



SÄHKE-hanke
Tekninen mallintaminen
SÄHKE-metatietojen XML Schema

Versio ja pvm	Laatinut	Tarkpvm	Tarkastanut	Hyvppvm	Hyväksynyt
2.0 / 04.02.2005	Anneli Rantanen	15.02.2005	Markus Merenmies	18.02.2005	Ohjausryhmä



04.02.2005

Muutoshistoria

Versio ja pvm	Laatija	Muutoksen kuvaus
2.0 / 04.02.2005	Anneli Rantanen	TE:n toimittama versio
1.9T / 25.01.2005	Anneli Rantanen	KA:n ehdottamat muutokset 21.12.2004
1.8T / 21.12.2004	Anneli Rantanen	Versiomuutos
1.4T / 14.12.2004	Anneli Rantanen	Versiomuutos
1.0 / 30.11.2004	Anneli Rantanen	Ylä- ja alaotsikot korjattu
0.2 / 30.04.2003	Heikki Ellonen	Korjattu johtoryhmän kokouksessa 23.4.2003 havaittuja virheitä.
0.1 / 17.02.2002	Heikki Ellonen	Ensimmäinen versio

SÄHKE-metatietojen XML Schema

Sisällysluettelo

1 Lukuohje.....	3
2 Nimiavaruudet (namespaces).....	5

LIITTEET:

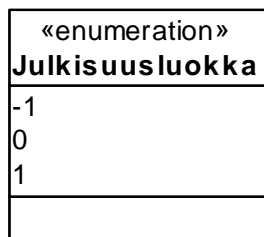
Liite 1a: Aineiston_metatiedot.xsd.

04.02.2005

1 Lukuohje

XML Schema on W3C:n standardi, jonka avulla määritetään rakenteisten dokumenttien rakenne ja jonka avulla voidaan tarkistaa toteuttaako jokin XML-dokumentti kyseisen määrittelyn (validointi).

Tämän lukuohjeen tavoitteena on opastaa lukijaansa hahmottamaan, miten edellisissä vaiheissa määritellyt metatiedot kuvataan XML Scheman avulla. Pääsääntöisesti jokaisesta UML:n kaavion luokasta on XML Schemaan luotu oma tietotyyppinsä, joka on tapauksesta riippuen joko yksinkertainen (simpleType) tai monimutkainen (complexType). Esimerkkinä muunnoksena on julkisuusluokan muuttuminen PublicityClassType-rakenteeksi (Kuva 1). Näiden perustyyppien avulla on luotu monimutkaisemmat tietotyypit.



```
<xs:simpleType name="publicityClassType"
  final="restriction">
  <xs:annotation>
    <xs:documentation xml:lang="fi">
      Hyväksytyt julkisuusluokat</xs:documentation>
    </xs:annotation>
    <xs:restriction base="xs:string">
      <xs:enumeration value="-1"/>
      <xs:enumeration value="0"/>
      <xs:enumeration value="1"/>
    </xs:restriction>
  </xs:simpleType>
```

Kuva 1 UML kaavion luokan muuntuminen XML Schemaan.

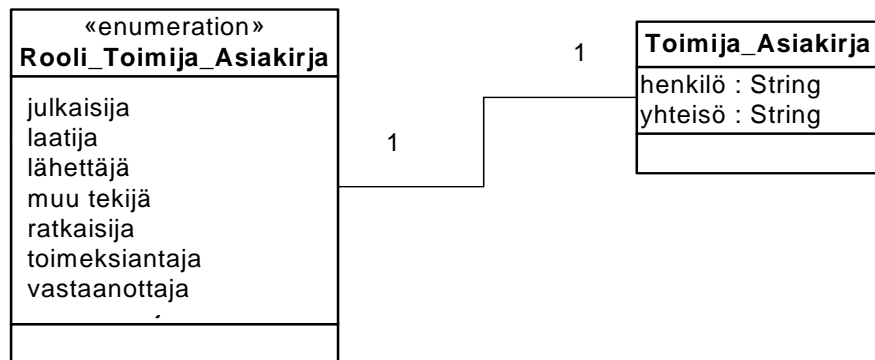
XML Scheman juurena toimii elementti ”metadataOfArchive”, jonka alle on sitten kasattu arkistonmuodostajasta lähtevä hierarkia. Hierarkia on rakennettu siten, että jokaiselle pääkäsitteelle (arkistonmuodostaja, arkisto, ryhmä, asia, toimenpide ja asiakirja) kuvautuvat kaikki siihen liittyvät käsitteet. Käytännössä tämä tarkoittaa sitä, että ylemmälle hierarkiatasolle määritelyt

04.02.2005

elementti (esim. ”julkisuus”) esiintyy myös alemmilla tasoilla elementtinä, vaikka tietosisältö olisi sama.

Tietoteknisistä syistä UML-kuvausta ei kuitenkaan ole muunnettu ylläolevan pääsäännön mukaan orjallisesti, vaan lähtökohdaksi on otettu, että XML skeema olisi ”XML-näkökulmasta” katsottuna mahdollisimman joustava ja XML-taitoiselle helposti avautuva. Merkittävimmät muutokset liittyvät seuraavaan tilanteeseen:

- Käsitteeseen (esim. toimija_asiakirja) liittyvä arvojoukko ei useimmiten esiinny omana rakenteena, vaan siitä on usein muodostettu attribuutti (Kuva 2).



```
<xs:complexType name="agentForDocumentType" mixed="false">
  <xs:all>
    <xs:element name="personalName" type="xs:string"/>
    <xs:element name="corporateName" type="xs:string"
      minOccurs="0"/>
  </xs:all>
  <xs:attribute name="isSecret" type="xs:IDREF" use="optional"/>
  <xs:attribute name="role" use="required">
    <xs:simpleType>
      <xs:restriction base="xs:string">
        <xs:enumeration value="publisher"/>
        <xs:enumeration value="mainCreator"/>
        <xs:enumeration value="pender"/>
        <xs:enumeration value="contributor"/>
        <xs:enumeration value="solver"/>
        <xs:enumeration value="mandator"/>
        <xs:enumeration value="receiver"/>
      </xs:restriction>
    </xs:simpleType>
  </xs:attribute>
</xs:complexType>
```

Kuva 2 UML-luokan "Toimija_Asiakirjalle" muunnos XML Schemaan



2 Nimiavaruudet (namespaces)

Nimiavaruudet on XML-teknologia, jonka avulla voidaan varmistua siitä, että käytettävät käsitteet ovat yksikäsitteisiä. Namespacen käyttö XML Schemoissa on perusteltua silloin, kun yhden XML Scheman sisältä viitataan toisissa XML Schemoissa esiintyviin käsitteisiin.

XML Schemassa aineiston_metatiedot.xsd (liite 1a) nimiavaruus käsitettä on käytetty määrittämään schemassa esiintyvä kohdenimiavaruus, jolloin käytetyt käsitteet ovat tässä tapauksessa Kansallisarkiston määrittämiä. Kohdenimiavaruuden tunnus on:

```
xs:schema targetNamespace=http://www.narc.fi/sahke/aineiston_metatiedot20  
xmlns:narc="http://www.narc.fi/sahke/aineiston_metatiedot20"
```