    <xsd:element name="surname" type="xsd:string" />  
  </xsd:sequence>  
  <xsd:attribute name="gender" type="xsd:string"  
    default="female"/>  
</xsd:complexType>
```

```
<?xml version="1.0"?>
<xsd:schema xmlns:xsd="http://www.w3.org/2001/XMLSchema"
xmlns:target="http://www.someone.net/name"
targetNamespace="http://www.someone.net/name"
elementFormDefault="qualified" >
  <xsd:element name="header">
    <xsd:complexType>
      <xsd:sequence>
        <xsd:element name="name">
          <xsd:complexType>
            <xsd:sequence>
              <xsd:element name="title" type="xsd:string"
maxOccurs="1"/>
              <xsd:element name="surname" type="xsd:string" />
              <xsd:element name="firstname" type="xsd:string" />
            </xsd:sequence>
          </xsd:complexType>
        </xsd:element>
        <xsd:element name="address">
          <xsd:complexType>
            etc..
          </xsd:complexType>
        </xsd:element>
      </xsd:sequence>
    </xsd:complexType>
  </xsd:element>
</xsd:schema>
```

Dokumenttiesimerkki

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

```
<name
```

```
  xmlns="http://www.someone.net/name"
```

```
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
```

```
  xsi:schemalocation="http://www.someone.net/name name3.xsd">
```

```
  <firstname>Cindy</firstname>
```

```
...
```

```
</name>
```


GML schema -määrittely GML.XSD

```
<?xml version="1.0" encoding="UTF-8"?>
<xsd:schema targetNamespace="http://www.opengis.net/gml"
xmlns:xlink="http://www.w3.org/1999/xlink" xmlns:gml="http://www.opengis.net/gml"
xmlns:sch="http://www.ascc.net/xml/schematron" xmlns="http://www.w3.org/2001/XMLSchema"
xmlns:xsd="http://www.w3.org/2001/XMLSchema" elementFormDefault="qualified"
attributeFormDefault="unqualified" version="3.0.1">
  <xsd:annotation>
    <xsd:appinfo source="urn:opengis:specification:gml:schema-xsd:gml:v3.0.1">
      gml.xsd
    </xsd:appinfo>
    <xsd:documentation>
      Copyright (c) 2002 OGC, All Rights Reserved. Top level GML schema
    </xsd:documentation>
  </xsd:annotation>
  <!-- ===== -->
  <xsd:include schemaLocation="dynamicFeature.xsd"/>
  <xsd:include schemaLocation="topology.xsd"/>
  <xsd:include schemaLocation="coverage.xsd"/>
  <xsd:include schemaLocation="coordinateReferenceSystems.xsd"/>
  <xsd:include schemaLocation="observation.xsd"/>
  <xsd:include schemaLocation="defaultStyle.xsd"/>
  <!-- ===== -->
</xsd:schema>
```

XML Schema: kehitysnäkymät

- Olemassaoleva määrä jo suuri, tulee olemaan pitkään tärkeä
- Sallii vapauksia määrittelyssä:
 - Ei vaadi datatyyppimäärittelyä – tärkeä
 - Versioiden hallinta suunniteltava itse
 - Laajennettavuus: sallii laajennukset mutta ei tarkkaan määrittele niiden käyttöä
- Tuki työkaluissa jo hyvä, esim. Visual Studio .NET

Linkit XML:ssä

- XML-standardin linkitystapa
 - ei kiinteää linkkielementtiä
 - ID ja IDREF-tyyppiset attribuutit identifioivat linkkinä toimivan elementin nimiavaruuden kautta
 - ainoastaan samassa dokumentissa
- ID-arvot aina yksikäsitteisiä
- `<!ELEMENT chapter (...)>`
- `<!ATTLIST chapter target ID #REQUIRED>`

XLink

- XLink = XML Linking Language
 - mekanismi yhden kohteen sisäisiin tai kahden ja useamman kohteen välisiin linkkeihin.
- yksinkertaiset linkit (suunnilleen kuin HTML:ssä)
- laajennetut linkit
 - kaksisuuntaiset linkit
 - tyypitetyt linkit
 - myös yhdestä moneen ja monesta-moneen suhteet
 - linkit myös read-only-tiedostoista ja -tiedostoihin
 - linkkien seurantajärjestys

XLink, esimerkki

Yksinkertaisen linkin esimerkki

```
<?xml version="1.0"?>
<ad xmlns:xlink= "http://www.w3.org/1999/xlink">
<p>Älä vain koskaan vieraile
<omalinkki xlink:type="simple"
           xlink:href="http://frogstar.com">
näillä sivuilla!
</omalinkki>
I told you! </p>
<p> toinen mainos </p>
</ad>
```

Yllä vapaasti valittu elementti identifioidaan Xlink-nimiavaruudessa määritellyn type-attribuutin avulla.

XLink, esimerkki 2

Yksinkertaisen linkin esimerkki

```
<my:crossReference
  xlink:href="students.xml"
  xlink:role="http://www.ex.com/studentlist"
  xlink:title="Student List"
  xlink:actuate="onLoad"
  xlink:show="embed">
  Current List of Students
</my:crossReference>
```

href = kohteen osoite

role = osoite, jossa pidempi selitys kohteesta

title = kohteen lyhyempi selitys

actuate - seurataanko linkkiä heti (onLoad) vai myöh.
(onRequest)

show - miten tulos esitetään (replace, new, embed)

XLink, esimerkki 3

Laajennetun linkin esimerkki

```
<extendedlink xlink:type="extended">
  <loc      xlink:type="locator" xlink:href="..."
            xlink:label="parent" xlink:title="p1" />
  <loc      xlink:type="locator" xlink:href="..."
            xlink:label="parent" xlink:title="p2" />
  <loc      xlink:type="locator" xlink:href="..."
            xlink:label="child"  xlink:title="c1" />
  <loc      xlink:type="locator" xlink:href="..."
            xlink:label="child"  xlink:title="c2" />
  <loc      xlink:type="locator" xlink:href="..."
            xlink:label="child"  xlink:title="c3" />

  <go      xlink:type="arc" xlink:from="parent"
            xlink:to="child" xlink:actuate="onRequest"
            xlink:show="replace"
            xlink:arcrole="#details.txt"/>
</extendedlink>
```

XLink, esimerkki 3 jatkuu

Laajennetun linkin esimerkki

locator-tyyppiset elementit (type="locator") - sisältävät kumpikin yhden kohteen

href - kohteen osoite

label - linkin tunnus

title - linkin lyhyt selitys

arc-tyyppiset (type="arc") sisältävät ohjeen reitityksestä; missä järjestyksessä linkit käydään läpi

from - mistä kohteesta voi mennä

to - mihin kohteeseen

arcrole - selitys reitityksestä