

Paikkatiedot ja Web-standardit

Ossi Nykänen

Tampereen teknillinen yliopisto (TTY),
Digitaalisen median instituutti (DMI),
Hypermedialaboratorio
W3C Suomen toimisto



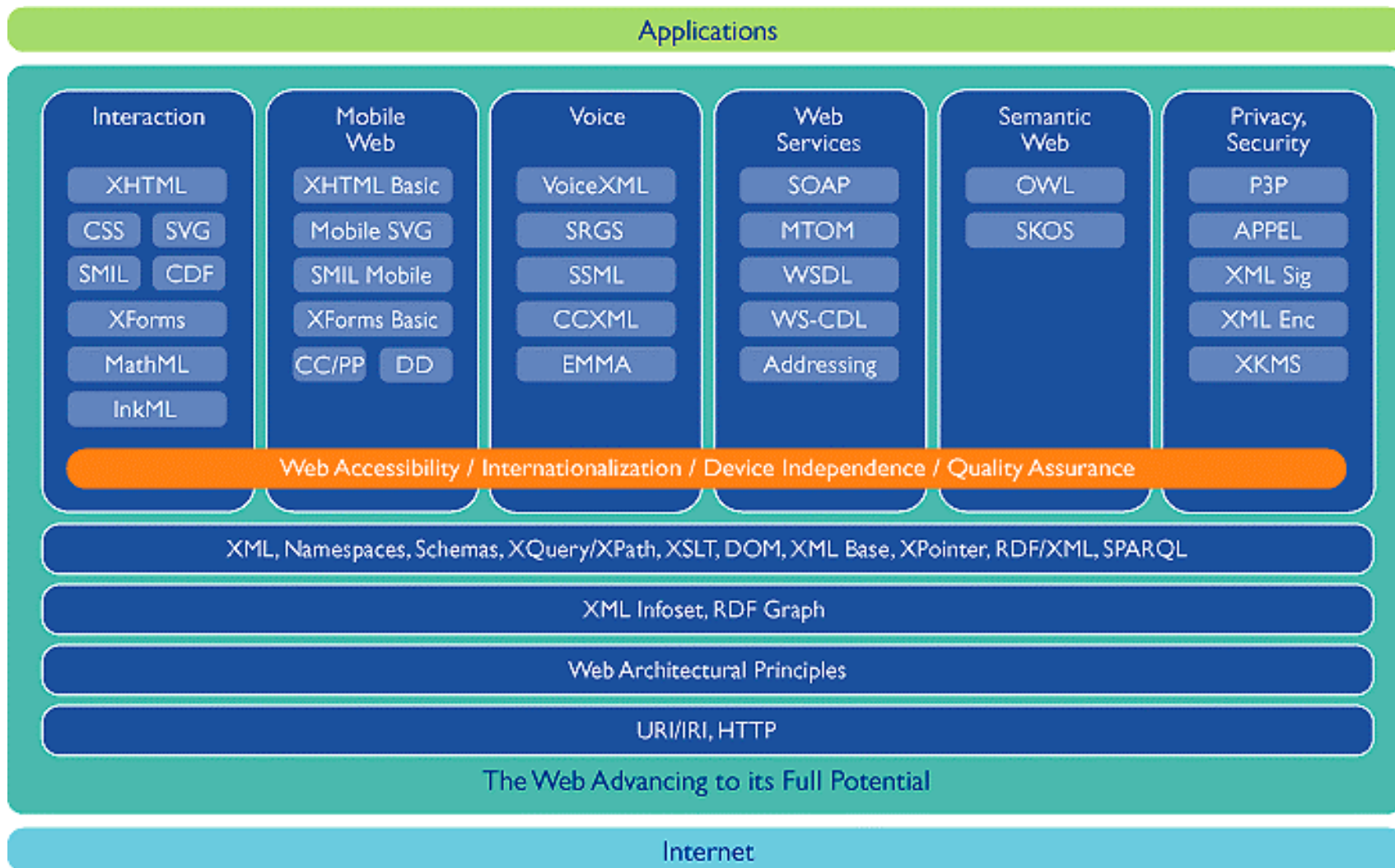
- Hyvin lyhyt versio:
 - World Wide Web Consortium (W3C) asettaa suosituksia universaaleista Web-tekniikoista. Tekniikoilla on sovelluksia myös paikkatietosovelluksissa, esim. tiedon esittämiseen, tarkistamiseen ja semanttiseen yhdistelyyn liittyen
- Pidempi versio, otsikoita:
 - W3C ja universaalit Web-standardit
 - tieto ja viestiminen: XML
 - tietolähteiden semanttinen yhdistely: Semanttinen Web (SW)
 - lopuksi

**World Wide Web Consortium
(W3C) ja universaalit
Web-standardit**

- W3C kehittää yhteensopivia teknologioita ja siten **ohjaa Webin kehittymistä täyteen mittaansa**
 - ...asettamalla teknisiä **suosituksia** (esim. HTML, XML, WAI)
- 3 päätstoa, 15 aluetstoa, n. 400 jäsenorganisaatiota



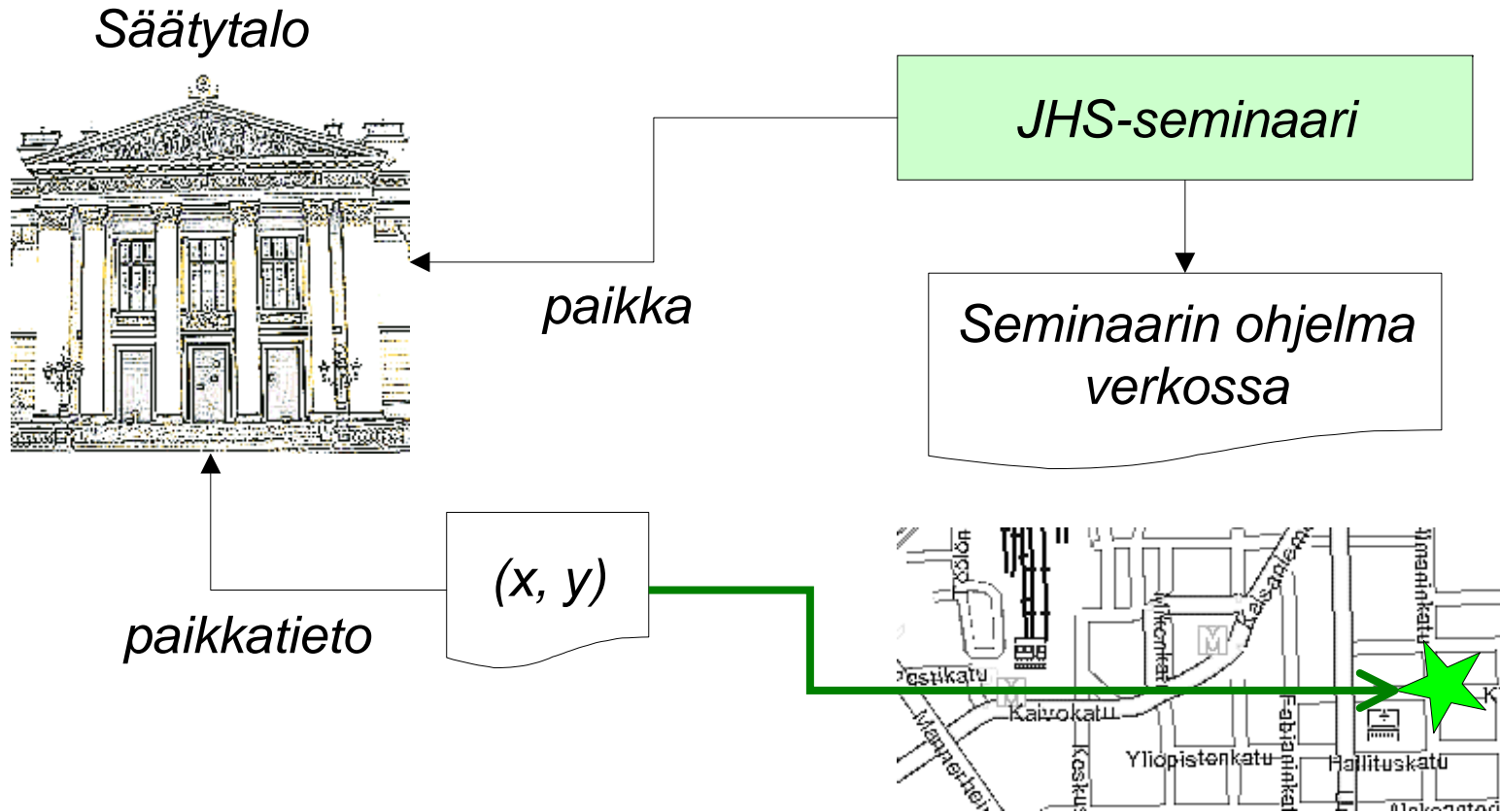
Jäseneksi? <http://www.w3c.tut.fi/joining.html>



(Paikka)tiedot ja Web-standardit

Esimerkki eri teknologioiden soveltamisesta

- "Halutaan kuvata ja yhdistää JHS-seminaarin paikkatieto ja karttasovellukseen liittyvä paikkatieto"



Lähestymistapa #1: XML ja tekstiformaatit

- Viestinvälitys edellyttää viestiformaatista sopimista
 - tietorakenne
 - tyyppimäärittely
- XML-suositus (*Extensible Markup Language*) määrittelee miten tietoa kirjoitetaan ja kuvataan täsmällisesti, esim.

```
<?xml version="1.0" encoding="iso-8859-1"?>
<location xmlns="http://www.example.org/location/"
  venueID="http://www.valtioneuvosto.fi/venue/saatyvalo">

  <desc xml:lang="fi">Säätytalo Helsingin kartalla.</desc>
  <coords map="http://www.freemaps.fi/hki/023002-200">
    <x>120</x>
    <y>155</y>
  </coords>
</location>
```

- XML-tekstiformaattiin nojautuminen tarjoaa välineitä tietorakenteiden mekaaniseen tarkistamiseen, esim.

```
..<xsd:element name="coords">
  <xsd:complexType>
    <xsd:sequence>
      <xsd:element name="x" type="xsd:string"/>
      <xsd:element name="y" type="xsd:string"/>
    </xsd:sequence>
    <xsd:attribute name="map" type="xsd:anyURI" use="required"/>
  </xsd:complexType>
</xsd:element> ...
```

- Tyypitieto on tärkeää käsittelijän näkökulmasta
 - Rakenne ok? Syötettiinkö pakollinen tieto? Onko x kokonaisluku?
 - Tietojen "järkevyyden" tarkistaminen ei toki tyhjentävästi onnistu

- Selkeä ja helppo ratkaisu
 - tieto teknisesti helposti hallittavassa muodossa
 - varsinainen haaste: sisällöistä sopiminen
- Suoraviivainen toteutus
 - erityisesti, XML-perheeseen on saatavilla hyviä **ohjelmointirajapintoja** ja **kyselykieliä** (esim. DOM ja XQuery)
- Mutta, keskeinen pulma: Koska tietorakenteet ovat **esitystavaltaan monimutkaisia**, eri lähteistä peräisin olevan tiedon semanttinen yhdistely saattaa olla työlästä

**Lähestymistapa #2:
SW ja tiedon semanttinen
integrointi**

- Tietojen semanttinen yhdistely onnistuu helpommin kun siirrytään käyttämään myös **yhteistä tietomallia**
 - ...RDF (*Resource Description Format*)
- Toki vieläkin tarvitaan tekstiformaatteja ja syötetietojen tarkistusta, mutta...
- **Mallinnuksen** ja erityisesti **tiedon julkaisun** idea on erilainen:
 - mallinnus tapahtuu (loogisten) **objektien**, **luokkien** ja **predikaattien** varassa
 - ...jotka identifioidaan toistuvasti käytettävien **URI-nimien** avulla
 - tietojen julkaisu esim. RDF/XML tai N3-formaatissa
- Tuloksena **semanttinen verkko** (ns. RDF-graafi)

Intuitiivinen esimerkki

Kohteen (tunniste)nimi

<http://www.valtioneuvosto.fi/venue/saatyalo>

x:venue

JHS:n julkaisema tieto

<http://www.jhs-suositukset.fi/seminaari2005-10-26>

x:homePage

<http://www.jhs-suositukset.fi/suomi/jhs-seminaari2005-10-26>

gis:venueID

gis:coords *gis:mapRef*

gis:x *gis:y*

155

120

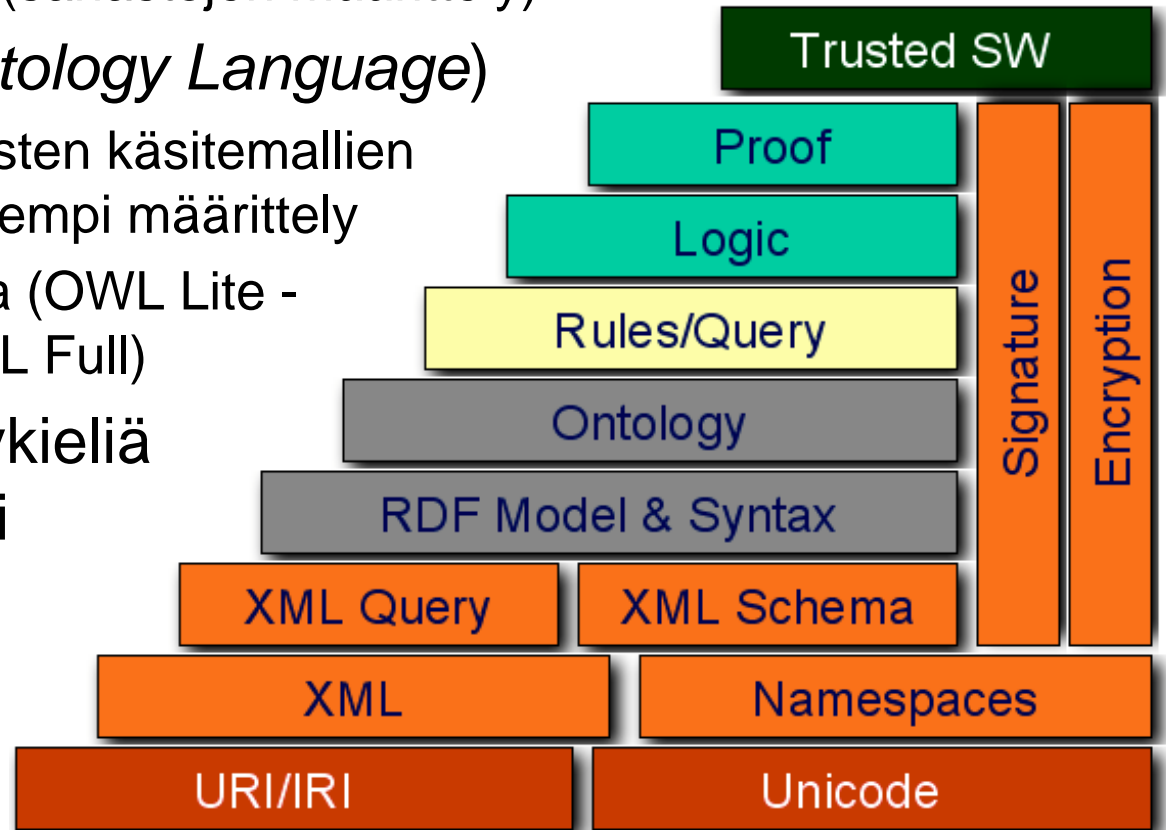
Kohteiden saatavilla oleva paikkatieto



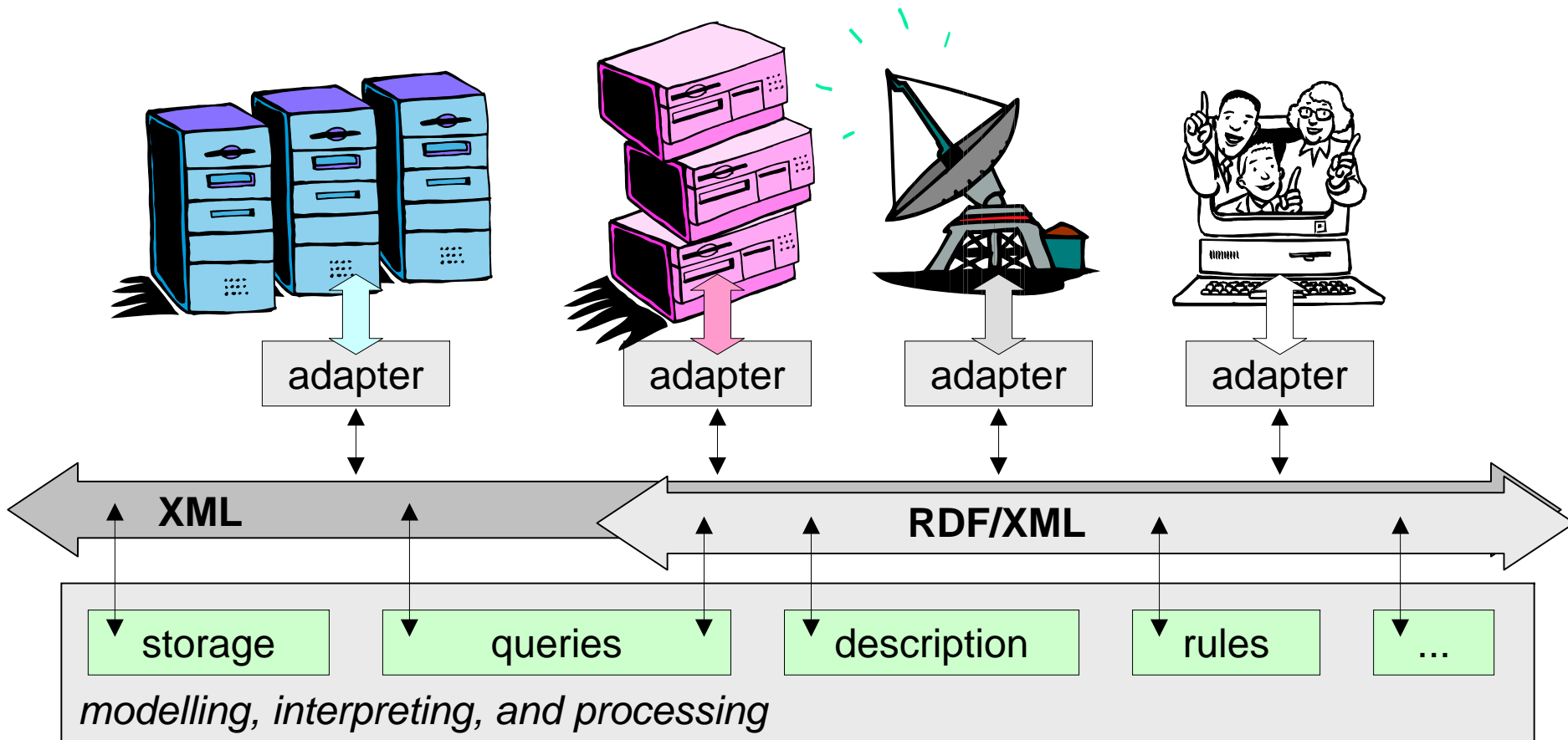
<http://www.freemaps.fi/hki/023002-200>

Saatavilla oleva karttatieto

- **RDF** (*Resource Description Format*)
 - yhtenäinen ja yksinkertainen tietomalli (semanttinen verkko)
 - RDF-skeemat (sanastojen määrittely)
- **OWL** (*Web Ontology Language*)
 - sovelluskohtaisten käsittemallien yksityiskohtaisempi määrittely
 - useita versioita (OWL Lite - OWL DL - OWL Full)
- *De facto* kyselykieliä yms. (standardi kehitteillä)



- Yleisemmin tarkasteltuna kyse on "vain" eri tietolähteiden yhdistelyyn ja tiedon tarkoituksenmukaiseen käsittelyyn liittyvistä menetelmistä



Lopuksi

- **World Wide Web Consortium (W3C)** kehittää Web-standardeja
 - työ luo yhteensopivaa perustaa erityyppisten sovellusten ja menetelmien rakentamiseen ("*enabling technologies*")
- Erityisesti, Web-tekniikat tarjoavat yleispätevän tavan myös ***tiedon*** esittämiseen ja käsittelyyn
 - vrt. "XML vs. Semanttinen Web"
- Kun tietolähteitä on sovelluksessa useita, nousevat semanttinen integroinnin näkökohdat keskeiseen rooliin

Jäikö joku W3C-asia mietityttämään?

Allekirjoittaneen tavoittaa helposti:

<http://www.w3c.tut.fi>

Ossi@w3.org



- W3C
 - <http://www.w3.org/> (*kotisivu*)
 - <http://www.w3c.tut.fi/> (*W3C Suomen toimiston kotisivu*)
- W3C-aktiviteetit
 - <http://www.w3.org/Consortium/Activities> (*yleiskuva*)
 - <http://www.w3.org/2001/sw> (*Semanttinen Web*)
- W3C:n suositukset, raportit ja teknistä tietoa
 - <http://www.w3c.org/TR/> (*kaikki tekniset dokumentit*)
 - <http://www.w3c.tut.fi/report-archive.html> (*yleisesityksiä suomeksi*)
- Jäseneksi!
 - <http://www.w3.org/Consortium/Prospectus>
 - <http://www.w3c.tut.fi/joining.html>