

JHS 143 Liite 1: SOVELTAMISOHJE USEAMMAN TIETOJÄRJESTELMÄN VÄLISEEN TIEDONSIIRTOON

Esimerkkinä Suomi.fi-integroinnit

Sisällysluettelo

1	Johdanto	2
2	Integroinnin lähtökohta	2
	2.1 Taustaa Suomi.fi-portaalin integroinneista	2
	2.2 Yleistä integrointiratkaisuista	2
3	Integrointitekniikka	3
	3.1 XML-siirtotekniikka	3
	3.2 Tietojen käsittely ja muutos- ja virhetilanteet.....	3
4	JHS 143 -suositus ja integroitavien tietojen määrittely	4
	4.1 Tietojen määrittely	4
	4.2 Muistilista tietojen määrittelyyn.....	4
	4.3 Siirtoon liittyvät yleiset tiedot	5
5	Suomi.fi-portaalin integroinnissa siirrettävät tiedot	5
	5.1 Siirtotiedosto A.....	5
	5.1.1 Siirtotiedosto A:n DTD.....	6
	5.1.2 Esimerkki siirtotiedostosta A	6
	5.2 Siirtotiedosto B.....	7
	5.2.1 Siirtotiedosto B:n DTD.....	11
	5.2.2 Esimerkki siirtotiedostosta B	11

1 JOHDANTO

Tässä asiakirjassa kuvataan, miten Lomake.fi- ja Finlex-verkkopalveluiden sisältöjen integrointi Suomi.fi-portaaliin on toteutettu yhdenmukaisten metatietojen avulla. Asiakirja sisältää integroitavien metatietojen yksityiskohtaiset kuvaukset. Asiakirja toimii esimerkkinä ja ohjeena JHS 143 -suosituksen soveltamisesta tietojärjestelmiä integroitaessa.

Asiakirjan luettuaan lukijalle muodostuu:

- Yleiskuva Suomi.fi-portaaliin tuotavien sisältöjen integrointivasta.
- Käsitys siitä, kuinka JHS 143 -suositusta on hyödynnetty integrointia suunniteltaessa.
- Ymmärrys siitä, mitä asioita integrointia suunniteltaessa on tarkennettava metatietojen osalta.

2 INTEGROINNIN LÄHTÖKOHTA

2.1 Taustaa Suomi.fi-portaalin integroinneista

Suomi.fi-portaali on julkishallinnon yhteinen verkkopalvelu, jonne on koottu kansalaisille keskeistä, arkielämään ja julkishallintoon liittyvää tietoa. Tieto on osaksi tekstiä portaalin sivuilla, mutta on suureksi osaksi linkityksiä muihin julkishallinnon verkkopalveluihin.

Suomi.fi-portaalin tarjoamiin linkkeihin liittyy niiden kohdesivuja kuvaavaa tietoa. Tällaista kuvaavaa tietoa on esimerkiksi kohdesivun lyhyt kuvaus, sivusta vastaava organisaatio tai sivun verkko-osoite. Näitä tietoja nimitetään linkin metatiedoiksi.

Jotta kansalaisen portaalista saama tieto olisi mahdollisimman monipuolista, vuonna 2003 käynnistettiin portaalissa kehityshanke, jossa integroitiin Lomake.fi-verkkopalvelun asiointipalvelut ja lomakkeet sekä Otakantaa.fi-verkkopalvelun keskustelut osaksi portaalia. Näin kansalainen sai portaalin sivulta paitsi tiettyyn aiheeseen liittyvää tietoa ja linkkejä tietoa sisältäville sivuille, myös linkkejä aiheeseen liittyviin asiointipalveluihin, lomakkeisiin ja keskusteluihin.

Otakantaa.fi-palvelun toiminnan hiljennyttyä integrointi keskusteluihin poistettiin, mutta vuoden 2005 keväällä samalla konseptilla integroitiin uusi verkkopalvelu: oikeusministeriön ylläpitämä Finlex-säädöstietokanta.

Portaalista on ilmestymässä uusi versio vuoden 2006 aikana. Tähän versioon integrointeja on määrää entisestään syventää mm. aiheuokittelua tarkentamalla. Tapaa, jolla portaaliin integroidaan sisältöjä, ei peruseriaatteiltaan ole muutettu uudessa versiossa. Uudessa versiossa integroitavien tietojen määrää on kuitenkin entisestään kasvatettu ja metatiedoista käytetyt nimitykset ja niiden väliset yhteydet on päivitetty vastaamaan uusinta JHS 143 -suosituksen versiota.

Seuraavissa luvuissa kuvataan tarkemmin Suomi.fi-portaalin integroinnin pääperiaatteita ja yksityiskohtia etenkin metatietojen käsittelyn osalta. Asiointipalveluihin ja lomakkeisiin viitataan jatkossa lyhyemmällä nimellä "lomakkeet".

2.2 Yleistä integrointiratkaisuista

Kun asiakirjojen tietoja halutaan jakaa usean eri tietojärjestelmän kesken, on päätettävä se taso, jolla integrointia tehdään.

Asiakirjat voivat sijaita yhdessä tietojärjestelmässä tai niistä voidaan pitää kopioita useammassa tietojärjestelmässä. Asiakirjoja voidaan ylläpitää osittain tai kokonaan molemmissa järjestelmissä. Asiakirjoista voidaan haluta siirtää järjestelmästä toiseen vain pieni osa tiedoista tai kaikki tiedot.

Päädyttiinpä em. ratkaisuihin mihin hyvänsä, metatietojen siirto järjestelmien välillä on oleellinen osa integrointia. Tehokkaaseen asiakirjahallintaan sisältyy aina asiakirjojen haku ja tunnistus niiden metatietojen avulla.

Suomi.fi-portaalin tapauksessa asiakirjoja (lomakkeet ja lait) ylläpidetään omilla tietojärjestelmissään (Lomake.fi-alusta, Finlex), mutta niitä voidaan hakea myös Suomi.fi-portaalin kautta.

Lomakkeet ja lait käyttävät kuitenkin osittain samoja luokitustapoja kuin Suomi.fi-portaalin muut linkit. Näiden luokitusten ylläpito tapahtuu Suomi.fi-portaalissa ja luokituksiin viittaaminen yksittäisistä lomakkeista ja laeista suoritetaan Lomake.fi-alustalla ja Finlex-järjestelmässä.

Mikäli integrointi kohdistuu useaan järjestelmään, tällaista järjestelmien yhteistä luokittelujen hallintaa kannattaa miettiä, etenkin keskeisten, muuttuvien ja paljon eri vaihtoehtoja sisältävien luokitusten osalta.

3 INTEGROINTITEKNIikka

3.1 XML-siirtotekniikka

Suomi.fi-portaalissa lomakkeiden ja lakien integrointi on toteutettu XML-tiedostosiirtoina. Lomakkeiden ja lakien tiedot siirretään järjestelmästä toiseen XML-muotoisina sovituin aikavälein.

XML-muotoon perustuvat tekniikat ovat yleisiä sellaisten etenkin sellaisten tietojärjestelmien integroinneissa, joiden taustalla olevat tekniset ratkaisut eroavat toisistaan tai joista vastaavat erilliset tahot. Tiedostosiirtoihin päädytään yleensä tilanteessa, jossa asiakirjan tietojen ei tarvitse olla keskenään reaaliaikaisesti ajan tasalla eri järjestelmissä, vaan jossa tietty päivitysväli voidaan asettaa.

Suomi.fi-portaalissa tiedot päivittyvät järjestelmästä toiseen pääosin kerran vuorokaudessa. Kerran vuorokaudessa Lomake.fi-alusta ja Finlex hakevat luokitustiedot Suomi.fi-portaalista ja portaali hakee kerran vuorokaudessa luokitellut lomakkeet ja lait Lomake.fi-alustalta ja Finlexistä.

Tiedostosiirto tapahtuu HTTP-pyyntöillä Internetin yli. Järjestelmät muodostavat XML-muotoisen listauksen integroitavista tiedoista, kun HTTP-pyyntö kohdistetaan järjestelmäkohtaisesti tiettyyn verkko-osoitteeseen.

3.2 Tietojen käsittely ja muutos- ja virhetilanteet

Kukin tietojärjestelmä käsittelee omalla tavallaan HTTP-pyyntöillä hakemansa tiedot. Lomake.fi-alusta ja Finlex tallentavat luokitustiedot omiin tietokantoihinsa. Suomi.fi-portaali tallentaa lomakkeiden ja lakien metatiedot omaan järjestelmäänsä. Vanhassa versiossa Suomi.fi-portaali tallensi lomakkeiden ja lakien metatiedot ainoastaan hakuindeksiin, mutta vuonna 2006 julkaistavassa uudessa versiossa lomakkeiden ja lakien tietojen perusteella luodaan tietueita portaalin julkaisujärjestelmään.

Aina kun järjestelmät hakevat siirrettävät tiedot, ne analysoivat tietojen sisällön.

Siirtotiedoston rakenteen oikeellisuus tarkastetaan siirtotiedoston rakenteen määrittelevää DTD-tiedostoa vasten. Näin varmistetaan siitä, että tiedosto on tullut kokonaisuutena ja että sen koneellinen käsittely on mahdollista.

Järjestelmät tarkistavat muutokset edelliseen siirtoon verrattuna: muuttuneet tiedot päivitetään järjestelmään, uudet tiedot lisätään ja poistetut tiedot poistetaan. Siirtotiedoston tiedot ja järjestelmään tallennetut tiedot pystytään yhdistämään toisiinsa tunnisteiden avulla.

Uudessa versiossa tehdään lisäksi tarkistus, että mikäli siirtotiedoston elementeistä (luokituksista, laiesta tai lomakkeista) on muuttunut tietty määrä (10 %), päivitystä ei tehdä ollenkaan, koska siirtotiedoston muodostumisessa on todennäköisesti tapahtunut virhe.

Suomi.fi-portaali lähettää lisäksi sähköpostiviestin ylläpitäjille, mikäli siirto ei onnistunut.

4 JHS 143 -SUOSITUS JA INTEGROITAVIEN TIETOJEN MÄÄRITTELY

4.1 Tietojen määrittely

Suomi.fi-portaaliin on kehitetty oma tapansa liittää ulkopuolisia palveluita. Tietojärjestelmien integroinneissa tyypillistä on, että niihin liittyy paljon eri osapuolia ja toimijoita ja vaihtoehtoisia tapoja toimia on lukuisia. Yhteisen näkemyksen aikaansaaminen vie aikaa ja vaivaa. Tällaisessa toimintaympäristössä yleiset suositukset ja tavat toimia nousevat arvoonsa. Asiakirjoja integroitaessa JHS 143 -suositus muodostaa hyvän pohjan erityisesti siinä vaiheessa kun pohditaan, mitä tietoja järjestelmien välillä siirretään ja mitä nimityksiä niistä käytetään.

Helpoin tilanne on, kun suositusta ei tarvitse sen kummemmin tulkita: asiakirjoja ylläpitävään järjestelmään on tallennettu metatiedot jo suosituksen mukaisilla sisällöillä ja nimillä. Oli tilanne kuitenkin mikä hyvänsä, integrointia suunniteltaessa on hyvä kiinnittää huomiota siihen, mikä tieto integroitavassa järjestelmässä vastaa samaa tietoa toisessa järjestelmässä.

Tietojen välille halutaan yleensä suunnitella rakenteisuutta eli ryhmitellä tietoja hierarkisesti toisiinsa tai viitata toisesta tiedosta toiseen. Lisäksi usein halutaan määritellä myös siirrettävien tietojen järjestys toisiinsa nähden, mm. jotta siirtotiedostoiesimerkkejä voitaisiin tarkastaa suunnitteluvaiheessa helposti silmämääräisesti.

Kaiken kaikkiaan integroitavien tietojen määrittelyyn ei riitä pelkkä XML-muotoisen rakenteen määrittely esimerkiksi DTD-muotoisena tiedostona, vaan XML-tiedoston sisällön selitys on tehtävä elementtikohtaisesti myös sanallisesti.

Suomi.fi-portaalissa siirtotiedostoista on tehty DTD-kuvausten lisäksi myös esimerkit XML-muotoisista tiedostoista sekä sanalliset kuvaukset tiedostojen sisällöistä. Esimerkit on tehty kaikista siirtotiedostoista: Suomi.fi:n luokituksista, Lomake.fi:n lomakkeista ja Finlexin laeista. Koska lomakkeiden ja lakien integroinneissa käytetään samaa integrointitapaa, on esimerkit ja sanalliset kuvaukset koottu yhteen määrittelyyn. Määrittelyssä on sitten kuvattu elementtikohtaisesti rivi riviltä XML-siirtotiedoston sisältö palvelukohtaisin esimerkein.

4.2 Muistilista tietojen määrittelyyn

1) Mitä tietoja siirretään?

Esimerkiksi: otsikko, kuvaus, aihe, tiedostomuoto ...

2) Mitä nimitystä tiedoista käytetään kussakin järjestelmässä?

Kirjaa ylös vastaavuudet. Käytä siirron kuvauksessa yhteisinä nimityksinä JHS 143 -suosituksen nimityksiä.

3) Mitkä tiedot on pakollisia, mitkä vapaaehtoisia?

Esimerkiksi otsikko on aina oltava, mutta kuvaus on vapaaehtoinen. Jos tieto puuttuu, mitä tehdään? Huomaa myös, että suosituksessa on kannanottoja tietojen pakollisuuteen. Suosituksessa pakollisuus on määritelty erikseen asiakirjahallinnollisen näkökulman ja julkaisunäkökulman osalta.

4) Millä tiedolla on monta arvoa, millä vain yksi?

Esimerkiksi aiheita on useita, mutta kuvauksia vain yksi. Suosituksessa tähän on otettu kantaa kunkin metatiedon osalta kohdassa "Toistettavuus".

5) Millä tiedoilla on yhteys toisiinsa ja minkälainen yhteys on kyseessä?

Esimerkiksi asiakirjaan voi liittyä liitteitä. Aiheet ja asiasanat voivat olla kytköksissä toisiinsa.

6) Mitkä on tietojen tarkat sisällöt?

Esimerkiksi kuinka pitkä kuvaus voi olla? Tai mikä on kuvauksen merkitys? Mitä vakioarvoja voi jollain tiedolla olla? Hyvä asia on myös kirjata, missä kaikkialla tietoa hyödynnetään. Sisällöllisiä asioita ei pysty kuvaamaan DTD:ssä ja suositus liikkuu tältä osin yleisellä tasolla, joten älä ohita tätä vaihetta!

7) Varmista, että tiedoilla on tarvittavat tunnisteet!

Jos tiedoilla on yhteyksiä toisiinsa siirtotiedoston sisällä tai järjestelmien välillä, varmista, että siirrossa tiedoista välitetään myös niitä vastaavien tietueiden tunnisteet. Nämä ovat useimmiten järjestelmien tallentamia tietueiden yksilöllisiä ID-tunnisteita.

4.3 Siirtoon liittyvät yleiset tiedot

Siirtotiedostoa voi ajatella omana asiakirjanaan, johon liittyy myös metatietoja.

Siirtotiedostoon voi määritellä omiksi elementeikseen esimerkiksi tiedoston tekijän, julkaisijan ja muun toimijan kuten järjestelmätoimittajan.

Muodostamisajankohdasta on hyötyä erityisesti, jos siirtotiedostoa ei muodosteta pyyntöhetkellä tai jos halutaan esimerkiksi kehitys- tai testausvaiheessa tallentaa siirtotiedostoja väliaikaisesti verkkolevyille tai tulosteiksi.

5 SUOMI.FI-PORTAALIN INTEGROINNISSA SIIRRETTÄVÄT TIEDOT

Suomi.fi-portaalin integroinnissa on suunniteltu kaksi XML-tiedostoa:

- Siirtotiedosto A, joka kuvaa muihin järjestelmiin siirrettävät luokitukset. Luokitukset ovat aiheita, alaihteita ja asiasanoja. Tässä tiedostossa ei ole luokituksia käsitelty asiakirjoina, koska se ei ole mielekäästä.
- Siirtotiedosto B, joka kuvaa muista järjestelmistä siirrettävät tiedot. Nämä ovat tällä hetkellä lomakkeita tai lakeja, mutta näitä voidaan lisätä tarvittaessa. Lisäksi on huomattava, että siirtotiedosto B:n kautta Lomake.fi-alusta ja Finlex tarjoavat XML-muodossa keskeistä sisältöään myös muiden mahdollisten järjestelmien käyttöön. Siirtotiedosto B on laadittu noudattamaan JHS 143 -suositusta.

Tässä luvussa selitetään vielä tarkemmin siirtotiedostojen sisältöä ja annetaan siirtotiedostojen rakenne DTD-kuvauksena sekä suppeat esimerkit XML-muotoisista tiedostoista. Esimerkeissä tietueiden määrää on rajattu, jotta ne säilyisivät helppolukuisina.

Tietojen pakollisuus on tässä yhteydessä pääsääntöisesti määritelty sen perusteella, onko tieto integroinnin tuloksen kannalta oleellinen.

5.1 Siirtotiedosto A

Siirtotiedosto A sisältää seuraavat tiedot (suluissa DTD:ssä käytetty nimitys):

- Yleiset tiedot: otsikko (*title*), laatijaorganisaatio (*creator*), julkaisijaorganisaatio (*publisher*), muu organisaatio (*contributor*), tiedoston muodostusajankohta (*dataGathered*), muoto (*fileFormat*), lähde (*source*)
-

- Aiheet (*topic*, *type=primary*) aakkosjärjestyksessä siten, että kieliversiot (*version*) on ryhmiteltyä yhteen. Aiheen tiedot: aiheen etusivun tunniste (*id*), otsikko (*version*), kieli (*language*)
- Ala-aiheet (*subtopic*) aakkosjärjestyksessä aiheiden alla siten, että kieliversiot (*version*) on ryhmiteltyä yhteen. Ala-aiheen tiedot: aiheen alisivun tunniste (*id*), otsikko (*version*), kieli (*language*)
- Kohderyhmät (*topic*, *type=audience*) aakkosjärjestyksessä tiedoston lopussa siten, että kieliversiot (*version*) on ryhmiteltyä yhteen. Kohderyhmän tiedot: kohderyhmäsivun tunniste (*id*), otsikko (*version*), kieli (*language*)
- Aiheen asiasanat (*keyword*) kunkin aiheen alla siten, että kieliversiot (*version*) on ryhmiteltyä yhteen. Asiasanan tiedot: asiasanan tunniste (*id*), asiasanan kieliversion tunniste (*id*), asiasanan kieliversion nimi (*version*), asiasanan huomautus (*annotation*)

5.1.1 Siirtotiedosto A:n DTD

```
<!ELEMENT documentTopics (title,agent,date,format,source,topic+)>
<!ELEMENT title (#PCDATA)>
<!ELEMENT agent (creator,publisher,contributor)>
<!ELEMENT date (dataGathered)>
<!ELEMENT format (fileFormat)>
<!ELEMENT source (#PCDATA)>
<!ELEMENT fileformat (#PCDATA)>
<!ELEMENT dataGathered (#PCDATA)>
<!ELEMENT creator (corporateName)>
<!ELEMENT publisher (corporateName)>
<!ELEMENT contributor (corporateName)>
<!ELEMENT corporateName (#PCDATA)>
<!ELEMENT fileFormat (#PCDATA)>
<!ELEMENT topic (version+,keyword*,subtopic*)>
<!ELEMENT subtopic (version+)>
<!ELEMENT keyword (version+,annotation?)>
<!ELEMENT version (#PCDATA)>
<!ELEMENT annotation (#PCDATA)>

<!ATTLIST topic
type NMTOKEN #REQUIRED>
<!ATTLIST version
id NMTOKEN #REQUIRED
language NMTOKEN #REQUIRED>
<!ATTLIST keyword
id NMTOKEN #REQUIRED>
```

5.1.2 Esimerkki siirtotiedostosta A

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE documentTopics SYSTEM "documentTopics.dtd" >
<documentTopics>
  <title>Aiheet ja asiasanat</title>
  <agent>
    <creator>
      <corporateName>Suomi.fi-toimitus/Valtiovarainministeriö</corporateName>
    </creator>
    <publisher>
      <corporateName>Valtiovarainministeriö</corporateName>
    </publisher>
```

```
<contributor>
  <corporateName>Onesta Solutions</corporateName>
</contributor>
</agent>
<date>
  <dataGathered>2005-12-15T22:00</dataGathered>
</date>
<format>
  <fileFormat>xml</fileFormat>
</format>
<source>http://www.suomi.fi</source>
<topic type="primary">
  <version id="1" language="fi">Asuminen ja rakentaminen</version>
  <version id="20" language="sv">Boende och byggande</version>
  <version id="40" language="en">Housing and construction</version>
  <keyword id="24">
    <version id="7" language="fi">asumismuodot</version>
    <version id="1006" language="sv">boendeformet</version>
    <version id="2006" language="en">housing types</version>
  </keyword>
  <keyword id="23">
    <version id="6" language="fi">asuminen</version>
    <version id="1005" language="sv">boende</version>
    <version id="2005" language="en">housing</version>
  </keyword>
  <keyword id="25">
    <version id="11" language="fi">asumisoikeus</version>
    <version id="1010" language="sv">bostadsrätt</version>
  </keyword>
  <subtopic>
    <version id="23" language="fi">Asuminen</version>
    <version id="24" language="sv">Boende</version>
    <version id="25" language="en">Housing</version>
  </subtopic>
</topic>
<topic type="primary">
  <version id="3" language="fi">Kirjasto- ja tietopalvelut</version>
</topic>
<topic type="audience">
  <version id="10" language="fi">nuoret</version>
  <version id="12" language="sv">ungdomar</version>
</topic>
</documentTopics>
```

5.2 Siirtotiedosto B

Siirtotiedosto B sisältää seuraavat tiedot (suluissa DTD:ssä käytetty nimitys)::

- Siirtotiedostoon liittyvät yleiset tiedot.
- Lomakkeen tai lain perustiedot (*document*). Lomakkeissa nämä tiedot ovat tallennettuina yhteisesti ns. äitielementtiin koskien kaikkia saman lomakkeen eri lomakemuotoja.
- Lomakkeen tai lain lisätiedot (*documentChild*). Lomakkeissa nämä tiedot ovat tallennettuina ns. lapsielementteihin eli erikseen kullekin saman lomakkeen lomakemuodolle. Jokaiseen lomakkeeseen liittyy siis useita lomakemuotoja eli lapsielementtejä, esimerkiksi Word- ja PDF-muodot lomakkeista. Kullakin lailla sen sijaan on ainoastaan yksi lapsielementti.

Yleiset tiedot:

Tieto	DTD:ssä	Selitys
Nimeke	<i>title</i>	Tiedostoon liittyvä yleisotsikko. Pakollinen elementti.
Toimija	<i>agent</i>	Pakollinen elementti.
Tekijä	<i>creator</i>	Aina Suomi.fi-toimitus. Pakollinen elementti.
Julkaisija	<i>publisher</i>	Se taho, jolla on aineistoon julkaisuoikeus. Pakollinen elementti.
Muu toimija	<i>contributor</i>	Järjestelmätoimittaja. Pakollinen elementti.
Aikamääre	<i>date</i>	Pakollinen elementti.
Kokoamis aika	<i>dataGathered</i>	Tiedoston muodostamisajankohta. Pakollinen elementti.
Formaatti	<i>format</i>	Pakollinen elementti.
Tiedostomuoto	<i>fileFormat</i>	Siirtotiedoston tallennusmuoto. Aina "xml". Pakollinen elementti.
Lähde	<i>source</i>	Aineistoon liittyvä yleinen verkko-osoite. Pakollinen elementti.

Lomakkeen tai lain perustiedot:

Tieto	DTD:ssä	Selitys
Asiakirjan tunniste	<i>id</i>	Tunnisteen perusteella tietue voidaan nopeasti paikallistaa mahdollisissa ongelmatilanteissa. Tietueet järjestetään tiedostoon tunnisteen mukaan. Pakollinen attribuutti. Lomakkeissa tunnisteena on kielikoodi sekä äitielementin tunniste. Laeissa paitsi paikallistamiseen liittyvä, myös tietueiden järjestämisen mahdollistava tieto. Laeissa tieto on muotoa llvvvnnnn, jossa ll=kielitunnus, vvv=säädöksen antopäivämäärän vuosi, nnnn=säädöksen järjestysnumero vuoden sisällä. Esim. "FI20040799".
Kieli	<i>language</i>	ISO-639 -standardin mukainen kaksikirjaiminen kielikoodi. Kirjoitetaan standardin mukaisesti pienin kirjaimin, esim. "fi". Jokainen kieliversio on oma tietueensa. Lomakkeen äitielementillä ei ole kieliversioita Lomake.fi-alustalla, mutta siirtotiedostoon äitielementistä muodostetaan kieliversiot lapsielementtien mukaan. Äitielementin tiedot siis toistetaan kieliversioiden määrän verran. Kieli on pakollinen elementti.
Identifiointitunnus	<i>identifier</i>	Lomaketunnus tai säädöstunnus. Pakollinen elementti.
Nimeke	<i>title</i>	Lomakkeen tai lain otsikko. Pakollinen elementti.
Vaihtoehtoinen nimeke	<i>alternative</i>	Lomakkeen tai lain kansanomaisen kutsumanimi, joka halutaan esittää varsinaisen nimen rinnalla. Kutsumanimiä voi olla useita. Vapaaehtoinen elementti.
Toimija	<i>agent</i>	Pakollinen elementti.
Tekijä	<i>creator</i>	Lomakkeen laatinut taho tai lain valmistelua koordinoanut hallinnonala. Tekijöitä on ainoastaan yksi. Tekijästä annetaan nimi (<i>corporateName</i>) ja lomakkeen ollessa kyseessä tekijälle annettu lyhennetyyppinen tunnus (<i>corporateld</i>). Pakollinen elementti.
Muu toimija	<i>otherAgent</i>	Lomakkeen vastaanottava organisaatio. Muita toimijoita voi olla enintään

		yksi. Sisällytetään tiedostoon ainoastaan silloin, kun muu toimija on eri kuin tekijä. Laeilla ei ole ollenkaan tätä tietoa. Vapaaehtoinen elementti.
Kattavuus	<i>coverage</i>	Vapaaehtoinen elementti
Alueellinen	<i>spatial</i>	Lomakkeilla on tämä elementti, mikäli kyseessä on jonkin kunnan lomake tai yleinen kuntalomake. Muissa tapauksissa tätä elementtiä ei esiinny. Vapaaehtoinen elementti.
Alueellisen tiedon tunniste	<i>id</i>	Kunnan virallinen numero. Jos kyseessä on kaikkien kuntien yhteiskäyttöön tarkoitettu lomake, arvona on 0. Pakollinen attribuutti.
Kuvaus	<i>description</i>	Tarkentava kuvausteksti otsikolle. Maksimipituus 255 merkkiä. Mikäli kuvaus on pidempi, katkaistaan 252 merkkiin ja lisätään loppuun 3 pistettä. Mikäli kuvaus on tyhjä, elementtiä ei sisällytetä tietueeseen. Hyödynnetään esim. Suomi.fi:n vapaasanahaun tulossivulla otsikon yhteydessä. Vapaaehtoinen elementti.
Aihe	<i>subject</i>	Aihe, pakollinen asiasana tai vapaa asiasana. Pakollinen elementti. Lomakkeeseen tai lakiin liitetään aiheen etu- ja alasivuja, joilla lomakkeen tai lain on määrä näkyä Suomi.fi-portaalissa. Aiheet määräytyvät tämän perusteella. Aiheita voi olla useita ja niitä tulee olla vähintään yksi. Vaikka aiheet määriteltäisiin lomakkeelle tai laille siten, että kaikki kieliversiot valitaan kerralla, siirtotiedostoon tuotetaan viittaukset aiheisiin kieliversioittain.
Aiheen tunniste	<i>id</i>	Suomi.fi:stä saatava tunniste. Pakollinen attribuutti.
Pakollinen asiasana	<i>controlled keyword</i>	Ne aiheet (<i>subject</i>), joille on määritelty tämä attribuutti, ovat lomakkeelle tai laille määriteltyjä Suomi.fi:n asiasanoja. Asiasanoja voi olla useita ja niitä on oltava vähintään yksi.
Pakollisen asiasanan tunniste	<i>id</i>	Suomi.fi:stä saatava tunniste. Pakollinen attribuutti.
Vapaat asiasanat	<i>uncontrolled keyword</i>	Ne aiheet (<i>subject</i>), joille on määritelty tämä attribuutti, ovat asiasanoja, joihin ei liity tunnistetta. Näitä ovat lomakkeelle määrätty kyseisen kieliset vapaat asiasanat, jotka annetaan kukin omassa elementissään. Lailla ei ole vapaita asiasanoja. Vapaaehtoinen elementti.
Kohdeyleisö	<i>audience</i>	Kohdeyleisöjä voi olla useita. Tällä hetkellä kohdeyleisöjä ovat "Nuorille" ja "Ikäihmisille". Vapaaehtoinen elementti.
Kohdeyleisön tunniste	<i>id</i>	Suomi.fi:stä saatava tunniste. Pakollinen attribuutti.
Suhde	<i>relation</i>	Vapaaehtoinen elementti.
Viittaa	<i>references</i>	Viittaussuhde. Lomakkeeseen liittyvät muut lomakkeet eli viitteet. Arvona lomakkeen tunniste ts. viittaavan lomakkeen document-elementin id-attribuutti. Vapaaehtoinen elementti.

Lomakkeen tai lain lisätiedot:

Tieto	DTD:ssä	Selitys
Asiakirjan tyyppi	<i>documentType</i>	Lomakemuodon tyyppi. Jokin tietyistä tekstimuotoisista vakioarvoista. Vakioarvot ja niiden selitykset: <ul style="list-style-type: none"> • xml = verkkolomake • print = tulostettava lomake

		<ul style="list-style-type: none"> • info = ohjetiedosto • external = ulkoinen asiointipalvelu • list = lomakkeiden koontisivu • other = muu <p>Asiakirjan tyyppi on lomakkeilla pakollinen elementti. Muilla vapaaehtoinen elementti.</p>
Formaatti	<i>format</i>	Lomakkeilla pakollinen, muilla vapaaehtoinen elementti.
Tiedostomuoto	<i>fileFormat</i>	<p>Lomakemuodon tiedostomuoto. Jokin tietyistä tekstimuotoisista vakioarvoista.</p> <p>Vakioarvot ja niiden selitykset:</p> <ul style="list-style-type: none"> • xml = verkkolomake • pdf = pdf-lomake • word = Word-lomake • rtf = rtf-lomake • excel = Excel-lomake • link = ulkoinen linkki • other = muu <p>Tiedostomuoto on lomakkeilla pakollinen, muilla vapaaehtoinen elementti.</p>
Koko	<i>extent</i>	Lomaketiedoston koko kieliversioon sopivalla kielellä ilmaistuna. Suomenkielisissä lomaketiedoissa kilotavu-yksikkö on ilmaistu lyhenteellä "kt", ruotsinkielisissä "kbyte" ja englanninkielisissä "kB". Tieto näytetään lomakkeen yhteydessä suoraan merkkijonona Suomi.fi-portaalin lomaketietosivulla. Vapaaehtoinen elementti.
Aikamääre	<i>date</i>	Vapaaehtoinen elementti.
Saatavillaoloaika	<i>available</i>	Aika, jolloin lomake on Suomi.fi:ssä julkaistuna. Annetaan SFS-EN 28601-standardin mukaisessa muodossa. Vapaaehtoinen elementti.
Oikeudet	<i>rights</i>	Vapaaehtoinen elementti.
Oikeuden tyyppi	<i>rightsType</i>	<p>Tunnistautuminen ja maksullisuus. Tunnistautumistapoja voi olla erilaisia kuten "sirukortti" tai "Nordea". Suomi.fi hyödyntää tietoa ainoastaan mainitsemalla lomakkeen yhteydessä vakiotekstin (esim. suomeksi "Vaatii tunnistautumisen") riippumatta tunnistautumistyyppistä. Jos lomake on maksullinen, sille on oma numeromuotoinen vakioarvonsa. Lailla ei ole tätä tietoa.</p> <p>Vakioarvot ja niiden selitykset:</p> <p>1 = sirukortti 2 = tunnistus.suomi.fi (Vetuma) 3 = tunnistus.fi (Katve) 10 = Nordea 11 = Osuuspankki 12 = Sampo 13 = Sp/Pop 0 = muu tunnistus 99 = maksullinen palvelu</p> <p>Oikeuden tyyppi on vapaaehtoinen elementti.</p>

Lähde	source	URI-muotoinen osoite tietueen näytettävälle Suomi.fi-portaalin ulkoiselle sivulle. Pakollinen elementti.
-------	--------	--

5.2.1 Siirtotiedosto B:n DTD

```

<!ELEMENT documentIndex (title,agent,date,format,source,document+)>
<!ELEMENT title (#PCDATA|alternative)*>
<!ELEMENT alternative (#PCDATA)>
<!ELEMENT agent (creator,publisher?,contributor?,otherAgent?)>
<!ELEMENT creator (corporateName, corporateID?)>
<!ELEMENT publisher (corporateName)>
<!ELEMENT contributor (corporateName)>
<!ELEMENT corporateID (#PCDATA)>
<!ELEMENT otherAgent (agentRole,corporateName)>
<!ELEMENT corporateName (#PCDATA)>
<!ELEMENT agentRole (#PCDATA)>
<!ELEMENT date (dataGathered|available)>
<!ELEMENT dataGathered (#PCDATA)>
<!ELEMENT available (#PCDATA)>
<!ELEMENT format (fileFormat,extent?)>
<!ELEMENT fileFormat (#PCDATA)>
<!ELEMENT extent (#PCDATA)>
<!ELEMENT source (#PCDATA)>
<!ELEMENT document
(language,identifier,title,agent,coverage?,description?,subject+,audience*,
relation?, documentChild+)>
<!ELEMENT language (#PCDATA)>
<!ELEMENT identifier (#PCDATA)>
<!ELEMENT coverage (spatial?)>
<!ELEMENT spatial (#PCDATA)>
<!ELEMENT description (#PCDATA)>
<!ELEMENT subject (#PCDATA)>
<!ELEMENT audience (#PCDATA)>
<!ELEMENT documentChild (documentType?,format?,date?,rights?,source)>
<!ELEMENT documentType (#PCDATA)>
<!ELEMENT rights (rightsType*)>
<!ELEMENT rightsType (#PCDATA)>
<!ELEMENT relation (references*)>
<!ELEMENT references (#PCDATA)>

<!ATTLIST document id NMTOKEN #REQUIRED>
<!ATTLIST spatial id NMTOKEN #REQUIRED>
<!ATTLIST subject id NMTOKEN #IMPLIED>
<!ATTLIST subject type CDATA #IMPLIED>
<!ATTLIST audience id NMTOKEN #REQUIRED>

```

5.2.2 Esimerkki siirtotiedostosta B

```

<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE documentIndex SYSTEM "documentIndex.dtd" >
<documentIndex>
<title>Lomakkeet</title>
<agent>

```

```
<creator>
<corporateName>Suomi.fi-toimitus/Valtiovarainministeriö</corporateName>
</creator>
<publisher>
<corporateName>Valtiovarainministeriö</corporateName>
</publisher>
<contributor>
<corporateName>Elma Oyj Electronic Trading</corporateName>
</contributor>
</agent>
<date>
<dataGathered>2005-10-21T14:15</dataGathered>
</date>
<format>
<fileFormat>xml</fileFormat>
</format>
<source>http://www.lomake.fi</source>
<document id="FI794">
<language>fi</language>
<identifier>Vrk 0.44 fi+sv.pdf</identifier>
<title>
Pyyntö avioliiton esteiden tutkinnasta sekä vakuutus ja ilmoitus avioliiton
esteiden tutkintaa varten
<alternative>Avioesteen tutkintapyyntö</alternative>
</title>
<agent>
<creator>
<corporateName>Väestörekisterikeskus</corporateName>
<corporateID>VRK</corporateID>
</creator>
<otherAgent>
<agentRole>receiver</agentRole>
<corporateName>Maistraatit</corporateName>
</otherAgent>
</agent>
<coverage>
<spatial id="091">Helsinki</spatial>
</coverage>
<description>
Ennen vihkimistä avioliiton esteet tutkitaan. Tutkinnan voi suorittaa maistraatti
tai se evankelis-luterilainen tai ortodoksinen seurakunta, johon toinen
vihittävästä kuuluu. (AL 10 §).
</description>
<subject id="105">Parisuhde</subject>
<subject type="controlled keyword" id="583">avioliitto</subject>
<subject type="uncontrolled keyword">vihkiminen</subject>
<subject type="uncontrolled keyword">häät</subject>
<audience id="16">
nuoret
</audience>
<relation>
<references>FI123
</references>
<references>FI234
</references>
```

```
</relation>
<documentChild>
<documentType>print</documentType>
<format>
<fileFormat>word</fileFormat>
<extent>34 kt</extent>
</format>
<date>
<available>2005-12-21T12:00/2006-12-02T12:00</available>
</date>
<rights>
<rightsType>1</rightsType>
<rightsType>5</rightsType>
</rights>
<source>
https://lomake.fi/ec/index.cgi/forminfo?s=kRIXEA6FXY388Ji&lomake_id=794&query=maistraatti&lang=fi&return=search_forms
</source>
</documentChild>
</document>
</documentIndex>
```
