

Mikä on schema?

Kertaus: oikein muodostettu dokumentti

XML-dokumentti on oikein muodostettu (well-formed), jos se noudattaa muutamia perussääntöjä (yksi juurielementti, oikein nimetyt elementit, ei erikoismerkkejä jne.). Oikein muodostettu –määrittely on perustaso, joka tietorakenteen on toteutettava, jotta sitä voisi kutsua XML:ksi.

Kertaus: validi dokumentti

XML-dokumentin voidaan sanoa olevan validi silloin, kun siitä löytyy kaikki ne elementit ja attribuutit, joita dokumentilla oletetaan olevan. Lisäksi elementtien ja attribuuttien on noudatettava niille asetettuja sääntöjä.

Esimerkki

Olemme sopineet, että juomalasi-elementissä on oltava attribuutti valmistusvuosi sekä elementit valmistaja ja tilavuus. Valmistusvuosi on numeerinen tieto välillä 1800-2100, valmistaja on merkkijono ja tilavuus on desimaaliluku. Seuraava XML-dokumentti on oikein muodostettu (well-formed). XML-dokumentti ei kuitenkaan ole validi, koska tilavuus-elementti ei ole numeerinen, vaan se sisältää kirjaimia.

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

<juomalasi valmistusvuosi="2003">
  <valmistaja>Juomalasitehdas Oy</valmistaja>
  <tilavuus>2.0 dl</tilavuus>
</juomalasi>
```

Schema

Schema on DTD:n lisäksi toinen XML-standardin mukainen tapa määrittellä XML-dokumentin rakenne. Schema-määrittelyt poikkeavat DTD-määrittelyistä pääosin siinä, että schema-määrittelyillä on mahdollista määrittellä useampia tietotyyppisiä ja tietotyyppien arvoja. Siinä missä DTD:llä pystytään korkeintaan määrittelemään, että elementin tyyppi on PCDATA (mikä tahansa merkkijono), schemalla pystytään esimerkiksi määrittelemään, että elementin tyyppi on numero väliltä 0-10.

XML-dokumentissa voidaan tarvittaessa määrittellä jokaiselle elementille, mitä schema-määrittely se noudattaa. Tällöin schema-määrittelyä tukevat työkalut osaavat automaattisesti tarkistaa, onko XML-dokumentti validi. Noudatettava schema määrittellään asettamalla elementille attribuutit

```
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"  
xsi:schemaLocation= "[namespace] [schema-tiedosto]"
```

Mikäli elementille ei ole määritelty nimiavaruutta, määritellään käytettävä schema attribuuteilla

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

Esimerkki

Määritellään, että juomalasi-elementti noudattaa `juomalasi.xsd`-määritystä.

```
<?xml version="1.0" encoding="ISO-8859-1" ?>  
  
<juomalasi valmistusvuosi="2003"  
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"  
  xsi:noNamespaceSchemaLocation="juomalasi.xsd">  
  
  <valmistaja>Juomalasitehdas Oy</valmistaja>  
  <tilavuus>2.0 dl</tilavuus>  
</juomalasi>
```

Scheman rakenne

Schema kuvaa yhden tietorakenteen

Schemalla kuvataan aina yksi tietorakenne. Yhden scheman ei välttämättä tarvitse kuvata koko XML-dokumenttia. XML-dokumentin eri elementeissä voidaan käyttää eri schemeja.

Esimerkki

Yritystä kuvaavassa XML-dokumentissa on elementti yrityksen yleiset tiedot sekä elementti työntekijöiden tiedot. Yrityksen yleisiä tietoja voidaan kuvata yhdellä schemalla, työntekijöiden tietoja toisella schemalla.

Schema yleisrakenne

Schema on itsessään XML-dokumentti ja noudattaa siten kaikkia oikein muodostetun (well-formed) XML-dokumentin sääntöjä. Juurielementtinä schema-kuvauksessa on aina `xs:schema`-niminen elementti. Schema-kuvauksen perusrakenne näyttää tällöin seuraavalta:

```
<?xml version="1.0" encoding="UTF-8"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema">
  [dokumentin sisällön kuvaus]
</xs:schema>
```

Varsinainen scheman sisältö, eli kuvaus XML-tietorakenteesta, koostuu yleensä pääosin kolmenlaisista määrittelyistä: elementtimäärittelyistä, simpleType-määrittelyistä ja complexType-määrittelyistä. Käytännössä schema-kuvausten luominen on hyvin pitkälle näiden kolmen määrittelyn luomista ja yhdistelemistä.

Element

Mitä elementti tarkoittaa schema-kuvauksissa?

Element on yksinkertaisin schemaan laitettava määrittely. Se kuvaa jotakin elementtiä, jonka tyyppi on jokin schema-standardiin määritellyistä tietotyypeistä (esimerkiksi merkkijono tai kokonaisluku) tai jokin itse määritelty tietotyyppi (esimerkiksi henkilötunnus tai osoite). Elementtimäärittelyyn ei voi enää kuvata attribuutteja tai lapsielementtejä tai rajoituksia sallituille arvoille; nämä asiat kuvataan simpleType- ja complexType-määrittelyillä.

Miten elementti kuvataan schemassa?

Elementti kuvataan schemassa syntaksilla

```
<xsd:element name="[elementin nimi]" type="[elementin tyyppi]" />
```

Esimerkki

Yksinkertaisessa henkilön tietoja kuvaavassa XML-tietorakenteessa on tiedot nimi ja ikä. Nimi on merkkijono ja ikä on kokonaisluku. XML-muodossa henkilön tiedot näyttäisivät seuraavilta:

```
<henkilö>
  <nimi>Matti Meikäläinen</nimi>
  <ikä>34</ikä>
</henkilö>
```

Henkilön tietojen rakenteen määrittelevä schema taas näyttäisi seuraavalta:

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema">
  <xs:element name="nimi" type="xs:string" />
  <xs:element name="ikä" type="xs:integer" />
</xs:schema>
```