

Ontologianmäärittyskieli OWL

Hansi Keijonen

Seminaariraportti
HELSINGIN YLIOPISTO
Tietojenkäsittelytieteen laitos

Helsinki, 13. maaliskuuta 2013

Tiedekunta — Fakultet — Faculty		Laitos — Institution — Department	
Matemaattis-luonnontieteellinen		Tietojenkäsittelytieteen laitos	
Tekijä — Författare — Author			
Hansi Keijonen			
Työn nimi — Arbetets titel — Title			
Ontologianmäärittelyskieli OWL			
Oppiaine — Läroämne — Subject			
Tietojenkäsittelytiede			
Työn laji — Arbetets art — Level		Aika — Datum — Month and year	Sivumäärä — Sidoantal — Number of pages
Seminaariraportti		13. maaliskuuta 2013	19
Tiivistelmä — Referat — Abstract			
<p>Semanttisessa webissä ei ole ainoastaan dokumentteja ihmisten luettavaksi, vaan myös tietoa dokumenteissa olevien asioiden merkityksistä. Näitä merkityksiä voidaan tulkita ja tuottaa myös koneellisesti. Merkityksiä, semantiikkaa, voidaan ilmaista ontologioilla. Tietojenkäsittelytieteessä ontologialla tarkoitetaan dokumenttia, jossa kerrotaan asioiden välisistä yhteyksistä. Ontologiat ovat kehittyneen semanttisen webin perusta. OWL (Web Ontology Language) on ilmaisuvoimainen mutta kuitenkin koneellisesti pääteltävissä oleva ontologianmäärittelyskieli. Tässä seminaariraportissa esitellään OWL-kieli sekä muita teknologioita ja kieliä, kuten RDF ja RDF Schema, joiden konstruktioita OWL käyttää.</p>			
Avainsanat — Nyckelord — Keywords			
semanttinen web, RDF, RDF Schema, OWL			
Säilytyspaikka — Förvaringsställe — Where deposited			
Muita tietoja — Övriga uppgifter — Additional information			

Sisältö

1	Johdanto	3
2	Teknologiat ja kielet, joihin OWL perustuu	4
2.1	URI ja XML	4
2.2	Tiedon esittäminen RDF-kolmikoilla	5
2.3	Yksinkertaiset ontologiat RDF Schemalla	6
2.3.1	Luokka ja yksilö	7
2.3.2	Aliluokka	7
2.3.3	Ominaisuus ja aliominaisuus	7
2.3.4	Rajoitukset ominaisuuden sovellusalueessa ja arvojoukossa	8
3	Kehittyneitä ontologioita OWL:llä	8
3.1	OWL:n kolme alikieltä	9
3.2	OWL-ontologian rakenne	10
3.2.1	Nimiavaruudet	10
3.2.2	Otsikkotiedot	10
3.2.3	Yksinkertaiset luokat ja aliluokat	11
3.2.4	Luokan ilmentymät	12
3.2.5	Ominaisuudet	12
3.2.6	Aliominaisuudet	13
3.2.7	Kardinaalisuusrajoitteet	13
3.2.8	Ominaisuuden transitiivisuus, symmetrisyys, funktionaalisuus ja käänteisfunktionaalisuus	14
3.2.9	Arvorajoitteet luokkaominaisuuksissa	14
3.2.10	Kompleksiset luokat	15
3.2.11	Ontologioiden yhdistäminen	16
3.3	OWL 2	17
3.4	Päätelmiä	18
	Lähteet	19

1 Johdanto

Semanttinen web on visio tulevaisuuden webistä, jossa informaatiolle annetaan eksplisiittinen merkitys mahdollistaen näin koneiden kyky prosessoida ja yhdistellä webissä olevaa tietoa [7]. Vuonna 2001 Tim Berners-Lee, James Hendler ja Ora Lassila julkaisivat artikkelin "the Semantic Web", jossa he luonnehtivat semanttista webiä seuraavasti:

"Semanttinen web ei ole erillinen web vaan laajennos tämänhetkiseen webiin, jossa informaatiolle on annettu hyvin muotoiltu merkitys mahdollistaen koneiden ja ihmisten paremman yhteistyön.[suomennos kirjoittajan]

Suurin osa tämän päivän webin sisällöstä on tarkoitettu ihmisten luettavaksi sekä tulkittavaksi. Kone pystyy tulkitsemaan esimerkiksi html-tiedoston ja esittämään dokumentin siinä määritellyllä tavalla mutta se ei *ymmärrä* dokumentin sisällön *merkitystä*, semantiikkaa [2].Tämä rajoittaa esimerkiksi haut internetissä olevista dokumenteista yksinkertaiseksi hakusanojen etsimiseksi. Sen sijaan jos hakukoneet ymmärtäisivät asioiden merkityksen ja niiden välillä vallitsevat yhteydet, olisivat hakutulokset tarkempia ja sisältäisivät mahdollisesti laajennettuja hakuja alkuperäisen asian ympäriltä [5].

Semanttisella webillä on mahdollisuuksia myös verkkokaupankäynnissä, jossa myyjäagentit ja ostaja-agentit voivat kommunikoida keskenään tuotetietojen pohjalta luotujen ontologioiden avulla [5]. Myös eri toimijoiden tuottamien web-palveluiden koostamisessa semanttisen webin teknologioilla on keskeinen rooli: palveluiden tuottajat voivat kuvata tarjoamansa palvelun jolloin palveluita kokoavat sovellukset voivat niitä hyödyntää tehokkaasti [5]. Maailmanlaajuinen tietoverkko voisi muuttua dokumenttien verkosta tiedon verkoksi [2].

Semantiikkaa voidaan webissä ilmaista ontologioilla. Tietojenkäsittelytieteessä ontologialla tarkoitetaan dokumenttia, jossa kerrotaan asioiden välisistä yhteyksistä [2]. Ontologioiden luomiseksi pitää olla menetelmiä käsitteiden, käsitteiden ominaisuuksien sekä käsitteiden välisten suhteiden määrittämiseen [2].

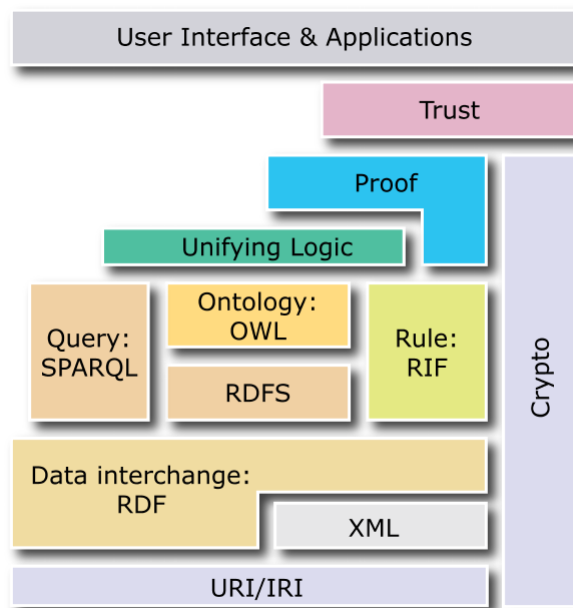
World Wide Web Consortium W3C on määritellyt joukon standardeja (suosituksia) kielille ja sanastoille, joilla merkitysten määrittäminen voidaan toteuttaa. Tässä artikkelissa selvitetään pääasiassa OWL-kielen periaatteita. Lyhenne OWL tulee sanoista Web Ontology Language ja sille on W3C:n suositus standardiksi vuodelta 2004. Uudempi suositus on OWL 2:lle vuodelta 2012.

Jotta olisi mahdollista ymmärtää OWL:n toimintaperiaate, on käytävä soveltuvien osien läpi myös teknologiat, joihin se perustuu. Tässä artikkelissa esitellään OWL:in lisäksi RDF ja RDF Schema, joiden konstruktioihin OWL vahvasti nojaa. Lopuksi selvitetään lyhyesti OWL:n laajennoksen OWL2:n tuomat lisäominaisuudet.

Artikkelin esimerkeissä käsitellään autoalan termistöä. RDF:n osalta tyydytään ainoastaan kuvaamaan joitain resursseja, mutta RDF Scheman ja OWL:n luvuissa pyritään määrittämään todellisia semantiikkoja aiheen piiristä. Esimerkeistä on poistettu nimiavaruusmäärittelyt pl. esimerkit, jotka nimenomaisesti keskittyvät nimiavaruuksiin. Tämä esimerkkien yksinkertaistamisen vuoksi.

2 Teknologiat ja kielet, joihin OWL perustuu

W3C tarjoaa suosituksia teknologioista ja kielistä semanttisen webin toteuttamiseen. Kuvassa 1 on semanttisen webin teknologiapino. Osa teknologioista on jo todellisuutta ja käytössä, osa vasta ideatasolla. Jokainen kerros hyödyntää alemman kerroksen toteuttamia palveluita. Seuraavissa kappaleissa käydään läpi kaavion teknologioita ja kieliä alhaalta ylöspäin kohti OWL:ää. Jokaisesta teknologiasta ja kielestä käsitellään tarkemmin ne konstruktiot, jotka ovat olennaisia ja käytössä myös OWL:ssä.



Kuva 1: Semanttisen webin toteutukseen tarvittavat teknologiat, kielet ja konseptit. Lähde: http://commons.wikimedia.org/wiki/File:Semantic_Web_Stack.png

2.1 URI ja XML

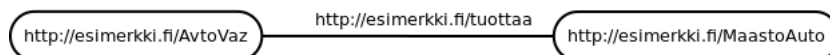
Semanttisessa webissä määritettyjä luokkia, ilmentymiä, ominaisuuksia ja ominaisuuksien arvoja kutsutaan resursseiksi [2]. Jotta sekaannusta jo määri-

tettyjen resurssien sekä uusien määritysten kanssa ei syntyisi, identifioidaan kaikki resurssit yksilöllisesti URI(Unified Resource Identifier):n avulla. URI:n avulla voidaan viitata mihin tahansa resurssiin webissä. Useimmiten URIna toimii perinteinen URL(Unified Resource Locator)-osoite [2]. IRI (Internationalized Resource Identifier) on ainoastaan merkistölaajennos URI:in.

Semanttisen webin kuvaukset toteutetaan useimmiten XML-kielillä. XML-kieli on notaatio monimutkaisen rakenteisen tiedon esittämiseen ja tarjoaa näin standardoidun mallin tiedon vaihtamiseen prosessoijien välillä. Tärkeä XML:n ominaisuus on nimiavaruudet, jotka mahdollistavat ja helpottavat resurssien identifiointia URI:en avulla. Nimiavaruuksien käyttöä selvitetään tarkemmin OWL-ontologioiden yhteydessä.

2.2 Tiedon esittäminen RDF-kolmikoilla

Resource Description Framework RDF on kieli webissä olevien resurssien kuvaamiseen [6]. RDF perustuu resurssien identifiointiin URI:lla ja näiden kuvaamiseen ominaisuuksilla ja ominaisuuksien arvoilla. Tämä mahdollistaa yksinkertaisten lausumien esittämisen verkkoina, joissa resurssit ja ominaisuuksien arvot ovat solmuja ja ominaisuudet verkon kaaria [6].



Kuva 2: RDF-kolmikko joka kuvaa yksinkertaisen lausuman. Jokainen kolmikon solmu ja kaari on identifioitu URI:lla.

Kuvassa 2 on kuvattu yksinkertainen *RDF-kolmikko*. Kolmikon subjekti, predikaatti ja objekti kertovat, että "AvtoVAZ tuottaa maastoautoja". Subjekti on siis asia, jota kuvataan, predikaatti on ominaisuus, jolla kuvataan ja objekti on ominaisuuden arvo [6]. Jokainen solu ja kaari on esimerkissä identifioitu URI:lla. Objekti voi olla myös literaali, jolloin sitä ei identifioida erikseen, mutta formaatti voidaan määritellä esimerkiksi XML Scheman datatyypin avulla [6].

Yleisin tapa esittää kolmikko on XML-notaatio. Myös muut tavat ovat mahdollisia, kuten esimerkiksi *JSON* ¹ ja *turtle* ². Alla on esitetty kuvassa 2 esitetyn kolmikon XML-notaatio:

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

```
<rdf:rdf
```

```
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:ex="http://www.esimerkki.fi/">
```

¹<http://www.json.org/>

²<http://www.w3.org/TeamSubmission/turtle/>

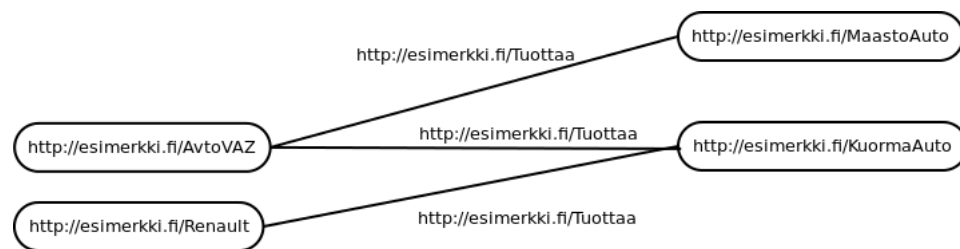
```

    <rdf:Description rdf:about="http://esim.fi/AvtoVAZ">
      <ex:Tuottaa rdf:resource="http://esim.fi/Maastoauto"/>
    </rdf:Description>
  </rdf:RDF>

```

Esimerkistä näkee, kuinka subjekti (AvtoVaz) ja objekti (Maastoauto) ovat identifioitu eksplisiittisellä URI:lla. Predikaattiin (Tuottaa) viitataan myös URI:lla, mutta nimiavaruuden kautta. Esimerkistä ilmenee RDF:n kolme peruskonstruktiota: *rdf:Description* ilmoittaa, että kyseessä on kuvaus, *rdf:about* viittaa subjektiin, *rdf:resource* viittaa resurssiin, tässä tapauksessa ominaisuuden arvoon.

Voimme laajentaa kuvan 2 esimerkkiä lisäämällä RDF-kolmikoita verkkoomme. AvtoVAZ tuottaa myös kuorma-autoja samon kuin Renault. Kun myös nämä kolmikot lisätään kuvan 2 kaavioon, alkaa pieni mutta informatiivinen verkko syntyä (kuva 3). Tästä verkosta voisimme tehdä jo hakuja, kuten "listaa yritykset, jotka tuottavat kuorma-autoja".



Kuva 3: RDF-kolmikot muodostavat verkon.

RDF tarjoaa melko alkeellisen tavan esittää lausumia, jotka muodostavat haut mahdollistavan verkon. RDF toteuttaa myös joukon muita ominaisuuksia, kuten *säiliöitä* (container) tiedon säilömiseen sekä *kokoomatietorakenteita* (collections) asioiden listaamiseen [6]. Näitä käsitellään niiltä osin kuin ne ovat relevantteja OWL-ontologioiden muodostamisessa.

2.3 Yksinkertaiset ontologiat RDF Schemalla

Vaikka RDF:llä voidaan kuvata resursseja ominaisuuksien avulla ja määrittää näin resurssien välisiä suhteita, se ei tarjoa keinoja määrittää itse luokkia, ominaisuuksia tai niiden välisiä suhteita ja hierarkioita [3]. Toisin sanoen, voimme määrittää, että "auton merkki on Lada", mutta emme voi määrittää esimerkiksi mitä arvoja ominaisuus "merkki" voi saada. Tämän mahdollistaa RDF:n laajennos RDF Schema.

RDF Schemalla on mahdollista määrittää resurssien joukkoja *luokiksi* (class). Luokalla voi olla *aliluokkia* (subclass) ja luokkien jäseniä kutsutaan *ilmentymiksi* (instance) tai *yksilöiksi* (individual). Luokkien ilmentymien välisiä suhteita kutsutaan *ominaisuuksiksi* (property) [3]. Ominaisuus siis

asettaa kahden ilmentymän välille suhteen. Myös ominaisuuksilla voi olla *aliominaisuuksia* (subproperty). RDF Schema on RDF:n semanttinen laajennos joka ei kuitenkaan tarjoa sanastoa asioiden ymmärtämiseen, vaan se tarjoaa työkaluja sanastojen luomiseen [3]. Näiden elementtien avulla voidaan määrittää melko yksinkertaisia luokkien ja suhteiden hierarkkisten järjestelmien kuvauksia, (kevyt)ontologioita.

Seuraavissa kappaleissa esitellään niitä RDF Scheman peruskonstruktioita, joita myös OWL käyttää toteutuksessaan lähes sellaisenaan. *RDF Schema määrittelee myös suurehkon joukon muita konstruktioita, mutta niiden käsittely on OWL:n esittelyn kannalta tässä tarpeetonta.* Huomionarvoista on, että RDF Schema -dokumentit ovat RDF-dokumentteja, jotka sisältävät kummankin kielen primitiivejä.

2.3.1 Luokka ja yksilö

Resurssit voidaan jaotella luokkiin. Luokkien jäseniä nimitetään luokan ilmentymiksi tai yksilöiksi[3]. Luokat ovat myös itse resursseja. Luokka määritellään `rdfs:Class` elementissä ja identifioidaan URI:lla. `rdf:type`-elementillä ilmaistaan, että resurssi kuuluu määrättyyn luokkaan [3]. Esimerkiksi AvtoVAZ on luokan Autonvalmistaja jäsen:

```
<rdfs:Class rdf:ID="AvtoVAZ" />
  <rdf:type rdf:resource="Autonvalmistaja"/>
</rdfs:Class>
```

2.3.2 Aliluokka

`rdfs:subClassOf`-elementillä voidaan ilmaista, että kaikki jonkin luokan jäsenet ovat myös jonkin toisen luokan jäseniä [3]. Esimerkiksi luokka Autonvalmistaja on luokan Liikkeyritys aliluokka:

```
<rdfs:Class rdf:ID="Autonvalmistaja" />
  <rdf:subClassOf rdf:resource="Liikkeyritys"/>
</rdfs:Class>
```

Kaikki yksilöt, jotka kuuluvat luokkaan Autonvalmistaja kuuluvat myös luokkaan Liikkeyritys.

2.3.3 Ominaisuus ja aliominaisuus

Resursien välisiä suhteita kuvataan ominaisuuksilla. Esimerkiksi "Lada 1200 A:n valmistusmaa on Venäjä". Voidaan ajatella, että resurssien Lada1200A ja Venäjä välinen suhde on ominaisuus Valmistusmaa. Ominaisuus Valmistusmaa on ominaisuuden Autonkuvailija aliluokka tai paremminkin aliominaisuus. Tällaisia konstruktioita kuvataan RDF Schemassa termillä `subPropertyOf`:


```

<rdf:Property rdf:about="Valmistusmaa">
  <rdfs:subPropertyOf rdf:resource="Autonkuvailija" />
</rdf:Property>

<rdf:Description rdf:about="Lada1200A">
  <Valmistusmaa rdf:resource="Venäjä" />
</rdf:Description>

```

Esimerkissä on määritetty ensiksi ominaisuus *Valmistusmaa*, joka on ominaisuuden *Autonkuvailija* aliominaisuus. Seuraavaksi kuvataan yksilöä *Lada1200A* ominaisuudella *Valmistusmaa*, joka saa arvokseen resurssin *Venäjä*. *subPropertyOf* ilmaisee määritelmällisesti, että resurssit joiden välistä suhdetta kuvataan jollain ominaisuudella, voidaan kuvata myös tämän suhteen *yliominaisuudella*, josta ko. suhde on periytynyt [3]. *Lada1200A*n ja *Venäjän* suhdetta voidaan kuvailla ominaisuudella *Valmistusmaa*, mutta myös ominaisuudella *Autonkuvailija*.

2.3.4 Rajoitukset ominaisuuden sovellusalueessa ja arvojoukossa

RDF Schemassa on mahdollista määrittää ominaisuuksille rajoituksia sen suhteen minkä luokkien jäsenten välistä suhdetta ominaisuus voi kuvata. Ensinnäkin voidaan rajoittaa ominaisuuden sovellusaluetta (domain). Esimerkiksi ominaisuuden *Valmistusmaa* sovellusalue voidaan rajata koskemaan ainoastaan luokkaa *Tuote*. Samoin ominaisuuden saamat arvot (range) voidaan rajata olemaan ainoastaan luokan *Valtio* ilmentymiä:

```

<rdf:Property rdf:about="Valmistusmaa">
  <rdfs:domain rdf:resource="Tuote" />
  <rdfs:range rdf:resource="Valtio" />
</rdf:Property>

```

Sovellusalueen rajausta määritetään *rdfs:domain* -elementillä ja arvojoukon rajausta *rdfs:range* -elementillä.

3 Kehittyneitä ontologioita OWL:llä

Kehittyneempi tapa ilmaista ontologioita on OWL-ontologiat. Kuten aiemmin todettiin, OWL tulee sanoista Web Ontology Language. Vaikka määritelmässä on sana language, kieli, on OWL-ontologiat ymmärrettävä ennemminkin sanastoina, joita on kuvattu RDF-kielellä. Joukko semanttisen webin parissa työskenteleviä ohjelmistokehittäjiä on luonnehtinut osuvasti RDF-kolmikoiden ja OWL-ontologioiden eroa verkkokurssisivullaan ³. Tekijät lähestyvät eroa siten, että verrattaan RDF:ää ja OWL:ia perinteiseen

³<http://www.linkeddatatools.com/>

relaatiotietokantaan. RDF-kolmikot on tapa tallettaa tietoa olioiden ominaisuuksista samalla tavalla kuin relaatiotietokannan taulun riveillä tallennetaan rakenteista tietoa tietokantaolioista. Jokaista riviä relaatiotietokannassa yksilöi yksilöivä avain kun taas RDF-kolmikoissa avaimen virkaa hoitaa URI. Relatiotietokannoissa tietokantaolion attribuuttien suhteita ilmaistaan taulurakenteilla ja tietokantaolioiden suhteita toisiin tietokantaoloihin ilmaistaan viitteillä taulujen välillä. Vastaavasti OWL-ontologiat kuvaavat ja jäsentävät samalla tavalla RDF-kolmikoilla ilmaistua tietoa: ontologialla voidaan kuvailla monipuolisesti luokkia ja niiden ilmentymiä sekä ilmentymien suhteita toisiin ominaisuuksien avulla. RDF on kuin kieli jolla ilmaistaan lausumia asioista kun taas OWL-ontologia on sanasto, jonka avulla lausumien merkitykset ymmärretään.

OWL tarjoaa ontologioiden määrittelyyn [1]:

- *hyvin määritellyn kieliopin* , jotta ontologiat olisivat koneluettavissa
- *hyvin määritellyn semantiikan*, jolla voi ilmaista merkityksiä tarkasti ja konsistentisti
- *tuen koneelliselle päättelylle*, jotta esim. ontologioiden eheys voidaan tarkastaa automaattisesti
- *riittävästi ilmaisuvoimaa* ilmaisemaan kaikki tarvittavat merkitykset
- *miellyttävän ilmaisutavan*, jotta työskentely olisi sujuvaa

Ideaalisesti OWL on RDF:n ja RDF Scheman laajennos [1]. OWL käyttää RDF:n luokkia ja ominaisuuksia lisäten niihin omia laajennoksiaan. RDF Schemassa on joitain hyvin vahvoja konstruktioita, kuten `rdf:Class` (kaikkien luokkien ylliluokka) sekä `rdf:Property` (kaikkien suhteiden ylliluokka).

Näiden primitiivien ilmaisuvoima yhdistettynä OWL:n tarjoamaan laajennoksiin on ristiriidassa sen tavoitteen kanssa, että ontologiat olisivat koneellisesti pääteltävissä. Tämä tasapainotila mielessäpitäen on määriteltä kolme OWL:n alikieltä sen perusteella, painotetaanko ilmaisuvoimaa vai koneellista päättelyä [1].

3.1 OWL:n kolme alikieltä

W3C:n Web Ontology Working Group on määritellyt OWL:lle kolme alikieltä, joiden on tarkoitus toteuttaa eri aspektit (ilmaisuvoima, koneellinen päättely), joita ontologioiden kuvaamiskieleltä vaaditaan [7]. Alikielet ovat ilmaisuvoiman mukaisesti kasvavassa järjestyksessä:

- *OWL Lite* tarjoaa ainoastaan yksinkertaisen luokitteluhierarkian ja yksinkertaiset rajoitteet [7]. Kardinaalisuusrajoitteet ovat ainoastaan muotoa 0 ta 1. Työkalutuen tarjoaminen on OWL Litelle helpompaa kuin ilmaisuvoimaisemmille versioille [7].

- *OWL DL* on tarkoitettu niille käyttäjille, jotka haluavat mahdollisimman hyvän ilmaisukyvyn siten, että ontologia on koneellisesti pääteltävissä ja kaikki päätelmät tehtävissä järjestetystä ajassa [7]. *OWL DL* tarjoaa kaikki kielen konstruktiot, mutta niitä voi käyttää tietyin rajoituksin, esimerkiksi luokka voi olla monen luokan aliluokka mutta ei voi olla samalla luokan ilmentymä. *DL* tulee sanoista Description Logics, deskriptiivinen logiikka, joka on eräs logiikan ala ja on *OWL:n* perusta [7].
- *OWL Full* on tarkoitettu niille käyttäjille, jotka haluavat maksimaalisen ilmaisukyvyn välittämättä siitä, onko ontologiat enää koneellisesti pääteltävissä [7]. Toisin kuin *OWL DL:ssä*, luokka voi olla kokoelma yksilöitä (instansseja) samalla kuin luokka itsessään on jonkin luokan instanssi. *OWL Full* mahdollistaa jo olemassa olevien ontologioiden laajentamisen. On epätodennäköistä, että mikään ohjelmisto pystyy täydellisesti päättelemään *OWL Full* ontologioita [7].

3.2 OWL-ontologian rakenne

Kaikki *OWL*-ontologiat ovat *RDF*-dokumentteja ja niiden yleisin serialisointitapa on *XML* [9].

3.2.1 Nimiavaruudet

OWL-dokumentissa tulee määritellä nimiavaruudet (namespace). Nimiavaruuksien avulla voidaan ratkaista mm. samannimisten elementtien aiheuttamia tulkintaongelmia sekä kertoa lukijalle (koneelle tai ihmiselle) konteksti, jonka mukaan elementtien tageja tulee tulkita. *OWL*-ontologiassa nimiavaruudet määritellään *rdf:RDF* -tagin sisään. Alla olevassa esimerkissä on eräs mahdollinen nimiavaruusmäärittely .

```
<rdf:RDF
  xmlns:owl ="http://www.w3.org/2002/07/owl#"
  xmlns:rdf ="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:xsd ="http://www.w3.org/2001/XMLSchema#">
```

Esimerkkimäärittelyssä nimiavaruudet on määritetty *owl:-*, *rdf:-* ja *rdfs:-* etuliitteille kertomaan, että näillä etuliitteillä varustetut tagit edustavat *OWL:n*, *RDF:n* ja *RDF Scheman* termistöä. Esimerkkiontologiassa käytetään myös *XMLSchema*-datatyyppiä (*xsd:*), joten myös niiden nimiavaruus tulee määrittää.

3.2.2 Otsikkotiedot

OWL-ontologian otsikkotiedoissa voidaan kertoa yleisiä asioita kuten versiotietoa, kommentteja ja ontologian nimi. Tärkeä ominaisuus on mahdollisuus

tuoda toisen tahon määrittelemiä ontologioita itse määriteltävän ontologian käyttöön [9]. Otsikkotiedot määritellään `owl:Ontology`-elementiksi:

```
<owl:Ontology>
  <rdfs:label>Esimerkkiontologia</rdfs:label>
  <rdfs:comment>Esimerkin voimaa</rdfs:comment>
  <owl:priorVersion>
    rdf:resource="http://www.esimerkki.com/vanhempi#"
  </owl:priorVersion>
  <owl:imports>rdf:resource="http://esimerkki.fi/lentokoneet.owl"</owl:imports>
</owl:Ontology>
```

Esimerkkiotsikossa perustietojen kertomisen lisäksi tuodaan ontologian käyttöön lentokoneet-ontologia.

3.2.3 Yksinkertaiset luokat ja aliluokat

OWL:n, samoin kuin RDF Scheman yksi perusajatus on, että on olemassa asioiden joukkoja eli luokkia [9]. Luokan jäseniä sanotaan myös sen ilmentymiksi. OWL-ontologiassa kaikki ilmentymät ovat myös luokan `owl:Thing` jäseniä ja kaikki käyttäjän määrittämät luokat `owl:Thing`in aliluokkia [9]. Myös `owl:Nothing` on määritelty. Luokilla on myös aliluokkia. Kaikki aliluokan jäsenet kuuluvat myös ylliluokkaansa [9]. Luokkamäärittelyt tapahtuvat hyvin samaan tapaan kuin RDF Schemassa, ainoastaan elementin nimi on `owl:Class`. Aliluokan määrittely on suoraan RDF Schemasta:

```
<owl:Class rdf:ID="Auto">
<owl:Class rdf:ID="Ajoneuvo">

<owl:Class rdf:ID="Auto">
  <rdfs:subClassOf rdf:resource="Ajoneuvo"/>
</owl:Class>
```

Esimerkissä on määritetty kaksi luokkaa `Auto` ja `Ajoneuvo` sekä määritelty edellinen jälkimmäisen aliluokaksi.

OWL:ssä on myös mahdollista määritellä luokka siten, että luokan jäsenet voivat kuulua johonkin luokkaan mutta eivät missään tapauksessa toiseen ilmoitettuun luokkaan [9]:

```
<owl:Class rdf:about="Ajoneuvo">
  <owl:disjointWith rdf:resource="Elain"
</owl>
```

Luokan `Ajoneuvo` ja sen mahdollisten aliluokkien jäsenet eivät voi siis kuulua samaan aikaan luokkaan `Elain` tai sen mahdollisiin aliluokkiin. Voidaan siis päätellä, että jos ilmentymä `Lada1200A` on jonkun `Ajoneuvo` -luokan aliluokan jäsen, se ei voi olla esimerkiksi luokan `Kissa` jäsen jos `Kissa` on määritetty luokan `Elain` aliluokaksi.

3.2.4 Luokan ilmentymät

Ilmentymän kuuluminen johonkin luokkaan ilmaistaan samalla tavalla kuin RDF Schemassa `rdf:type` -konstruktiolla [9]:

```
<owl:Thing rdf:ID="AvtoVAZ"/>

<owl:Thing rdf:about="AvtoVAZ">
  <rdf:type rdf:resource="Autonvalmistaja"/>
</owl:Thing>
```

Esimerkissä määritetään `Autonvalmistaja`-luokan ilmentymä `AvtoVAZ`. Luokka `Autonvalmistaja` on pitänyt määrittää toisaalla, jotta siihen voidaan näin viitata.

3.2.5 Ominaisuudet

Ominaisuuksien avulla voidaan kuvailla asioita, kuten esimerkiksi "auton merkki on Lada". Luodaan siis suhde kulkuväline auton ja Lada-brändin välille. Ominaisuudet jaetaan OWL:ssä luokkaominaisuuksiksi (Object Property) ja datatyyppiominaisuuksiksi (Data Type Property) sen perusteella liittääkö ominaisuus yhteen kaksi luokan ilmentymää vai luokan ilmentymän ja RDF-literaalin tai XML Schema -datatyyppin [9].

OWL:ssä luokkaominaisuudet määritellään kuten mitkä tahansa luokat, mutta käyttäen `ObjectProperty` -konstruktiota. Määrittelyssä kerrotaan minkä luokan ilmentymiin ominaisuus on sovellettavissa sekä mitä arvoja (luokan ilmentymiä) ominaisuus voi saada. Nämä rajoitteet ilmaistaan RDF Schemasta tutuilla `rdfs:domain`- ja `rdfs:range`-elementeillä [9]:

```
<owl:ObjectProperty rdf:ID="Merkki">
  <rdfs:domain rdf:resource="Auto"/>
  <rdfs:range rdf:resource="Autonvalmistaja"/>
</owl:ObjectProperty>
```

Esimerkissä on määritelty ominaisuus `Merkki`. Kyseisellä ominaisuudella voi kuvata ainoastaan luokan `Auto` ilmentymiä ja joka voi ainoastaan saada arvokseen ainoastaan luokan `Autonvalmistaja` ilmentymiä.

Datatyyppiominaisuus liittää yhteen luokan ilmentymän ja arvoliteraalin, RDF-literaalin tai XML Schema -datatyyppin arvon. Myös datatyyppiominaisuuksien määrittelyssä kerrotaan minkä luokan ilmentymiin ominaisuus on sovellettavissa sekä minkä tyyppisiä arvoja ominaisuus voi saada [9]. Rajoitteet tehdään samoilla elementeillä kuin luokkaominaisuuden määrittelyssä:

```
<owl:DatatypeProperty rdf:ID="Valmistusvuosi">
  <rdfs:domain rdf:resource="Auto"/>
  <rdfs:range rdf:resource="&xsd:gYear"/>
<owl:datatypeProperty>
```

Esimerkissä määritetään ominaisuus `Valmistusvuosi`, jolla voidaan kuva-
ta luokan `Auto` ilmentymiä ja joka voi saada arvokseen XML Schema -
standardissa määritetyn `gYear`-tyypin arvon. XML-Schema -standardi mää-
rittelee 19 eri datatyyppiä. Kun käytetään määritettyjä datatyyppejä, XML-
parseri pystyy tarkistamaan, onko annetut arvot sallittuja ja esitystapa
skeeman mukainen [10].

3.2.6 Aliominaisuudet

Myös ominaisuuksille voidaan määrittää aliominaisuuksia samaan tapaan
kuin luokille voidaan määrittää aliluokkia. Aliominaisuus toteutetaan `subPropertyOf`-
elementillä, joka on toteutettu jo RDF Schemassa [9]. Alla olevassa esimer-
kissä lisätään ominaisuuteen `Valmistusvuosi` määritys, että se on ominai-
suuden `AutonKuvailija` aliominaisuus:

```
<owl:Class rdf:ID="AutonKuvailija">

<owl:DatatypeProperty rdf:ID="Valmistusvuosi">
  <rdfs:subPropertyOf rdf:resource="AutonKuvailija">
  . . .
</owl:datatypeProperty>
```

3.2.7 Kardinaalisuusrajoitteet

OWL antaa mahdollisuuden määrätä kardinaalisuusrajoitteita ominaisuuksien arvoille kun ominaisuutta sovelletaan tietyssä kontekstissa [9]. Voidaan esimerkiksi määrätä, että luokka `Ajoneuvo` on joukko asioita, joilla on aina vähintään kaksi `Renkaat`-ominaisuutta. Toisin sanoen, määritellään *anonyymi aliluokka* [9] luokalle `Ajoneuvo`. Koska anonyymi aliluokka on määritetty `Ajoneuvo`-luokkamäärittelyn sisällä, on kaikki `Ajoneuvon` ilmentymät myös määritetyn aliluokan ilmentymiä. Alla oleva esimerkki selventää asiaa:

```
<owl:Class rdf:ID="Ajoneuvo">
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="Renkaat">
        <owl:minCardinality rdf:datatype="xsd:nonNegativeInteger">2
      </owl:minCardinality>
    </owl:Restriction>
  </owl:subClassOf>
</owl:Class/>
```

Anonyymi aliluokka määritetään `subClassOf`-konstruktiolla. Aliluokalle määritetään rajoite `owl:Restriction`-elementillä ja kardinaalisuusrajoite tässä tapauksessa `minCardinality`-elementillä. Kardinaalisuusrajoite määritetään

koskemaan ominaisuutta **Renkaat** onProperty-elementillä. Kardinaalisuusrajoitteita voi määrätä myös **maxCardinality**-elementillä, joka määrää suhteen yläkardinaliteetin, sekä **Cardinality**-elementillä, joka määrää tarkan arvon mikä on luokan ja ominaisuuden suhde. Voidaan esimerkiksi määrittää luokka **Moottoripyörä** siten, että sillä voi olla tasan kaksi **Renkaat** -ominaisuutta.

3.2.8 Ominaisuuden transitiivisuus, symmetrisyys, funktionaalisuus ja käänteisfunktionaalisuus

Ominaisuuden voi tarvittaessa määrittää transitiiviseksi, symmetriseksi, funktionaaliseksi tai käänteisfunktionaaliseksi implisiittisen päättelyn helpottamiseksi [1]:

- *owl:TransitiveProperty* määrittää transitiivisen ominaisuuden, kuten "isompi kuin", "pidempi kuin"[1]. Myös **rdfs:SubClassOf** on transitiivinen ominaisuus.
- *owl:SymmetricProperty* määrittää symmetrisen ominaisuuden, kuten "työkaveri", "sisar"[1].
- *owl:FunctionalProperty* määrittää funktionaalisen ominaisuuden, jolla on korkeintaan yksi uniikki arvo määritettäväänsä kohden, kuten "ikä", "nimi", "pituus"[1]
- *owl:InverseFunctionalProperty* määrittää ominaisuuden, jonka arvo ei voi olla sama kahdella määritettävällä oliolla. Esimerkiksi ominaisuus **Sarjanumero** autojen yhteydessä ei voi olla sama kahdella autolla. Kun tiedetään yksi sarjanumero, voidaan päätellä mitä autoilmentymää tarkoitetaan [1].

Esimerkissä määritetään ominaisuus **Sarjanumero** käänteisfunktionaaliseksi:

```
<owl:DatatypeProperty rdf:ID="Sarjanumero">
  <rdf:type rdf:resource="&owl:InverseFunctionalProperty" />
  <rdfs:domain rdf:resource="Auto"/>
  <rdfs:range rdf:resource="&xsd:positiveInteger" />
</owl:datatypeProperty>
```

3.2.9 Arvorajoitteet luokkaominaisuuksissa

Edellisessä kappaleessa esitellyt tavat määrittää rajoitteita ominaisuuksille toimivat globaalilla tasolla, ts. ne ovat voimassa kaikissa sovellustapauksissa [9]. Seuraavat rajoitteet, kuten myös kardinaalisuusrajoitteet, ovat voimassa ainoastaan niissä luokissa, joiden sisällä rajoitteet määritetään [9]. Esimerkiksi on mahdollista määrätä, että luokan **Auto** kontekstissa ominaisuuden **Omistaja** arvona voi olla vain luokan **Henkilo** jäseniä. Tämä saadaan aikaan **owl:allValuesFrom** -elementillä:

```

<owl:Class rdf:ID="Auto">
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="Omistaja"/>
      <owl:allValuesFrom rdf:resource="Henkilo"/>
    </owl:Restriction>
  </rdfs:subClassOf>
</owl:Class>

```

Esimerkissä määritettiin, että auton omistaja on aina Henkilo-luokan jäsen. OWL mahdollistaa myös hieman kevyemmän rajoitteen, jossa määrätään, että ominaisuuden *jonkin* arvon tulee olla määrätyn luokan jäsen [9]. Voidaan määrittää, että autojen edellisistä omistajista ainakin yksi on luokan Henkilo jäsen:

```

. . .
      <owl:onProperty rdf:resource="EdellinenOmistaja"/>
      <owl:someValuesFrom rdf:resource="Henkilo"/>
. . .

```

Rajoite toteutetaan owl:someValuesFrom-elementillä.

On myös mahdollista määrätä ominaisuuden arvoksi täsmällisesti joku olemassa oleva resurssi [9]. Tällöin aliluokkamäärittelyyn lisätään owl:hasValue -elementti:

```

. . .
      <owl:onProperty rdf:resource="Omistaja"/>
      <owl:hasValue rdf:ID="Hansi"/>
. . .

```

3.2.10 Kompleksiset luokat

Kompleksiset luokat on erittäin vahva ja ilmaisuvoimainen OWL:n konstruktio [9]. Ne tarjoavat mahdollisuuden määrittää luokkausekkeitä, joissa käytetään joukko-opista tuttuja operaattoreita *yhdistettä*, *leikkausta* ja *komplementtia*. Esimerkiksi voimme määrittää luokan Lada-auto, joka on luokan Auto ja sellaisten asioiden, jotka ovat jollain tapaa Lada, leikkaus:

```

<owl:Class rdf:ID="Lada-auto">
  <owl:intersectionOf rdf:parseType="Collection">
    <owl:Class rdf:about="Auto"/>
    <owl:Class rdf:about="Lada"/>
  </owl:intersectionOf>
</owl:Class>

```


Esimerkissä leikkaus-operaattori on ilmaistu `owl:intersectionOf` -elementillä. `rdf:parseType="Collection"` ilmoittaa jäsentäjälle, että tämän elementin sisältämät elementit tulee tulkita lueteltuna kokoelmana [9].

Vastaavasti voidaan määrittää luokkien yhdiste käyttämällä elementtiä `owl:unionOf` [9]:

```
<owl:Class rdf:ID="Maansiirtokone"/>
  <owl:unionOf rdf:parseType="Collection">
    <owl:Class rdf:about="Kauhakuormaaja"/>
    <owl:Class rdf:about="Dumpperi"/>
  </owl:intersectionOf>
</owl:Class>
```

Esimerkissä luokkaan `Maansiirtokone` kuuluu ilmentymät, jotka ovat joko `kauhakuormaajia` tai `dumppereita` tai peräti molempia.

Komplementtiluokka määritellään elementillä `owl:complementOf` [9]. Komplementtiluokan jäseniä määrittää se, että ne eivät missään tapauksessa ole komplementiksi ilmoitetun luokan jäseniä. Komplementtiluokalla on helppo esimerkiksi määrittää luokka, `Manuaalivaihteisto`, joka määritelmällisesti on luokan `Automaattivaihteisto` komplementti:

```
<owl:Class ref:about="Manuaalivaihteisto">
  <owl::complementOf rdf:resource="Automaattivaihteisto">
</owl:Class>
```

Esimerkkimäärittäyksessä luokkaan `Manuaalivaihteisto` kuuluvat vain ne ilmentymät, jotka eivät kuulu luokkaan `Automaattivaihteisto`.

3.2.11 Ontologioiden yhdistäminen

Jotta ontologioiden luominen olisi mielekästä, tulee olla kyky yhdistää ontologioita toisiinsa laajempien ontologioiden aikaansaamiseksi [9]. Ontologian otsikkotiedoissa voidaan import-lasuseella tuoda jonkun toisen jo olemassa olevan ontologian määrittäykset oman ontologian käyttöön. OWL tarjoaa elementtejä, joilla tuoduissa ontologioissa määritettyjä voi käyttää apuna omissa määrittäyksissä [9]. `owl:equivalentClass` -elementillä voidaan oma luokkamäärittäys määrätä vastaamaan jotain olemassaolevaa määrittäystä:

```
<owl:Class rdf:ID="Pienkone">
  <owl:equivalentClass rdf:resource="&lentokoneet;Yksitasokone"/>
</owl:Class>
```

Esimerkissä siis oletetaan, että otsikkotiedoissa on import-määrittäys ontologialle `lentokoneet`. Oma määrittäys luokalle `Pienkone` asetetaan vastaamaan tuodun ontologian määrittäystä luokalle `Yksitasokone`.

OWL mahdollistaa myös määrittää yksittäisen ilmentymän vastaamaan toista ilmentymää [9]. Tämä tehdään `sameAs`-konstruktiolla:

```

<owl:Class rdf:ID="HansinSuosikkiAuto">
  <owl:sameAs rdf:resource="Lada1200A"/>
</owl:Class>

```

Tällä konstruktiolla ei ole suurtakaan käyttöä. Ainoa, mitä tässä ilmaistaan on, että Hansin suosikkiauto on Ladan klassikkomalli. Käänteinen ominaisuus `sameAs`-konstruktiolle on `differentFrom` [9]. Sen käyttö on samaan tapaan suoraviivaista eikä se trivialiteettien lisäksi määritä mitään oleellista.

Käyttökelpoisempi konstruktio on `AllDifferent`, jolla voidaan määrittää kokoelma ilmentymiä olemaan erillisiä toistensa suhteen [9]:

```

<owl:AllDifferent>
  <owl:distinctMembers rdf:parseType="Collection">
    <Autonvalmistaja rdf:about="AvtoVAZ"/>
    <Autonvalmistaja rdf:about="Renault"/>
    <Autonvalmistaja rdf:about="Daimler-Benz"/>
  </owl:distinctMembers>
</owl:AllDifferent>

```

Esimerkki kertoo, että jokainen Autonvalmistajan ilmentymä viittaa eri ilmentymään.

3.3 OWL 2

OWL 2 on laajennos OWL-kieleen, jolle on annettu W3C-suositus vuonna 2009 [8]. OWL 2:n ominaisuudet ovat suurilta osin samat kuin OWL 1:ssä, osa ainoastaan eri nimillä esitettynä. OWL 1:llä tarkoitetaan tässä OWL:n vuoden 2004 suositusta standardiksi, jota tässä artikkelissa on käsitelty. Myös OWL 2:n ytimen muodostaa XML ja RDF ja mikä tärkeintä, OWL 2 on taaksepäin yhteensopiva OWL 1:n kanssa. Jokainen OWL 1 -ontologia on siis kelvollinen OWL 2 -ontologia [8]. OWL 2 tekee pääasiassa samat asiat kuin OWL 1 hieman erilaisella kieliopilla. Se tarjoaa kuitenkin myös joitain uusia toiminnallisuuksia [8]:

- *avaimet*, joka mahdollistaa sen, että ominaisuuksia (tai ominaisuuksien joukkoa) voidaan käyttää avaimena tunnistamaan luokan ilmentymää.
- *ominaisuusketjut*, jotka mahdollistavat ominaisuuksien ketjutuksen. Voidaan esimerkiksi määrittää ominaisuus "Isovanhempi" kahden "Vanhempi-ominaisuuden ketjuksi
- *uusia datatyyppejä*, jotka mahdollistavat ominaisuuksille annettavien numeeristen arvojen rajoittamisen jollekin välille. Esimerkiksi luokalle "Teini-ikäinen" voidaan määrätä, että "Ikä-ominaisuus voi saada ainoastaan arvoja väliltä [13..19].

- *asymmetriset, refleksiiviset ja poissulkevat ominaisuudet.* Asymmetrisyys on vastakohta OWL 1:n symmetrisyydelle, reflektiivinen ominaisuus viittaa aina itseensä ja poissulkevat ominaisuudet eivät voi olla voimassa samaan aikaan.
- *uusia kommentointimahdollisuuksia,* jotka mahdollistavat uusia tapoja kommentoida ontologioita.

3.4 Päätelmiä

Kun semanttisen webin teknologioita upotetaan web-sovelluksiin, puhutaan usein web 3.0:sta [4]. Kuitenkaan web 3.0:n konseptit eivät ole vielä leikanneet läpi koko verkkomaailman web 2.0:n tavoin. (Web 2.0:na käsitetään web-sivut, jotka eivät enää ole stattisia dokumentteja vaan niissä käyttäjät ovat vuorovaikutuksessa toistensa kanssa ja luovat omaa sisältöään sivuille [11]). Web 3.0. ei ole synnyttänyt samanlaisia ihmisten arkipäivään kiinteästi nivoutuneita web-palveluita kuin Facebook, Flickr tai Twitter. Tämä ei kuitenkaan tarkoita, etteikö semanttinen web ja web 3.0 kehittyisi edelleen.

W3C Semantic Web Action tekee tutkimustyötä semanttisen webin alueella ja on vastuussa jo olemassa olevista semanttisen webin suosituksista standardeiksi. Semanttinen web on tutkimuskohteena myös useissa yliopistoissa maailmanlaajuisesti. Esimerkiksi suomalainen tutkimusryhmä Semantic Computing Research Group (SeCo) Aalto Yliopistosta tutkii koneellisesti pääteltävien semantiikkojen mahdollisuuksia semanttisen webin kontekstissa.

Myös kuluttajille suunnattuja semanttisen webin palveluita on tarjolla. Esimerkkinä voidaan mainita tripit.com, joka muodostaa matkasuunnitelman automaattisesti palveluun lähetettyjen tilausvahvistusten pohjalta. Ei-kaupallinen DBpedia-projekti⁴ pyrkii luomaan Wikipedian tietojen pohjalta kattavan linkitetyn datan tietokannan. DBpedia tarjoaa tietoa Wikipedian lisäksi monesta eri lähteestä koottuna ja mahdollisuuden tehdä kyselyjä, jotka hyödyntävät Wikipedian ja muiden lähteiden tietoja yhdellä kertaa.

Semanttisen webin teknologioilla on mahdollisuudet mullistaa koko tiedon haun ja tuottamisen mallit webissä.

⁴<http://dbpedia.org/About>

Lähteet

- [1] Antoniou, Grigoris ja Harmelen, Frank Van: *Web Ontology Language: OWL*. Handbook on ontologies, 2(September):91–110, 2009. <http://www.springerlink.com/index/10.1007/978-3-540-92673-3>.
- [2] Berners-Lee, Tim, Hendler, James ja Lassila, Ora: *The Semantic Web*. Scientific American, 284(5):34–43, 2001. <http://www.nature.com/doifinder/10.1038/scientificamerican0501-34>.
- [3] Brickley, Dan ja Guha, R V: *RDF Vocabulary Description Language 1.0: RDF Schema*. W3C Recommendation, 2009(10 February 2004), 2004. <http://www.w3.org/TR/rdf-schema/>.
- [4] Hendler, J: *Web 3.0 Emerging*, 2009. <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4755170>.
- [5] Horrocks, I: *From SHIQ and RDF to OWL: the making of a Web Ontology Language*. Web Semantics Science Services and Agents on the World Wide Web, 1(1):7–26, 2003. <http://linkinghub.elsevier.com/retrieve/pii/S1570826803000027>.
- [6] Manola, Frank ja Miller, Eric: *RDF Primer*. W3C Recommendation, 10(February):1–107, 2004. <http://www.w3.org/TR/rdf-primer/>.
- [7] McGuinness, Deborah L ja Van Harmelen, Frank: *OWL Web Ontology Language Overview*. W3C recommendation, 10(February):1–22, 2004. <http://www.w3.org/TR/owl-features/>.
- [8] Ontology, Web, Document, Language, Recommendation, Latest, Group, O W L Working ja Reserved, All Rights: *OWL 2 Web Ontology Language Document Overview*. October, 2(October):1–12, 2009. <http://www.w3.org/TR/owl2-overview/>.
- [9] Smith, Michael K, Welty, Chris ja McGuinness, Deborah L: *OWL Web Ontology Language Guide*. W3C Recommendation, 10(February):1–46, 2004. <http://www.w3.org/TR/owl-guide/>.
- [10] Wikipedia: *XML Schema (W3C)*, 2012. [http://http://en.wikipedia.org/wiki/XML_Schema_\(W3C\)](http://http://en.wikipedia.org/wiki/XML_Schema_(W3C)), [Online; noudettu 8.3.2013].
- [11] Wikipedia: *Web 2.0*, 2013. http://http://http://en.wikipedia.org/wiki/Web_2.0, [Online; noudettu 13.3.2013].