

Kristiina Huhtanen

**ONTOLOGIOIDEN KEHITTÄMINEN
LAINSÄÄDÄNTÖTYÖTÄ VARTEN**

Tietojärjestelmätieteen
pro gradu -tutkielma
25.11.2003

Jyväskylän yliopisto
Tietojenkäsittelytieteiden laitos
Jyväskylä

TIIVISTELMÄ

Huhtanen, Kristiina Marianne

Ontologioiden kehittäminen lainsäädäntötyötä varten / Kristiina Huhtanen

Jyväskylä: Jyväskylän yliopisto, 2003.

120 s.

Pro gradu -tutkielma

Tässä tutkielmassa tarkastellaan ontologioiden kehittämistä suomalaista lainsäädäntötyötä varten. Ontologialla tarkoitetaan yhteisesti sovittua ja muodollisesti esitettyä käsitteistöä, joka on tarkoitettu tietokoneen tulkittavaksi. Tutkielman tavoitteena on kartoittaa, miten lainsäädäntöprosessin tietojärjestelmien yhteentoimivuutta tukeva ontologia voitaisiin kehittää. Tarkoituksena on keskittyä erityisesti kehittämisen eri vaiheisiin.

Lainsäädäntötyöhön liittyy lukuisia tietojärjestelmiä, jotka on toteutettu eri aikoina toisistaan irrallisina käyttäen erilaisia tekniikoita ja loogisia rakenteita. Tämä vaikeuttaa tietojen hakemista ja hyödynnettävyyttä. Ongelma voidaan ratkaista osittain integroimalla järjestelmät ontologioiden avulla semanttisella tasolla. Tutkimusaiheen käsittely perustuu aikaisemmista tutkimuksista, kirjallisuudesta ja kohdeympäristöstä koottuun tietoon. Tutkielman aluksi on perehdytty ontologioihin ja niiden suunnitteluun. Tämän jälkeen on kuvattu sekä EU:n että Suomen lainsäädäntöprosessia. Lopuksi on pohdittu ongelmaa ja esitetty yhden keskeisen tietojärjestelmien integroimista tukevan ontologian kehittämistä. Ehdotus etenemistavaksi sisältää suunnittelu-, määrittely- ja toteutusvaiheet, joihin sisältyy suosituksia muun muassa lähestymistavoista.

AVAINSANAT: ontologia, ontologioiden kehittäminen, lainsäädäntötyö

ABSTRACT

Huhtanen, Kristiina Marianne

Developing Ontologies for Legislative Work / Kristiina Huhtanen

Jyväskylä: University of Jyväskylä, 2003.

120 p.

Master's Thesis

This thesis examines the development of ontologies for the Finnish legislative work. An ontology is understood as a collectively agreed and formally represented conceptualization. The aim of the study is to find out how an ontology supporting interoperability of information systems related to the Finnish legislative process could be developed. The main focus is on different stages of the development process.

The legislative work involves several information systems that have been implemented separately at different times using various techniques and logical structures. This makes it difficult to access and exploit information. The problem can be partially solved by integrating systems with ontologies at the semantic level. This study is based on observations from scientific research, literature and target environment. At first basic knowledge about ontologies and their development is explained. After that legislative processes of European Union and Finland are described. Finally the development of one central ontology that supports the integration of information systems is suggested. The proposal for a development process contains some recommendations and following stages: planning, specification and implementation.

KEYWORDS: ontology, development of ontology, legislative work

SISÄLLYS

1 JOHDANTO	6
2 ONTOLOGIAT JA NIIDEN TARKOITUS.....	11
2.1 Ontologian määritelmä ja perusidea	11
2.2 Ontologian rakenne	14
2.3 Ontologioiden luokittelua	19
2.4 Ontologioiden käyttötapoja	22
2.5 Esimerkkejä ontologioista	24
3 ONTOLOGIOIDEN SUUNNITTELU	29
3.1 Suunnittelun periaatteita ja lähestymistapoja	29
3.2 Ontologioiden kehittämismenetelmiä	34
3.3 Ontologioiden esityskieliä	38
3.3.1 Perinteiset esityskielet	39
3.3.2 Web-kielet.....	41
3.3.3 Web-perustaiset esityskielet	43
3.4 Ontologiatyökaluja	46
4 EU:N LAINSÄÄDÄNTÖTYÖ	51
4.1 Eurooppaoikeus, yhteisön oikeus ja säädösten muodot	52
4.2 Lainsäädäntöprosessi.....	54
4.2.1 Vireilletulo- ja valmisteluvaiheet	57
4.2.2 Käsittely- ja päätöksentekovaiheet	59
4.2.3 Täytäntöönpano.....	61
4.3 Tietojärjestelmät	62
4.4 Ontologia- ja sanastotyö EU:n alueella	64
5 SUOMEN LAINSÄÄDÄNTÖTYÖ	66
5.1 EU-asioiden käsittely Suomessa	66
5.1.1 Valmistelu ja käsittely valtioneuvostossa	69
5.1.2 Käsittely eduskunnassa	71
5.2 Suomen kansallinen lainsäädäntöprosessi	73
5.2.1 Valmistelu ja hallituksen esityksen käsittely valtioneuvostossa ...	74
5.2.2 Lakiesityksen käsittely eduskunnassa	75
5.2.3 Hyväksytyjen lakien käsittely valtioneuvostossa	76
5.3 Tietojärjestelmät	77
5.4 Ontologia- ja sanastotyö suomalaisessa julkishallinnossa	79

6 LAINSÄÄDÄNTÖPROSESSIN JÄRJESTELMIEN YHTEENTOIMIVUUTTA	
TUKEVAN ONTOLOGIAN KEHITTÄMINEN	83
6.1 Nykytilan ongelmia ja tavoitetila	83
6.2 Tietojärjestelmien ontologiaperustainen integrointi	85
6.3 Ehdotus etenemistavaksi.....	90
6.3.1 Suunnitteluvaihe	91
6.3.2 Määrittelyvaihe	95
6.3.3 Toteutusvaihe.....	98
6.4 Ehdotetun etenemistavan arviointi.....	102
7 YHTEENVETO	104
LÄHDELUETTELO.....	107

1 JOHDANTO

Ontologioita koskeva tutkimus on tällä hetkellä suosittua, laaja-alaista ja aktiivista. Ontologioiden merkitys on huomioitu monilla tutkimusalueilla, joita ovat muun muassa tietämyksen esittäminen ja hallinta (ks. esim. Sowa 2000), älykäs tiedon yhdistäminen (ks. esim. Wiederhold 1996), tiedonhallinta (ks. esim. Kim 2000) ja tiedonhaku (ks. esim. Benjamins & Fensel 1998). Ontologiat ovat hyvin ajankohtaisia myös semanttisen webin standardien kehittyessä nopeasti (ks. esim. Hyvönen 2002). Tutkimuksen lisäksi ontologioita on sovellettu myös käytännössä. Guarino (1998, 3) luettelee esimerkkeinä seuraavat alat: luonnollisen kielen kääntäminen, lääketiede, konetekniikka, tuotetiedon standardointi, elektroninen kaupankäynti ja digitaaliset kirjastot. Käsitteistö, menetelmät, formaalisuus ja ontologian kohteena olevan sovellusalueen laajuus vaihtelevat käytännön sovellutuksissa.

Tämän tutkielman kannalta kiinnostavin ontologioiden tutkimusalue on tietojärjestelmien yhteentoimivuus, joka on tullut yhä tärkeämmäksi viime aikoina organisaatioiden yrittäessä maksimoida tehokkuuttaan ja tietovarantojensa hyödynnettävyyttä. Organisaatiomuistit tallennetun tiedon ja sisäisten prosessien muodossa kasvavat räjähdysmäisellä vauhdilla sekä yksityisellä että julkisella sektorilla. Huomattava osa organisaatioissa tallennetusta tiedosta on hajaantunut erilaisiin dokumentteihin, tietojärjestelmiin ja sovellusohjelmiin. Useissa tapauksissa mahdollisuudet hakea, jakaa ja hyödyntää tehokkaasti tätä tietoa ovat rajoitettuja. Tilanne aiheuttaa tehottomuutta organisaatioprosesseissa, sillä työntekijöiden työaika kuluu huomattava määrä tiedon etsimiseen. Lisäksi asioita joudutaan tekemään moneen kertaan, jos etsittyjä dokumentteja ja tietoja ei löydykään mistään. Tällainen skenaario on tuttu monissa organisaatioissa.

Tämän tutkielman tekemistä ovat motivoineet erityisesti ongelmat suomalaisessa lainsäädäntötyössä. Lainsäädäntötyöllä tarkoitetaan tässä yhteydessä toimintaa, jonka perimmäisenä tavoitteena on tarkoituksenmukaisten lakien säätäminen. Tämän toiminnan tehokkuutta vähentävät eri toimijoiden eli valtioneuvoston ja eduskunnan heterogeeniset ja hajautetut tietojärjestelmät. Heterogeenisyys aiheutuu tietojärjestelmien erilaisista tekniikoista ja loogisista rakenteista, jotka ovat seurausta eri aikoina koordinoimattomasti toteutetuista tietojärjestelmien kehittämisprojekteista. Järjestelmien heterogeenisuus vaikeuttaa tietojen jakamista, saatavuutta ja hyödynnettävyyttä eri toimintayksiköissä ja niiden välillä. Tietojärjestelmien ja niihin liittyvien tietovarantojen integrointi yhteentoimivuuden saavuttamiseksi vaatii taustalla olevien mallien ja käsitteistöjen tuomista esiin ja esittämistä eksplisiittisesti. Tätä käsitteistön eksplisiittistä esittämistä kutsutaan ontologiaksi (Gruber 1993, 199).

Ontologiat tarjoavat yhteisen ymmärryksen tietystä asiasta tai alasta, joka edelleen mahdollistaa kommunikaation ihmisten ja tietokonejärjestelmien välillä. Ontologiat ovat monikäyttöisiä, sillä ne pystyvät kuvaamaan minkä tahansa alan semantiikkaa. Ontologioiden avulla järjestelmät voidaan integroida semanttisella tasolla, jolloin yhteentoimivuus saavutetaan sopimalla eri tietojärjestelmien käyttämien käsitteiden yhteisistä merkityksistä. Ontologioiden kehittämisellä tarkoitetaan tässä tutkielmassa ontologioiden suunnittelua ja luomista.

Tutkielma on tehty osana RASKE2-projektia <<http://www.it.jyu.fi/raske/>>, jossa kehitetään vuosina 2003-2006 menetelmiä organisaatioverkostojen hajallaan olevien tietojärjestelmien ja web-palveluiden integrointiin metatietojen

standardoinnin avulla. Tutkimuksessa paneudutaan erityisesti suomalaisen lainsäädäntöprosessin integrointitarpeisiin ja semanttisen webin tekniikoiden käyttöönottoon. Projektin tutkimus toteutetaan Jyväskylän yliopistossa, käytännön kehitysratkaisut suunnitellaan ja toteutetaan valtioneuvostossa ja eduskunnassa. Hanke on jatkoa aikaisemmille RASKE- ja EULEGIS-projekteille, joissa eduskunta, eräät ministeriöt ja Jyväskylän yliopisto kehittivät yhteistyössä uusia ratkaisuja sähköisten asiakirjojen hallintaan.

RASKE-projektissa (Rakenteisten ASiakirjojen KEhittäminen) (ks. esim. Salminen 2003) kehitettiin vuosina 1994-1998 menetelmiä dokumenttien standardoimiseen. Menetelmiä kehitettiin ja testattiin eduskunnan dokumenttien SGML-standardoinnin yhteydessä. RASKE-projektin työtä jatkettiin vuosina 1998-2000 EULEGIS-projektissa (European User Views to Legislative Information in Structured Form) (ks. esim. Lyytikäinen, Tiitinen & Salminen 2003). EULEGIS-projektin aikana kehitettiin prototyyppi käyttöliittymästä, joka yhdisti useita Euroopan oikeudellisia tietokantoja. Käyttöliittymässä XML-muotoinen kontekstuaalinen metatieto esitettiin käyttäjille graafisina tietomalleina ja tekstihakua tuettiin ontologian avulla.

Tämän tutkielman tavoitteena on kartoittaa, miten suomalaisen lainsäädäntöprosessin tietojärjestelmien yhteentoimivuutta tukeva ontologia voitaisiin kehittää. Tutkielma pyrkii toimimaan alustavana ontologian kehittämisprosessin suunnitelmana ja kehittämistyötä koskevien valintojen perustana. Ontologian kehittämisen tarkastelussa painotetaan erityisesti eri vaiheiden ja niihin liittyvien tehtävien tunnistamista. Tutkielma rajoittuu ontologian kehittämisprosessin kuvailemiseen ja kartoittamiseen. Ontologian varsinainen toteuttaminen ja käyttöönotto rajataan tämän tutkielman ulkopuolelle. Tutkimusmenetelmänä on käytetty tapaustutkimusta (engl. case

study), sillä ilmiötä on pyritty kuvailemaan ja aineisto on rajattu yhteen tapaukseen. Tutkimuksen tietolähteitä ovat olleet tieteelliset artikkelit, kirjallisuus, RASKE- ja EULEGIS-projektien raportit, kohdeympäristöön liittyvät raportit ja julkaisut, RASKE2-projektin johtoryhmän kokoukset, asiantuntijoiden epämuodolliset haastattelut ja Internet-lähteet.

Tutkielman luvussa 2 selvitetään ontologian määritelmä, perusidea ja rakenneosat sekä erilaisia luokitteluperusteita ja käyttötapoja. Lisäksi kuvataan lyhyesti muutamaa erityyppistä ontologiaa ja tunnetuimpia lakiontologioita. Luvussa 3 tarkastellaan ontologioiden suunnittelun periaatteita ja lähestymistapoja sekä kartoitetaan kehittämismenetelmiä, esityskieliä ja työkaluja. Luvussa 4 tutustutaan lyhyesti Euroopan unionin lainsäädäntötyöhön ja siihen liittyviin järjestelmiin sekä EU:n alueen ontologia- ja sanastotyöhön. Luvussa 5 kuvataan suomalaista lainsäädäntötyötä sekä EU-asioiden valmistelun että kansallisesta näkökulmasta. Lisäksi esitellään lyhyesti suomalaisen lainsäädäntötyön tietojärjestelmiä ja julkishallinnon ontologia- ja sanastotyötä. Luku 6 kartoittaa nykytilan ongelmia ja tavoitetilaa sekä tietojärjestelmien ontologiaperustaista integrointia. Lisäksi ehdotetaan tapaa, jolla edetä ontologian kehittämisprosessissa. Luvussa 7 esitetään tutkielman yhteenveto ja kartoitetaan jatkotutkimusmahdollisuuksia.

Tutkielman keskeisenä tuloksena ehdotetaan yhden keskeisen tietojärjestelmien integroimista tukevan ontologian kehittämistä ja etenemistapaa kehittämisprosessille. Ehdotus etenemistavaksi sisältää suunnittelu-, määrittely- ja toteutusvaiheet, joihin sisältyy erilaisia tehtäviä. Suunnitteluvaiheessa päätetään ontologian tarkoitus ja laajuus, mallinnetaan aihealuetta ja hankitaan aineistoa. Määrittelyvaiheessa kartoitetaan käsitteitä ja muita ontologian rakenneosia sekä tehdään niistä kuvioita ja taulukoita.

Toteutusvaiheessa koodataan ontologia esityskielellä ja arvioidaan valmista ontologiaa. Vaiheisiin on sisällytetty suosituksia tehtävistä valinnoista, kuten esimerkiksi käsitteiden määrittelyyn käytettävästä strategiasta.

Tutkimuksesta on hyötyä eduskunnalle ja valtioneuvostolle, koska sitä voidaan soveltaa suoraan kehitettäessä ontologiaa suomalaista lainsäädäntötyötä varten. Etenemistavan ehdotuksen yleisyydestä johtuen se soveltuu käytettäväksi myös muilla aihealueilla, tosin hieman epäsuoremmin. Tiedeyhteisölle tutkielmasta on hyötyä siten, että se kokoaa yksiin kansiin tietoa ontologioista ja niiden suunnittelusta.

2 ONTOLOGIAT JA NIIDEN TARKOITUS

Tässä luvussa selvitetään, mitä ontologiat ovat ja mistä ne rakentuvat. Lisäksi kartoitetaan ontologioiden luokittelutapoja ja käyttötapoja. Lopuksi esitellään erityyppisiä ontologioita ja tutustutaan tunnetuimpiin lakiontologioihin.

2.1 Ontologian määritelmä ja perusidea

Filosofian alueella ontologia (engl. ontology) tarkoittaa olemassaolon ja olemassaolevan tutkimista (ks. esim. Grüninger & Lee 2002, 39). Tietokoneiden, tietokantojen ja tekoälyn aikakaudella ontologialle on syntynyt uusi merkitys, mutta yksimielisyyttä määritelmästä ei ole vielä saavutettu ja toisistaan poikkeavia tulkintoja on siten useita (Guarino 1998).

Tässä tutkielmassa tukeudutaan Gruberin määritelmään, jonka monet ajattelevat parhaiten luonnehtivan ontologian ydinolemusta (esim. Fensel 2000). Gruberin (1993, 199) mukaan *ontologia* on formaali, eksplisiittinen määrittely yhteisestä käsitteistöstä (engl. "An ontology is a formal, explicit specification of a shared conceptualization."). Toisin sanoen ontologia määrittelee ja esittää täsmällisesti tietyn aihealueen tai ilmiön käsitteet ja niiden väliset suhteet. Täsmällinen tarkoittaa sitä, että määrittelyn on oltava täysin yksiselitteinen ja myös koneellisesti tulkittavaksi soveltuva. Ontologian tarkoituksena on mahdollistaa tiedon tallentaminen, jakaminen, yhdistäminen ja yhteiskäyttö yhteisesti määritetyllä tavalla. Näin estetään yksilön käsityksen korostuminen sekä ontologian luomisen että käyttämisen aikana.

Geerts ja McCarthy (2000, 4) ovat määritelleet kolme ontologioille ominaista piirrettä: Ontologian tavoitteena on esittää käsitteistö, joka on jaettavissa ja

uudelleenkäytettävissä, ja jossa tietyn sovelluksen ominaispiirteet jätetään huomiotta. Ontologia kattaa kaikki aihealueen (engl. domain) sovellukset, ei vain yhtä. Ontologia sisältää selkeästi eriteltyjä ja rajoitettuja määrittelyjä tietämyksestä sekä sääntöjä lisätietämyksen päättelystä varten.

Holsapplen ja Joshin (2002, 42) mukaan ontologia on yhteinen kieli, jonka avulla voidaan jakaa ja käyttää uudelleen tietämystä mielenkiinnon kohteena olevasta ilmiöstä. Grüninger ja Lee (2002, 40) ovat samaa mieltä todeten, että alkujaan ontologiatekniikoiden kehittämisen motiivina oli tekoälyn alalla tarve jaettavissa ja uudelleenkäytettävissä oleville tietämyskannoille. Informaation ja tietämyksen jakaminen ja uudelleenkäyttö mahdollistuvat ontologioiden avulla, koska yhteisen kielen käyttämisellä voidaan taata, että kaikki ovat samaa mieltä käsitteistä, luokista, ominaisuuksista ja rajoitteista (Hendler 2001, 32).

Fenselin (2000, 1) mukaan ontologioilla on tietämyksen organisoinnissa samantapainen tehtävä kuin tietokantakaavioilla. Hänen mukaansa ne kuitenkin myös eroavat toisistaan seuraavilla tavoilla:

- Ontologioiden määrittelykieli on syntaktisesti ja semanttisesti rikkaampi kuin yleiset tietokantojen toteuttamisen lähestymistavat.
- Ontologian kuvaava informaatio koostuu puolirakenteisista luonnollisen kielen teksteistä eikä taulukkomuotoisesta tiedosta.
- Ontologian pitää olla jaettu ja yhteisymmärrykseen perustuva käsitteistö, koska sitä käytetään tiedon jakamiseen ja vaihtamiseen. Tietokantakaavio taas voi olla vain yhden tietokannan käytössä.
- Ontologia on aihealuateoria, ei tietovaraston rakenne.

Ontologioita ei pidä myöskään sekoittaa tietämyskantoihin. Ontologia voi tarjota perusrakenteen, jonka ympärille tietämyskanta voidaan rakentaa.

Ontologia siis sisältää peruskäsitteet tietyn aihealueen kuvaamiseen, mutta ei itse tietämystä. Tietämiskanta taas käyttää kyseisiä käsitteitä esittämään sitä, mikä on totta tietyssä kontekstissa. (Swartout & Tate 1999, 19) Siten lainsäädännön ontologia voi sisältää määrittelyt termeille ”vankeusrangaistus” ja ”oikeudenistunto”, mutta ei väitteitä siitä, että tietty henkilö on tuomittu vankeusrangaistukseen tietyssä oikeudenistunnossa. Tietämiskanta taas voi sisältää tämänkaltaisia tietoja.

Ontologiat voivat sisältää linkkejä toisiinsa sillä, mitä useammat ihmiset käyttävät yhteisiä käsitteitä, sitä helpommin informaatiota voidaan jakaa ja käyttää uudelleen (Hendler 2001, 31-32). XML-perustaisten nimiavaruuksien (Bray, Hollander & Layman 1999) käyttö mahdollistaa nämä ontologioiden väliset viittaukset (Maedche & Staab 2001, 79). Esimerkiksi ontologiassa A koulu voi tarkoittaa rakennusta, ontologiassa B organisaatiota ja C-ontologiassa voidaan käyttää englanninkielen school-sanaa. Nimiavaruuksien määrittelymisen avulla nämä ontologiat voidaan yhdistää ilman päällekkäisyyksiä.

Ontologioiden loogisten ja järjestelmällisten rakenteiden ansiosta niistä on hyötyä useilla eri tutkimusalueilla. Järjestelmät, joihin liittyvät tiedon hakeminen, esittäminen tai mallintaminen, hyötyvät ontologioista. Tällaisia ovat muun muassa erilaiset tiedonhakujärjestelmät, digitaaliset kirjastot ja Internetin hakuohjelmat (Chandrasekaran, Josephson & Benjamins 1999, 23). Nämä järjestelmät tarvitsevat ontologioita tiedon järjestämiseen ja hakuprosessin ohjaamiseen. Ontologiat nähdään hyödyllisinä myös erilaisia ongelmanratkaisumenetelmiä käyttävissä sovelluksissa, sillä ontologiaan voidaan kohdistaa kyselyjä ja sitä voidaan selata (Guarino 1998, 10).

Kiinnostus ontologioihin on kasvanut, kun tutkijat ovat havainneet tiedon jakamisen ja uudelleenkäytön tarpeellisuuden (Hendler 2001, 31-32). Ontologiat mahdollistavat tämän entistä tehokkaammin. Tähän asti yksi suurimmista esteistä tiedon tehokkaalle jakamiselle ja uudelleenkäytettävyydelle on ollut eri järjestelmien tietyistä aihealueesta muodostaman käsitteistön ja mallinnuksen erilaisuus (Maedche & Staab 2001, 74). Tietyn aihealueen pohjalta muodostettua ontologiaa voidaan käyttää eri järjestelmissä, koska ontologian semantiikka on asiayhteydestä ja lukijasta riippumatonta (Chandrasedaran ym. 1999, 20).

2.2 Ontologian rakenne

Ontologian rakenne on riippuvainen ontologian kuvaamiseen käytetystä ontologiakielestä ja sen formalismista. Suppeat kielet tarjoavat mahdollisuuden käyttää vain yksinkertaisimpia rakenteita, kun taas laajemmat kielet mahdollistavat monimutkaisempienkin rakenteiden käyttämisen. Seuraavassa esitetyt rakenneosat ovat kokoelma ontologioissa nykyisin käytettävistä rakenteista, joita yksikään ontologiakieli ei tue täydellisesti. Toisaalta kokoelma ei myöskään ole täydellinen vaan tarpeiden ja keinojen kehittyessä ontologioille syntyy uusia rakenteita. Tässä kohdassa ei esitellä ontologioiden esityskieliä, sillä niitä kartoitetaan kohdassa 3.3.

Cómez-Pérez ja Corcho (2002, 55) ovat listanneet rakenneosia, joiden avulla ontologiatietämystä voidaan määritellä. Näitä komponentteja ovat käsitteet (engl. concepts), suhteet (engl. relationships), funktiot (engl. functions), aksioomat (engl. axioms) ja ilmentymät (engl. instances). Seuraavassa kyseisten rakenneosien tarkemmassa esittelyssä on käytetty lähteenä Cómez-Pérezin ja Corchon (2002, 55-57) artikkelia. Esimerkit on pyritty keksimään itse koko tutkielmassa.

Käsite voi olla esimerkiksi tehtävän, toiminnon, strategian tai päättelyprosessin kuvaus. Käsitteet voivat olla abstrakteja tai konkreettisia, peruskäsitteitä tai yhdistettyjä, todellisia tai kuvitteellisia. Käsitteitä voidaan kutsua myös luokiksi (engl. classes), olioiksi (engl. objects) tai kategorioiksi (engl. categories). Käsitteet muodostavat taksonomian eli luokkahierarkian (engl. class hierarchy), jossa on alaluokkia (engl. subclass) ja yläluokkia (engl. superclass). Hierarkiassa alimmaisena ovat ilmentymät (engl. instance), jotka ovat kyseessä olevaan luokkaan liittyviä yksilöitä. Käsitteille voidaan määritellä ominaisuuksia eli attribuutteja (engl. attributes) ja niiden välillä voidaan ilmaista suhteita (engl. relations). Perinnöllisyyden (engl. inheritance) kautta attribuutti voi periä alaryhmille, jolloin ilmentymälle määritellään kyseiselle attribuutille jokin arvo (engl. value).

Käsitteeseen liittyy yleensä yksi tai useampia *attribuutteja*. Attribuutit tunnetaan myös nimillä lokerot (engl. slots) tai ominaisuudet (engl. properties). Attribuutteja on olemassa neljää eri tyyppiä:

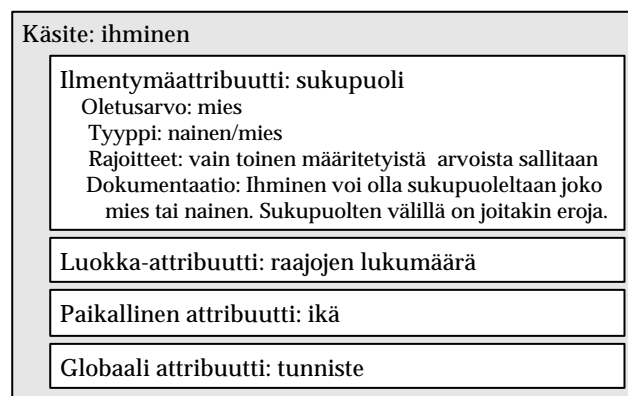
- Ilmentymäattribuuttien arvot voivat vaihdella käsitteen eri ilmentymien välillä. Esimerkiksi käsitteen ihminen ilmentymäattribuutti voisi olla sukupuoli. Tällöin ilmentymien Minna ja Mikko sukupuoli-attribuutin arvot olisivat nainen ja mies, eli toisistaan eroavat.
- Luokka-attribuutti on ilmentymäattribuutin vastakohta. Sen saamat arvot riippuvat aina käsitteestä. Tällöin kaikki käsitteen ilmentymät saavat saman luokka-attribuutin arvon. Esimerkiksi ihminen-käsitteen luokka-attribuutti voisi olla raajojen lukumäärä. Nyt sekä Minna että Mikko saavat saman arvon: neljä.
- Paikalliset attribuutit ovat käsitteen sisäisiä ominaisuuksia, jotka eivät vaikuta muihin käsitteisiin eivätkä periydy luokkahierarkiassa. Tämä

mahdollistaa samannimisten attribuuttien määrittämisen eri käsitteille.

Esimerkiksi käsitteillä ihminen ja hyönteinen voi molemmilla olla paikallinen attribuutti ikä, mutta sen yksikkönä toisella on vuosi ja toisella kuukausi. Kaikille käsitteille ei kyseistä attribuuttia tarvitse lainkaan määrittellä. Esimerkiksi käsitteelle lumi ei voi määrittää ikää.

- Globaalit attribuutit ovat yhteisiä kaikille käsitteille. Jos käsitteelle lumi määritetään globaali attribuutti koostumus, sama attribuutti määräytyy automaattisesti myös kaikille muille käsitteille.

Attribuuttien saamia arvoja voidaan määrätä tai rajoittaa eri tavoilla. Oletusarvo (engl. default value) määrää attribuutin arvon. Voidaan esimerkiksi olettaa, että tuotteen alennus on 0 prosenttia, jos sitä ei ole muuksi määritelty. Tyyppi (engl. type) määrää arvon tietotyypin esimerkiksi merkkijonoksi tai kokonaisluvuksi. Kardinaalisuusrajoitteet (engl. cardinality constraints) määrittävät arvojen minimi- ja maksimimäärät. Esimerkiksi tuotteella voi olla vain yksi hinta, mutta värejä voidaan määrittellä yhdestä viiteen. Dokumentaatio (engl. documentation) on luonnollisella kielellä esitetty attribuutin määritelmä. Seuraavassa kuviossa 1 on havainnollistettu aikaisempien esimerkkien avulla käsitteen erilaisia attribuuttityyppejä.

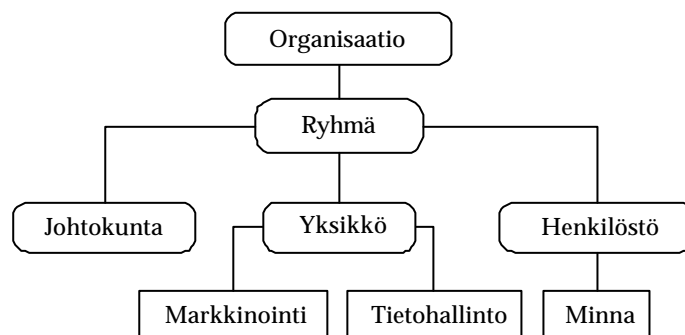


KUVIO 1. Esimerkki käsitteestä ja sen erilaisista attribuuttityypeistä

Ilmentymät voidaan jakaa kolmeen eri ryhmään. Käsitteen ilmentymä on perusilmentymätyyppi. Esimerkiksi käsitteellä ostaja voi olla ilmentymä Mikko. Fakta (engl. fact) on suhteen ilmentymä, joka on tosi. Esimerkkinä faktasta todetaan, että Mikko ostaa Sonyn stereot. Väite (engl. claim) voidaan esittää käsitteen ilmentymän faktasta. Väite ei välttämättä pidä paikkaansa. Esimerkiksi 'Mikko sanoo, että Minna on ostanut Sonyn stereot'.

Taksonomiat määrittelevät käsitteiden jaottelun ja ryhmittelyn hierarkioiksi. Hierarkiat muodostavat puumaisia rakenteita. Liikuttaessa puussa ylöspäin päädytään aina edellistä laajentavaan käsitteeseen. Vastaavasti kuljettaessa alaspäin tulee vastaan tarkentavia termejä. Taksonomioita voidaan käyttää hyväksi myös periytyvien ominaisuuksien hallinnassa.

Ontologian rakennetta voidaan havainnollistaa esimerkkiontologialla, jonka yläkäsitteenä on organisaatio. Organisaation alaluokkana on ryhmä, joka jakautuu johtokunta-, yksikkö- ja henkilöstö-alaluokkiin. Yksikkö-luokan yhtenä ilmentymänä on tietohallinto. Henkilöstö-käsitteen ilmentymänä on Minna. Kuviossa 2 on pyritty havainnollistamaan esimerkkiontologian käsitehierarkiaa.



KUVIO 2. Esimerkki käsitehierarkiasta

Ontologioissa esiintyy neljä taksonomista pääryhmää. Alaluokka (engl. subclass of) on yleisin taksonomiaryhmä. Se kertoo käsitteen B olevan käsitettä A tarkentava. B on siis A:n alaluokka. Hajoitelma (engl. disjoint decomposition) on joukko täysin erillisiä käsitteitä, jotka kaikki ovat jonkin käsitteen alaluokkia. Hajoitelman ei tarvitse sisältää kaikkia mahdollisia kyseisen käsitteen alaluokkia. Täydellinen hajoitelma (engl. exhaustive subclass decomposition) vastaa hajoitelmaa, mutta sisältää kaikki mahdolliset kyseisen käsitteen alaluokat. Esimerkiksi täydellinen hajoitelma ministeriö-luokan alaluokista sisältäisi kaikki 13 ministeriötä, mutta hajoitelma esimerkiksi vain oikeusministeriön, puolustusministeriön ja ympäristöministeriön. Ei-alaluokka (engl. not subclass of) kertoo, että käsite B ei ole käsitteen A:n alaluokka. Useissa ontologiakielissä tätä taksonomiaryhmää ei ole lainkaan, sillä alaryhmämäärittelyn puuttuessa oletetaan automaattisesti, että käsite ei ole alaluokka.

Ontologia ei ole pelkästään sanasto jostain tietystä aihealueesta, vaan ontologia määrittelee myös käsitteiden keskinäisiä *suhteita*. Esimerkiksi lauseet "on osa" ja "koostuu" kuvaavat binäärisiä suhteita. Tarkoituksena onkin, että ontologioissa tieto on hyvin loogisesti järjestettyä. Suhteista käytetään myös nimitystä roolit. Kuvion 2 esimerkki-ontologiassa voitaisiin määritellä suhde, jonka mukaan kaikki henkilöstöön kuuluvat työskentelevät yksikössä. Tällöin voidaan esimerkiksi määritellä, että Minna työskentelee tietohallintoyksikössä. *Funktiot* taas ovat suhteiden erikoistapauksia. Ne ovat toiminnallisia kaavoja, jotka yhdistävät joukon syötteitä ja niistä saatavan lopputuloksen. Esimerkiksi voidaan määritellä, että tiettyä tuotetta on oltava varastossa tietty määrä. Tällöin funktiolla voidaan selvittää, onko varastossa liian vähän tätä tuotetta. Funktio palauttaa vastauksen totta (engl. true), jos näin on.

Kaikkein yksinkertaisimmillaan ontologia kuvaa käsitteiden hierarkian, jossa käsitteet liittyvät toisiinsa yleistyssuhteiden (engl. subsumption relationships) kautta. Monimutkaisemmissa tapauksissa lisätään sopivia aksioomia. *Aksioomat* ovat malleja, joiden avulla esitetyt lauseet ovat aina totta. Niitä voidaan käyttää ontologian sisältämien tietojen tarkistamiseen, tiedon rajoittamiseen tai uuden tiedon päättelyyn. Aksioomat voidaan määrittellä 1. tai 2. kertaluvun logiikalla. Lisäksi aksioomat voivat olla itsenäisiä tai sulautettuja. Itsenäiset aksioomat ovat muista elementeistä riippumattomia, nimettyjä aksioomia. Sulautetut aksioomat on liitetty jonkin toisen elementin, kuten käsitteen tai suhteen, määritelmän sisään. Esimerkkiontologiaan voitaisiin lisätä esimerkiksi seuraava aksiooma: "Kaikkien henkilöstön jäsenten tulee olla suorittanut perehdytyskoulutus ennen vakituisen kulkukortin saamista". Tällöin työntekijä, jolla ei ole merkintää suoritetusta perehdytyskoulutuksesta, ei voi saada vakituista kulkukorttia.

2.3 Ontologioiden luokittelua

Ontologioiden rakenneosien esittämisen jälkeen nousee kysymys siitä, millaisia ontologioita voidaan esittää näiden rakenteiden avulla. Uschold ja Grüninger (1996, 98) ovat erottaneet neljä ontologiatyyppiä niiden formaaliuden perusteella. *Erittäin vapaamuotoinen* (engl. highly informal) ontologia on esitetty luonnollisella kielellä. *Puoliksi vapaamuotoinen* (engl. semi-informal) ontologia on ilmaistu rajoitetulla ja rakenteistetulla luonnollisen kielen muodolla eli käyttäen kaavoja (engl. patterns). Näiden kahden tyyppin osalta ei täyty Gruberin määritelmän (ks. kohta 2.1) formaalisuusehto, mutta joissakin muissa määritelmissä esimerkiksi pelkkää sanastoa voidaan pitää ontologiana. *Puoliformaali* (engl. semi-formal) ontologia on määritelty keinotekoisella ja formaalisti määritetyllä kielellä. *Tiukan formaali* (engl. rigorously formal)

ontologia on määritelty kielellä, joka sisältää muun muassa formaalia semantiikkaa ja teoreemoja.

Yhteenvedon voidaan sanoa, että vapaamuotoiset ontologiat sisältävät käsitteitä, luokkia, suhteita käsitteiden välillä ja ominaisuuksia, jotka kuvaavat käsitteitä. Formaali ontologiat lisäävät aksioomia ja rajoitteita vapaamuotoisiin ontologioihin. Esimerkiksi WordNet (Miller 1995) on vapaamuotoinen ontologia, koska se on lähinnä taksonomia ja esitetty luonnollisella kielellä. Cyc (Lenat & Guha 1990) taas on formaali ontologia, sillä se on määritelty formaalilla CycL-kielellä ja sisältää aksioomia. WordNet- ja Cyc-ontologioita on esitelty tarkemmin kohdassa 3.5.

Guarino (1998, 9-10) on käyttänyt eri luokitteluperustetta ja jaotellut ontologiat yleistettävyyden perusteella eri kategorioihin. *Ydinontologiat* (engl. top-level ontologies) kuvaavat erittäin yleisiä käsitteitä, kuten aika, tapahtuma ja kohde, jotka ovat riippumattomia tietyistä ongelmista tai aihealueista. *Aihealueontologioiden* (engl. domain ontologies) käsitteistö rakentuu yleisen aihealueen ympärille, kuten lainsäädäntö tai autot. *Tehtäväontologioiden* (engl. task ontologies) käsitteistö puolestaan rakentuu yleisen metodin, tehtävän tai käyttötarkoituksen ympärille, kuten lakien säätäminen tai myyminen. Aihealueontologiat ja tehtäväontologiat voidaan luoda erikoistamalla ydinontologian termejä. *Sovellusontologia* (engl. application ontology) kuvaa käsitteitä, jotka riippuvat sekä tietystä aihealueesta että tehtävästä. Se voidaankin luoda aihealue- ja tehtäväontologioiden pohjalta, esimerkkinä tällaisesta sovellusontologiasta voisi olla lainsäädäntötyö tai autojen myyminen.

Ontologioita voidaan jaotella formaaliuden ja yleistettävyyden lisäksi myös muilla tavoilla. Esimerkiksi Fensel (2000, 8) luokittelee ontologiat viiteen eri tyyppiin seuraavalla tavalla:

- Aihealueontologiat tallentavat tietylle aihealueelle tärkeän tietämyksen.
- Metatieto-ontologiat, kuten Dublin Core, ovat sanastoja, jotka sisältävät tietoja tietolähteiden sisällöstä.
- Yleisontologiat pyrkivät tallentamaan yleisluontoista tietoa maailmasta, sisältäen peruskäsitteitä, kuten aika, tila ja tapahtuma. Siten ne ovat päteviä useilla aihealueilla.
- Kuvaavat ontologiat eivät sitoudu mihinkään tiettyyn aihealueeseen. Ne sisältävät kuvaavia käsitteitä kertomatta, mitä pitäisi esittää. Esimerkiksi Gruberin (1993) Frame Ontology on kuvaava ontologia, joka määrittelee käsitteitä, kuten kehykset, attribuutit ja rajoitteet, sallien tietämyksen esittämisen oliokeskeisellä tai kehysperustaisella tavalla.
- Menetelmä- ja tehtäväontologiat sisältävät termejä, jotka liittyvät johonkin tiettyyn menetelmään tai tehtävään.

Hyvönen (2002, 14-15) on luokitellut ontologiatyyppejä sovellusalueiden perusteella: Tiedeontologioita hyödynnetään esimerkiksi biologian ja elektroniikan aloilla. Liiketoimintaontologioita käytetään muun muassa tuotteiden ja liiketoimintamallien esittämiseen. Taiteen ja käsitöiden esittämiseen luodaan kulttuurionologioita. Metatieto-ontologioilla ilmaistaan dokumenttien julkaisutietoja. Metakäsitteiden kokoelmia, kuten metakäsitteitä, käytetään ontologioiden itsensä kuvaamiseen. Dynaamisten ilmiöiden ontologioita ovat esimerkiksi tehtävien, prosessien ja palveluiden kuvaamiseen käytettävät ontologiat.

2.4 Ontologioiden käyttötapoja

Ontologioilla on useita erilaisia käyttötapoja. Kommunikointi (ks. esim. Grüninger & Lee 2002), tietämyksen organisointi ja hallinta (ks. esim. Grüninger & Lee 2002), tiedonhaku (ks. esim. Fensel 2000), koneellinen päättely (ks. esim. Grüninger & Lee 2002), yritysten välinen elektroninen kaupankäynti (ks. esim. Fensel 2000) ja verkkokauppa (ks. esim. Fensel 2000) ovat niistä tärkeimpiä. Seuraavassa on lyhyesti käyty läpi näitä käyttötapoja.

Ihmisten on kyettävä kommunikoimaan keskenään. Erilaisista tarpeista ja taustakontekstista johtuen ihmisillä voi olla laajalti eroavia näkökantoja ja olettamuksia siitä, mitkä ovat samoja asioita. Kommunikoinnin osapuolet käyttävät erilaista sanastoja, jotka sisältävät vaihtelevia, päällekkäisiä ja/tai yhteensopimattomia käsitteitä. Tästä seuraava jaetun ymmärtämyksen puute johtaa puutteelliseen viestintään ihmisten välillä. Ontologiat vähentävät käsitteellistä sekaannusta ja auttavat saavuttamaan jaetun ymmärtämyksen. Ontologioiden avulla ihmiset, joilla on erilainen tausta ja erilaiset näkökulmat, voivat kommunikoida paremmin keskenään.

Ihmisten lisäksi myös tietojärjestelmien tulee kommunikoida keskenään. Ontologioita voidaan käyttää helpottamaan ja tehostamaan tietojärjestelmien yhteentoimivuutta (engl. interoperability), sillä ne eivät välttämättä ymmärrä toisiaan oikein keskinäisessä vuorovaikutuksessa. Eri tietojärjestelmät voivat käyttää erilaisia tunnisteita kuvaamaan samaa käsitettä tai toisaalta samalla käsitteellä voidaan tarkoittaa eri asioita (Hovy 2003, 47). Lisäksi ongelmia aiheuttavat muun muassa erilaiset tietorakenteet ja tunnistekäytännöt. Järjestelmien toimiessa itsenäisesti näistä ei ole haittaa, mutta ongelmia syntyy kun yritetään yhdistää eri järjestelmät tai jakaa informaatiota niiden välillä. Jos

ohjelma haluaa verrata tai yhdistää informaatiota kahden tietojärjestelmän välillä, sen tarvitsee tietää, että kahta termiä käytetään tarkoittamaan samaa asiaa tai saman käsitteen eri merkitykset. Ontologian kautta ohjelma saa tämän tiedon.

Tietämyksen organisoinnin ja hallinnan tekevät ongelmalliseksi organisaatioiden tietovarastot, jotka sisältävät valtavia määriä monimuotoista, hajautettua ja useissa eri muodoissa tallennettua informaatiota. Ontologioita hyödyntävät dokumenttikannat vähentävät ongelmia. Dokumenttikantaa voivat käyttää yrityksen eri yksiköt, kuten myyntiosasto ja tekninen osasto. Eri käyttäjillä voi kuitenkin olla aiheeseen liittyen oma käsitteistönsä, esimerkiksi myyntihenkilö ei välttämättä ymmärrä jonkin ominaisuuden teknistä nimeä. Tällöin joka yksikön käsitteistölle voidaan luoda oma ontologia, jotka ovat kytköksissä siten, että sanojen käännökset kunkin yksikön "kielelle" tapahtuvat automaattisesti.

Ontologioilla voidaan parantaa hakujen tarkkuutta, sillä hakujärjestelmä voi etsiä vain niitä dokumentteja, jotka viittaavat täsmälliseen käsitteeseen monimerkityksisten avainsanojen sijasta. Esimerkkinä ontologiaa hyödyntävästä portaalista on OntoWeb <<http://www.ontoweb.org/>>, joka palvelee ontologioista kiinnostunutta akateemista ja teollista yhteisöä. Lisäksi päättelysääntöjä sisältävät ontologiat, jotka mahdollistavat informaation yhdistämisen monesta eri lähteestä, tarjoavat käyttäjälle mahdollisuuden saada sellaisia hakutuloksia, jotka eivät perinteisillä menetelmillä ole mahdollisia. Automaattista koneellista päättelyä varten ontologiakielen on pystyttävä erittelemään käsitteet ja niiden väliset suhteet tarpeeksi formaalisti (Hendler 2001). Esimerkiksi katuosoitetietoon voidaan soveltaa sääntöä, että jos siihen

liittyy jokin postinumero ja postinumeroon kaupunki, niin silloin katuosoite sijaitsee postinumeron osoittamassa kaupungissa.

Elektroninen kaupankäynti ja verkkokauppa hyötyvät ontologioista, koska yhteisesti jaetun aihealueontologian, kuten IT- ja elektroteknikkateollisuuden RosettaNetin (RosettaNet.org 2003), avulla sähköisen kaupan agentit pystyvät jakamaan tietoa ja automatisoimaan toimintaansa. Ontologialla voidaan kuvata moninaiset tuotteet ja agentit voivat käyttää sitä navigoimaan ja etsimään automaattisesti haluttua informaatiota.

2.5 Esimerkkejä ontologioista

Kohdassa 2.3 selvitettiin erilaisia ontologioiden luokittelutapoja. Tässä alaluvussa esitellään muutama ontologia esimerkkinä erityyppisistä ontologioista. WordNet ja Cyc ovat molemmat yleisontologioita, mutta WordNet on vapaamuotoinen ja Cyc formaali. TOVE taas on esimerkki tehtävä- ja (KA)² aihealueontologiasta. Lisäksi esitellään harvoista lakiontologioista tunnetuimpia, jotka ovat McCarty'n LLD (Language for Legal Discours), Stamperin NORMA, Valentin toiminnallinen lakiontologia sekä van Kralingenin ja Visserin kehysperustainen lakiontologia. Seuraavassa on esitelty näitä ontologioita lyhyesti.

WordNet (Miller 1995) on kehitetty Princetonin yliopiston kognitiivisen tieteen laitoksella. WordNet <<http://www.cogsci.princeton.edu/~wn/>> on laaja ja sovellusalueeriippumaton, luonnollisella kielellä toteutettu ontologia englannin kielen sanoista ja niiden suhteista. WordNet kattaa reilusti yli 100 000 sanaa, jotka on linkitetty toisiinsa sadoin tuhansin eri tavoin. WordNet perustuu kielen leksikaalisiin rakenteisiin ja sen taustalla ovat psykokieli-tieteelliset

teoriat ihmisen kielellisestä muistista. Hakemisto sisältää substantiiveja (esim. book, car), verbejä (esim. read, drive), adjektiiveja (esim. green, big) ja adverbejä (esim. particularly, quickly), joita on organisoitu yhteen synonyymikäsitteiksi siten, että niistä muodostuu sanastokäsite. Käsitteiden välille on määritelty suhteita. Jokaisesta sanasta on saatavilla muun muassa sen määritelmä, synonyymit, vastakohtat, yläkäsitteet, osa-kokonaisuussuhteet ja paljon muuta.

WordNetille on tunnusomaista epäformaali määrittely. Käsitteiden ja kielellisten suhteiden merkitykset on annettu luonnollisella kielellä, mikä rajoittaa automaattisen päättelyn tekoa (Fensel 2000, 9-10). WordNetin laadukkaasta tasosta kertovat ympäri maailman valmisteilla olevat sanastot, jotka käyttävät sitä mallinaan. Esimerkiksi arvostetulla Oxford English -sanakirjalla on oma versionsa WordNetistä seuraavassa www-osoitteessa: <<http://www.oed.com/learn/WordNet/index.cfm>>.

Cyc (Lenat & Guha 1990) on sovellusalueeriippumaton ontologia, jonka tavoitteena on formalisoida ja saattaa tietokoneiden käsiteltäväksi inhimillistä arkitietämystä. Se on toteutettu CycL-kielellä ja sisältää formaaleja aksiomia. Cyc-järjestelmän tietämuskanta koostuu yli 100 000 aksiomaattisesti määritellystä käsitteestä. Cyc ryhmittelee käsitteet mikroteorioiksi kokonaisuontologian rakentamiseksi. Mikroteoriat voivat kuvata vaikkapa joukkoa (kuten ihmiset), yksittäistä kohdetta (kuten henkilö, esimerkiksi Matti Jokunen), sanan merkitystä (esimerkiksi henkilö-sana) tai roolimäärettä (kuten omistaa). Mikroteorioiden avulla pystytään esittämään tietämyksen kontekstisidonnaisuus. Cyc tietää esimerkiksi, että puut ovat yleensä ulkona, ja että nestettä sisältävää lasia on kannettava oikeinpäin. Järjestelmän kehitys jatkuu edelleen. Web-käyttöliittymän <<http://www.opencyc.org/>> avulla voi selata Cyc-ontologian noin 3000 ylempään tason käsitettä.

TOVE (Grüniger & Fox 1994) on esimerkki tehtäväontologiasta. Ontologia tukee yrityksen integroimista tarjoamalla jaettavan esityksen tietämyksestä. *TOVE*-projektin (TOronto Virtual Enterprise) päämääränä on luoda yleispätevä ja uudelleenkäytettävä tietomalli, joka tarjoaa yrityksille jaetun käsitteistön, jota jokainen toimija voi yhteisesti ymmärtää ja käyttää. Tietomallin on tarkoitus määritellä jokaisen käsitteen merkitys mahdollisimman tarkasti ja lisätä aksioomia, joiden avulla ontologia pystyy päättämään vastauksia kysymyksiin. Lisäksi tietomalli määrittelee symboliikan käsitteiden kuvaamiseen graafisesti.

Knowledge Annotation Initiative of Knowledge Acquisition Community (KA)² (Benjamins & Fensel 1998) oli tapaustutkimus kehittämisprosessista, jossa tavoitteena oli kehittää ontologia heterogeeniselle ja maailmanlaajuiselle tutkimusyhteisölle. Projektissa luotiin tietämyksen hankinta -yhteisön (engl. knowledge acquisition) ontologia. Koska ontologian tulisi kiteyttää yhteisymmärrykseen perustuvaa tietämystä, useat tutkijat tekivät yhteistyötä eri sijaintipaikoissa ontologian luomiseksi. Tällä tavalla varmistettiin, että suurin osa yhteisön tutkijoista hyväksyy ontologian. (KA)² ontologia koostuu seitsemästä toisiinsa liittyvästä ontologiasta: organisaatio-, projekti-, henkilö-, tutkimusaihe-, julkaisu-, tapahtuma- ja tutkimustulosontologioista. Näistä varsinkin tutkimusaiheontologia on erityinen aihealueelle.

LLD-kieli eli Language for Legal Discourse (McCarty 1989) on oikeastaan kuvaava kieli eikä ontologia, mutta se selkeästi ilmaisee oikeudellisen aihealueen yleisen käsitteistön. *LLD*:n peruskomponentteja ovat itsenäiset kaavat (engl. formulae) ja säännöt (engl. rules). Kaavat ovat predikaattisuhteita, joilla ilmaistaan tosiasiallisia väitteitä. Säännöt muodostetaan yhdistämällä itsenäisiä kaavoja loogisilla sidoksilla. Yhdessä kaavoilla ja säännöillä voidaan

ilmaista 1. kertaluvun logiikan lausekkeita. Ajallisia väitteitä ilmaistaan tilojen (engl. states) avulla. Tapahtumat (engl. events) taas ilmaisevat tilojen muutoksia. Toiminta (engl. action) on suhde tekijän (engl. actor) ja tapahtuman välillä.

NORMA-formalismi (Stamper 1991) perustuu kahteen olettamukseen: (1) tietämystä ei ole olemassa ilman tietäjää ja (2) tietäjän tietämys riippuu hänen käyttäytymisestään *NORMA*a käyttäen maailman kohteet kuvataan käyttäytymisen perusteella. Ontologiset pääkäsitteet ovat seuraavat: agentit (engl. agents), käyttäytymismuuttajat (engl. behavioural invariants) ja toteutumaiset (engl. realisations). Agentti hankkii tietämystä ja muuntaa maailmaa toimiansa kautta. Agentti-käsite voidaan laajentaa koskemaan muun muassa ryhmiä, yrityksiä tai valtioita. Käyttäytymismuuttaja on kuvaus tilanteesta, jonka piirteet pysyvät muuttumattomina. Tilanne ilmaisee jotakin tietämystä maailmasta, kuten kohteita (esim. kuppi, kitara) tai asioiden tilaa (esim. kävellä, maksaa). Tilanteen toteutuma määritellään agentin ja käyttäytymismuuttujan yhdistelmänä. Esimerkki toteutumasta voisi olla esimerkiksi "Mikko kävelee".

Toiminnallinen lakiontologia eli *Functional Ontology of Law* (Valenten 1995) pitää oikeudellista järjestelmää välineenä, jonka avulla voi muuttaa yhteiskuntaa tiettyyn, sosiaalisten tavoitteiden määrittämään suuntaan. Sen päätoimintona on sosiaaliseen käyttäytymiseen reagointi. Tämä päätoiminto voidaan hajottaa kuuteen toimintoon, joista jokainen vastaa ontologian lakitietämyksen kategorialla. Näitä kategorioita ovat normatiivinen tietämys (engl. normative knowledge, maailmantietämys (engl. world knowledge), vastuutietämys (engl. responsibility knowledge), reagoiva tietämys (engl. reactive knowledge), metaoikeudellinen tietämys (engl. meta-legal knowledge) ja luova tietämys (engl. creative knowledge). Normatiivinen tietämys määrittää

sosiaalisen käyttäytymisen standardit. Maailmantietämys kuvaa säännösteltävää maailmaa. Vastuutietämys joko laajentaa tai rajoittaa agentin vastuuta käyttäytymisestään. Reagoiva tietämys määrittelee, miten pitää reagoida, jos agentti rikkoo normeja. Metaoikeudellinen tietämys on tietoa oikeudellisesta tietämyksestä, tai tietoa, joka viittaa muuhun tietämykseen. Luova tietämys sallii uusien oikeudellisten kohteiden luomisen.

FBO-ontologia eli *Frame-based Ontology* on yhdistelmä van Kralingenin (1995) käsitteellisestä ja Visserin (1995) formaalista kehysperustaisesta lakiontologiasta. FBO tekee eron yleisen oikeudellisen ontologian (engl. generic legal ontology) ja säädöksiin erikoistuneen ontologian (engl. statute-specific ontology) välillä. Yleinen oikeudellinen ontologia on ontologian uudelleenkäytettävä osio. Se jakaa oikeudellisen tietämyksen kolmeen erilliseen kohteeseen: normeihin (engl. norms), tekoihin (engl. acts) ja käsitteiden kuvauksiin (engl. concept descriptions). Ontologia määrittelee jokaiselle kohteelle kehysrakenteen, joka listaa kaikki kohteelle olennaiset attribuutit. Normit ovat yleisiä sääntöjä ja käyttäytymisperiaatteita, joita tulee noudattaa. Teot edustavat maailman tiloihin vaikuttavia muutoksia. Käsitteiden kuvaukset määrittävät aihealueen käsitteiden merkitykset. Säädöksiin erikoistunut ontologia muodostuu predikaattisuhteista, joita käytetään täydentämään normien, tekojen ja käsitteiden kuvausten terminologiaa. Tätä ontologiaa ei voi uudelleenkäyttää muilla oikeudellisilla aihealueilla.

Tämän luvun tarkoituksena oli antaa peruskäsitys ontologioista. Seuraavassa luvussa syvennyttään tarkastelemaan ontologioiden suunnittelua, muun muassa siihen liittyviä periaatteita ja menetelmiä.

3 ONTOLOGIOIDEN SUUNNITTELU

Tässä luvussa perehdytään aluksi ontologioiden suunnittelun periaatteisiin ja lähestymistapoihin. Niiden jälkeen selvitetään lyhyesti, millaisilla menetelmillä, esityskielillä ja työkaluilla ontologioita voidaan kehittää.

3.1 Suunnittelun periaatteita ja lähestymistapoja

Ontologioiden kehittämisen päämääränä on formalisoida ja mallintaa aihealueen olennaiset semanttiset piirteet. Ontologian kehittämisen aikana aihealueen tai ilmiön pääkäsitteet tuodaan selkeästi esille, määritellään ja kuvataan. Selkeät määrittelyt vähentävät käyttäjien välisistä erilaisista käsityksistä ja sanastoista johtuvia epäselvyyksiä.

Seuraavassa on lyhyesti listattu joitakin suunnittelukriteerejä ja -periaatteita, jotka ovat osoittautuneet hyödyllisiksi ontologioiden kehittämisessä:

- Täsmällisyys ja objektiivisyys (engl. clarity and objectivity). Ontologian tulee sisältää määrittelyjen termien merkitykset. Määrittelyjen tulee olla objektiivisiä, täydellisiä ja formaaleja. Kaikki merkitykset tulee dokumentoida luonnollisella kielellä. (Gruber 1995, 909)
- Johdonmukaisuus (engl. coherence). Ontologian tulee sallia päättelyt, jotka ovat yhdenmukaisia määritelmien kanssa. Aksiomien tulee olla loogisesti yhtenäisiä. (Gruber 1995, 909)
- Laajennettavuus (engl. extendibility). Ontologia tulee toteuttaa siten, että sitä voidaan laajentaa eli määritellä uusia termejä ilman, että olemassaolevia määritelmiä tarvitsee muuttaa. (Gruber 1995, 909-910)

- Minimaalinen riippuvuus koodaustavasta (engl. minimal encoding bias). Käsitteistö pitää määritellä tietämystasolla ilman riippuvuutta tietystä symbolisen tason koodaustavasta. (Gruber 1995, 910)
- Minimaalinen ontologinen sitoutuminen (engl. minimal ontological commitment). Ontologian pitää esittää mahdollisimman vähän väitteitä mallinnettavasta maailmasta, jotta ontologiaan sitoutuneet osapuolet voivat vapaasti erikoistaa ontologiaa tarvittaessa. (Gruber 1995, 910)
- Modulaarisuus (engl. modularity). Uudelleenkäytön mahdollistamiseksi ontologiat tulisi muodostaa pienistä moduuleista, joilla on korkea sisäinen yhtenäisyys ja mahdollisimman vähän kytköksiä moduulien välillä. (Bernaras, Laresgoiti & Corera 1996)
- Hierarkioiden monipuolisuus (engl. diversification of hierarchies). Hierarkioiden tulee olla monipuolisia lisätäkseen päättelymekanismien tehokkuutta. (Arpírez, Gómez-Pérez, Lozano & Pinto 1998)
- Sisaruskäsitteiden minimaalinen semanttinen etäisyys (engl. minimal semantic distance between sibling concepts). Samankaltaiset käsitteet tulee yhdistää, ja muiden luokkien tulee erota toisistaan. (Arpírez ym. 1998)

Jokelan (2001, 40) mukaan on tärkeää huomata ontologian kehittämisen iteratiivinen luonne ja tarve peilata ympäristössä tapahtuvia muutoksia ontologiaan. Hänen kokemuksensa mukaan ontologian kehittäminen on vaikeaa ja vaatii jatkuvia muutoksia ennen kuin ontologiasta tulee hyödyllinen. Vaikka tyydyttävä ontologia tietyllä aihealueella olisikin saatu luotua, aihealue vääjäämättömästi tulee muuttumaan ajan kuluessa. Nämä muutokset, kuten uusien käsitteiden ilmaantuminen tai vanhojen käsitteiden semantiikan muuttuminen, tulee huomioida myös ontologiassa. Jos ontologiaa ei mukauteta ympäristön muutoksiin, se vähitellen rapistuu eikä enää pysty kunnolla kiteyttämään aihealueen semantiikkaa.

Holsapple ja Joshi (2002, 42) muistuttavat, että ontologian luomisessa ja soveltamisessa on erotettava toisistaan käsitteet ja niiden väliset suhteet määrittelevä ontologia sekä kyseessä olevia käsitteitä ja suhteita koskevat kokemusperäiset tosiasiat. Tosiasiat eivät ole osa ontologiaa, vaikkakin ne jäsenyvät sen mukaan. Ontologia voi esimerkiksi sisältää tiedon siitä, että maksamatonta laskua seuraa karhukirje, mutta ei tietoa siitä, milloin se lähetetään. Todellisuudessa karhukirje lähetetään yleensä kuukauden sisällä, mutta tästä voi olla poikkeaviakin tapoja.

Ontologioiden kehittämisessä voidaan käyttää erilaisia lähestymistapoja. Geerts ja McCarthy (2000, 6) jaottelevat nämä lähestymistavat kolmeen ryhmään: pragmaattiseen (engl. pragmatic), teoreettiseen (engl. theoretic) ja empiiriseen (engl. empirical) lähestymistapaan. Pragmaattisessa lähestymistavassa määritellään ontologisia rakenteita ratkaisemalla ongelmia. Esimerkiksi yritysontologia TOVE on luotu tarkastelemalla erityisiä aihealueeseen liittyviä ongelmia (Grüninger & Fox 1994). Teoreettinen lähestymistapa johtaa käsitteistöä olemassa olevista teorioista. Tätä lähestymistapaa on hyödynnetty muun muassa REA-viitekehysessä, jossa on käytetty laskentatoimen ja talouden teorioita yritysontologian luomiseen. (Geerts & McCarthy 2000). Empiirisiä menetelmiä käytetään esimerkiksi testattaessa ihmisen kognitioon sisältyviä ontologisia teorioita (Smith & Mark 1999). Näitä samoja menetelmiä voitaisiin mahdollisesti hyödyntää rakennettaessa ontologioita tietojärjestelmiä varten.

Holsapplen ja Joshin (2002, 43-45) erilaisen näkökulman mukaan ontologioiden kehittämiseen on olemassa viisi lähestymistapaa: inspiroiva (engl. inspirational), induktiivinen (engl. inductive), deduktiivinen (engl. deductive), yhdistävä (engl. synthetic) ja yhteistoiminnallinen (engl. collaborative)

lähestymistapa. Niiden lähtökohtia on lyhyesti esitelty taulukossa 1. Myös eri lähestymistapojen yhdistelmät ovat mahdollisia.

TAULUKKO 1. Lähestymistapoja ontologioiden kehittämiseen (mukailten Holsapple & Joshi 2002)

Lähestymistapa	Lähtökohta
Inspiroiva	Ontologia kehitetään ratkaisemaan aihealueessa havaittu ongelma. Yksittäinen suunnittelija suunnittelee ontologian oman näkemyksenä mukaisesti.
Induktiivinen	Ontologia suunnitellaan havainnoimalla ja analysoimalla aihealueen tiettyä yksittäistä tapausta. Tuloksena saatavia määrittelyjä sovelletaan muihin saman aihealueen tapauksiin.
Deduktiivinen	Yleisiä periaatteita sovelletaan soveltuvin osin tiettyyn tapaukseen tarkoitetun ontologian suunnitteluun. Tämä vaatii muun muassa käsitteiden suodattamista ja räätälöintiä aihealueen mukaisiksi.
Yhdistävä	Suunnittelija tunnistaa perusjoukon ontologioita, joista kukin sisältää osittaisen aihealueen määrittelyn. Kyseiset piirteet kustakin ontologiasta yhdistetään ja tuotetaan yhtenäinen ontologia.
Yhteistoiminnallinen	Suunnittelu tapahtuu useiden osapuolien yhteistyönä. Ontologia on useista eri näkökannoista aikaansaatu yhteisymmärrys. Siksi suunnittelun aikana on käytettävä yhteistyön organisointiin tarkoitettua menetelmää. Suunnittelun lähtökohtana voidaan käyttää hyväksi aikaisemmin luotua ontologiaa.

Eri lähestymistavoilla on sekä hyviä että huonoja puolia. Inspiroiva lähestymistapa on ehkä liian tapauskohtainen, epäluotettava ja epäkäytännöllinen, mutta toisaalta sillä saa tuotettua ainutkertaisia ja innovatiivisia ontologioita. Induktiivisen lähestymistavan kautta luotu ontologia sopii mahdollisesti hyvin yksittäiseen tapaukseen, mutta ei välttämättä ole helposti yleistettävissä. Deduktiivinen lähestymistapa edellyttää suunnittelijan näkökulmasta riippuvaa yleisen määrittelyn olemassaoloa ja valintaa. Myös yhdistävä lähestymistapa on hyvin suunnittelijakeskeinen, koska lopputulokseen vaikuttaa vahvasti suunnittelijan taidot yhdistää eri ontologioita. Yhteistoiminnallinen lähestymistapa on siitä erilainen, että se ottaa

huomioon erilaisia näkökulmia ja toteuttaa iteratiivista työskentelytapaa. Kyseisen lähestymistavan onnistunut soveltaminen on kiinni osapuolien luonteesta, osallistumisen asteesta ja vastaavan suunnittelijan kyvystä hallita yhteistyöprosessia. (Holsapple & Joshi 2002, 45)

Hendler (2001, 31-32) suosittaa inkrementaalista ontologioiden kehittämistä, joka osittain vastaa Holsapplen ja Joshin (2002, 44) esittämää yhdistävää lähestymistapaa. Inkrementaalisisessa kehittämisessä ontologia kootaan valitsemalla sopivat osat valmiista ontologioista ja lisäämällä vielä siihen tarvittavat lisäosat. Myös Hovy (1998, 536) kirjoittaa samantyyilisestä tavasta luoda ontologia, mutta hän aloittaisi laajalla ydinontologialla ja sitten lisäisi siihen systemaattisesti muita ontologioita ja käsitteitä. Edellä mainitut lähestymistavat edellyttävät ontologiakirjastojen luomista ja käyttämistä. Muutama tällainen ontologioiden säilytyspaikka on esitelty kohdassa 3.4.

Ontologioiden onnistunut hyödyntäminen edellyttää yhteisymmärrystä eri osapuolten kesken. Sekä Holsapple ja Joshi (2002, 43) että Grüninger ja Lee (2002, 40) painottavatkin *ontologisen sitoutumisen* (engl. ontological commitment) tärkeyttä. Sillä tarkoitetaan useiden osapuolien (ihmisten, organisaatioiden ja tietojärjestelmien) sopimusta omaksua tietty ontologia kun kommunikoidaan mielenkiinnon kohteena olevasta alueesta. Kun ontologinen sitoutuminen puuttuu, on vaikeaa kommunikoida selkeästi aihealueesta ja hyötyä muiden tietämyksestä. Siksi Holsapplen ja Joshin (2002, 43) mielestä toimenpiteet sitoutumisen varmistamiseksi ovat elintärkeitä ontologioiden suunnittelussa. Sitoutuminen toteutuu, kun mahdolliset käyttäjät kokevat ontologian sisältämät määrittelyt riittävän valmiiksi, oikeiksi, selviksi ja ytimekkäiksi.

3.2 Ontologioiden kehittämismenetelmiä

Ontologioiden kehittämiseen liittyy monia käytännön vaikeuksia. Terminologian standardointi on hankalaa, aikaa vievää ja kallista johtuen eri tahojen erilaisista tarpeista ja mieltymyksistä. Ontologioilla on myös taipumusta muodostua laajoiksi, jolloin niiden hallinta voi olla vaikeaa, kun ajan kuluessa tapahtuu muutoksia (Hyvönen 2002, 16). Näiden vaikeuksien voittamiseksi on luotu erilaisia kehittämismenetelmiä, jotka perustuvat ontologioiden rakentamisesta saatuihin kokemuksiin.

Wangin ja Xun (2000, 673) menetelmä sisältää neljä vaihetta, jotka ovat (1) aihealueen analysointi ja tietämysketjujen (engl. knowledge chains) rakentaminen sekä (2) ontologian rakenteistaminen, (3) formalisointi ja (4) arviointi. Ensimmäisessä vaiheessa hankitaan ja luokitellaan tietämystä aihealueesta. Toisessa vaiheessa luodaan tietämyssanakirja sekä tunnistetaan peruskäsitteet, attribuutit ja arvot. Kolmannessa vaiheessa luodaan ontologia ontologiakielen avulla. Viimeisessä vaiheessa ontologiaa arvioidaan, jotta sitä voitaisiin edelleen parantaa.

Kim (2000) on kehittänyt menetelmän olemassa olevan datan päälle rakennettavan ontologian luomiseen. Hänen menetelmässään tehdään ensin skenaarioita, joiden avulla kuvataan ontologiaa kehittävän organisaation tehtävät ja niihin liittyvät ongelmat. Sen jälkeen todetaan ontologialle asetettavat vaatimukset ja poimitaan skenaarioista avainsanoja ja -lauseita. Seuraavassa vaiheessa analysoidaan pääkäsitteitä ja otetaan huomioon ontologiaan vaikuttava konteksti. Kaikkien näiden vaiheiden tuloksia käytetään ontologian määrittelemiseen, jonka jälkeen niitä käytetään myös luodun ontologian arvioimiseen.

Holsapple ja Joshi (2002, 45-46) ovat määritelleet menetelmän, jonka avulla voidaan kehittää ontologia yhteistyössä monen eri osapuolen välillä. Heidän menetelmänsä jakautuu neljään eri päävaiheeseen: valmistelu (engl. preparation), ankkurointi (engl. anchoring), iteratiivinen parantaminen (engl. iterative improvement) ja soveltaminen (engl. application). Valmisteluvaiheessa määritellään suunnittelukriteerit sekä määritellään ontologian rajaehdot ja arviointiperiaatteet. Ankkurointivaiheessa määritellään alustava ontologia, joka toimii yhteistoiminnan alkusysäyksenä. Iteratiivisen parantamisen vaiheessa muodostetaan erilaisia keskusteluryhmiä, selvitetään osapuolten kommentit ja kritiikki ontologiaa kohtaan sekä muokataan ontologiaa saadun palautteen perusteella. Näitä tehtäviä toistetaan kunnes saavutetaan yhteisymmärrys. Soveltamisvaiheessa demonstroidaan ontologian käyttämistä.

Cyc-ontologian luomiseen käytetty menetelmä (Lenat & Guha 1990) sisältää kolme vaihetta. Ensimmäinen vaihe koostuu manuaalisesta tietämyksen osien kokoamisesta. Toinen ja kolmas vaihe muodostuvat uuden yleistajuisen tietämyksen hankkimisesta käyttäen luonnollista kieltä tai koneellisia työkaluja. Vaiheet eroavat siten, että toinen vaihe tehdään pääasiassa ihmisvoimin, mutta kolmas vaihe täysin koneellisesti.

Uscholdin ja Kingin menetelmä (Uschold & King 1995) käsittää neljä toimintoa: ontologian tarkoituksen tunnistaminen, ontologian rakentaminen, arviointi ja dokumentointi. Rakentamisvaiheeseen sisältyvät tietämyksen kokoaminen, koodaaminen ja muiden ontologioiden liittäminen kehitettävään ontologiaan. Pääkäsitteiden tunnistamiseksi on olemassa kolme strategiaa: Ylhäältä alaspäin -strategiassa (engl. top-down approach) abstraktit käsitteet määritellään ensin ja erikoistetaan tarkemmiksi. Alhaalta ylöspäin -strategiassa (engl. bottom-up approach) tarkat käsitteet määritellään ensin ja yleistetään sen jälkeen

abstraktimmiksi. Keskeltä molempiin suuntiin lähtevässä strategiassa (engl. middle-out approach) tunnistetaan ensin tärkeimmät käsitteet, joita sitten yleistetään ja erikoistetaan hierarkian molempiin suuntiin.

Grüningerin ja Foxin menetelmä (Grüninger & Fox 1995) on saanut innoituksensa tietämysjärjestelmien suunnittelusta, jossa käytetään 1. kertaluvun logiikkaa. Heidän menetelmässään tunnistetaan ensin pääskenaariot eli mahdolliset sovellukset, joissa ontologiaa tullaan käyttämään. Sen jälkeen joukkoa luonnollisen kielen kysymyksiä, joita kutsutaan pätevyyskysymyksiksi (engl. competency questions), käytetään määrittelemään ontologian laajuus. Näitä kysymyksiä ja niiden vastauksia käytetään sekä peruskäsitteiden ja niiden attribuuttien että suhteiden ja aksioomien määrittelemiseen. Tässä erittäin formaalissa menetelmässä kyseiset ontologian rakenneosat ilmaistaan 1. kertaluvun logiikalla.

KACTUS-projektissa esitellyssä menetelmässä (Bernaras ym. 1996) ontologian rakentaminen perustuu sovelluksen tietämyskantaan käyttäen alhaalta ylöspäin -strategiaa käsitteiden määrittelemiseen. Mitä useampia sovelluksia rakennetaan, sitä yleisempiä ontologioista tulee, ja sitä kauemmaksi ontologia siirtyy tietämyskannasta. Toisin sanoen ontologian kehittäminen kannattaa heidän mukaansa aloittaa tietämyskannan rakentamisella tiettyyn sovellukseen. Myöhemmin, kun tarvitaan uutta tietämyskanta samankaltaisella aihealueella, tietämyskanta yleistetään ontologiaksi ja sovitetaan molempiin sovelluksiin.

Sensus-ontologiaan perustuva menetelmä (Swartout, Patil, Knight & Russ 1996) hyödyntää ylhäältä alaspäin -lähestymistapaa aihealuekohtaisten ontologioiden johtamiseksi laajoista ontologioista. Menetelmässä tunnistetaan ydinkäsitteet, jotka ovat olennaisia tietylle aihealueelle. Nämä käsitteet linkitetään

manuaalisesti laaja-alaiseen ontologiaan, joka tässä tapauksessa on yli 70 000 käsitettä sisältävä Sensus. Sen jälkeen kaikki käsitteet, jotka ovat Sensus-ontologian taksonomiassa juuren ja latvan välisellä reitillä, sisällytetään kehitettävään ontologiaan. Jos olennainen käsite ei ole tullut esille, se lisätään manuaalisesti ja edellinen vaihe suoritetaan uudelleen kunnes yhtään käsitettä ei puutu. Tämä lähestymistapa edistää tietämyksen jaettavuutta, sillä samaa pohjaontologiaa käytetään kehitettäessä ontologioita erityisille aihealueille.

Madridin teknisen yliopiston tekoälylaboratorio on kehittänyt *Methontology*-menetelmän (Fernández, Cómez-Pérez & Juristo 1997), joka mahdollistaa ontologioiden luomisen tietämystasolla. Menetelmän avulla voidaan kehittää ontologioita joko tyhjästä tai käyttäen hyväksi muita ontologioita. *Methontology*-menetelmän kehittämisprosessi jakautuu kolmeen osaluokeseen, jotka ovat hallinnointi, kehittäminen ja tukitoiminnot. Hallinnointi koostuu suunnittelusta, kontrolloinnista ja laadunvalvonnasta. Varsinainen ontologian kehittämisvaihe koostuu määrittelystä, käsitteellistämisestä, formalisoinnista, toteutuksesta ja ylläpidosta. Tukitoimintoja, eli aineiston keruuta, arviointia, yhdistämistä ja dokumentointia, tehdään koko prosessin ajan. Lisäksi menetelmä määrittelee tekniikat eri tehtävien toteuttamiseksi sekä vaiheiden lopputuotteet ja niiden arviointikriteerit.

On-To-Knowledge-menetelmä (Staab, Schnurr, Studer & Sure 2001) perustuu käyttäjäskenaarioiden analysointiin. Menetelmä muodostuu seuraavista vaiheista: aloitusvaihe, jalostusvaihe, arviointivaihe ja ontologian ylläpito. Aloitusvaiheessa määritellään ontologian vaatimukset ja pätevyyskysymykset, tutkitaan mahdollisia uudelleenkäytettäviä ontologioita sekä luodaan ontologian ensimmäinen versio. Jalostusvaiheessa luodaan kypsä ja

sovelluskohtainen ontologia. Arviointivaiheessa tarkistetaan vaatimukset ja pätevyyskysymykset sekä testataan ontologiaa sovellusympäristössä.

3.3 Ontologioiden esityskieliä

Ontologioita voidaan määritellä erilaisilla ontologiakielillä, joiden terminologia ja semantiikka voivat poiketa toisistaan huomattavasti. Käytettävän esityskielen ja formalismin valinnalla on suuri merkitys ontologian rakentamisessa. Kieli määrää, millaisia asioita ontologiassa kyetään esittämään ja millaisissa sovelluksissa sitä pystytään hyödyntämään.

Perinteisillä ontologiakielillä ontologia voidaan kuvata formaalisti logiikalla tai jollakin verkkoformalismilla, kuten kehyksillä. Internetin suosion myötä on luotu myös ontologiakieliä, joita kutsutaan yleensä web-perustaisiksi ontologiakieliksi. Kyseiset kielet ovat suppeampia, mutta helppokäyttöisempiä kuin perinteiset ontologiakielet. Web-perustaiset ontologiakielet ovat yhdistäneet sekä perinteisten ontologiakielten että web-kielten, eli XML:n (Bray, Paoli, Sperberg-McQueen & Maler 2000) ja RDF:n (Lassila & Swick 1999), piirteitä. XML- ja RDF-kieliä voidaan käyttää myös itsenäisesti joidenkin ontologisten piirteiden kuvaamiseen.

Seuraavassa esitellään lyhyesti tunnetuimpia ontologiakieliä jaoteltuna perinteisiin ja web-perustaisiin esityskieliin. Myös web-kieliä on kuvattu lyhyesti.

3.3.1 Perinteiset esityskielet

Perinteiset ontologioiden esityskielet on kehitetty ennen www:n yleistymistä. Ne voidaan jakaa seuraavasti (Corcho, Fernández-López & Cómez-Pérez 2003, 54; Fensel 2000, 62):

- Laajennettuun 1. kertaluvun predikaattilogiikkaan (engl. first-order predicate logic) perustuvat kielet, joita ovat muun muassa CycL ja KIF.
- Kehyksiin (engl. frames) perustuvat kielet, joita ovat muun muassa Ontolingua, Frame Logic ja OCML.
- Kuvauslogiikkaan (engl. description logics) perustuvat kielet, kuten Loom, jotka kuvaavat tietämystä käsitteillä ja roolirajoituksilla, joita käytetään luokittelujen automaattiseen johtamiseen.

CycL (Lenat & Guha 1990) on formaali kieli, jonka syntaksi on johdettu 1. kertaluvun predikaattilogiikasta. Tosin CycL laajentaa 1. kertaluvun logiikkaa käyttämällä 2. kertaluvun käsitteitä. CycL-kielen sanasto koostuu muun muassa seuraavista termeistä: semanttiset vakiot, ei-atomiset termit, muuttujat, numerot ja merkkijonot. Termit yhdistetään CycL-lausekkeilla, muodostaen lopulta suljettuja CycL-virkkeitä, joissa ei ole vapaita muuttujia. Yhdessä CycL-virkkeet muodostavat tietämuskannan.

KIF eli Knowledge Interchange Format (Genesereth & Fikes 1992) on 1. kertaluvun logiikkaan perustuva kieli, joka käyttää etumerkintäsyntaksia (engl. prefix notation). Se määrittelee käsitteet, attribuutit sekä suhteet ja joitakin suhdepäätelyn ominaisuuksia. KIF kehitettiin tietämyksen vaihtamiseen eri ohjelmoijien eri aikoina ja eri kielillä toteuttamien tietokonejärjestelmien välillä. Sitä voidaan käyttää myös ontologioiden esittämiseen ja vaihtamiseen.

Ontolingua (Farquhar, Fikes & Rice 1997) pohjautuu KIF-kieleen ja Frame Ontology -tekniikkaan (Gruber 1993), joka mahdollistaa *Ontolinguan* käyttämisen kehysten avulla. Frame Ontology määrittelee ontologioille tyypillisiä termejä, joista kukin on oma kehyksensä. Esimerkiksi luokka, ilmentymä, alaluokka ja arvotyyppi ovat Frame Ontologyn kehyksiä. Koska KIF on ilmaisukykyisempi kuin Frame Ontology, *Ontolingua* antaa mahdollisuuden sijoittaa KIF-lauseita suoraan Frame Ontologyn määrittelyjen sisään. *Ontolinguaa* voi siis käyttää kolmella eri tavalla: kielenä voi olla joko KIF, Frame Ontology tai niiden yhdistelmä.

Frame Logic tai FLogic (Kifer, Lausen & Wu 1995) yhdistää kehyksiä ja 1. kertaluvun logiikkaa. Sillä voi esittää käsitteitä, taksonomioita, binäärisiä suhteita, funktioita, ilmentymiä, aksioomia ja päättelysääntöjä. Lisäksi sen ominaisuuksiin kuuluvat esimerkiksi periytyminen, monimuotoiset tyypit ja kyselymenetelmät. Frame Logic -kielen Ontobroker-päätelykonetta voi käyttää rajoitteiden tarkistamiseen ja uuden informaation päättelemiseen.

OCML (Motta 1998) on toiminnallinen mallinnuskieli, jonka avulla voidaan ilmaista käsitteitä, suhteita, funktioita, sääntöjä ja ilmentymiä. Toiminnallisuuteen kuuluvat esimerkiksi loogiset päätelyoperaatiot, funktioiden arvioinnit ja väittämien todistaminen. Näitä ominaisuuksia voidaan käyttää hyväksi myös *OCML:n* kanssa yhteensopivassa *Ontolingua*ssa.

Loom (MacGregor 1999) on sekä ontologiakieli että kehitysympäristö, joka luotiin tukemaan mallipohjaisten sovellusten rakentamista ja ylläpitoa. Loom-kieli perustuu kuvauslogiikkaan ja muodostuu säännöistä sekä faktoista. Kehitysympäristön ytimenä toimii luokittelijamoottori. Se tutkii annettuja muuttujia ja niitä koskevia sääntöjä luoden niiden pohjalta tietoa kuvaavia

suhdeverkkoja. Näitä verkkoja käytetään avuksi esimerkiksi erilaisia hakuja tehtäessä.

3.3.2 Web-kielet

XML-kieli eli Extensible Markup Language (Bray, Paoli, Sperberg-McQueen & Maler 2000) on W3C:n kehittämä standardi *www*-resurssien rakenteen kuvaamiseen. *XML*-kielellä käyttäjät voivat muun muassa määrittellä omia elementtejä, attribuutteja ja tietorakenteita. *XML*-kielessä itsessään ei ole erityispiirteitä ontologioiden määrittelemiseen, mutta Fenselin (2000, 72) mukaan ontologiseen mallintamiseen voitaisiin mahdollisesti käyttää *XML*:n DTD-määritystä (Document Type Definition) tai *XML Schema* -kieltä (Fallside 2001). Esimerkiksi DTD määrittelee elementit, attribuutit ja rakenteiden sallitun sisäkkäisyyden, mikä voidaan rinnastaa ontologian määrittelemiseen. Ne eivät kuitenkaan ole kovin ilmaisuvoimaisia tapoja esittää ontologioita. *XML*-kieltä on käytetty *web*-perustaisten ontologiakielten perustana, koska se tarjoaa syntaksin määrittelemiseen yksinkertaisen ja tehokkaan tavan. Lisäksi *XML*:n hyötynä on helposti ihmisten luettavissa oleva koodattu tieto.

RDF eli Resource Description Framework (Lassila & Swick 1999) on W3C:n kehittämä malli ja kieli *www*-resurssien metatiedon kuvaamiseen. *RDF* mahdollistaa rakenteistetun metatiedon koodaamisen, jakamisen ja uudelleenkäyttämisen. *RDF*:n tietomalli perustuu semanttisiin verkkoihin ja keskeisiä käsitteitä ovat resurssit (engl. resources), ominaisuudet (engl. properties) ja lauseet (engl. statements). Resurssit ovat mitä tahansa asioita, joihin voidaan liittää tunniste. Resurssia voidaan kuvata yhdellä tai useammalla ominaisuudella ja niiden arvoilla. Lause koostuu resurssista, ominaisuudesta ja ominaisuuden arvosta. Näitä lauseen osia kutsutaan myös subjektiksi,

predikaatiksi ja objektiksi. Ominaisuuden arvo voi olla myös resurssi, jolloin voidaan muodostaa lauseketjuja.

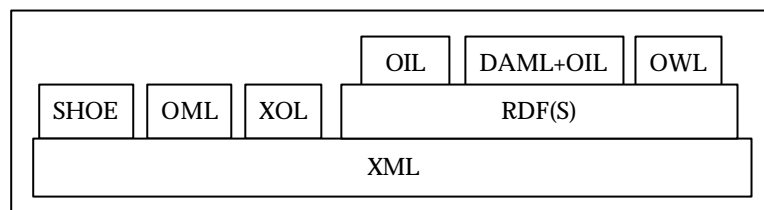
Esimerkkinä RDF-kielestä voitaisiin kuvata tilanne, jossa Matti Jokunen on tehnyt raportin. Tällöin lauseen subjektina eli resurssina on raportti, joka yksilöidään URI:n (Uniform Resource Identifier) avulla tunnisteella <http://www.koti.fi/raportti>. Lauseen predikaattina eli ominaisuutena on tekijä, jonka arvona eli objektina on Matti Jokunen. Sama asia voidaan esittää RDF-kielen syntaksilla seuraavasti:

```
<rdf:RDF>
  <rdf:Description about="http://www.koti.fi/raportti">
    <rdf:Tekija>Matti Jokunen</rdf:Tekija>
  <rdf:Description>
</rdf:RDF>
```

RDF Schema (Brickley & Guha 2003), lyhenteeltään RDFS, on W3C:n kehittämä kieli RDF-skeemojen määrittämiseen. RDFS:n tietomalli perustuu kehyksiin ja tarjoaa mekanismin ominaisuuksien ja resurssien välisten suhteiden määrittämiseen. RDFS:n tehtävänä on myös mahdollistaa luokkien ja rajoitteiden määrittely sekä luokkahierarkioiden käyttö. Fenselin (2000, 74) mukaan RDF Schema -kieltä voidaan suoraan käyttää ontologian kuvaamiseen, koska sillä voidaan ilmaista käsitteitä, suhteita, ominaisuuksia ja rajoitteita. RDFS tarjoaa siis syntaksin ontologioiden esittämiseen, mutta se on ilmaisuvoimaltaan melko rajoitettu, koska siitä puuttuu muun muassa tapa kuvata loogisia aksioomia.

3.3.3 Web-perustaiset esityskielet

Internetin räjähdysmäinen kasvu on luonut tarpeen tuoda ontologiat ja niiden mahdollistamat sovellukset myös www-ympäristöön. Ontologioiden web-perustaisten esityskielen suhteet XML:ään ja RDF(S):ään on esitetty kuviossa 3. RDF(S) on merkintätapa, jolla tarkoitetaan RDF:n ja RDF Scheman yhdistelmää.



KUVIO 3. Web-perustaiset kielet: suhdekaavio (mukaillen Corcho ym. 2003, 55)

SHOE-kieli eli Simple HTML Ontology Extension (Luke & Heflin 2000) on kehitetty Marylandin yliopistossa HTML:n laajenuksena liittämään tietokoneiden luettavissa olevaa semanttista tietämystä HTML-dokumentteihin. *SHOE* muodostuu joukosta HTML:ää laajentavia tunnisteita. *SHOE*-syntaksi on muokattu yhteensopivaksi myös XML-kielen kanssa. Tunnisteiden avulla voidaan sekä luoda ontologioita että käyttää niiden määrittämiä selittämään www-sivujen semantiikkaa. Omien määrittämiensä lisäksi *SHOE*-ontologiaa voidaan laajentaa liittämällä siihen toinen ontologia ja periyttämällä sen ominaisuudet.

Washingtonin yliopistossa kehitetty *OML* eli Ontology Markup Language (Kent 1999) perustuu osittain *SHOE*-kieleen. *OML*-kieli on jaettu neljään tasoon: Core, Simple, Abbreviated ja Standard. Ytimenä toimii muihin tasoihin sisältyvä

OML Core. Simple OML:n esitystapoina voi olla RDF(S) tai XML Schema. Abbreviated- ja Standard-tasoissa käytetään lisäksi muun muassa käsitekaavioita.

XOL-kieli eli XML-based Ontology Exchange Language (Karp, Chaudhri & Thomere 1999) kehitettiin alunperin bioinformatiikan tarpeisiin. XOL-kieli on kehitetty Ontolingua- ja OML-kielten pohjalta. Sen syntaksi pohjautuu OML:n tavoin XML:ään eikä Lisp-kieleen kuten Ontolinguaassa. XOL-kielen semantiikka perustuu OKBC-sovellusrajapintaan (Open Knowledge Base Connectivity), joka vastaa melko pitkälti Ontolinguan semantiikkaa. OKBC-sovellusrajapinta mahdollistaa kehyspohjaisten tietämysjärjestelmien käytön (ks. esim. Chaudhri, Farquhar, Fikes, Karp & Rice 1998). XOL-kielestä on pyritty tekemään yleispätevä ontologiakieli, jota voidaan käyttää minkä tahansa ontologian kuvaamiseen. XOL-kielessä tarvitaan vain yksi XML:n DTD, jonka avulla pystytään kuvaamaan kaikki ontologiat.

Esimerkki XOL-kielestä (mukaillen Karp, Chaudhri & Thomere 1999, 6):

<pre> <class> <name>henkilö</name> </class> <slot> <name>ikä</name> <domain>henkilö</domain> <value-type>integer</value-type> <numeric-max>150</numeric-max> </slot> <individual> <name>Minna</name> <type>henkilö</type> <slot-values> <name>ikä</name> <value>28</value> </slot-values> </individual> </pre>	<pre> <!-- luokan määrittely --> <!-- luokan nimenä on henkilö --> <!-- attribuutin määrittely --> <!-- attribuutin nimenä on ikä --> <!-- attribuuttia käytetään henkilö-luokassa --> <!-- attribuutin arvon tietotyypin määrittely --> <!-- ylärajan määrittely kokonaisluku-tietotyypille --> <!-- ilmentymän määrittely --> <!-- ilmentymän nimenä on Minna --> <!-- ilmentymä kuuluu henkilö-luokkaan --> <!-- ilmentymän attribuuttien määrittely --> <!-- ilmentymä perii henkilö-luokan ikä-attribuutin --> <!-- attribuutti saa arvokseen Minnan iän --> </pre>
--	--

OIL-kielen eli *Ontology Inference Layerin* (Fensel, van Harmelen, Horrocks, McGuinness & Patel-Schneider 2001) taustalla on kolme eri lähestymistapaa. Siinä yhdistyvät kuvauslogiikka, kehyspohjaiset menetelmät sekä *www*-standardit XML ja RDF. Kuvauslogiikka mahdollistaa formaalin esitystavan sekä tehokkaat päättelyominaisuudet. Kehyspohjainen lähestymistapa tarjoaa käyttäjille intuitiivisen mallinnus- ja käyttöympäristön. XML- ja RDF-standardien käyttö puolestaan takaa laajat käyttömahdollisuudet Internet-ympäristössä. *OIL* muodostuu neljästä tasosta: Core, Standard, Instance ja Heavy *OIL*, joiden toiminnallisuus ja monimutkaisuus kasvavat ylöspäin edettäessä. Core *OIL* on lähes kokonaan RDFS-syntaksin mukainen. Standard *OIL* laajentaa Core *OIL*:ia termimäärittelyillä, joiden avulla ontologian semantiikasta saadaan yksiselitteinen. Kolmas taso Instance *OIL* laajentaa edellisiä tarjoamalla tietokantatoimintoja. Vielä määrittelemättömän Heavy *OIL*-tason tavoitteena on tarjota laajennettuja esitystapoja ja loogista päättelyä.

DAML+OIL (McGuinness, Fikes, Hendler & Stein 2002) on *DAML*-kielen (*DARPA Agent Markup Language*) ja edellä esitellyn *OIL*-kielen yhdistelmä. Kielen kehittäjänä toimii Yhdysvaltojen puolustusministeriön alainen *DARPA* (*Defence Advanced Research Projects Agency*). *DAML+OIL* rakentuu RDF(S)-syntaksin päälle ja on hyvin samankaltainen kuin *OIL*. Eroja löytyy lähinnä RDF-syntaksin käytössä sekä ominaisuuksien ja osien nimeämisessä. Lisäksi on olemassa muutamia muita eroavuuksia; *OIL* esimerkiksi määrittelee konkreettiset tietotyypit, jotka puuttuvat kokonaan *DAML+OIL*-määrittelystä.

OWL eli *Web Ontology Language* (McGuinness & van Harmelen 2003) on W3C:n web-ontologiatyöryhmän kehitteillä oleva ehdotus viralliseksi semanttisen webin ontologiakieleksi. Se pohjautuu *DAML+OIL*-kieleen. *OWL*:stä on olemassa täydellisen *OWL*:n (*OWL Full*) lisäksi *OWL Lite* ja *OWL*

DL. OWL Lite on karsittu versio, joka sisältää lähinnä luokitteluhierarkian ja yksinkertaisia rajoitteita. OWL DL sisältää kaikki OWL-kielen rakenteet, mutta niitä voi käyttää vain tiettyjen rajoitteiden salliessa.

3.4 Ontologiatyökaluja

Ontologioiden kehittämismenetelmien ja esityskielen lisäksi tarvitaan vielä työkaluja, joiden avulla ontologioita voidaan luoda. Hyvösen (2002, 15) mukaan ontologioiden kehittämiseen tarvitaan seuraavia työkaluja:

- Editoreja, joilla ontologiakuvaukset luodaan manuaalisesti tai puolimanuaalisesti.
- Annotointityökaluja, joilla linkitetään tietolähteet metatietoon.
- Päättelytyökaluja ontologioiden oikeellisuuden tarkistamiseen ja hallintaan.
- Ontologiakirjastoja, joiden avulla luodaan uusia ontologioita olemassa olevista ontologioista.

Tässä tutkielmassa kuitenkin keskitytään editoreihin, joilla ontologioita voidaan luoda, selata, muokata, ylläpitää ja käyttää. Niistä keskeisimpiä esitellään seuraavassa lyhyesti. Lisäksi esitellään lyhyesti tunnetuimpia ontologiakirjastoja.

Ontolingua (Farquhar ym. 1997) on työkalu, jonka avulla voi rakentaa, julkaista ja jakaa ontologioita. Sen avulla on tarkoitus kehittää ontologioiden uudelleenkäyttöä. Ontolinguassa on web-pohjainen käyttöliittymä, jonka avulla voi selata valmiita ontologioita tai tehdä oman ontologian editorilla. Ontolingua kääntää ontologian automaattisesti muille sisältökielille, esimerkiksi KIF-, Loom- tai Prolog-kielille.

Ontosaurus (Swartout ym. 1996) on web-selain Loom-tietämuskantojen selaamiseen. Sen avulla voi myös editoida rajoitetusti Loomin tietämuskantoja. Ontosaurusin tietämysten esityskieli on Loom, joka on sekä kieli että ympäristö älykkäiden sovellusten rakentamiseen. Loomin ydin on tietämysten esittämisjärjestelmä, jota käytetään tuottamaan deduktiivinen tuki Loom-kielen sisällölle. Tietämysten sisältö Loomissa koostuu määritelmistä, säännöistä, faktoista ja oletussäännöistä. Päätelyjärjestelmään kuuluvat eteenpäin ketjutus, semanttinen yhdistäminen ja oliosuuntautunut teknologia totuusarvojen ylläpitoon.

WebOnto (Domingue 1998) on täysin Internetin kautta käytettävä ontologiaeditori. Se on suunniteltu tukemaan yhteiskäyttöä ontologioiden luomisessa, selaamisessa ja editoinnissa. WebOnton käyttämä kieli ontologioiden mallintamiseen on OCML (Operational Conceptual Modeling Language). WebOnton ominaisuuksiin kuuluvat rakennekaavioiden tallentaminen sekä luokkien, suhteiden ja sääntöjen selkeä esittäminen. Erityinen WebOnton lisäominaisuus on mahdollisuus yhteistyöhön lähetys- ja vastaanottotoimintojen avulla.

Ontolingua-palvelin, *Ontosaurus* ja *WebOnto* ovat kaikki tutkimustoimintaan kehitettyjä työkaluja, joilla on vahva suhde tiettyyn kieleen. Niiden jälkeen on kehitetty monipuolisempia ontologiatyökaluja. Esimerkiksi *Protégé2000*, *WebODE* ja *OntoEdit* on rakennettu tukemaan useimpia ontologioiden elinkaaren toimintoja. Niissä on kattava, komponenttiperustainen arkkitehtuuri, johon voidaan lisätä uusia moduuleja helposti lisätoiminnallisuuksien tarjoamiseksi. Lisäksi nämä työkalut ovat kieliriippumattomia.

Protégé2000 (Grosso, Eriksson, Fergerson, Gennari, Tu & Musen 1999) on itsenäinen sovellus, joka perustuu avoimeen lähdekoodiin. Työkalu avustaa kehittäjiä laajojen elektronisten tietämuskantojen luomisessa. Protégén avulla voi luoda, selata ja muokata aihealueontologioita kehysperustaisella esitystavalla, mikä on yhteensopiva OKBC-tietämysmallin kanssa. Kehittäjät voivat räätälöidä graafista tietämyksenhankintatyökalua suoraan järjestelemällä lomakkeiden graafisia olioita, jotka on liitetty ontologian jokaiseen luokkaan ilmentymien johtamiseksi. Tietämuskantoja voidaan tallentaa useissa muodoissa, kuten CLIPS-pohjaisella syntaksilla ja RDF-muodossa. Protégé-työkaluun saa lisättyä toiminnallisuutta lisäohjelmien (engl. plug-in) avulla.

WebODE-työkalua (Arpírez, Corcho, Fernández-López & Gómez-Pérez 2001) käytetään www-palvelimen tavoin www-käyttöliittymän kautta. Työkalu sisältää useita toimintoja muun muassa ontologioiden dokumentointiin, analysointiin ja yhdistämiseen. WebODE tukee haluttaessa Methontology-menetelmää. Myös *OntoEdit*-työkalulla on kattava ja joustava arkkitehtuuri, joka mahdollistaa ontologioiden selaamisen ja muokkaamisen sekä lisättävät toiminnallisuudet (Corcho ym. 2003, 49-50).

Protégé2000, WebODE ja OntoEdit sallivat DAML+OIL- ja RDF(S)-kielillä luotujen ontologioiden tuonnin ja viennin. Semanttisen Webin standardien kehittyessä on kuitenkin luotu työkaluja myös web-perustaisia ontologioita varten. *OILEd* on DAML+OIL-ontologioiden editori. *DUET*-työkalun avulla tietokantojen suunnittelijat ja järjestelmien kehittäjät voivat mallintaa ontologiat UML-mallinnuskielen (Unified Modeling Language) avulla ja sitten kääntää ne DAML+OIL-kielelle. (Corcho ym. 2003, 50)

Osa erilaisilla ontologiakielillä ja työkaluilla eri tahoilla kehitetyistä ontologioista on kerätty uudelleenkäytön helpottamiseksi ontologiakirjastoihin (engl. ontology library). Kirjastot eivät ole kaikenkattavia ja usein ne keskittyvät vain yhdellä kielellä tai työkalulla esitettyihin ontologioihin. Kaikilla kielillä ja työkaluilla ei kuitenkaan ole omaa ontologiakirjastoa, mikä vaikeuttaa niiden käyttöasteen vertailua. Pienempiä tai kaupallisesti arvokkaita ontologioita kirjastoista ei luonnollisestikaan löydy. Kirjastot ovat kuitenkin hyvä lähtökohta etsittäessä sopivaa pohjaa oman ontologian rakentamiselle.

Tunnetuimpia ontologiakirjastoja ovat (Corcho, Fernández-López & Cómez-Pérez 2001):

- DAML Ontology Library. Sisältää yli 250 ontologiaa toteutettuna DAML- ja DAML+OIL-kielillä. Vapaasti käytettävissä osoitteessa <<http://www.daml.org/ontologies/>>.
- Ontolingua Server Repository. Sisältää yli 50 ontologiaa toteutettuna Ontolingua-kielillä. Käytettävissä osoitteessa <<http://www-ksl-svc.stanford.edu:5915/>>.
- SHOE Ontologies. Sisältää 12 ontologiaa toteutettuna SHOE-kielillä. Vapaasti käytettävissä osoitteessa <<http://www.cs.umd.edu/projects/plus/SHOE/onts/>>.
- WebODE. Ontologiat on toteutettu RDF(S)- ja OIL-kielillä. Käytettävissä rekisteröinnin jälkeen osoitteessa <<http://delicias.dia.fi.upm.es/webODE/>>.
- WebONTO. Ontologioiden toteutuskielenä on OCML. Vapaasti käytettävissä osoitteessa <<http://kmi.open.ac.uk/projects/webonto/>>.
- OntoSaurus. Ontologioiden toteutuskielenä toimii LOOM. Käytettävissä rekisteröinnin jälkeen osoitteessa <<http://www.isi.edu/isd/ontosaurus.html>>.

Tässä tutkielmassa on tähän mennessä keskitytty ontologioihin. Luvussa kaksi käsiteltiin ontologioihin liittyviä perusasioita ja luvussa kolme paneuduttiin ontologioiden kehittämiseen liittyviin asioihin. Seuraavaksi keskitytään sovellusalueeseen eli lainsäädäntötyöhön. Ensin luvussa neljä perehdytään Euroopan unionin lainsäädäntötyöhön ja sen jälkeen luvussa viisi Suomen lainsäädäntötyöhön. Tämän tutkielman varsinainen sovellusalue ei kata EU:n lainsäädäntötyötä, mutta sitä esitellään silti, koska se vaikuttaa monella tavalla suomalaisen lainsäädäntötyöhön ja siten myös sitä varten kehitettävän ontologiaan.

4 EU:N LAINSÄÄDÄNTÖTYÖ

Euroopan unionin luomaa oikeusjärjestystä kutsutaan eurooppaoikeudeksi, jolla on omat erityispiirteensä, johtavat periaatteensa ja oikeuslähteensä. Euroopan unionin lainsäädäntöprosessi käsittää erilaisia toimijoita, vaiheita, säädöstyyppejä, päätösmenettelyitä, asiakirjoja ja tietojärjestelmiä. Tässä luvussa käsitellään kuitenkin EU:n lainsäädäntötyötä vain hyvin yleisellä tasolla. Ensimmäiseksi selvitetään eurooppaoikeutta ja säädösten muotoja. Tämän jälkeen kartoitetaan lainsäädäntötyön toimijoita ja varsinaista lainsäädäntöprosessia. Lopuksi tarkastellaan lainsäädäntöprosessiin liittyviä tietojärjestelmiä sekä EU:n alueen ontologia- ja sanastotyötä.

Euroopan unionin lainsäädäntötyöhön tutustutaan tässä tutkielmassa, koska Suomi on osa EU:ta ja siten osa EU:n lainsäädäntöprosessia. Lisäksi EU:n tasolla toteutetuista tietojärjestelmäratkaisuksista ja niiden kehittämiskokemuksista voi olla hyötyä luotaessa ontologioita suomalaista lainsäädäntöprosessia varten. Ontologian toteuttamisessa on otettava myös huomioon tietojärjestelmien integroiminen EU:n tietojärjestelmiin tulevaisuudessa.

Tämän luvun keskeisimpiä tietolähteitä ovat olleet RASKE-projektin raportti *EU-lainsäädäntöasiakirjat Suomessa* (Tiitinen, Salminen ja Lyytikäinen 1997), Euroopan komission julkaisema *Yhteisön oikeuden perusteet* (Borchardt 2000), Kempin *Suomi Euroopan unionissa – perusteos* (2002) sekä Joutsamon, Aallon, Kailan ja Maunun kirja *Eurooppaoikeus* (2000). Tämän luvun kuviot on mallinnettu RASKE-projektissa kehitetyillä menetelmillä (Salminen 2003), koska niiden avulla kuvioista saadaan selkeitä ja helposti ymmärrettäviä.

4.1 Eurooppaoikeus, yhteisön oikeus ja säädösten muodot

Euroopan unioni (EU) on Maastrichtin sopimuksella (vuonna 1993) perustettu itsenäinen kansainvälinen organisaatio. Euroopan unioni ei ole korvannut Euroopan yhteisöjä (EY), vaan täydentää niitä uusilla menettelytavoilla ja yhteistyön muodoilla. Siten Euroopan unionin muodostaa kolme pilaria: Euroopan yhteisöt, yhteinen ulko- ja turvallisuuspolitiikka sekä yhteistyö oikeus- ja sisäasioissa. (Borchardt 2000, 8)

Yhteisön oikeudella tarkoitetaan EU:n ensimmäisen pilarin oikeusjärjestelmää kun taas eurooppaoikeudella viitataan kolmen pilarin kokonaisuuteen (Joutsamo, Aalto, Kaila & Maunu 2000, 17). Yhteisön oikeus on EY:n luoma oma ja riippumaton oikeusjärjestelmä, jolle kansallinen lainsäädäntö on alisteista EY:n toimialoilla. Oikeusjärjestelmän perustana on lainsäädäntövallan ja muiden valtiollisten toimivaltojen rajattu siirtäminen jäsenvaltioilta yhteisön toimielimille yhteisön päämäärien toteuttamiseksi (Eerola, Mylly & Saarinen 2000, 19). Ylikansallisia piirteitä ei sisälly unionisopimuksen toiseen ja kolmanteen pilariin. Näihin kahteen pilariin sisältyy ohjelmia ja aiejulistuksia, jotka toteutetaan hallitusten välisenä yhteistyönä.

Oikeusjärjestelmän toteuttamiseen, valvontaan ja kehittämiseen liittyy joukko menettelytapoja. Näitä ovat esimerkiksi oikeuslähteet, joilla tarkoitetaan oikeuden alkuperää ja vakiintunutta muotoa (Borchardt 2000, 58). Yhteisön oikeuden lähteitä ovat primaarioikeus (perustamissopimukset), johdettu oikeus, kansainväliset sopimukset, oikeuskäytäntö ja jäsenvaltioiden väliset sopimukset. Oikeuslähteistä johdettu oikeus vaikuttaa eniten Suomen lainsäädäntötyöhön, sillä se on lakeihin verrattavien, sitovien oikeusnormien muodostama kokonaisuus (Kemppinen 2002, 19). Johdettu lainsäädäntö kattaa

kaikki yhteisöjen toimialat. Johdetun oikeuden säädöstyyppejä ovat sitovat asetukset, direktiivit ja päätökset sekä ei-sitovat suositukset ja lausunnot (EY:n perustamissopimus 2002, 249 artikla).

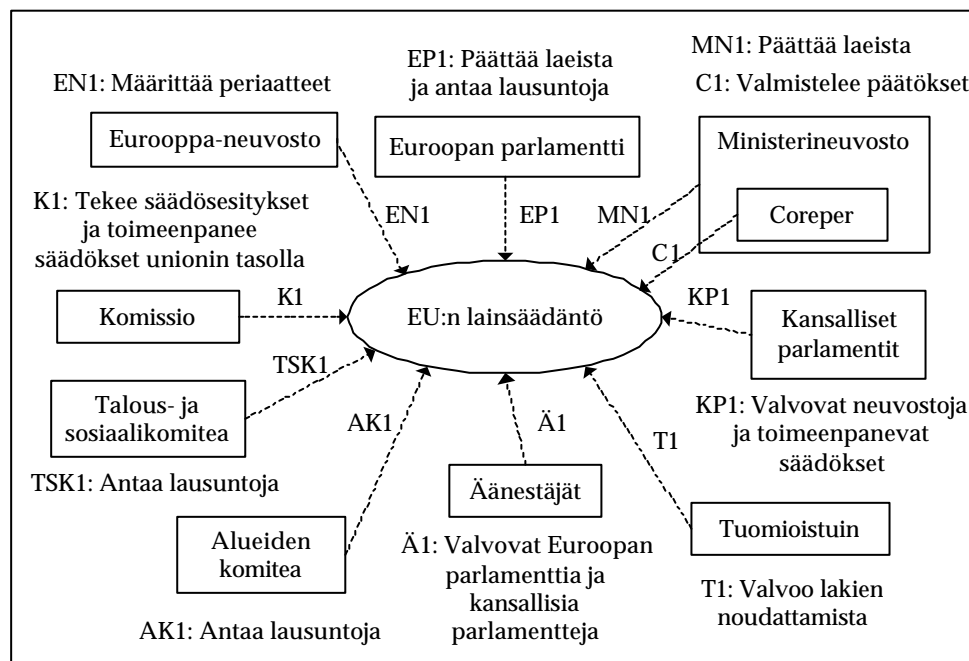
Asetuksia voi luonnehtia ylikansalliseksi lainsäädännöksi ja niitä voi vaikutustensa puolesta verrata kansalliseen lakiin. Asetuksia sovelletaan suoraan ja ne sitovat kaikkia EU:n jäsenvaltioita ilman kansallista täytäntöönpanevaa lainsäädäntöä (EY:n perustamissopimus 2002, 249 artikla). Jäsenvaltion on kumottava niissä asioissa antamansa ristiriitaiset säädökset ja määräykset, joista unioni antaa asetuksen (Kemppinen 2002, 19). Direktiivissä asetetaan tietty tavoite, joka jäsenvaltioiden pitää toteuttaa määrätyssä ajassa. Kansalliset viranomaiset saavat kuitenkin itse valita muodot ja keinot (EY:n perustamissopimus 2002, 249 artikla). Direktiivit on pantava täytäntöön kansallisessa lainsäädännössä esimerkiksi säätämällä laki kunkin jäsenvaltion menettelytapojen mukaisesti.

Päätöksen soveltamisala on yksilöity, se on kaikilta osiltaan velvoittava ja sitä sovelletaan sellaisenaan kohderyhmään (EY:n perustamissopimus 2002, 249 artikla). Päätökset eivät siten edellytä kansallista täytäntöönpanevaa lainsäädäntöä. Useimmiten päätöksiä käytetään silloin, kun asetuksella tai direktiivillä säädettyä asiaa on täydennettävä yksityiskohtaisemmilla määräyksillä (Kemppinen 2002, 20). Suositukset ja lausunnot eivät ole sitovia eli niistä ei aiheudu vastaanottajalle oikeudellisia velvoitteita (EY:n perustamissopimus 2002, 249 artikla). Suosituksilla ja lausunnoilla on kuitenkin poliittista ja moraalista merkitystä (Borchardt 2000, 70). Toimielin yleensä antaa suosituksia omasta aloitteestaan kun taas lausuntoja toimielin antaa yleensä toisen toimielimen pyynnöstä (Kemppinen 2002, 20).

4.2 Lainsäädäntöprosessi

EU:n lainsäädäntöprosessiin tutustuttaessa on muistettava, että unioni voi päättää ainoastaan sellaisista asioista, jotka kuuluvat sille perussopimusten nojalla. Jäsenvaltiot ovat antaneet tähän suostumuksensa ratifioimalla sopimukset. Lisäksi päätöksentekoon liittyvän läheisyysperiaatteen mukaisesti unionin tasolla päätetään vain sellaisista asioista, joita ei tarkoituksenmukaisemmin voida hoitaa jäsenmaissa (Tiitinen, Salminen & Lyytikäinen 1997, 3).

Johdetun oikeuden normien säätämisen prosessin rungon muodostavat Eurooppa-neuvosto, komissio, ministerineuvosto, pysyvien edustajien komitea (Coreper), parlamentti, tuomioistuin, talous- ja sosiaalikomitea sekä alueiden komitea kuvion 4 osoittamalla tavalla.



KUVIO 4. EU:n lainsäädäntötyön toimijat (mukailen Raunio 1999, 22)

Eurooppa-neuvosto <<http://ue.eu.int/>> on jäsenvaltioiden valtionpäämiesten ja pääministereiden sekä komission puheenjohtajan välinen kokous. Eurooppa-neuvoston tehtävänä on muun muassa antaa unionille virikkeitä toiminnan kehittämiseksi ja linjata keskeisiä tulevaisuuden suuntaviivoja. (Kemppinen 2002, 23) *Euroopan komissio* <<http://europa.eu.int/comm/>> on ylikansallinen ja riippumaton toimielin, joka vastaa lainsäädäntöä koskevien aloitteiden valmistelusta, käyttää neuvoston delegoimaa itsenäistä päätösvaltaa ja valvoo säädösten toimeenpanoa. *Ministerineuvosto* <<http://ue.eu.int/>> on Euroopan unionin varsinainen lainsäätävä ja tärkein päätöksentekijä, sillä yksikään EU-laki ei voi tulla voimaan vastoin sen tahtoa. Neuvostossa ovat edustettuina jäsenvaltioiden hallitukset. (Tiitinen ym. 1997, 3-4)

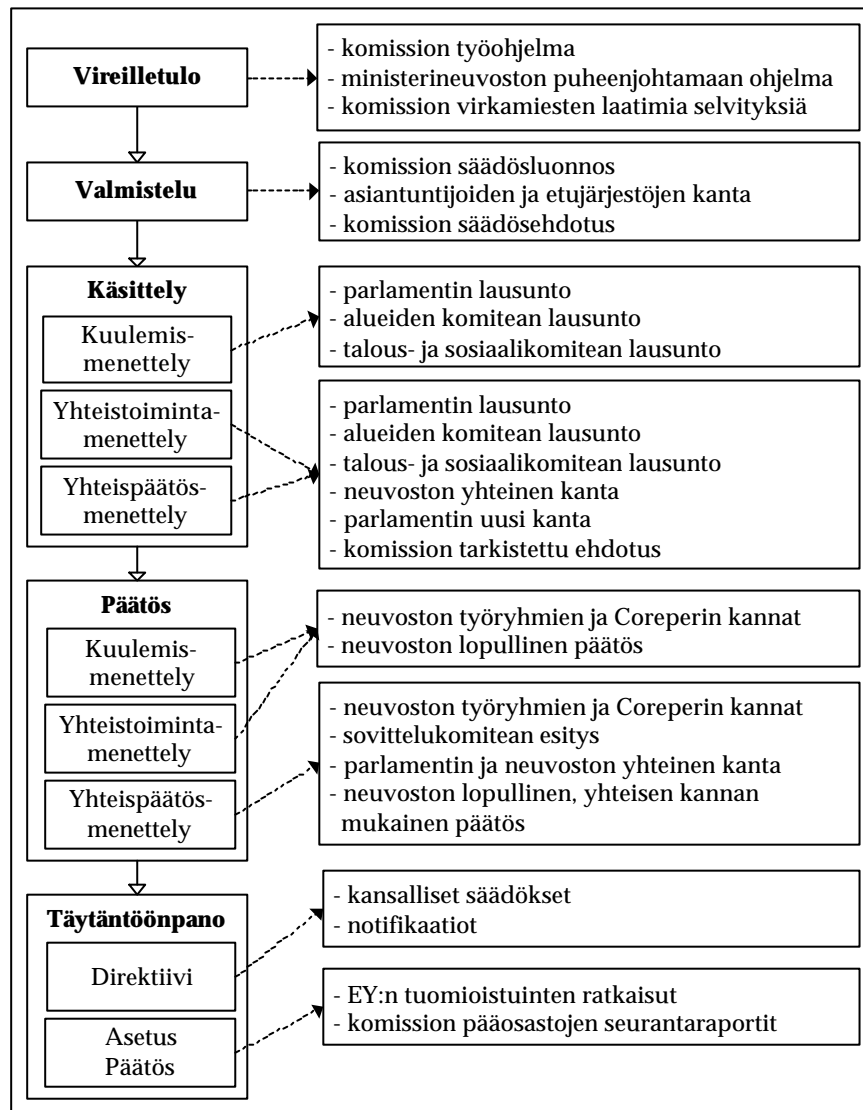
Neuvoston päätökset valmistellaan työryhmissä ja komiteoissa. Tärkein virkamieskomitea on *Coreper* eli jäsenvaltioiden pysyvien edustajien komitea (Borchardt 2000, 42). EU-suurlähettiläistä muodostuva *Coreper II* käsittelee kaikki poliittisesti kiistanalaiset asiat ja EU-suurlähettiläiden sijaisista muodostuva *Coreper I* teknisluontoiset asiat. *Coreper* pyrkii ratkaisemaan erilaisia työryhmissä ratkaisematta jääneitä ongelmia ja tuottaa joko valmiita ratkaisuja tai muutamaaan vaihtoehtoon karsittuja ratkaisumalleja ministerien päätettäväksi. (Kemppinen 2002, 31)

Euroopan parlamentti <<http://www.europarl.eu.int/>> on unionin kansanedustuselin, joka edustaa jäsenvaltioiden kansalaisia. EY:n perustamissopimuksen (2002, 192 artikla) mukaan parlamentti osallistuu lakien säätämiseen, antaa lausuntoja ja valvoo komissiota. *Euroopan yhteisöjen tuomioistuin* <<http://curia.eu.int/>> valvoo päätösten laillisuutta ja tulkitsee perussopimuksia luoden samalla ylikansallista EU:n lainsäädäntöä. Tuomioistuimen yhteydessä toimii erillinen *ensimmäisen asteen tuomioistuin*.

Tietyissä asioissa kannellaan aluksi ensimmäisen asteen tuomioistuimeen, josta sitten voi valittaa varsinaiseen tuomioistuimeen. (Kemppinen 2002, 44)

Talous- ja sosiaalikomitea <<http://www.ces.eu.int/>> on pysyvä neuvoa-antava toimija, joka muodostuu työnantaja- ja työntekijäjärjestöistä sekä muista eturyhmistä. *Alueiden komitea* <<http://www.cor.eu.int/>> on riippumaton neuvoa-antava toimija, jonka jäsenet ovat alueellisten ja paikallisten yhteisöjen viranomaisia. Kyseiset komiteat antavat lausuntoja niiden toimivaltaan kuuluvissa asioissa. Lausunnot eivät kuitenkaan sido neuvostoa tai parlamenttia. *Kansalliset parlamentit* valvovat hallitusten työskentelyä ministerineuvostossa ja Eurooppa-neuvostossa. *Kansalliset* vaikuttavat kahden parlamentin kautta: valitsemalla edustajat kansallisiin parlamenteihin sekä Euroopan parlamenttiin. (Eerola ym. 2000, 55)

Toimijat osallistuvat tehtäviensä mukaisesti lainsäädäntöprosessiin. Lainsäädäntöprosessi on monimutkainen tapahtumasarja, joka voidaan jakaa vireilletulo-, valmistelu-, käsittely-, päätöksenteko- ja täytäntöönpanovaiheisiin. Kuviossa 5 on esitetty nämä vaiheet tiivistetysti.



KUVIO 5. EU:n lainsäädäntöprosessin vaiheet (mukaillen Tiitinen ym. 1997, 12)

Seuraavassa on kuvattu lyhyesti lainsäädäntöprosessia vaiheittain.

4.2.1 Vireilletulo- ja valmisteluvaiheet

Yleisimpiä säädöshankkeen vireilletulon alullepanijoita ovat kansainvälisistä sopimuksista johtuvat velvoitteet, voimassaolevan lainsäädännön muutostarve,

jäsenvaltioiden pyynnöt, neuvoston päätökset, neuvoston kanssa hyväksytyt ohjelmat ja elinkeinoelämän vaatimukset (Kemppinen 2002, 39-40). Komission esitysten valmistelun pohjana on vuotuinen työohjelma, joka sisältää ohjelmavuoden toiminnan keskeiset painopistealueet. Myös ministerineuvoston puheenjohtajana toimiva jäsenvaltio esittää työohjelman niistä hankkeista, joita se pyrkii puheenjohtajakaudellaan ajamaan. (Joutsamo ym. 2000, 194)

Työohjelmien lisäksi vireilletulovaiheeseen kuuluvat komission virkamiesten laatimat selvityksiä. Vihreiksi kirjoiksi kutsutuissa keskusteluasiakirjoissa esitetään komission tulevaa toimintaa yksittäisen hankkeen osalta. Niiden tarkoituksena on saada aikaan keskustelua ja käynnistää unionin tasolla neuvottelut. Yksittäisten isojen hankkeiden toteuttamiseen tarvittavat konkreettiset lainsäädäntötoimet komissio esittää kootusti niin sanotuissa valkoisissa kirjoissa. (Joutsamo ym. 2000, 194)

Varsinainen säädösvalmistelu tapahtuu komission pääosastoissa. Toimeksiannon valmistelutyöhön ryhtymisestä antaa päällikkö, komissaari tai komission yleisistunto. Komission sisäiseen säädösvalmisteluun kuuluu muun muassa hankkeen käynnistäminen toimivaltaisessa pääosastossa sekä hankkeen koordinointi muiden pääosastojen kanssa. (Joutsamo ym. 2000, 195) Kun on kyse laajakantoisesta hankkeesta, tehdään asiasta ensin suunnitelma, joka perustuu selvityksiin (vihreät ja valkoiset kirjat). Suunnitelman perusteella komission virkamiehet laativat säädösluonnoksen. (Eerola ym. 2000, 67-68)

Komission ulkoiseen valmisteluun kuuluu muun muassa eri tahojen kuuleminen ja tausta-asiakirjojen julkaisu. Komissio voi kuulla yksittäisiä asiantuntijoita tai komission kutsumia asiantuntijakomiteoita. Lisäksi komissio kuulee eri jäsenvaltioiden etujärjestöjä sekä erilaisia painostus- ja eturyhmiä.

Kuulemiskierroksen jälkeen komission virkamiehet laativat tarkistetun säädösehdotuksen, joka etenee asianomaisen komissaarin henkilökohtaisista avustajista koostuvaan kabinettiin. Jos kabinetti hyväksyy luonnoksen, se lähetetään vielä arvioitavaksi ja tiedoksi muihin mahdollisesti osallisiin kabinetteihin. (Joutsamo ym. 2000, 195)

Virallisen päätöksen säädösehdotuksesta tekee kaikkien komissaarien muodostama collegio, joka voi hyväksyä, hylätä tai palauttaa säädösehdotuksen uudelleen valmisteltavaksi. Jos enemmistö komissaareista hyväksyy luonnoksen, siitä tulee virallinen ehdotus, joka lähetetään eteenpäin oikeusperustan mukaisesti joko vain neuvostolle tai sekä parlamentille että neuvostolle. (Joutsamo ym. 2000, 140-141, 196)

4.2.2 Käsittely- ja päätöksentekovaiheet

Käsittely- ja päätöksentekovaiheessa asian etenemistapa riippuu päätöksentekomenettelyistä, joista tärkeimmät ovat kuulemis-, yhteistoiminta- ja yhteispäätösmenttelyt. Lainsäädäntömenettelyt eroavat toisistaan sen mukaan, mikä on parlamentin rooli päätöksenteossa. Valittava menettelytapa riippuu siitä perustamissopimuksen määräyksestä, jonka nojalla säädös annetaan. Borchardt (2000, 72-82) on kattavasti esitellyt eri lainsäädäntömenettelyjä, joita ei tässä tutkielmassa käsitellä kovin tarkasti.

Kuulemismenettelyssä parlamentti, talous- ja sosiaalikomitea sekä alueiden komitea antavat komission säädösehdotusta koskevat lausunnot, jonka jälkeen esitys menee neuvoston ratkaistavaksi. Kuulemismenettelyä käytetään esimerkiksi maataloutta, verotusta ja kilpailulainsäädäntöä koskevassa päätöksenteossa. Neuvoston kanta ehdotukseen muodostuu kaikissa

lainsäädäntömenettelyissä siten, että esitys menee ensin jäsenvaltioiden virkamiehistä koostuvaan neuvoston työryhmään. Työryhmän jälkeen asia menee Coreperiin, jossa ratkaistaan asiat, joista työryhmässä ei ole saavutettu sovintoa. Sen jälkeen säädöksestä päätetään ministerineuvostossa, jossa ratkaistaan puolestaan asiat, joista Coreperissa ei ole saatu sovittua. (Tiitinen ym. 1997, 12-13)

Yhteistoiminta- ja yhteispäätösmenettelyssä komission säädösehdotus käy myös lausuntokierroksella parlamentissa, talous- ja sosiaalikomiteassa sekä alueiden komiteassa. Lausuntojen pohjalta neuvosto julkaisee korjatun ehdotuksen eli niin sanotun neuvoston yhteisen kannan, jonka jälkeen asia palaa parlamenttiin. Parlamentilla on kolmen kuukauden määräajassa mahdollisuus yrittää muuttaa säädösehdotusta tai ehdottaa sen hylkäämistä. Jos parlamentti hyväksyy neuvoston yhteisen kannan tai ei tee päätöstä asiassa, voi neuvosto päättää ehdotuksen hyväksymisestä. Jos parlamentti esittää muutoksia yhteiseen kantaan, komissio ottaa niihin kantaa. Neuvosto voi päättää asiasta, jos komissio hyväksyy muutosehdotukset. Jos komissio hylkää ne, neuvosto voi tehdä ratkaisun joko vastoin komission tai parlamentin kantaa. Yhteistoimintamenettelyä käytetään enää oikeastaan vain joissakin Euroopan talous- ja rahaliittoa koskevissa asioissa. (Tiitinen ym. 1997, 13)

Parlamentilla on yhteispäätösmenttelyä käytettäessä veto-oikeus eli se voi hylätä neuvoston yhteisen kannan. Jos näin tapahtuu, asiaa käsitellään erityisessä sovittelukomiteassa. Mikäli parlamentti ei sen jälkeenkään hyväksy ehdotusta, se raukeaa ja käsittely päättyy siihen. Yhteispäätösmenttelyä käytetään muun muassa sisämarkkinoita, kuluttajansuojaa, tutkimusohjelmia ja henkilöiden vapaata liikkumista koskevassa päätöksenteossa. (Tiitinen ym. 1997, 13)

4.2.3 Täytäntöönpano

Neuvoston yksin sekä neuvoston ja parlamentin yhdessä antamat säädökset julkaistaan Euroopan yhteisöjen virallisessa lehdessä (EYVL), sen L-sarjassa (Eerola ym. 2000, 69). Euroopan yhteisöjen virallinen lehti julkaistaan joka arkipäivä Euroopan unionin 11 virallisella kielellä <suomenkielinen: http://europa.eu.int/eur-lex/fi/search/search_oj.html>. Lehti koostuu L-sarjasta (lainsäädäntö) ja C-sarjasta (tiedonannot, ilmoitukset ja valmisteilla olevat säädökset) sekä täydennysosasta (julkisia hankintoja koskevat ilmoitukset).

Myös neuvoston tai komission asetukset sekä näiden toimielinten antamat, kaikille jäsenvaltioille osoitetut direktiivit julkaistaan EYVL:ssä. Säädöksissä, asetuksissa ja direktiiveissä on tavallisesti maininta voimaantulopäivästä: ellei mainintaa ole, ne tulevat voimaan kahdentenakymmenentenä päivänä siitä, kun ne on julkaistu. Muut direktiivit ja päätökset annetaan tiedoksi niille, joille ne on osoitettu. Ne tulevat voimaan, kun ne on annettu tiedoksi. (Joutsamo ym. 2000, 211)

Komissio panee annetut säädökset täytäntöön yhteisön tasolla tehden asetuksia ja päätöksiä. Jäsenvaltioiden edustajat pystyvät valvomaan täytäntöönpanoa komitologiamenettelyssä. Yhteisön oikeuden täytäntöönpano tapahtuu kuitenkin tarpeellisilta osiltaan keskeisesti jäsenvaltioissa (ks. EU-asioiden valmistelu Suomessa kohdassa 5.1). Kansallisen täytäntöönpanon eli lakien ja määräyksien muuttamisen on tapahduttava säädöksissä ilmaistun määräajan puitteissa. Jäsenmaat lähettävät komissiolle ilmoituksen eli notifikaation säädösten voimaan saattamisesta. (Kemppinen 2002, 20) Säädösten käytännön toimeenpanoon liittyvät myös EY:n tuomioistuinten ratkaisut, jotka voivat vaikuttaa lain soveltamiseen. Komissiossa säädösten toimeenpanoa valvotaan

pääosastoilla ja valvontaan liittyvistä ohjeiden ja selvitysten antamista. (Tiitinen ym. 1997, 13-14)

4.3 Tietojärjestelmät

EU:n lainsäädäntötyö laajuuden ja monimutkaisuuden vuoksi on kehitetty useita tiedottamiseen liittyviä palveluja, jotta tärkeät tiedonlähteet olisivat jatkuvasti ja helposti saatavilla. Näiden palvelujen tarkoitus on tarjota kattava kuva EU:n lainsäädännöstä ja päätöksenteosta vireilletulovaiheesta päätöksentekoon saakka. Seuraavassa on lyhyesti kuvattu tärkeimpiä EU:n lainsäädäntötyön tietojärjestelmiä EU:n lainsäädännön keskeisten osapuolien (2003) pohjalta.

Celex <http://europa.eu.int/celex/htm/celex_fi.htm> on maksullinen tietokantapalvelu, joka tarjoaa käyttäjille kattavat hakutoiminnot jatkuvasti kasvavaan lähdemateriaaliin, kuten perustamissopimukseen, lainsäädäntöön, kansainvälisiin sopimukseen, oikeuskäytäntöön, valmisteleviin asiakirjoihin ja parlamentin jäsenten kysymyksiin. Toisiinsa liittyvissä teksteissä on hyperlinkkejä, jotka johtavat esimerkiksi myöhemmin tehtyihin muutoksiin, aikaisempiin asiakirjoihin ja jopa kansalliseen täytäntöönpanevaan lainsäädäntöön.

EUR-Lex <<http://europa.eu.int/eur-lex/fi/index.html>> tarjoaa ajan tasalla olevaa tietoa sekä voimassa olevasta Euroopan unionin lainsäädännöstä että uudesta, juuri hyväksytystä lainsäädännöstä. EUR-Lex sisältää myös Euroopan unionin virallisen lehden (EUVL), sopimuskokoelman, lakiehdotukset ja oikeuskäytännön. *Pre-Lex* <<http://europa.eu.int/prelex/apcnet.cfm?CL=fi>> on toimielinten välisten menettelyjen tietokanta, jonka avulla voidaan seurata

komission ja muiden toimielinten välisen päätöksenteon tärkeimpiä vaiheita. Komission pääsihteeristön ylläpitämä tietokanta ei itse sisällä asiakirjoja vaan linkit sähköisinä saataviin teksteihin.

OEIL <<http://wwwdb.europarl.eu.int/dors/oeil/en/default.htm>> on Euroopan parlamentin lainsäädäntötietokanta, joka kattaa toimielinten väliset lainsäädäntömenettelyt, toimielinten toimet ja Euroopan parlamentin istuntojaksot. Tiedot esitetään vain englanniksi ja ranskaksi. *CURIA* <<http://curia.eu.int/fi/index.htm>> on Euroopan yhteisöjen tuomioistuimen ja yhteisöjen ensimmäisen oikeusasteen tuomioistuimen monikielinen verkkosivu. Sivun avulla voidaan hakea tuomioistuimien tuomioita ja määräyksiä.

Tiivistelmät keskeisistä asiakirjoista ja lainsäädäntömenettelyistä kullakin Euroopan unionin toiminta-alalla sekä linkit kokoteksteihin löytyvät ilmaiseksi *SCADPlus*-palvelusta <http://europa.eu.int/scadplus/scad_fi.htm>. *RAPID* on tietokanta <<http://europa.eu.int/rapid/start/welcome.htm>>, joka sisältää komission, neuvoston ja muiden toimielinten lehdistötiedotteet sekä komissaarien puheet.

Euroopan unionin omien tietojärjestelmien lisäksi on olemassa muitakin lainsäädäntöön liittyviä järjestelmiä, kuten NATLEX, joita voidaan käyttää hyväksi lainsäädäntötyössä. Natlex (Natlex 2003) on Kansainvälisen työjärjestön ILO:n tietokanta <<http://natlex.ilo.org/>>, joka sisältää viitteitä yli 50 000 työelämään, sosiaaliturvaan ja ihmisoikeuksiin liittyvään kansalliseen säädökseen noin 175 maasta. Viitteet sisältävät kattavat bibliografiset tiedot, tiivistelmän ja avainsanat. Eri lakien väliset suhteet osoitetaan ja niiden välillä on linkitys. Viitteet kirjoitetaan joko englanniksi, ranskaksi tai espanjaksi, jotka

ovat ILO:n työskentelykieliä. Tietokannassa on myös yli 300 kokonaista lakitekstiä.

4.4 Ontologia- ja sanastotyö EU:n alueella

Euroopan unionissa ja sen jäsenvaltioissa tehdään vaihtelevassa määrin erityyppistä ontologia- ja sanastotyötä. Tässä esitellään tunnetuimpina esimerkkeinä EU:n ontologia- ja sanastotyöstä ParlML-hanke ja Eurovoc-tesaurus. Lisäksi kuvataan lyhyesti tunnettua Ison-Britannian e-GIF-viitekehystä.

ParlML (ParlML 2000) on Euroopan parlamentin alainen hanke, jonka tavoitteena on määrittellä yhteinen XML-perheen standardeihin perustuva sanasto lainsäädännöllisten ja parlamentaaristen tekstien rakenteistamiseen. Yhteisen, puhutuista kielistä riippumattoman merkkauskielen etuja olisivat muun muassa informaation saatavuuden ja vaihtamisen helpottaminen sekä puhuttujen kielten asettamien rajoitusten poistuminen.

Eurovoc (Eurovoc 2003) on monikielinen tesaurus, joka kattaa kaikki Euroopan yhteisöjen toimialat. Sen avulla voidaan indeksoida yhteisöjen toimielinten asiakirjajärjestelmien ja niiden käyttäjien asiakirjat. Tesaurusta käyttävät tällä hetkellä Euroopan parlamentti, Euroopan yhteisöjen virallisten julkaisujen toimisto, Euroopan kansalliset parlamentit ja hallinnot, aluehallinnot ja eräät eurooppalaiset järjestöt. Eurovoc on saatavana osoitteessa <<http://europa.eu.int/celex/eurovoc/>> Euroopan unionin 11 virallisella kielellä (espanja, tanska, saksa, kreikka, englantia, ranska, italia, hollanti, portugali, suomi ja ruotsi).

Ison-Britannian e-GIF-viitekehys eli e-Government Interoperability Framework (e-GIF 2003) auttaa parantamaan julkisia palveluita liittämällä niitä yhteen ja tekemällä ne helpommin yleisön saavutettavaksi. e-GIF määrittelee julkishallinnon teknisen menettelytavat ja vaatimukset, jonka tarkoituksena on saavuttaa tietojärjestelmien yhteentoimivuus ja yhtenäisyys koko julkisella sektorilla. e-GIF-viitekehys käyttää hyväksi Internet- ja www-standardeja. Esimerkiksi XML- ja XML Schema -kielet ovat datan yhdistämisen ja esittämisen ydinstandardeja. Kaikkien Ison-Britannian julkisen sektorin järjestelmien tulee olla yhteensopivia e-GIF-viitekehysten kanssa vuoteen 2005 mennessä.

Tämän luvun päätarkoituksena oli kuvata lyhyesti EU:n lainsäädäntötyötä. Seuraavassa luvussa tarkastellaan varsinaista sovellusaluetta eli Suomen lainsäädäntötyötä. Tarkoituksena on tulevaisuudessa integroida lainsäädäntöprosessin tietojärjestelmät ontologioiden avulla. Koska järjestelmät on kehitetty tukemaan tiettyjä prosesseja, on tärkeää ymmärtää nämä toimintaa ohjaavat prosessit hyvin.

5 SUOMEN LAINSÄÄDÄNTÖTYÖ

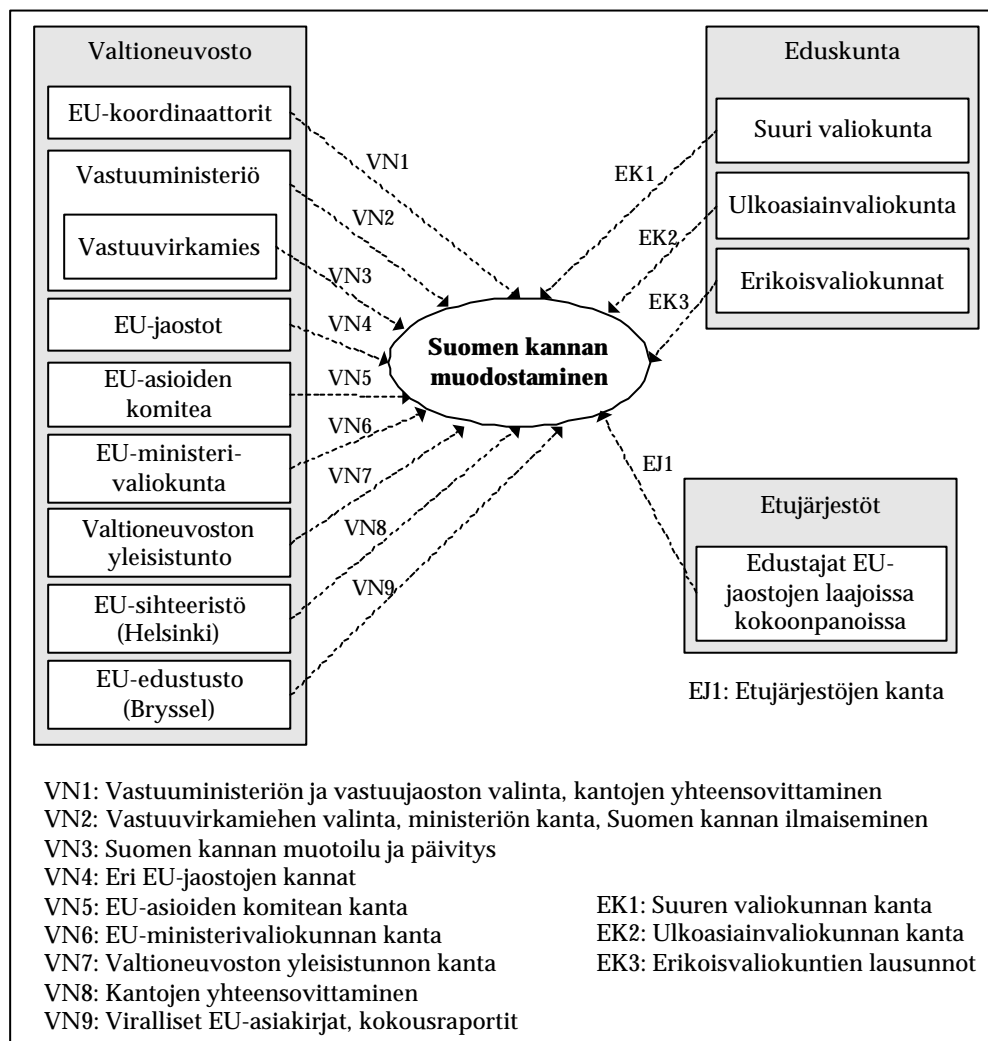
Suomen jäsenyys Euroopan unionissa edellyttää sekä kansallisten kantojen muodostamista unionin säädöshankkeisiin että menettelyjä unionissa hyväksytyjen säädösten täytäntöönpanoa varten. Tämän lisäksi Suomessa laaditaan kansallista lainsäädäntöä, jossa pitää myös ottaa Euroopan yhteisön säädösten vaikutus huomioon. Tässä luvussa selvitetään ensin EU-asioiden valmistelua Suomessa. Siihen liittyvät toimivallan jakaminen sekä Suomen kannan muodostaminen valtioneuvostossa ja eduskunnassa. Sen jälkeen selvitetään Suomen kansallista lainsäädäntötyötä lyhyesti vaiheittain. Luvun lopuksi perehdytään valtioneuvoston ja eduskunnan käytössä oleviin lainsäädäntöprosessin tietojärjestelmiin sekä suomalaisen julkishallinnon ontologia- ja sanastotyöhön.

Tärkeimpiä lähteitä tätä lukua koostettaessa ovat olleet RASKE-projektin raportti *EU-lainsäädäntöasiakirjat Suomessa* (Tiitinen, Salminen ja Lyytikäinen 1997), eduskunnan kanslian julkaisu *Suomen eduskunta* (2000), Mattilan artikkeli *Valtioneuvosto – Suomen EU-politiikan määrittelijä* (2000) sekä Joutsamon, Aallon, Kailan ja Maunun kirja *Eurooppaoikeus* (2000). Myös tämän luvun kuviot on mallinnettu johdonmukaisesti RASKE-projektin menetelmillä (Salminen 2003). RASKE2-projektissa mallinnetaan Suomen lainsäädäntöprosessia dokumenttien ja tietojärjestelmien näkökulmasta. Mallintamisen tuloksiin ei tässä luvussa kuitenkaan ole voitu tukeutua, koska työ on vielä kesken.

5.1 EU-asioiden käsittely Suomessa

Euroopan unionin jäsenvaltiona Suomi osallistuu täysivaltaisesti unionin päätöksentekoon. Kun EU:ssa käynnistyy lainsäädäntöhake, on Suomen kanta

hankkeeseen määriteltävä. Kansallisesti valmisteltu kanta on välttämätön perusta Suomen edustajille heidän osallistuessaan päätöksentekoon EU:n toimielimissä, kuten neuvoston työryhmissä, Coreperissa ja ministerineuvostossa. Kuviossa 6 on esitetty Suomen kannan muodostamiseen keskeisesti vaikuttavat osapuolet. Pääasialliset osapuolet ovat valtioneuvosto ja eduskunta, mutta niiden lisäksi myös etujärjestöt osallistuvat kannan muodostamiseen valtioneuvoston käsittelyn ohessa.



KUVIO 6. EU-asioiden käsittelyn osapuolet (mukaillen Tiitinen ym. 1997, 15)

Suomen perustuslain mukaan (93 §) valtioneuvosto vastaa EU:ssa tehtävien päätösten kansallisesta valmistelusta ja päättää niihin liittyvistä Suomen toimenpiteistä, jollei päätös vaadi eduskunnan hyväksymistä. Suomen perustuslakiin on sisällytetty säännökset myös eduskunnan osallistumisesta valmisteluun, jotka velvoittavat valtioneuvoston informoimaan eduskuntaa asioiden valmistelusta EU:ssa (Suomen eduskunta 2000, 59).

Taulukossa 2 esitetään EU-tason ja suomalaisen säädösvalmistelun eteneminen yleisellä tasolla. Tosin taulukko antaa osittain harhaanjohtavan kuvan valmistelun askeleittaisesta etenemisestä, sillä tosiasiasa monet valmisteluvaiheet ovat käynnissä samanaikaisesti sekä Brysselissä että Helsingissä.

TAULUKKO 2. Säädösvalmistelu EU:ssa ja Suomessa (mukaillen Mattila 2000)

EU-tason valmistelu	Valmistelu Suomessa
Komissio aloittaa säädösvalmistelun	Vastuuministeriö aloittaa kansallisen valmistelun
Komissio neuvottelee eturyhmien ja kansallisten virkamiesten kanssa	Vastuuministeriö osallistuu neuvotteluihin ja muotoilee Suomen alustavaa kantaa
Komissio päättää ehdotuksen sisällön	Virkamiehet päättävät alustavasta kannasta. Eduskuntaa informoidaan asiasta.
Komissio julkistaa ehdotuksen.	Ministeriöiden välisissä keskusteluissa päätetään Suomen tarkempi kanta.
Neuvoston työryhmät ja Euroopan parlamentti aloittavat ehdotuksen käsittelyn.	Vastuuvirkamies osallistuu työryhmätyöskentelyyn ja esittää Suomen kannan sekä raportoi edistymisestä Suomessa. Eduskunnan suurta valiokuntaa informoidaan edistyksestä ja se voi antaa ohjeita Suomen linjasta. Tarvittaessa Suomen kantaa voidaan muuttaa. EU-edustusto esittää tarvittaessa Suomen tavoitteet suomalaisille Euroopan parlamentin jäsenille.
Coreper ja neuvosto käsittelevät ehdotuksen.	Suomen edustajat osallistuvat Coreperin toimintaan. Pysyvä edustaja, ministeri ja virkamiehet hiovat Suomen kantaa. EU-ministerivaliokunta päättää ministerin liikkumavarasta neuvostossa. Suuri valiokunta kuulee vastuuministeriä.

Seuraavassa on tarkasteltu erikseen Suomen kannan valmistelua ja käsittelyä valtioneuvostossa sekä käsittelyä eduskunnassa.

5.1.1 Valmistelu ja käsittely valtioneuvostossa

Valtioneuvostolla tarkoitetaan tiukassa merkityksessä “pääministeriä ja tarvittavaa määrää muista ministereitä” (Suomen perustuslaki 60 §). Laajemmassa merkityksessä valtioneuvostoon voidaan katsoa kuuluvan pääministerin ja ministereiden lisäksi myös heidän johtamansa ministeriöt, pääministerin alainen valtioneuvoston kanslia sekä oikeuskansleri virastoineen. Tässä kirjoituksessa käsitteellä valtioneuvosto viitataan lähinnä tähän sanan laajempaan merkitykseen.

Kansallinen EU-lainsäädännön valmistelutyö lähtee useimmiten liikkeelle siitä, että komissio ilmoittaa aloittavansa säädösvalmistelun tietyllä alalla ja viestittää siitä kansallisiin vastuuministeriöihin eli kyseisen alan toimivaltaiseen ministeriöön (ks. Valtioneuvoston ohjesääntö 2003, 3 luku). Vastuuministeriö vastaa asioiden seurannasta, valmistelusta sekä Suomen kannan määrittelystä. (Joutsamo ym. 2000, 227) Vastuuministeriössä asiaa alkaa hoitaa se virkamies, jonka toimenkuvaan kyseinen asia kuuluu. Kuultuaan komission suunnitelmista hän neuvottelee ministeriönsä johdon kanssa vaadittavista toimenpiteistä ja Suomen alustavasta kannasta asiaan. Samaan aikaan vastuuvirkamies osallistuu myös Brysselissä asiasta käytäviin neuvotteluihin ja raportoi asian edistymisestä Suomessa. (Mattila 2000, 139-140)

Vastuuministeriö toimittaa asianomaiselle EU-jaostolle käsiteltäväksi määrittelemänsä Suomen kannat. Jaostojen päätehtävänä on koordinoida ja sovittaa ministeriöiden, keskusvirastojen ja muiden asianosaisten kantoja

käsiteltävän asian suhteen. EU-jaostot on koottu tiettyjen politiikkasektoreiden ympärille, esimerkkeinä ympäristö-, liikenne- ja teollisuusjaostot. Jaostot muodostuvat sille kuuluvan asia-alueen virkamiehistä sekä ministeriöistä että keskusvirastoista. Jaostot voivat kokoontua joko suppeassa tai laajassa kokoonpanossa. Suppeassa kokoonpanossa asioiden käsittelyyn osallistuvat vain virkamiehet, kun taas laaja kokoonpano sisältää myös eturyhmien nimittämiä edustajia esimerkiksi elinkeinoelämän, ammattiyhdistysliikkeen tai maatalouden järjestöistä. (Mattila 2000, 140-141)

Mikäli vastuuministeriön määrittelemään Suomen kantaan on huomautettavaa, eikä asiasta päästä yhteisymmärrykseen, voi jaosto pyytää asian käsittelyä EU-asioiden komiteassa. Siinä on edustus kaikista ministeriöistä, valtioneuvoston kansliasta, presidentin kansliasta, EU-sihteeristöstä, oikeuskanslerin virastosta, Suomen Pankista ja Ahvenanmaan maakuntahallituksesta. EU-asioiden komitean pääasiallinen tehtävä on toimia ministeriöiden välisenä koordinaatioelimenä, joka käsittelee kaikki asiat, joissa ei alemmilla tasoilla päästy yhteisymmärrykseen. (Mattila 2000, 141)

Mikäli asiassa ei päästä yhteisymmärrykseen EU-asioiden komiteassakaan, komiteassa edustettuna oleva taho voi pyytää asian käsittelyä EU-ministerivaliokunnassa. Muutoinkin ministerivaliokunnan käsiteltäväksi saatetaan merkitykseltään laajakantoiset EU-päätöksentekoa vaativat asiat. Jos vielä ministerivaliokunnassakaan ei päästä yhteisymmärrykseen tai jos kyseessä on hyvin merkittävä asia, annetaan Suomen kanta valtioneuvoston yleisistunnossa. Kun Suomen kanta on tullut lopulliseksi, toimivaltainen ministeriö laatii asiasta kirjallisen ohjeen. Ohje lähetetään EU-sihteeristöön, joka toimittaa sen edelleen EU-edustustolle. Ministeriöiden EU-

koordinaattorit tekevät yhteistyötä keskenään ja EU-sihteeristön kanssa kantojen yhteensovittamiseen liittyvissä kysymyksissä. (Tiitinen ym. 1997, 16)

EU-sihteeristö koordinoi koko Suomen EU-asioiden valmistelukoneiston toimintaa. EU-sihteeristö avustaa ministeriöitä ja muita valmisteluun osallistuvia tahoja kantojen yhteensovittamisessa. EU-sihteeristö huolehtii tiedonkulusta Suomessa sijaitsevien tahojen ja EU-edustuston välillä. Lisäksi EU-sihteeristö huolehtii päätöksentekoon liittyvästä asiakirjahallinnosta. (Mattila 2000, 142) Brysselissä Suomen valmistelua koordinoi Suomen pysyvä EU-edustusto, jota johtaa Suomen EU-suurlähettiläs. Edustuston tehtävänä on toimia yhteydenpitäjänä EU-tason elinten ja Suomessa sijaitsevan valmistelukoneiston välillä. EU-edustuston tärkeyttä korostaa erityisesti se, että EU-suurlähettiläs osallistuu Coreperin kokouksiin Suomen edustajana. (Joutsamo ym. 2000, 228)

5.1.2 Käsittely eduskunnassa

Eduskunnan osallistuminen EU-asioiden valmisteluun toteutuu valiokuntalaitoksen kautta. Eduskunnan EU-valiokuntana toimii suuri valiokunta lukuun ottamatta yhteisen ulko- ja turvallisuuspolitiikan asioita, joissa ulkoasiainvaliokunnalla on kyseinen tehtävä. Suuren valiokunnan kannanotot EU:n lainsäädäntöehdotuksiin valmistellaan asian sisällön mukaan toimivaltaisessa erikoisvaliokunnassa. (Kemppinen 2002, 30)

Valtioneuvoston on saatettava merkittävät lainsäädäntöhankkeet ja Suomen kanta niihin eduskunnan tietoon mahdollisimman pian komission annettua esityksen. Valtioneuvosto lähettää EU:sta saadun ehdotuksen eduskunnan puhemiehelle, joka lähettää sen suureen valiokuntaan sekä yhteen tai

useampaan erikoisvaliokuntaan. Erikoisvaliokunta käsittelee ehdotuksen samoin kuin kotimaisetkin hallituksen esitykset. Erikoisvaliokuntia ovat mm. laki-, valtiovarain-, liikenne- ja ympäristövaliokunnat. Erikoisvaliokunnassa kuullaan esittelevän ministeriön edustajia sekä muita asiantuntijoita ehdotuksesta. Erikoisvaliokunta antaa käsittelyn pohjalta lausunnon suurelle valiokunnalle. (Tiitinen ym. 1997, 17)

Suuri valiokunta perehtyy EU:n säädösehdotukseen, valtioneuvoston sitä koskevaan kirjelmään sekä erikoisvaliokunnan lausuntoon ja ilmaisee sitten eduskunnan kannanoton. Ennen kannanottoa suuri valiokunta kuulee tarvittaessa myös asianomaista ministeriä sekä häntä avustavia virkamiehiä ja asiantuntijoita. (Suomen eduskunta 2000, 60-61) Näiden pohjalta suuri valiokunta arvioi, onko valtioneuvoston esittämällä Suomen neuvottelukannalla eduskunnan tuki, vai olisiko kantaa muutettava. Kannanotto kirjataan valiokunnan pöytäkirjaan. Tarvittaessa suuri valiokunta voi antaa myös kirjallisen lausunnon valtioneuvostolle. Valtioneuvosto ottaa suuren valiokunnan ilmaiseman eduskunnan kannan huomioon neuvotellessaan ehdotuksesta EU:ssa. Suuren valiokunnan kannanotto ei kuitenkaan sido valtioneuvostoa oikeudellisesti. (Tiitinen ym. 1997, 17)

Ennen jokaista EU:n ministerineuvoston kokousta siihen osallistuvat valtioneuvoston jäsenet selvittävät suurelle valiokunnalle kokouksen asialistaa ja kantoja, jotka Suomen edustajien on tarkoitus esittää neuvostossa. Asianomainen ministeri antaa aina myös jälkikäteisselvityksen neuvoston kokouksen kulusta ja tuloksista suurelle valiokunnalle. Myös pääministeri antaa etu- ja jälkikäteen suurelle valiokunnalle tietoja Eurooppa-neuvoston kokouksista. Samoin toimitaan hallitusten välisissä konferensseissa esille tuleviin ja käsiteltyihin asioihin. Näin valtioneuvoston toimia valvotaan

parlamentaarisesti myös silloin, kun ratkaisevia päätöksiä tehdään EU:ssa. (Suomen eduskunta 2000, 61-62)

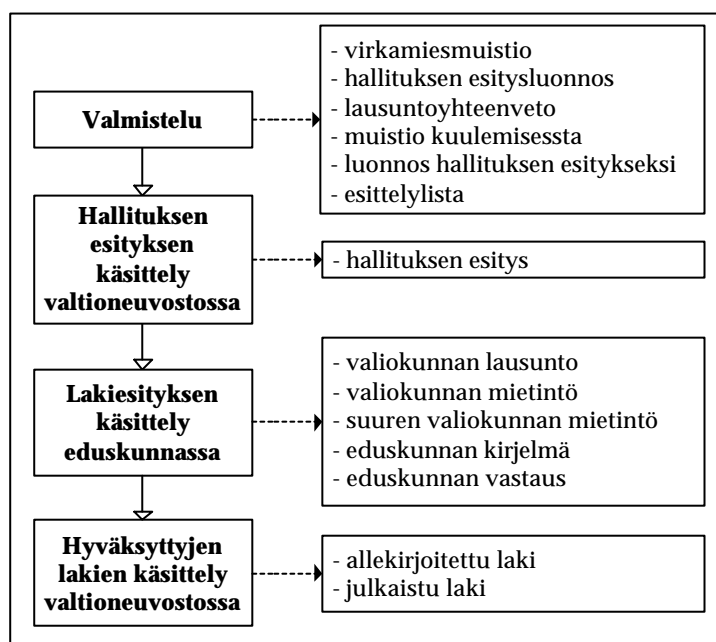
EU:n päätökset, ennen kaikkea direktiivit, vaativat usein uutta lainsäädäntöä Suomessa. Tällaisesta lainsäädännöstä päätetään normaalissa lainsäädäntöjärjestyksessä. Eräät unionin päätökset, esimerkiksi hallitustenvälisissä konferensseissa valmistellut unionin perussopimusten muutokset ja uusien maiden ottaminen unionin jäseniksi, edellyttävät eduskunnan hyväksymistä. Eduskunta soveltaa silloin Suomen perustuslakien säännöksiä. (Suomen eduskunta 2000, 63)

5.2 Suomen kansallinen lainsäädäntöprosessi

Kansallisesta näkökulmasta Suomen EU-jäsenyys merkitsee sitä, että kahta oikeusjärjestelmää, unionin ja kansallista, sovelletaan rinnakkain. EU:n asetukset tulevat välittömästi voimaan myös Suomessa, kun taas direktiivit edellyttävät, että Suomi toteuttaa siinä määritellyt tavoitteet kansallisesti. Tämä voi vaatia kansallisen lainsäädännön muuttamista. EU-lainsäädännön toimeenpanosta ja kotimaisista säädöksistä päätetään normaalissa lainsäädäntömenettelyssä.

Eduskunta edustaa Suomen kansaa ja käyttää lainsäädäntövaltaa. Hallitusvaltaa käyttävät tasavallan presidentti sekä valtioneuvosto, jonka jäsenten tulee nauttia eduskunnan luottamusta. Kukin ministeriö vastaa toimialallaan valtioneuvostolle kuuluvien asioiden valmistelusta ja hallinnon asianmukaisesta toiminnasta. (Suomen perustuslaki 2003, 3 §, 68 §) Tasavallan presidentti antaa ministeriöissä valmistellut hallituksen lakiesitykset eduskunnalle ja vahvistaa eduskunnan hyväksymät lait.

Kuviossa 7 on esitetty yksinkertaistetusti Suomen lainsäädäntöprosessin päävaiheet. Vaiheiden jakaminen on näkökulmakysymys, esimerkiksi Hallituksen esitysten laatimisohjeissa (2003, 80) lainsäädäntöprosessi on jaettu 16 osaan. Kuvion 7 vaiheita käytetään RASKE2-projektissa päätason mallina ja vaiheet esitetään myös tarkemmilla tasoilla omina kuvioinaan. Lainsäädäntöprosessin mallintaminen on vielä kesken RASKE2-projektissa, joten tuloksia julkaistaan aikaisintaan vuonna 2004.



KUVIO 7. Suomen lainsäädäntöprosessin päävaiheet (mukaiillen RASKE2 2003)

Seuraavassa on kuvattu lyhyesti lainsäädäntöprosessin eri vaiheita.

5.2.1 Valmistelu ja hallituksen esityksen käsittely valtioneuvostossa

Tässä alaluvussa on käytetty lähteenä Valtioneuvoston yhteisten asianhallintaprosessien kuvauksia ja suosituksia (2000, 83).

Lainsäädäntöprosessi voi saada alkunsa joko hallituksen esityksestä tai kansanedustajan lakialoitteesta. Käytännössä pääosa aloitteista tulee hallitukselta. Valmistelu tehdään yleensä siinä ministeriössä, jonka toimialaan asia kuuluu.

Hallituksen esityksen valmistelu voi saada alkunsa esimerkiksi aikaisemmasta valmistelusta, hallitusohjelmasta, eduskunnan vastaukseen sisältyvästä lausumasta, kansainvälisestä sopimuksesta, EY:n säädöksestä tai kehittämishankkeesta. Luonnokset hallituksen esityksiksi valmistellaan joko valtioneuvoston tai ministeriön asettamissa hankkeissa, komiteoissa tai työryhmissä, joissa on mukana eri hallinnonalojen, puolueiden ja muiden etutahojen edustajia. Valmistelu voidaan tehdä myös virkamiestyönä. Ministeriö vastaa oman hallinnonalansa lainsäädännöstä ja johtaa sen valmistelua.

Valmistelun jälkeen hallituksen esitystä käsitellään valtioneuvoston yleisistunnossa, jossa tehdään ratkaisuehdotus presidentille. Sen jälkeen presidentti tekee päätöksen hallituksen esityksen lähettamisestä eduskunnan käsittelyyn. Hallituksen esitykset sisältävät suppean kuvauksen pääasiallisesta sisällöstä, lain tavoitteet ja yleisperustelut, yksityiskohtaiset perustelut pykälittäin, voimaantuloa koskevan esityksen ja varsinaisen lakitekstin.

5.2.2 Lakiesityksen käsittely eduskunnassa

Tämän alaluvun koostamisessa on käytetty hyväksi Suomen eduskunta-julkaisua (2000, 39-40).

Hallituksen esitykseen tai kansanedustajan lakialoitteeseen sisältyvän lakiehdotuksen eduskuntakäsittely alkaa lähetekeskustelulla, joka käydään täysistunnossa. Tarkoituksena on antaa evästyksiä valiokuntatyöhön. Tässä vaiheessa ei tehdä lain sisältöä koskevia päätöksiä. Lähetekeskustelun jälkeen täysistunto päättää, mihin valiokuntaan hallituksen esitys lähetetään valmisteltavaksi. Valiokunta antaa hallituksen esityksestä mietintönsä, jossa esitetään valiokunnan kanta lakiehdotukseen. Valiokunta voi esittää lakiehdotuksen hyväksymistä muuttamattomana tai joissakin kohdin muutettuna; se voi myös ehdottaa lakiehdotuksen hylkäämistä. Valiokunnan enemmistön kanssa eri mieltä olevat edustajat voivat jättää mietintöön vastalauseensa. Lakiehdotukset käsitellään eduskunnan täysistunnossa kahdessa käsittelyssä.

Lakiehdotuksen ensimmäisessä käsittelyssä esitellään valiokunnan mietintö ja käydään siitä keskustelu (yleiskeskustelu) sekä päätetään lakiehdotuksen sisällöstä (yksityiskohtainen käsittely). Lakiehdotus voidaan ensimmäisen käsittelyn aikana lähettää suuren valiokunnan käsiteltäväksi. Toinen käsittely voi alkaa aikaisintaan kolmantena päivänä ensimmäisen käsittelyn päättymisestä, ja siinä päätetään ainoastaan lain hyväksymisestä tai hylkäämisestä. Lain sisältöön ei voi enää tällöin puuttua. Eduskunnan hyväksymä laki lähetetään valtioneuvostolle eduskunnan vastauksena. Kansanedustajan lakialoitteena vireille pantu eduskunnan hyväksymä laki lähetetään eduskunnan kirjelmänä.

5.2.3 Hyväksytyjen lakien käsittely valtioneuvostossa

Eduskunnan hyväksymä laki on esiteltävä valtioneuvostossa tasavallan presidentin vahvistettavaksi. Presidentti voi myös jättää lain vahvistamatta.

Tällöin laki palautetaan eduskuntaan uuteen käsittelyyn. Laki on tällöin hyväksyttävä asiasisällöltään muuttamattomana tai hylättävä. Jos eduskunta hyväksyy lain uudestaan, se tulee voimaan ilman vahvistusta. Tasavallan presidentin on kuitenkin allekirjoitettava laki. Jos lakia ei hyväksytä, se katsotaan rauenneeksi. (Suomen eduskunta 2000, 41)

Lain julkaisemisesta säädetään perustuslain 79 §:n 2 momentissa, jonka mukaan valtioneuvoston on julkaistava viipymättä presidentin ja asianomaisen ministerin allekirjoittama eli vahvistettu laki Suomen säädöskokoelmassa. Säädöskokoelma julkaistaan sekä painettuna että sähköisenä versiona internetissä. Laista tulee käydä ilmi, milloin se tulee voimaan. Lain osalta tästä säädetään perustuslain 79 §:n 3 momentissa. Lain voimaantuloajankohdan tulee siten ensisijaisesti käydä ilmi itse laista. Jollei lakia ole julkaistu viimeistään säädettynä voimaantuloajankohtana, se tulee voimaan julkaisemispäivänä.

5.3 Tietojärjestelmät

Monimutkaisen lainsäädäntöprosessin hallitsemiseksi on kehitetty erilaisia tietojärjestelmiä. Valtioneuvostolla ja eduskunnalla on omat järjestelmänsä. Juuri näitä järjestelmiä aiotaan integroida toisiinsa ontologian avulla.

Valtioneuvoston yhteisiä, lainsäädäntötyöhön liittyviä tietojärjestelmiä ovat PTJ (valtioneuvoston sähköinen päätöksentekojärjestelmä), EUTORI (EU-asiakirjojen ja -asioiden valmistelu- ja jakelujärjestelmä) ja HARE (valtioneuvoston hankerekisteri). Lähteenä seuraavissa kuvauksissa on käytetty Valtioneuvoston tietohallintostrategiaa vuosille 2003-2007 (2003, liite1)

Valtioneuvoston sähköinen päätöksentekojärjestelmä PTJ on tasavallan presidentin, valtioneuvoston yleisistunnon ja raha-asiainvaliokunnan päätöksentekoa ja siihen liittyvää esittelymenettelyä palveleva tietojärjestelmä. Sen avulla suoritetaan asiakirjojen jakelu, laaditaan pöytäkirjat sekä hoidetaan yhteydet eduskuntaan ja painatukseen.

EUTORI on valtioneuvoston kattava asianhallintajärjestelmä EU-asiakirjojen laatimista, käsittelyä ja arkistointia varten. EUTORI-järjestelmällä käsitellään periaatteessa kaikki erityyppiset EU-toimielinasiakirjat. EUTORI-järjestelmä sisältää sekä EU-toimielinten laatimat asiakirjat että kansalliset asiakirjat (esimerkiksi ohjeet, muistiot ja pöytäkirjat). EUTORI toimii myös EU-asioiden arkistona, jonka avulla virkamies voi laatia, kommentoida, jaella, vastaanottaa, hakea, rekisteröidä ja arkistoida asiakirjoja sekä hallinnoida kokouksia.

Valtioneuvoston hankerekisteri HARE on ministeriöiden yhteiskäyttöinen lainsäädäntö- ja valmisteluhankkeiden tietokanta. HAREsta on sekä Internet-versio <<http://www.hare.vn.fi>> että valtioneuvoston ja eduskunnan käytössä oleva intranet-HARE. HARE sisältää hankkeen perus-, tausta- ja tilannetiedot. Lisäksi se sisältää tiedot hankkeen asiakirjoista, julkaisuista, henkilöistä ja liittyvistä hankkeista.

Eduskunnan lainsäädäntötyöhön liittyviä tietojärjestelmiä ovat tekstiarkisto, Veps (valtiopäiväasioiden seurantajärjestelmä), Juva (julkaisuvaraston tietojärjestelmä) ja Tilta (tilaus-, laskutus- ja tilastointijärjestelmä). Seuraavien kuvauksien pohjana on käytetty Tiitisen, Salmisen ja Lyytikäisen julkaisua (1997, liite 3).

Eduskunnan tekstiarkisto on valtiopäivätyössä kertyvän aineiston elektroninen arkisto. Tekstiarkistoon tallennetaan ensin asiakirjojen käsikirjoitukset, jotka korvataan myöhemmin Editasta saatavilla lopullisilla versioilla. Asiakirjoihin liitetään tunnistetiedot, jotka liittävät ne vastaaviin Veps-dokumentteihin. Dokumenttien haku perustuu vapaaseen tekstihakuun, joka voidaan kohdistaa suoraan asiakirjan sisältöön ja/tai tunnistetietoihin. Valtiopäiväasioiden seurantajärjestelmä Veps sisältää viitetietoja valtiopäiväasioista ja niiden käsittelystä eduskunnassa. Vepsissä on myös valmiita yhteenvetoraportteja.

Julkaisuvaraston tietojärjestelmä Juva sisältää tiedot eduskunnan julkaisuvarastossa olevista asiakirjoista. Asiakirjoihin liittyviä määrätietoja saadaan esimerkiksi eri kieliversioiden mukaan ryhmiteltyinä. Juvan avulla paikannetaan julkaisut sekä seurataan julkaisuvaraston tilannetta. Valmiina raportteina voidaan tulostaa muun muassa asiakirjanimiketarrat, painatetilannemäärät ja karsintaluettelot. Tiltta on valtiopäiväasiakirjatuotannon tilaus-, laskutus- ja tilastointijärjestelmä, jonka avulla välitetään painotilaus ja painotyö eduskunnasta Editaan ja valtiopäiväasiakirjojen elektroninen versio Editasta eduskuntaan. Tiltan avulla voidaan seurata ja tilastoida vedosliikennettä. Järjestelmä sisältää myös laskun vastaanotto- ja tarkastustoiminnot.

5.4 Ontologia- ja sanastotyö suomalaisessa julkishallinnossa

Suomen julkishallinnossa on jo kehitetty ja kehitetään ontologioita ja sanastoja. Niiden merkitys on huomattu, kun heterogeenisten ja hajautettujen tietojärjestelmien aiheuttamien ongelmien ratkaisemista on tutkittu. Seuraavassa on kuvattu niistä muutamia, jotka ovat tämän tutkielman kannalta olennaisimpia.

Helsingin yliopistossa on meneillään tietojenkäsittelytieteen laitoksen Suomalaisen semanttisen webin ontologiat -tutkimushanke (Tekes 2003), jossa kehitetään pilottijärjestelmä kansallisesti merkittävien semanttisen webin ontologioiden hajautettua kehitystyötä varten. Tavoitteena on siirtyä tiedon indeksoinnissa ja haussa nykyisestä asiasanatekniikasta semanttisesti rikkaampaan ontologiatekniikkaan. Hankkeeseen osallistuu monia eri tahoja, kuten Kansalliskirjasto, Museovirasto ja TietoEnator.

Julkishallinnon XML-strategiassa (2003) esitetään XML-standardoinnin yleiseksi viitekehykseksi ebXML:ää (Electronic Business using Extensible Markup Language), jota rahoittavat UN/CEFACT ja OASIS. ebXML on avoin ja XML:ään pohjautuva kokoelma määrittelyjä, joiden tavoitteena on mahdollistaa eri järjestelmien yhteentoimivuus internetissä riippumatta yritysten paikasta tai koosta. ebXML sisältää muun muassa mallit eri osapuolille yhteisten tietojen määrittämiseksi ja nimeämiseksi yhteisellä tavalla sekä standardit toimintaprosessien kuvaamiseen ja järjestelmien välisiin liitäntöihin. ebXML:n avulla pyritään tukemaan erilaisten sanastojen ja tietorakenteiden määrittelyä ja käyttämistä erilaisissa yhteyksissä ja tilanteissa. ebXML:ää pidetään ainoana XML:n viitekehyksenä, joka samalla on yleinen ja joustava ja joka sopii hallintojen välisten järjestelmien liittämiseen.

Finnish Museums on-line (Hyvönen, Kettula, Raatikka, Saarela & Viljanen 2002) on Helsingin yliopiston tutkimusprojekti, jossa pyritään yhdistämään kahden museon tietokannat www:ssä. Museoiden tietokantojen sisällöt julkaistaan samalla tavalla kuten tavalliset WWW-sivut, mutta hakurobottia (engl. web crawler) käytetään hakemaan ja indeksoimaan sisältöä. Hakutoiminnot on sijoitettu palvelimelle. Tällaisesta tietokantojen yhdistämisestä on hyötyä siten, että eri osapuolet voivat julkaista ja hallita sisältöä riippumatta

palveluntarjoajasta. Museoiden tietokannat on tehty yhteentoimiviksi semanttisella tasolla. Syntaktisella tasolla käytössä on XML- ja RDF-kielet rakenteen ja metatiedon esittämiseksi. Semanttisella tasolla käytössä on RDF Schema -kieli.

YSA eli Yleinen suomalainen asiasanasto (2003) <<http://vesa.lib.helsinki.fi/ysa/>> on kaikkien alojen yleisimmän terminologian kattava indeksointi- ja hakusanasto. Helsingin yliopiston kirjaston ylläpitämästä sanastosta voi etsiä sanoille synonyymeja, rinnakkaistermejä ja samaan ryhmään kuuluvia asiasanoja. YSA on suomenkielinen ja kattaa noin 14 000 varsinaista asiasanaa ja noin 3 000 ohjaustermiä, jotka ohjaavat hyväksytyin asiasanan käyttöön. Aakkosellisen ja alanmukaisen hakemiston lisäksi YSA:an kuuluu maantieteellinen hakemisto. Eläimet, kasvit, sienet sekä urheilu- ja liikuntalajit on luetteloitu erillisessä liitetaulukossa. Allärs <<http://vesa.lib.helsinki.fi/allars/>> on ruotsinkielinen käänös YSAsta. YSA on tarkoitettu kaikenlaisen painetun ja elektronisen aineiston indeksointiin sekä aiheenmukaiseen tiedonhakuun.

Esimerkki YSA-sanaston hakutuloksesta:

Käytettävä asiasana:	asiakirjajulkisuus
Rinnakkaistermit:	julkisuus salassapitovelvollisuus tietosuoja
Kuuluu ryhmisiin:	[69] Oikeus. Lainsäädäntö [71] Hallinto. Organisaatiotutkimus. Julkinen hallinto
Ruotsinkielinen asiasana:	handlingsoffentlighet

Dublin Core (2002) on standardoitu metatietoformaatti, joka määrittelee joukon kenttä-arvo-pareja tallenteiden kuvailuun. Se ei esitä eri käsitteiden välisiä suhteita, ja on siksi ontologiana hyvin vaatimaton. Dublin Core pyrkii kattamaan vain tietyn osa-alueen, eikä siten kuvaa kaikkea olemassa olevaa.

Dublin Coressa on määritelty 15 elementtiä, joista kukin elementti voi esiintyä tarvittaessa useita kertoja yhden teoksen tiedoissa. Helsingin yliopiston kirjasto ylläpitää suomenkielistä Dublin Core -versiota ja -tallennusalustaa <http://www.lib.helsinki.fi/dublin_core/>.

Esimerkkejä Dublin Coren elementeistä (mukaillen Dublin Core 2002):

Nimi	Tunniste	Määritelmä
Aihe	subject	Tallenteen sisällön aihealueen kuvaus, yleensä asiasanoilla tai fraaseilla esitettyinä.
Identifikaatiotunnus	identifier	Tunnus, joka yksilöi tallenteen yksiselitteisesti.
Laji	type	Tallenteen kirjallisuuslaji, luonne tai genre. Tallenteen lajina voi olla esim. kotisivu, romaani, tekninen raportti tai muistio.

Dublin Coren avulla kuvattavat tallenteet voivat olla minkälaisia informaatioresursseja tahansa, myös painettuja julkaisuja. Dublin Core on kuitenkin tarkoitettu erityisesti verkkojulkaisujen, sähköisten asiakirjojen tai muiden elektronisten tallenteiden kuvailuun. Sähköisten asiakirjojen kuvailuun on tehty oma, Dublin Coreen perustuva formaatti, joka on ilmestynyt julkisen hallinnon tietohallinnon neuvottelukunnan (JUHTA) suosituksena. Tämä JHS 143 -suositus (2001) on julkishallinnon standardi metatiedoille. Dublin Core -standardin suomalaisen versioon ei ole tehty suuria muutoksia, mutta kuvailuformaatti on sovitettu terminologialtaan asiakirja-aineistolle sopivaksi. Lisäksi mukana on myös muutamia erityisesti asiakirjojen kuvailussa tarvittavia ominaisuuksia.

Tämän luvun aikana on tarkasteltu Suomen lainsäädäntöprosessia sekä EU-asioiden valmistelun että kansallisen lainsäädännön kannalta. Seuraavassa luvussa käsitellään lainsäädäntöprosessin järjestelmien yhteentoimivuutta tukevan ontologian kehittämistä

6 LAINSÄÄDÄNTÖPROSESSIN JÄRJESTELMIEN YHTEENTOIMIVUUTTA TUKEVAN ONTOLOGIAN KEHITTÄMINEN

Lainsäädäntötyöhön liittyviä ontologioita, kuten JHS 143 -suositus, on jo olemassa ja niitä myös kehitetään edelleen. Yhtenä mahdollisuutena tässä luvussa esitetään kuitenkin yhden keskeisen tietojärjestelmien integroimista tukevan ontologian kehittämistä. Tämä siksi, että yhden ontologian kehittäminen tiettyä tarkoitusta varten laajojen ja monikäyttöisten ontologioiden sijaan on helpompaa ja vaatii vähemmän resursseja. Kehittämistyössä voidaan silti hyödyntää aikaisempien ontologioiden rakentamisesta saatuja kokemuksia ja yhdistää niitä soveltuvin osin kehitettävään ontologiaan.

Luvun alussa pohditaan nykytilan ongelmia ja todetaan tavoitteet, joihin pyritään. Sen jälkeen tutustutaan tietojärjestelmien ontologiaperustaisiin integrointitapoihin. Sitten on vuorossa ehdotus etenemistavaksi, joka sisältää suosituksia liittyen ontologian kehittämisen eri vaiheisiin. Lopuksi arvioidaan tehtyä ehdotusta.

6.1 Nykytilan ongelmia ja tavoitetila

Lainsäädäntötyöhön liittyvien toimijoiden eli valtioneuvoston ja eduskunnan tietojärjestelmiä on toteutettu irrallisina, eri aikoina ja kulloinkin käytettävissä olevan tekniikan mukaisesti. Lainsäädäntötyön tietojärjestelmät ovat siten heterogeenisiä eli niiltä puuttuu yhtenäinen looginen rakenne (skeema, taulukot, tietokentät, nimet jne.) ja niiden tekninen perusta on erilainen. Lisäksi

käytössä olevat järjestelmät on rakennettu tukemaan tiettyä lainsäädäntöprosessin osaa eikä yhtenäisistä tietotarpeista lähtevän konseptin osana. Seurauksena itsenäisesti toteutetuista järjestelmistä on, että prosessit katkeilevat ja järjestelmien yhteensopimattomuus aiheuttaa tehottomuutta ja laadullisia ongelmia, kun asiakirjoja siirretään järjestelmästä ja esitystavasta toiseen.

Torniainen ja Valtari (2001, 11) toteavat, että useat tekijät edellyttävät eri toimijoilta yhä tiiviimpää yhteistyötä ja entistä tehokkaampia toimintatapoja. Muutospaineita aiheuttavat muun muassa julkisuuslainsäädäntö ja hyvä tiedonhallintatapa, jotka edellyttävät mahdollisimman helposti, taloudellisesti ja poikkihallinnollisesti käytettävissä olevia ajantasaisia tietovarantoja. Tarve toimielinten väliset rajat ylittävään yhteistyöhön ja verkostoitumiseen syntyy, kun käsiteltävät ja hoidettavat asiakokonaisuudet muuttuvat yhä monimutkaisemmiksi. Tämä tarkoittaa sitä, että hallinnollisesti itsenäisiltä toimijoilta vaaditaan toimintatapaa, jonka kautta ihmiset kykenevät toimimaan verkostoina ja jakamaan tehokkaasti ja joustavasti tietoa. Edellytyksenä tällaiselle toimintamallille ovat toimijoiden väliset yhteiset tietoprosessit ja niitä tukevat keskenään yhteensopivat tieto- ja viestintäjärjestelmät.

Tehokkaan ja joustavan tiedon käytön kannalta on tärkeää tietää, mistä järjestelmästä ja miten tieto on saatavissa. Eri tietojärjestelmät on kyettävä käyttäjän näkökulmasta sulauttamaan riittävän yhtenäisiksi tiedon löytämiseksi ja prosessoimiseksi. Tällöin järjestelmien kautta tuotetaan ja on saatavissa luotettavasti ja helposti löydettävästi tietoa työn tueksi. Lisäksi tieto kerätään ja tallennetaan vain kerran. Näin vältytään päällekkäiseltä työltä ja saavutetaan rationalisointihyötyjä. Tavoitteena on siis se, että lainsäädäntötyötä tuetaan yhteiskäyttöisillä, yhteentoimivilla, ajantasaisilla ja helposti käytettävissä

olevilla tietojärjestelmillä. Nykyiset irralliset järjestelmät on saatava tukemaan saumattomasti kokonaisia asiankäsittelyprosesseja. Valtioneuvoston ja eduskunnan välisten tietojärjestelmien yhteentoimivuuden lisäksi tavoitteena on yhteentoimivuus myös Euroopan unionin tasolla hyödynnettävien tietojärjestelmien kanssa. Tämä tavoite on kirjattu muun muassa Valtioneuvoston tietohallintostrategiaan (2003), jossa todetaan, että vuonna 2007 EU:n hallintoelinten valmistelujärjestelmät, kuten esimerkiksi Council Extranet, tulisi olla integroitu yhteentoimiviksi vastaavien Suomen lainsäädäntöön järjestelmien kanssa.

Lainsäädäntöön tietojärjestelmien integroimisessa on huomioitava monia eri asioita. Näitä ovat muun muassa Suomen kaksikielisyys, EU:n prosessit ja tietojärjestelmät sekä Valtion tietotekniikan rajapintasuosituksat. Ensin kuitenkin tarkastellaan sitä, miten tietojärjestelmät voidaan integroida ontologioiden avulla.

6.2 Tietojärjestelmien ontologiaperustainen integrointi

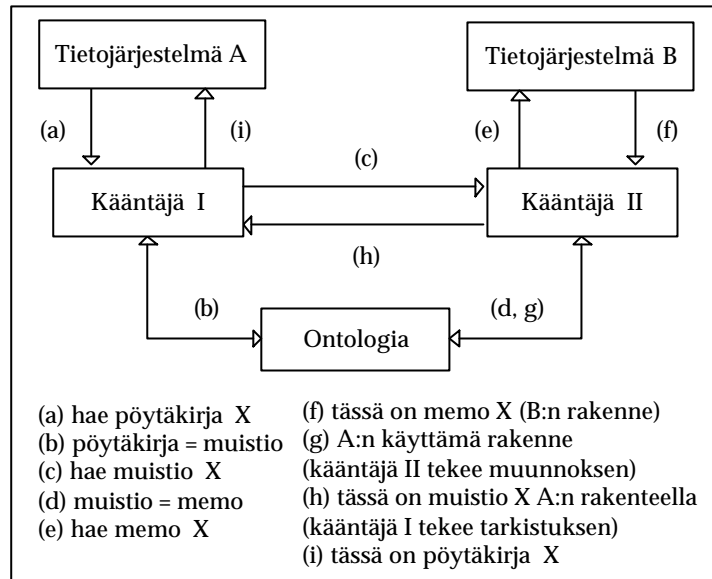
Yhteentoimivuus eli viestintä erillisten tietojärjestelmien kesken on saanut paljon huomiota viime aikoina, koska hajautettujen ja heterogeenisten tietojärjestelmäympäristöjen määrä sekä tarve vaihtaa informaatiota käyttäjien ja tietojärjestelmien kesken on lisääntynyt.

Tietojärjestelmien integroimisesta voidaan puhua eri tasoilla. Nämä tasot voidaan jakaa esimerkiksi tekniseen, syntaktiseen ja semanttiseen integrointiin. Teknisellä tasolla tietojärjestelmien integrointi toteutetaan verkkoteknologialla ja kommunikaatioprotokollilla, joiden avulla eri tietolähteet voivat

kommunikoida toistensa kanssa fyysisellä tasolla (esimerkiksi TCP/IP, HTTP, FTP). Fyysisen tietojen vaihtamisen lisäksi tietojärjestelmien täytyy sopia yleisestä syntaktista tiedon vaihtamiseksi. Syntaktisen tason integrointi onnistuu esimerkiksi XML:n avulla. Tämän jälkeen ongelmaksi jää eri tietojärjestelmien käsitteiden semanttinen liittäminen toisiinsa eli integrointi, sillä tietojärjestelmät eivät välttämättä käytä samanlaista käsitteistöä, vaikka ne käyttäisivätkin XML:ää tiedonvaihtoformaattinaan.

Hovyn (1998, 535) mukaan semanttisen tason integrointi voidaan toteuttaa liittämällä tietojärjestelmien käsitteet toisiinsa suoraan käyttäen vastaavuustaulukoita. Tämä tapa on kuitenkin hyvin työlästä. Semanttisen tason integrointi voidaan toteuttaa myös epäsuorasti käyttäen ontologiaa, jota voidaan käyttää tunnistamaan ja yhdistämään semanttisesti toisiaan vastaavat informaatiokäsitteet. Ontologioita voidaan käyttää tukemaan käännöksiä eri kielten ja esitystapojen välillä, jolloin tarvitaan vain käännöksiä alkuperäisestä kielestä ontologiakielelle.

Kuviossa 8 on havainnollistettu ontologian käyttämistä välittäjäkielenä. Esimerkissä XML-perustaisen tietojärjestelmä A:n haettavan dokumentin dokumenttityypistä käyttämä termi ”pöytäkirja” käännetään ohjelmiston avulla termiksi ”memo”, jota myös XML-perustainen tietojärjestelmä B käyttää samasta dokumenttityypistä. Tämä käänнос tapahtuu dokumenttityyppien ontologian kautta, jonka termi samalle käsitteelle on ”muistio”. Käänनों avulla voidaan hakea tietojärjestelmä B:stä fyysisesti löytyvää dokumenttia. Ontologia sisältää termien lisäksi tiedot myös dokumenttityyppien rakenneosista. Näin kääntäjä voi, dokumenttia lähetettäessä, tehdä koko XML-dokumentille XSLT-muunnoksen (Clark 1999), jolloin rakenneosien nimet muuttuvat tietojärjestelmä A:n ymmärtämiksi.



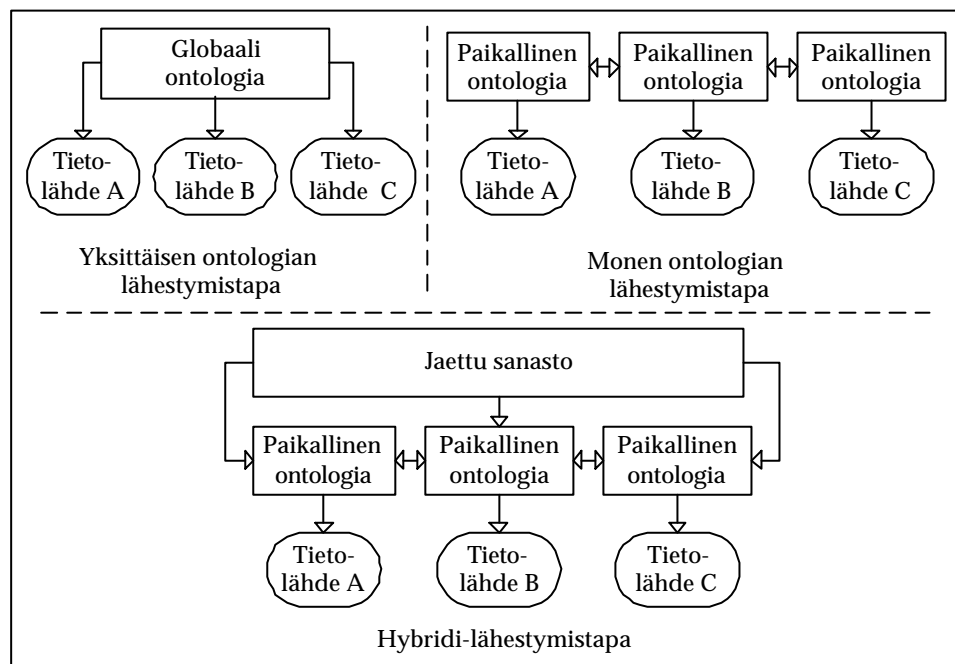
KUVIO 8. Esimerkki ontologiasta välittäjäkielenä

Lainsäädäntötyöhön liittyvien tietojärjestelmien yhteentoimivuuden ongelmana on erityisesti eri tietojärjestelmissä käytettävien tunnisteiden hajanaisuus, joka aiheutuu yhtenäisen tunnistekäytännön puuttumisesta. Ennen tietojärjestelmien integroimista on sovittava yhteisestä tunnistekäytännöstä, sillä ilman sitä tietojärjestelmät eivät voi ymmärtää toisten tietojärjestelmien dokumenteista käyttämiä tunnisteita ja niiden muotoja. Kuvan 9 esimerkki perustuu kuitenkin ajatukseen siitä, että tunnistekäytännöt olisi jo yhdenmukaistettu yksinkertaisuuden vuoksi.

Ontologian käyttäminen välittäjäkielenä vaatii vähemmän työtä kuin vastaavuustaulukoiden rakentaminen, sillä käännöksiä kielestä toiseen ei tarvitse tehdä suoraan tietojärjestelmien välillä. Ontologian käyttämisestä ei välttämättä saada vielä kahden tietojärjestelmän tapauksessa täyttä hyötyä, mutta useamman tietojärjestelmän tapauksessa hyödyt ovat selvät. Esimerkiksi jo neljän tietojärjestelmän yhdistämisessä jouduttaisiin tekemään termien

vastaavuustaulukoita yhteensä kuusi kappaletta. Käytettäessä ontologiaa voidaan tehdä käännökset vain tietojärjestelmän käyttämän kielen ja ontologiakielen välillä.

Ontologioita voidaan käyttää tietojärjestelmien integroimiseen kolmen eri lähestymistavan kautta, jotka ovat yksittäisen ontologian (engl. single ontology) ja monen ontologian (engl. multiple ontologies) lähestymistavat sekä hybridi-lähestymistapa (engl. hybrid) (Wache, Vögele, Visser, Stuckenschmidt, Schuster, Neumann & Hübner 2001, 109). Nämä lähestymistavat on esitelty kuviossa 9.



KUVIO 9. Ontologiaperustainen integrointi (mukaillen Wache ym. 2001, 109)

Yksittäisen ontologian lähestymistavassa yksi globaali ontologia tarjoaa jaetun sanaston semantiikan määrittelemiseen. Kaikki tietolähteet on liitetty tähän ontologiaan. Yksittäisen ontologian lähestymistapaa voidaan soveltaa

integrointiongelmiin, joissa kaikki integroitavaksi aiotut tietolähteet sisältävät lähes saman näkökulman aihealueesta. Muutoin minimaalisen ontologisen sitoutumisen (katso kohta 3.1) saavuttaminen tulee liian vaikeaksi. Tietolähteiden muutokset ovat ongelmakohta, sillä muutoksia voidaan joutua tekemään myös globaaliin ontologiaan ja liityntöihin muihin tietolähteisiin. (Wache ym. 2001, 109)

Monen ontologian lähestymistavassa jokainen tietolähde kuvataan omalla ontologiallaan. Tässä lähestymistavassa ei tarvita yhteistä ja minimaalista ontologista sitoutumista yhteen globaaliin ontologiaan. Tämä lähestymistapa myös yksinkertaistaa muutosten tekemistä yhteen tietolähteeseen sekä tietolähteiden lisäämistä tai poistamista. Yhteisen sanaston puuttuminen vaikeuttaa kuitenkin huomattavasti eri ontologioiden rinnastamista. Siksi tarvitsee määritellä ontologioiden välisiä liittymiä, jotka muun muassa tunnistavat semanttisesti vastaavat termit eri ontologioista. Näitä liittymiä on kuitenkin vaikea määritellä. (Wache ym. 2001, 109-110)

Hybridi-lähestymistavassa kuvataan jokaisen tietolähteen semantiikka omalla ontologialla. Tämän lisäksi rakennetaan globaali jaettu sanasto, joka sisältää aihealueen perustermit, joihin tietolähteiden ontologiat perustuvat. Hybridi-lähestymistavassa voidaan helposti lisätä uusia tietolähteitä. Aikaisemmin luotuja ontologioita ei kuitenkaan voida käyttää lainkaan hyväksi, koska tietolähteiden ontologioiden pitää viitata jaettuun sanastoon. (Wache ym. 2001, 110) Taulukossa 3 on listattu edellä esiteltyjen ontologiaperustaisten integroinnin lähestymistapojen höytyjä ja haittoja.

TAULUKKO 3. Integrointitapojen hyödyt ja haitat (mukaiillen Wache ym. 2001)

	Yksittäinen ontologia	Monta ontologiaa	Hybridi
Toteutus	helppoa	vaikeaa	keskivaikeaa
Semanttinen heterogeenisuus	samanlainen näkemys aihealueesta	heterogeeniset näkemykset	heterogeeniset näkemykset
Tietolähteiden lisääys ja poisto	globaalien ontologian mukauttaminen	uuden ontologian määrittäminen ja yhteydet muihin ontologioihin	uuden ontologian määrittäminen
Useiden ontologioiden rinnastaminen	(sisältää vain yhden ontologian)	vaikeaa, koska puuttuu yhteinen sanasto	yksinkertaista, koska ontologiat käyttävät yhteistä sanastoa

Seuraavassa esitetään ehdotus tavaksi, jolla lainsäädäntöprosessin tietojärjestelmien yhteentoimivuutta tukevan ontologian kehittämisessä voidaan edetä.

6.3 Ehdotus etenemistavaksi

Ontologioiden kehittämisessä pitää edetä pienin askelin, sillä kehittämistyö on haasteellista ja vaikeaa. Asia on oletettavasti näin varsinkin lainsäädäntötyön monimutkaisella ja laajalla alueella, jossa käytetään hyvin monenlaisia käsitteitä tällä hetkellä. Ontologian kehittämiseen vaikuttavat myös arvolataukset, jolloin voi olla vaikeaa päästä yhteisymmärrykseen ontologian määrittelyistä. Tällöin on keskityttävä tärkeimpien ontologian osasten etsimiseen ja pyrittävä pääsemään pääsy yksimielisyyteen suppeasta käsitteistöstä.

Ontologioiden kehittämiselle olennaisia päävaiheita ovat suunnittelu-, määrittely- ja toteutusvaiheet. Ne esiintyvätkin lähes jokaisessa kohdassa 3.2 esitellyssä kehittämismenetelmässä muodossa tai toisessa. Seuraavassa on

käyty läpi ehdotus ontologioiden kehittämisen etenemistavaksi näiden vaiheiden pohjalta.

6.3.1 Suunnitteluvaihe

Kaikkien kehittämishankkeiden läpivienti menestyksekkäästi on epävarmaa, jos niitä ei ole suunniteltu huolellisesti. Suunnittelu on vaikeaa, jos toiminnan päämäärä ei ole tiedossa. Ontologian kehittäjillä tulisikin olla selvä käsitys ontologian tarkoituksesta ja laajuudesta eli siitä, miksi ontologia halutaan luoda, ketkä ovat sen käyttäjiä sekä mihin sitä käytetään ja miten. Ontologian kehittäjien pitäisi myös päättää siitä, kuinka formaali ontologian pitää olla (kohdassa 2.3 on listattu neljä eri astetta). Tämä määräytyy suurelta osin ontologian tarkoituksen ja käyttäjien perusteella. Jos ontologian päätarkoituksena on esimerkiksi toimia jaettuna sanastona, joka helpottaa ihmisten välistä viestintää, voi vapaamuotoinen esitystapa riittää.

Ontologian tarkoituksen ja laajuuden selvittämistä varten tarvitaan aihealueen mallintamista ja muuta aineiston keruuta. Aihealueen mallintamista on tehty RASKE- ja RASKE2-projekteissa <<http://www.it.jyu.fi/raske/julkaisut.html>>. RASKE-projektissa dokumenttien analysointiin kehitettyjen mallintamismenetelmien (Salminen 2003) avulla on pystytty selvittämään Suomen lainsäädäntöprosessia ja sen vaiheita, vaiheisiin liittyviä dokumentteja ja tietojärjestelmiä sekä vaiheisiin osallistuvia organisaatioita ja henkilöitä. Mallintamistyö on tärkeää, jotta voidaan ymmärtää kattavasti lainsäädäntöprosessia, johon tuotettava ontologia on tarkoitus liittää. Kaikkea prosesseihin liittyvää ei kuitenkaan ole välttämättä tarkoituksenmukaista formalisoida ontologiassa. Esimerkiksi hallituksen esityksen valmisteluprosessiin sisältyvä kuulemisvaihe on hyvin laaja sisältäen

monenlaisia viestejä ja erityyppistä palautetta useilta eri tahoilta eri tilanteissa. Siten on tarpeen selvittää muun muassa se, missä lainsäädäntöprosessin vaiheessa ontologiaa aletaan hyödyntää.

Ontologian suunnittelussa ja kehittämisessä tarvittavan aineiston lähteitä ovat muun muassa aihealueen asiantuntijat, käsikirjat, raportit, kuviot, taulukot ja muut ontologiat. Aineiston keruuta tehdään periaatteessa koko kehittämisprosessin ajan, mutta suurin osa siitä tehdään samanaikaisesti suunnittelun ja määrittelyn kanssa. Aineiston hyödyntämiseen voi käyttää erilaisia tekniikoita, kuten aivoriihitoimintaa, haastatteluita ja tekstianalyysejä. Esimerkiksi aihealueen asiantuntijoiden haastatteluiden kautta voi kerätä ontologian kriteerejä ja alustavia käsitteitä, joista ontologian suunnittelu voidaan aloittaa. Haastatteluilla voi olla myös muuta hyötyä, sillä ne mahdollisesti lisäävät käyttäjien sitoutumista kehittämisprojektiin ja tuotettavaan ontologiaan. Tekstianalyysiä voi käyttää muun muassa avuksi käsitteiden, attribuuttien ja suhteiden kartoittamisessa kirjoista ja raporteista. Esimerkiksi Lainlaatijan oppaasta (1996 tai uudempi) voisi saada kartoitettua hyödyllistä aineistoa, sillä se sisältää tarkkoja tietoja lakeja laadittaessa noudatettavista periaatteista. Aineistoa kerätessä tulee mainita lähteet tarkasti ja kertoa hankintaprosessista sekä siinä käytetyistä tekniikoista.

Suomen lainsäädäntötyöhön osallistuu useita toimijoita (ks. luku 5), joten verkostoituminen ja toimivien kommunikaatioyhteyksien luominen suunnitteluvaiheessa näiden toimijoiden kesken ovat perusedellytyksiä ontologian kehittämistyön onnistumiselle. Lainsäädäntöprosessin kannalta myös Euroopan unioni on tärkeä yhteistyökumppani, koska suomalaisen lainsäädäntötyön tietojärjestelmiä tullaan integroimaan jossakin vaiheessa myös EU-tason tietojärjestelmiin. Yhteistyöverkoston laajentaminen kattamaan myös

tutkijayhteisön ja yritysmaailman edustajia on harkitseminen arvoinen asia tiedon ja osaamisen kartuttamisen vuoksi. Yliopistojen ja tutkimusinstituuttien tutkijoilla on kaikkein ajantasaisin tietämys ontologioista, koska he tekevät omia tutkimuksiaan ja seuraavat tiiviisti muiden suorittamia tutkimuksia. Esimerkiksi Helsingin ja Jyväskylän yliopistoissa tehdään ontologioihin liittyvää tutkimusta. Yritykset, kuten esimerkiksi TietoEnator, taas hallitsevat toteutuksen kannalta tärkeää käytännön osaamista.

Tarkoituksen ja laajuuden lisäksi suunnitteluvaiheessa pitää valita ontologioiden kehittämisen lähestymistapa. Holsapplen ja Joshin (2002) kohdassa 3.1. esittämistä lähestymistavoista oletettavasti soveltuvin on yhteistoiminnallinen lähestymistapa, sillä siinä pyritään kehittämään ontologia yhteistyönä eri osapuolten kesken. Jo aikaisemmin on todettu, että Suomen lainsäädäntötyöhön liittyy useita toimijoita, jolloin ontologian kehittäminen on pakostakin tehtävä yhteistyössä niiden kesken. Yhteistyössä tapahtuva kehittäminen lisää ontologista sitoutumista, joka on tärkeää ontologian hyödyntämisen kannalta. Ontologiaan sitoutuneet henkilöt käyttävät sitä, muut taas välttämättä eivät, jolloin ontologiasta saatava hyöty jää pieneksi ja investoinnit menevät osittain hukkaan.

Ontologioiden kehittämismenetelmät (ks. kohta 3.2) antavat ohjeita siitä, mitkä ovat kehittämisen päävaiheet ja mikä on niiden järjestys. Lisäksi ne usein kertovat, millaisilla tekniikoilla kyseiset vaiheet toteutetaan ja mitkä ovat niiden tuotoksia. Menetelmät auttavatkin toteuttamaan hallitusti ontologian kehittämisprosessin. Menetelmää valittaessa on kiinnitettävä huomiota kieli- ja aihealuesidonnaisuuksiin, jotka näkyvät muun muassa liitteestä 1. Myös menetelmän vaiheet ja toiminnot vaikuttavat valintaan, sillä joissakin menetelmissä tuetaan ontologian luomista tyhjästä, toisissa taas suositaan

aikaisempien ontologioiden hyödyntämistä. Kehittämismenetelmän valinnassa voidaan myös käyttää muita kriteerejä, kuten kypsyy, yksityiskohtaisuus ja helppo opittavuus.

Suunnitteluvaiheessa on päätettävä myös yksikielisestä tai monikielisestä toteutuksesta ja integroinnin lähestymistavasta. Lainsäädäntötyötä varten kehitettävä ontologia on oletettavasti toteutettava ainakin kaksikielisenä, sillä molempia virallisia kieliä, suomea ja ruotsia, käytetään lainsäädäntöprosessissa. Lisäksi ontologian toteuttamisesta myös englanninkielisenä voisi olla etua, sillä monikielisyys mahdollistaa kieliriippumattoman tiedonhaun ja ontologian hyödyntämisen www-palveluissa. Englanninkieli helpottaa myös liittymien rakentamista EU-tason järjestelmiin. Ontologiaperustaisen integroinnin lähestymistapoja on esitelty kohdassa 6.2. Lähestymistapa, jossa yksittäinen ontologia toimii välittäjäkielenä, vaikuttaa soveltuvimmalta. Kyseinen lähestymistapa on myös muihin verrattuna suhteellisen helppo toteuttaa ja vaatii siten vähemmän resursseja.

Suunnitteluvaiheen päämääränä on tuottaa ontologiasuunnitelma kirjoitettuna luonnollisella kielellä. Suunnitteludokumentin pitäisi sisältää ainakin tiedot aihealueesta, kehittäjistä, ontologian tarkoituksesta, formaalisuuden asteesta, laajuudesta (eli olennaisista käsitteistä) ja tietolähteistä. Seuraavassa on kuvattu esimerkki suunnitteludokumentin mahdollisesta sisällöstä:

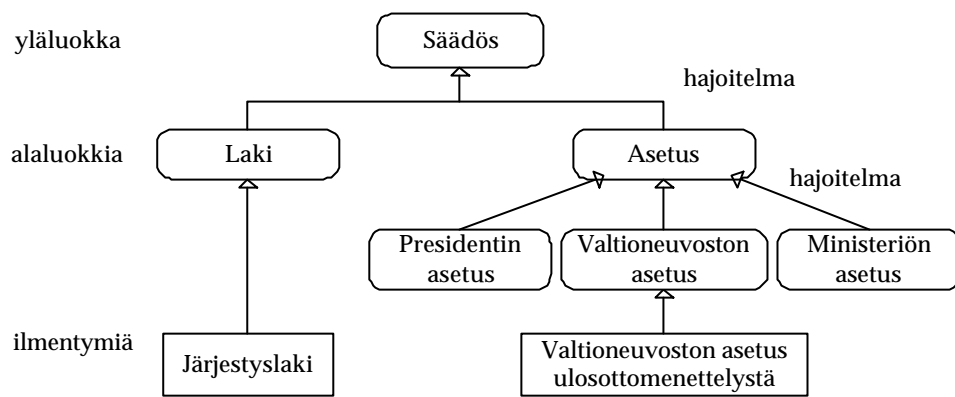
Aihealue	Suomen lainsäädäntötyö
Kehittäjät	Eduskunta ja valtioneuvosto
Ontologian tarkoitus	Lainsäädäntöprosessiin liittyvien tietojärjestelmien yhteentoimivuuden saavuttaminen
Formaalisuuden aste	Puoliformaali
Laajuus	Säädöskokoelma, säädös, laki, asetus, jne.
Lähteet	Lainlaatijan opas, jne.

6.3.2 Määrittelyvaihe

Aihealueen mallintamisen ja aineiston keruun jälkeen kehittäjillä on paljon materiaalia, joka pitää analysoida ja organisoida. Määrittelyvaiheen tarkoituksena on tunnistaa aihealueen keskeisimmät käsitteet ja niiden väliset suhteet sekä tuottaa kaikille ontologian rakenneosille tarkat ja yksiselitteiset tekstimäärittelyt. Lisäksi tarkoituksena on saavuttaa yksimielisyys kyseisistä määrittelyistä. Hankittu aineisto organisoidaan käyttämällä esitystapoja, jotka ovat riippumattomia ontologian toteuttamiseen käytettävästä esityskielestä ja työkalusta. Esitystapoina voi käyttää esimerkiksi taulukoita ja kuvioita, joita kehittäjien lisäksi myös aihealueen asiantuntijat pystyvät vaivatta ymmärtämään. Aineiston organisoinnissa ja määrittelyjen luomisessa tulee ottaa huomioon ja soveltaa erilaisia suunnittelukriteerejä, joita on lyhyesti esitelty kohdassa 3.1.

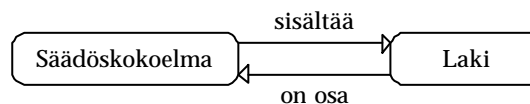
Käsitteiden kokoamiseen käytettäviä strategioita on esitelty kohdassa 3.2 Uscholdin ja Kingin (1995) kehittämismenetelmän yhteydessä. Ylhäältä alaspäin -strategian ongelmana ovat mahdolliset epätarkat korkean tason luokat. Alhaalta ylöspäin -strategian ongelmana taas saattaa päinvastoin olla liiallinen yksityiskohtaisuus. Kaikki yksityiskohtaisesti määritellyt käsitteet eivät loppujen lopuksi ole välttämättä tärkeitä lopullisen ontologian kannalta. Nämä ongelmat vaikeuttavat kehittämistyötä. Keskeltä molempiin suuntiin lähtevä strategia on suositeltavin, koska siinä määritellään ensin tärkeimmät keskitason peruskäsitteet. Kun näistä käsitteistä ja niiden määritelmistä on päästy yhteisymmärrykseen, voidaan vasta tarvittaessa siirtyä niiden erikoistamiseen tai yleistämiseen. Siten säästytään useasti turhalta työltä, jota muiden strategioiden käyttämiseen liittyy.

Määritellyistä käsitteistä, ilmentymistä ja attribuuteista luodaan sanasto, joka sisältää kaiken tarpeellisen tietämyksen aihealueesta. Sanaston valmistuttua käsitteitä ja muita ontologian rakennneosia ryhmitellään hierarkioiksi. Jokainen käsiteryhmä sisältää käsitteitä, jotka liittyvät läheisesti toisiinsa ja eroavat muista ryhmistä. Hierarkioista piirretään kuvioita, esimerkiksi kuviossa 10 on pyritty havainnollistamaan osaa mahdollisesta lainsäädäntötyön ontologian käsitehierarkiasta.



KUVIO 10. Osa mahdollista lainsäädäntötyön ontologian käsitehierarkiaa

Suhteista piirretään kuvioita, jotka sisältävät kaksi käsitettä ja niiden väliset suhteet. Kuviossa 11 on havainnollistettu käsitteiden välisten binääristen suhteiden esittämistä graafisesti.



KUVIO 11. Esimerkki binäärisen suhteen kuvaamisesta

Muun muassa käsitteet, ilmentymät, ilmentymäattribuutit, luokka-attribuutit, suhteet ja aksioomat kuvataan taulukoittain (ks. taulukko 4). Taulukot ja kuviot toimivat asiantuntijoiden ja kehittäjien välisen vuoropuhelun avustajina. Kehittäjät muokkaavat ontologiaa asiantuntijoilta saamansa palautteen avulla. Tehdyistä muutoksista on aiheellista pitää kirjaa, jotta muutokset voidaan tarvittaessa jäljittää myöhemmin. Ontologian rakenneosien kokoamisen, analysoimisen, määrittelemisen ja palautteen keräämisen vaiheita voidaan toistaa niin kauan kunnes kaikkien osapuolten kesken saavutetaan yhteisymmärrys siitä, että ontologia on tarpeeksi selkeä, johdonmukainen ja kattava.

TAULUKKO 4. Esimerkki käsitetaulukosta

Nimi	Kuvaus	Ilmentymät	Luokka-attribuutit	Ilmentymä-attribuutit	Suhteet
Säädös-kokoelma	Sisältää kaikki säädetyt lait.	--	--	- julkaisupäivä	- muodostuu
Säädös	Kattaa lait ja asetukset.	--	--	- voimaantulo - kohdealue	- on osa - jakautuu
Asetus	Presidentin, valtioneuvoston tai ministeriön antama säädös.	Asetus ulosottomenettelystä jne.	--	- säätäjä	- kuuluu

Määrittelyvaiheessa voi harkita aikaisempien ontologioiden hyväksikäyttöä, sillä ontologiat pyritään rakentamaan uudelleenkäytettäviksi tietämyksen jakamisen mahdollistamiseksi. Aikaisemmin luotujen ontologioiden määritelmien käyttäminen voi mahdollisesti nopeuttaa oman ontologian kehittämistä, sillä siinä ei tarvitse aloittaa työtä täysin tyhjästä. Kohdassa 3.4 sekä kohdissa 4.4 ja 5.4 on esitelty lakiontologioita ja EU-tason ja Suomen julkishallinnon ontologia- ja sanastotyötä juuri siksi, että niistä voisi olla apua kehitettäessä ontologiaa Suomen lainsäädäntötyötä varten. EU:n

lainsäädäntötyössä käytettävien järjestelmien (ks. kohta 4.3) toteutukseen voisi myös tutustua, sillä niissä on voitu hyödyntää ontologioita, jotka voisivat olla osittain käyttökelpoisia luotaessa ontologiaa Suomen lainsäädäntötyötä varten. Jos soveltuvia määritelmiä löytyy, mutta ne on luotu eri esityskielellä kuin omaan ontologiaan valittu kieli, voi mahdollisesti löytyä kääntäjäohjelmia, jotka muuntavat määrittelyt kielestä toiseen.

Ontologioiden tai niiden osien yhdistäminen omaan ontologiaan ei välttämättä ole kovin yksinkertaista. Siksi ontologisesti heterogeenisen tietämyksen integroimiseen on kehitetty omia menetelmiä (ks. esim. Farquhar, Fikes, Pratt & Rice 1995). Yhdistäminen tulisi myös dokumentoida ylläpidon ja jatkokehityksen mahdollistamiseksi. Dokumentoinnin voisi hoitaa vaikkapa taulukoiden avulla, joista kävisi ilmi lainatun käsitteen tiedot ja sitä vastaavan käsitteen tiedot omasta ontologiasta.

6.3.3 Toteutusvaihe

Toteutusvaiheessa koodataan edellisessä vaiheessa luotu ontologia tietyllä esityskielellä, jotta ontologiasta tulee myös tietokoneiden ymmärtämä. Toteutusvaiheen lisäksi tämän kohdan lopussa käsitellään lyhyesti myös arviointia, dokumentointia, ylläpitoa ja jatkokehitystä.

Yksinkertainen esimerkki koodauksen tuloksesta OWL-kielellä:

<pre><owl:Class rdf:ID="laki"> <rdfs:subClassOf rdf:resource="#säädös"/> <rdfs:label xml:lang="fi">laki</rdfs:label> <rdfs:label xml:lang="se">lag</rdfs:label> <rdfs:label xml:lang="en">law</rdfs:label> </owl:Class></pre>	<pre><!-- määritellään laki-niminen luokka --> <!-- laki on säädös-luokan alaluokka --> <!-- luokan nimi suomeksi --> <!-- luokan nimi ruotsiksi --> <!-- luokan nimi englanniksi --></pre>
---	---

Kun aiotaan luoda uusi ontologia, nousee esille useita kysymyksiä liittyen kehittämisprosessissa käytettävän esityskielen ja työkalun valintaan. Kielen valintaan vaikuttavat muun muassa ilmaisuvoima ja päättelymekanismit. Tehtäessä valintaa perinteisten ja web-perustaisten ontologiakielten välillä pitää ensin päättää, mitkä ovat tarpeet niiden suhteen. Esittäminen ja päättely ontologioiden perusrakenteilla, kuten käsitteillä, taksonomioilla ja binäärisillä suhteilla, ei ole yleensä tarpeeksi, jos halutaan tehdä ontologialla monimutkaisia päättelyitä. Tällöin valinta tulee tehdä perinteisten esityskielten joukosta, jotka mahdollistavat raskaampien ontologioiden luomisen esimerkiksi tietämuskantojen käyttöön.

Web-perustainen ontologiakieli on mahdollisesti riittävä lainsäädäntötyötä varten luotavan ontologian esittämiseen, sillä ontologian tarkoituksena on toimia välittäjäkielenä tietojärjestelmien välillä. Web-perustaiset ontologiakielet pohjautuvat XML-kieleen, joten tällöin noudatetaan automaattisesti Valtion tietotekniikan rajapintasuosituksia (2001), joissa painotetaan XML:n käyttöä julkishallinnossa. Kaikki web-perustaiset ontologiakielet eivät kuitenkaan mahdollista rakenneosien esittämistä ja päättelyä samalla tavalla. Ontologian rakenneosiin (ks. kohta 2.2) tutustumalla voi päätellä omat esitystarpeet ja sen perusteella valita sopivan kielen (ks. kohta 3.3.3).

Ilmaisuvoiman perusteella nykyiset web-perustaiset ontologiakielet voi asettaa heikommasta vahvempaan seuraavasti: XOL, RDF(S), SHOE, OML, OIL ja DAML+OIL. Päättelytuki on saatavilla RDF(S)-, SHOE-, OIL- ja DAML+OIL-kielille, mutta ei XOL- ja OML-kielille. (Cómez-Pérez & Corcho 2002) Ontologiakielen valinnassa voidaan käyttää myös erilaisia kriteerejä, kuten selkeys, yhteensopivuus standardien kanssa, siirrettävyys ja kehittämismenetelmien tuki. Kielen valinnassa voi käyttää apuna myös

olemassa olevien työkalujen tukea ja esitysformalismien tuttuutta. Esimerkiksi OIL- ja DAML+OIL-kieliä tukevat monet ontologioiden kehittämistyökalut, mutta joitakin kieliä ei tue yhtään työkalua. XOL ja SHOE käyttävät esitysformalisminaan kehyksiä, RDF(S) semanttisia verkkoja, OML käsitekarttoja sekä OIL ja DAML+OIL kuvauslogiikkaa (Cómez-Pérez & Corcho 2002).

Tietyn standardin valintaan liittyy aina riskejä, sillä esimerkiksi investoinnit voivat mennä hukkaan, jos valitun standardin kehitys lopetetaan ja jokin toinen standardi nousee vallitsevaan asemaan. W3C on arvostettu organisaatio, joka tuottaa laajaan käyttöön leviäviä standardeja. Siksi on suositeltavaa noudattaa sen ohjeistuksia riskien minimoimiseksi. W3C:n ontologioiden esittämiseen kehittämä OWL-kieli ei ole vielä ihan W3C-suosituksen asteella. Ontologian kehittämiseen voi kuitenkin mennä aikaa useampia vuosiakin, joten OWL voi olla esityskieltä valittaessa jo W3C:n suositus ontologiakieleksi. Toisaalta tulevaisuutta on vaikea ennustaa, joten silloin voi myös jokin toinen kieli olla suositeltavin.

Esityskielten lisäksi myös uusia työkaluja kehitetään ja vanhoja parannellaan. Siten tiettyä työkalua on mahdotonta suositella, mutta sen pitäisi tietenkin tukea valittua esityskieltä ja mahdollisesti myös käytettävää kehittämismenetelmää. Työkalua valittaessa on selvitettävä muun muassa ontologioiden tallentamistapa (tietokantaan vai tiedostona), päättelymoottori ja käännökset muihin ontologiakieliin. Ontologian toteuttamisympäristöön tulisi sisältyä ainakin editori, selain, hakutoimintoja, tarkistaja ja kääntäjiä. Editorin avulla voi lisätä, poistaa tai muokata määrittelyjä. Selaimen avulla voi selata ontologiaa ja ontologiakirjastoa. Hakutoiminnot hakevat haluttuja määrittelyjä, tarkastaja etsii virheitä ja ristiriitaisuuksia sekä kääntäjä takaa määrittelyjen

siirrettävyyden toisiin kieliin. Työkalun on hyvä mahdollistaa muun muassa päättelyiden tekemisen ja toiminnallisuuden lisäämisen.

Soveltuvan esityskielen ja työkalun valinnan jälkeen tuotetaan itse koodi, jonka aikana tulee huomioida tiivistetysti kohdassa 3.1. listatut suunnittelukriteerit. Koodin kommentointi ja dokumentointi ovat suositeltavia toimia ontologian ylläpitoa ja jatkokehitystä ajatellen. Toteutuksen jälkeen ontologiaa tulee arvioida valittujen kriteerien pohjalta. Yleisiä kriteerejä on esitetty kohdassa 3.1 eli niitä ovat esimerkiksi selkeys, johdonmukaisuus ja uudelleenkäytettävyys. Niitä ei kuitenkaan voi oikein mitata, joten kyseisiä kriteerejä on pyritty operationalisoimaan mitattavuuden saavuttamiseksi (ks. esim. Cómez-Pérez, Juristo & Pazos 1995). Ontologiaa päivitetään ja kehitetään edelleen tarvittaessa arvioinnin pohjalta. Arviointidokumentissa tulisi kertoa, kuinka arviointi on tehty, mitä tekniikoita ja tietolähteitä on käytetty sekä millaisia virheitä on löydetty.

Kattavan, yksityiskohtaisen ja selkeän dokumentoinnin puute estää ontologian jatkokehityksen, jakamisen ja uudelleenkäyttämisen. Siksi ontologia ja sen kehittäminen tulee dokumentoida mahdollisimman hyvin. Pelkkä koodin tallentaminen ei riitä, sillä sen tutkiminen on vaikeaa ja aikaa vievää. Lisäksi kehittäjien vaihtuessa katoaa paljon tietoa ontologiasta, jolloin uusien työntekijöiden on hyvin vaikeaa ryhtyä muokkaamaan ja uudistamaan ontologiaa pelkän koodin perusteella. Siksi jokainen ontologian kehittämisprosessin vaihe tulisi dokumentoida huolellisesti ja kattavasti.

Ylläpitovaiheessa ontologiaan tehdään tarvittavia muutoksia, lisäyksiä ja poistoja esimerkiksi reagointina ympäristössä tapahtuneisiin muutoksiin. Jatkokehittämisen tavoitteena on useiden kehittämiskierrosten avulla parantaa

ja tarkentaa ontologiaa. Jatkokehitys voidaan toteuttaa muun muassa hankkimalla lisääineistoa, rakentamalla demoja ja tekemällä käyttäjätestejä.

6.4 Ehdotetun etenemistavan arviointi

Tässä luvussa lähdettiin liikkeelle nykytilan ongelmien ja tavoitetilan hahmottamisesta. Perusteena olivat hajautetut ja heterogeeniset tietojärjestelmät, joiden yhteentoimivuutta pyritään tukemaan ontologian kehittämällä. Lainsäädäntötyöhön osallistuva henkilö olisi varmasti osannut nimetä enemmänkin nykytilan ongelmia ja vastaavasti tavoitetilan osa-alueita. Lainsäädäntötyöhön osallistumatta ja tietojärjestelmiä sen paremmin tuntematta kohdassa pyrittiin kuitenkin antamaan yleisellä tasolla kuva siitä ongelmakentästä, mikä on motivoinut tätä tutkielmaa. Tavoitetilan jälkeen luvussa selvitettiin tietojärjestelmien ontologiaperustaisia integrointitapoja. Tämä tehtiin melko yksinkertaistetusti ja vain pariin lähteeseen tukeutuen. Aiheesta olisi siis voinut kirjoittaa paljon enemmänkin, mutta lyhyt ja yksinkertainen esitys nähtiin tässä kohtaa tarkoituksenmukaisemmaksi.

Varsinaisessa ehdotuksessa etenemistavaksi pyrittiin kiteyttämään kaikkein olennaisimmat asiat lyhyeksi kuvaukseksi ontologian kehittämisprosessista. Ehdotuksessa esitettiin yhden keskeisen tietojärjestelmien integroimista tukevan ontologian kehittämistä. Tämä lähestymistapa ei kuitenkaan ole välttämättä kaikkein suositeltavin. Käytännön kehittämistyössä voidaan päätyä siihen, että luodaankin monia ontologioita, joiden välille tehdään kytkentöjä. Ontologian kehittämiselle olennaisia päävaiheita esitettiin kolme eli suunnittelu-, määrittely- ja toteutusvaiheet. Tämä on hyvin karkea vaihejako, sillä kehittämistyö pitää sisällään myös paljon muita toimintoja, jotka tässä tutkielmassa sisällytettiin esitettyjen vaiheiden sisään. Kehittämistyön

toteuttamisessa voi kuitenkin olla selvempää, että kaikki isommat tehtäväalueet pidetään omina vaiheinaan.

Tiettyä kehittämismenetelmää, ontologian esityskieltä tai työkalua ei voinut tässä tutkielmassa ehdottaa, sillä ne vielä kehittyvät ja muuttuvat ennen kuin lainsäädäntöön ontologiassa ollaan toteutusvaiheessa. Ehdotus etenemistavaksi on kirjoitettu tieteellisten artikkelien ja kirjallisuuden pohjalta, joten se ei välttämättä käytännössä osoittaudukaan parhaaksi etenemistavaksi kehittämisprosessissa. Lisäksi se on hyvin yleisen tason esitys, joka vaatii vielä lisäselvityksiä ennen kuin käytännön kehittämistyö voidaan aloittaa. Siksi kehittäjien tulee vielä muun muassa yhteistyöverkoston tietämyksen ja osaamisen perusteella arvioida ja suunnitella uudestaan muun muassa kehittämisen vaiheita sekä niihin liittyviä tehtäviä ja lähestymistapoja.

7 YHTEENVETO

Ontologioiden tutkiminen ja soveltaminen ovat laaja-alaisen kiinnostuksen kohteena tällä hetkellä. Yleisesti ollaan sitä mieltä, että ontologiat voivat parantaa kommunikointia, tietämyksen jakamista ja uudelleenkäyttöä. Hajautettujen ja heterogeenisten tietojärjestelmäympäristöjen suuri määrä on lisännyt kiinnostusta ontologioita kohtaan. Ontologioiden avulla voidaan määritellä tarkasti tietyn ilmiön tai aihealueen semantiikka, joka tekee niistä merkittäviä kommunikoinnin helpottajia ja tietojärjestelmien semanttisen yhteentoimivuuden mahdollistajia.

Tämän tutkimuksen tarkoituksena oli selvittää, miten lainsäädäntöprosessin tietojärjestelmien yhteentoimivuutta tukeva ontologia voitaisiin kehittää. Tarkoituksena oli keskittyä erityisesti kehittämisen eri vaiheisiin. Tutkimuskysymyksen ratkaisemiseksi perehdyttiin aluksi ontologioita ja lainsäädäntötyötä käsittelevään kirjallisuuteen. Kirjallisuuden pohjalta luotiin katsaus ontologioiden perusideaan, rakenteeseen sekä luokittelu- ja käyttötapoihin. Tämän jälkeen perehdyttiin ontologioiden suunnittelun periaatteisiin ja menetelmiin sekä ontologioiden esityskieliin ja työkaluihin. Lisäksi esiteltiin lyhyesti muutamaa erityyppistä ontologiaa ja tunnetuimpia lakiontologioita. Tutkielma jatkui EU:n lainsäädäntötyön kuvaamisella. Sen todettiin vaikuttavan olennaisesti Suomen lainsäädäntötyöhön, jota kartoitettiin seuraavaksi. Todettiin, että lainsäädäntöprosessin tunteminen on tärkeää, koska sitä tukemaan on luotu tietojärjestelmät, joiden integroimisen tukemiseksi ehdotettiin yhden keskeisen ontologian kehittämistä.

Ontologian kehittämisen tueksi ehdotettiin etenemistapaa, joka sisälsi suunnittelu-, määrittely- ja toteutusvaiheet. Suunnitteluvaiheen sisällöksi

suositeltiin muun muassa ontologian tarkoituksen ja laajuuden selvittämistä, aihealueen mallintamista, aineiston keruuta ja yhteistyöverkoston luomista. Määrittelyvaiheen tarkoituksiksi todettiin ontologian rakenneosien määrittelemisen, joista pitää myös päästä yksimielisyyteen. Rakenneosien määrittelemiseen on olemassa erilaisia strategioita ja kuvaustapoja, joista suositeltiin keskeltä molempiin suuntiin lähtevää strategiaa sekä taulukoiden ja kuvioiden piirtämistä. Määrittelyvaiheessa voidaan lisäksi yhdistää muiden ontologioiden ja sanastojen sisältöjä omaan ontologiaan. Toteutusvaiheessa edellisen vaiheen tulokset on tarkoitus koodata valmiiksi ontologiaksi tietyllä esityskielellä ja työkalulla. Esityskielestä ja työkalusta ei voitu antaa suoria suosituksia, koska ne tulevat vielä kehittymään ja muuttumaan. Joitakin suuntaviivoja niiden valinnassa käytettävistä kriteereistä kuitenkin listattiin.

Tutkimuksen rajoitteena on ontologioiden suunnittelun, lainsäädäntöprosessien ja integrointitapojen kuvaamisen teoreettisuus. Näistä kaikista tulisi saada lisää käytäntöön pohjautuvaa tietoa ja osaamista, jotta ehdotusta etenemistavaksi voitaisiin laajentaa, kehittää edelleen ja arvioida paremmin. Toisaalta taas ehdotus käsittelee asioita melko yleisellä tasolla, joten siitä voi olla hyötyä lainsäädäntötyön lisäksi myös muilla sovellusalueilla.

Kiinnostava jatkotutkimus voitaisiin toteuttaa lainsäädäntöprosessin järjestelmien yhteentoimivuutta tukevan ontologian käytännön kehittämistä. Siinä olisi mielenkiintoista verrata tehtyjä ratkaisuja ja niiden perusteluja tässä tutkielmassa tehtyihin suosituksiin. Tutkielman tuloksia voidaan soveltaa suomalaisen lainsäädäntötyön lisäksi myös muissa EU-maissa lainsäädäntötyöhön liittyvien tietojärjestelmien integroimisessa. Tällöin voitaisiin verrata eri maiden ontologioiden toteuttamista ja yhteensopivuutta.

Samankaltaisia tutkimuksia voitaisiin tehdä mahdollisesti muillakin sovellusalueilla.

Tässä tutkielmassa ei käsitellä ontologioita kovin tekniseltä kannalta, joten tutkimus teknisemmistä asioista voisi antaa erilaisen näkökulman aiheeseen. Myös erilaiset päättelysäännöt ja niihin liittyvät kielet yms. rajattiin tämän tutkielman ulkopuolelle, mutta lähitulevaisuudessa erilaisten standardien kehittyminen ja niihin liittyvät asiat voisivat osoittautua kiinnostavaksi tutkimusaiheeksi. Myös tässä tutkielmassa esitetyt menetelmät, kuvaustavat ja työkalut tulevat kehittymään jatkossa. Samoin uusia ilmaantuu ontologiatekniikoiden kehittyessä. Tällöin tarvittaisiin päivitystä tämän tutkielman tietoihin.

Ontologiat liittyvät myös agenttitekнологiaan ja web services -konseptiin. Semanttisen webin tekniikoiden kehittyessä voisi olla mielenkiintoista tehdä tutkimusta niiden ja ontologioiden suhteesta. Ontologioiden lisäksi myös lainsäädäntötyöhön saattaa tulla muutoksia. Vuonna 2004 tapahtuvalla EU:n laajenemisella voi olla vaikutuksensa myös lainsäädäntöprosessiin ja siihen liittyviin tietojärjestelmiin. Muutokset EU:n lainsäädäntötyössä voivat heijastua myös Suomessa tapahtuvaan EU-asioiden valmisteluun.

LÄHDELUETTELO

- Arpírez J.C., Corcho O., Fernández-López M. & Gómez-Pérez A. 2001.
 WebODE: A Scalable Ontological Engineering Workbench. Proceedings of
 the First International Conference on Knowledge Capture (KCAP'01).
 Victoria: ACM Press, 6-13.
- Arpírez J.C., Gómez-Pérez A., Lozano A. & Pinto H.S. 1998. (ONTO)2Agent: An
 Ontology-Based WWW Broker to Select Ontologies. Workshop on
 Applications of Ontologies and Problem-Solving Methods (ECAI'98).
 Brighton, United Kingdom, 16-24.
- Benjamins V.R. & Fensel D. 1998. The Ontological Engineering Initiative (KA)².
 Teoksessa N. Guarino (toim.) Formal Ontology in Information Systems.
 IOS Press.
- Bernaras A., Laresgoiti I. & Corera J. 1996. Building and Reusing Ontologies for
 Electrical Network Applications. Proceedings of the European Conference
 on Artificial Intelligence (ECAI'96), 298-302.
- Borchardt K-D. 2000. Yhteisön oikeuden perusteet. Tietoa Euroopasta –sarjan
 julkaisu. Euroopan komissio, Koulutuksen ja kulttuurin pääosasto.
 Luxemburg: Euroopan yhteisöjen virallisten julkaisujen toimisto.
- Bray T., Hollander D. & Layman A. 1999. Namespaces in XML [online]. W3C
 Recommendation [viitattu 21.11.2003]. Saatavilla [www-muodossa
 <http://www.w3.org/TR/REC-xml-names>](http://www.w3.org/TR/REC-xml-names).

Bray T., Paoli J., Sperberg-McQueen C. M. & Maler E. 2000. Extensible Markup Language (XML) 1.0 (Second Edition) [online]. W3C Recommendation [viitattu 07.09.2003]. Saatavilla [www-muodossa <http://www.w3.org/TR/REC-xml>](http://www.w3.org/TR/REC-xml).

Brickley D. & Guha R.V. 2003. RDF Vocabulary Description Language 1.0: RDF Schema [online]. W3C Working Draft [viitattu 25.09.2003]. Saatavilla [www-muodossa <http://www.w3.org/TR/rdf-schema/>](http://www.w3.org/TR/rdf-schema/).

Chandrasekaran B., Josephson J.R. & Benjamins R. 1999. What Are Ontologies and Why Do We Need Them? *IEEE Intelligent Systems* 14(1), 20-26.

Chaudhri V.K., Farquhar A., Fikes R., Karp P.D. & Rice J.P. 1998. Open Knowledge Base Connectivity (OKBC). Specification Document 2.0.3. Technical Report. SRI International and Stanford University (KSL).

Clark J. 1999. XSL Transformations (XSLT) [online]. W3C Recommendation [viitattu 21.11.2003]. Saatavilla [www-muodossa <http://www.w3.org/TR/xslt>](http://www.w3.org/TR/xslt).

Cómez-Pérez A. & Corcho O. 2002. Ontology Languages for the Semantic Web. *IEEE Intelligent System* 17(1), 54-60.

Cómez-Pérez A., Juristo N. & Pazos J. 1995. Evaluation and Assessment of Knowledge Sharing Technology. Teoksessa N.J. Mars (toim.) *Towards Very Large Knowledge Bases - Knowledge Building and Knowledge Sharing*. Amsterdam: IOS Press, 289-296.

- Corcho O., Fernández-López M. & Cómez-Pérez A. 2001. Technical Roadmap v1.0 [online]. Amsterdam: OntoWeb Consortium [viitattu 30.09.2003]. Saatavilla pdf-muodossa <<http://ontoweb.aifb.uni-karlsruhe.de/About/Deliverables/Deliverable111.pdf>>.
- Corcho O., Fernández-López M. & Cómez-Pérez A. 2003. Methodologies, Tools and Languages for Building Ontologies. Where is their Meeting Point? Data & Knowledge Engineering 46(1), 41-64.
- Domingue J. 1998. Tadzebao and WebOnto: Discussing, Browsing and Editing Ontologies on the Web. Proceedings of the 11th Knowledge Acquisition Workshop (KAW'98), Banff, Canada.
- Dublin Core 2002. Dublin Core metadataformaatin suomalainen versio [online]. Helsingin yliopiston kirjasto [viitattu 25.09.2003]. Saatavilla www-muodossa <http://www.lib.helsinki.fi/dublin_core/dc-sfs.html>.
- Eerola R., Mylly T. & Saarinen P. 2000. EU-oikeuden perusteet. Tampereen yliopisto, Julkisoikeuden laitos. Jyväskylä: Gummerus.
- e-GIF 2003. Government Interoperability Framework [online]. NHS Information Authority [viitattu 07.10.2003]. Saatavilla www-muodossa <<http://www.nhsia.nhs.uk/egif/>>.
- EU:n lainsäädännön keskeiset osapuolet 2003. EU:n julkaisutoimisto. Saatavilla www-muodossa <<http://europa.eu.int/eur-lex/fi/about/pa p/index.html>>.

Eurovoc 2003. Eurovoc-tesaurus [online]. Euroopan unionin julkaisutoimisto [viitattu 07.10.2003]. Saatavilla www-muodossa <<http://europa.eu.int/celex/eurovoc/>>.

EY:n perustamissopimus 2002. Euroopan yhteisön perustamissopimuksen konsolidoitu toisinto. Euroopan yhteisöjen virallinen lehti (45)C325. Saatavilla myös www-muodossa <http://europa.eu.int/eur-lex/fi/treaties/dat/EC_consol.html>.

Fallside D.C. 2001. XML Schema Part 0: Primer [online]. W3C Recommendation [viitattu 10.10.2003]. Saatavilla www-muodossa <<http://www.w3.org/TR/xmlschema-0/>>.

Farquhar A., Fikes R., Pratt W. & Rice J. 1995. Collaborative Ontology Construction for Information Integration. Technical Report KSL-95-63. Palo Alto, California: Stanford University, Knowledge Systems Laboratory.

Farquhar A., Fikes R. & Rice J. 1997. The Ontolingua Server: A Tool for Collaborative Ontology Construction. *International Journal of Human-Computer Studies* 46(6), 707-727.

Fensel D. 2000. *Ontologies: Silver Bullet for Knowledge Management and Electronic Commerce*. Berlin: Springer-Verlag. Saatavilla pdf-muodossa <<http://citeseer.nj.nec.com/413498.html>>.

- Fensel D., van Harmelen F., Horrocks I., McGuinness D.L. & Patel-Schneider P.F. 2001. OIL: an ontology infrastructure for the Semantic Web. *IEEE Intelligent Systems* 16 (2), 38 -45.
- Fernández M., Cómez-Pérez A. & Juristo N. 1997. METHONTOLOGY: From Ontological Art Towards Ontological Engineering. Proceedings of the AAAI'97 Spring Symposium Series on Ontological Engineering, Stanford, USA, 33-40.
- Geerts G.L. & McCarthy W.E. 2000. The Ontological Foundation of REA Enterprise Information Systems. Paper presented at the American Accounting Association Conference, Philadelphia.
- Genesereth M. & Fikes R. 1992. Knowledge Interchange Format. Technical Report Logic-92-1. Stanford University, Computer Science Department.
- Grosso W.E., Eriksson H., Ferguson R.W., Gennari J.H., Tu S.W. & Musen M.A. 1999. Knowledge Modeling at the Millennium (The Design and Evolution of Protégé2000). Proceedings of the Twelfth Workshop on Knowledge Acquisition, Modeling and Management (KAW'99). Banff, Canada.
- Gruber T.R. 1993. A Translation Approach to Portable Ontology Specifications. *Knowledge Acquisition* 5, 199-220.
- Gruber T.R. 1995. Toward Principles for the Design of Ontologies Used for Knowledge Sharing. *International Journal of Human-Computer Studies* 43(5-6), 907-928.

- Grüninger M. & Fox M.S. 1994. The Role Of Competency Questions in Enterprise Engineering. Proceedings of the Workshop on Benchmarking - Theory and Practice. Trondheim, Norway.
- Grüninger M. & Fox M.S. 1995. Methodology for the Design and Evaluation of Ontologies. Proceedings of the Workshop on Basic Ontological Issues in Knowledge Sharing. Montreal, Canada.
- Grüninger M. & Lee J. 2002. Ontology Applications and Design. Communications of the ACM 45(2), 39-41.
- Guarino N. 1998. Formal Ontology and Information Systems. Teoksessa N. Guarino (toim.) Formal Ontology in Information Systems. Proceedings of the First International Conference (FOIS'98), Trento, Italy. Amsterdam: IOS Press, 3-15.
- Hallituksen esitysten laatimisohteet 2003. HELO-työryhmän ehdotus. Helsinki: Oikeusministeriö.
- Hendler J. 2001. Agents and the Semantic Web. IEEE Intelligent Systems 16(2), 30-37.
- Holsapple C.W. & Joshi K.D. 2002. A Collaborative Approach to Ontology Design. Communications of the ACM 45(2), 42-47.
- Hovy E.H. 1998. Combining and Standardizing Large-Scale, Practical Ontologies for Machine Translation and Other Uses. Proceedings of the 1st

International Conference on Language Resources and Evaluation (LREC).
European Language Resources Association, Paris, 535-542.

Hovy E.H. 2003. Using an Ontology to Simplify Data Access. *Communications of the ACM* 46(1), 47-49.

Hyvönen E. 2002. *The Semantic Web – The New Internet of Meanings*.

Teoksessa E. Hyvönen (toim.) *Semantic Web Kick-Off in Finland – Vision, Technologies, Research, and Applications*. Helsinki Institute for Information Technology. Helsinki, Finland: HIIT Publications, 3-25.

Hyvönen E., Kettula S., Raatikka V., Saarela S. & Viljanen K. 2002. *Semantic Interoperability on the Web: Case Finnish Museums Online*. Teoksessa E. Hyvönen (toim.) *Towards the semantic web and web services*. Proceedings of XML Finland 2002 Conference, Helsinki, Finland, pp. 41-53.

JHS 143 -suositus 2001. *Asiakirjojen kuvailuformaatti [online]*. Julkisen hallinnon tietohallinnon neuvottelukunta [viitattu 15.10.2003]. Saatavilla [www-muodossa](http://www.muodossa)
<<http://www.intermin.fi/intermin/hankkeet/juhta/home.nsf/>>.

Jokela S. 2001. *Metadata Enhanced Content Management in Media Companies*. Dissertation, Computer Science and Engineering. Acta Polytechnica Scandinavica, Mathematics and Computing Series No. 114. Espoo: Finnish Academies of Technology.

Joutsamo K., Aalto P., Kaila H. & Maunu A. 2000. Eurooppaoikeus. Kolmas uudistettu painos. Helsinki: Kauppakaari, Lakimiesliiton kustannus.

Julkishallinnon XML-strategia 2003. Työryhmämuistioita 18/2003. Helsinki: Valtiovarainministeriö, Hallinnon kehittämisosasto.

Karp P.D., Chaudhri V.K. & Thomere J. 1999. XOL: An XML-Based Ontology Exchange Language. Technical Report 559. Menlo Park: Artificial Intelligence Center, SRI International.

Kemppinen R. 2002. Suomi Euroopan unionissa – perusteos. 4. uudistettu painos. Helsinki: Edita.

Kent R.E. 1999. Ontology Markup Language [online]. The Ontology Consortium [viitattu 10.10.2003]. Saatavana [www-muodossa](http://www.muodossa) <<http://www.ontologos.org/OML/OML%200.3.htm>>.

Kifer M., Lausen G. & Wu J. 1995. Logical Foundations of Object-Oriented and Frame-Based Languages. *Journal of the ACM* 42(4), 741-843.

Kim H. 2000. Developing Ontologies to Enable Knowledge Management: Integrating Business Process and Data Driven Approaches. AAI Workshop on Bringing Knowledge to Business Processes, Stanford, California.

Lainlaatijan opas 1996. Oikeusministeriö. Helsinki: Edita.

- Lassila O. & Swick R. 1999. Resource Description Framework (RDF): Model and Syntax Specification [online]. W3C Recommendation [viitattu 25.09.2003]. Saatavilla [www-muodossa < http://www.w3.org/TR/REC-rdf-syntax/>](http://www.w3.org/TR/REC-rdf-syntax/).
- Lenat D.B & Guha R.V 1990. Building Large Knowledge-Based Systems: Representation and Inference in the Cyc Project. Boston: Addison Wesley.
- Luke S. & Heflin J. 2000. SHOE 1.01. Proposed Specification [online]. University of Maryland [viitattu 22.09.2003]. Saatavilla [www-muodossa <http://www.cs.umd.edu/projects/plus/SHOE/spec.html>](http://www.cs.umd.edu/projects/plus/SHOE/spec.html).
- Lyytikäinen, Tiitinen & Salminen 2003. Unifying Access to Heterogeneous Document Databases Through Contextual Metadata. Teoksessa Becker S.A. (toim.) Effective Databases for Text and Document Management. Hersley, Pennsylvania: Idea Group Publishing, 93-107.
- Maedche A. & Staab S. 2001. Ontology Learning for the Semantic Web. IEEE Intelligent Systems 16(2), 72-79.
- MacGregor R. 1999. Retrospective on Loom. Loom Report [online]. University of Southern California [viitattu 22.09.2003]. Saatavilla [www-muodossa <http://www.isi.edu/isd/LOOM/papers/macgregor/>](http://www.isi.edu/isd/LOOM/papers/macgregor/).
- McCarty L.T. 1989. A Language for Legal Discourse. I. Basic Features. Proceedings of the Second International Conference on Artificial Intelligence and Law. Vancouver, Canada, 180-189.

- McGuinness D.L., Fikes R., Hendler J. & Stein L.A. 2002. DAML+OIL: An Ontology Language for the Semantic Web. *IEEE Intelligent Systems* 17(5), 72 -80.
- McGuinness D.L. & van Harmelen F. 2003. OWL Web Ontology Language Overview [online]. W3C Candidate Recommendation [viitattu 28.09.2003]. Saatavilla [www-muodossa <http://www.w3.org/TR/owl-features/>](http://www.w3.org/TR/owl-features/).
- Mattila M. 2000. Valtioneuvosto – Suomen EU-politiikan määrittelijä. Teoksessa T. Raunio & M. Wiberg (toim.) *EU ja Suomi. Unionijäsenyyden vaikutukset suomalaiseen yhteiskuntaan*. Helsinki: Edita, 135-150.
- Miller G.A. 1995. WordNet: A Lexical Database for English. *Communications of the ACM* 38(11), 39-41.
- Motta E. 1998. An Overview of the OCML Modelling Language. *Proceedings of the 8th Workshop on Knowledge Engineering: Methods and Languages (KEML'98)*. Karlsruhe, Germany.
- Natlex 2003. Guide to Using the NATLEX Database on National Labour and Social Security Law [online]. International Labour Organization (ILO) [viitattu 07.10.2003]. Saatavilla [www-muodossa <http://natlex.ilo.org/natlexhelp.htm>](http://natlex.ilo.org/natlexhelp.htm).
- ParlML 2000. A Common Vocabulary for Parliamentary Language [online]. European Parliament [viitattu 07.10.2003]. Saatavilla [pdf-muodossa <http://www.europarl.eu.int/docman/texts/TFDM\(2000\)0014EN0.pdf>](http://www.europarl.eu.int/docman/texts/TFDM(2000)0014EN0.pdf).

RASKE2 2003. RASKE2-projektin kotisivut [online]. Jyväskylä: Jyväskylän yliopisto [viitattu 13.11.2003]. Saatavilla www-muodossa <<http://www.it.jyu.fi/raske/>>.

RosettaNet.org 2003. RosettaNet - Lingua Franca for eBusiness [online]. RosettaNet consortium [viitattu 04.11.2003]. Saatavilla www-muodossa <<http://www.rosettanet.org/>>.

Salminen A. 2003. Document Analysis Methods. Teoksessa Bernie C.L. (toim.) Encyclopedia of Library and Information Science, Second Edition, Revised and Expanded. New York: Marcel Dekker, 916-927.

Smith B. & Mark D. 1999. Ontology with Human Subjects Testing: An Empirical Investigation of Geographic Categories. American Journal of Economic and Sociology 58(2), 246-265.

Sowa J.F. 2000. Knowledge Representation: Logical, Philosophical, and Computational Foundations. Pacific Grove, California: Brooks Cole Publishing. Osa julkaisusta saatavilla www-muodossa [viitattu 19.04.2003] <<http://users.bestweb.net/~sowa/ontology/index.htm>>.

Staab S., Schnurr H-P., Studer R. & Sure Y. 2001. Knowledge Processes and Ontologies. IEEE Intelligent Systems 16(1), 26-34.

Stamper R.K. 1991. The Role of Semantics in Legal Expert Systems and Legal Reasoning. Ratio Juris 4(2), 219-244.

Suomen eduskunta 2000. Eduskunnan kanslia. Toinen uudistettu painos.
Helsinki: Edita.

Suomen perustuslaki 731/1999. Valtion säädöstietopankki Finlex [online].
Helsinki: Oikeusministeriö [viitattu 08.09.2003]. Saatavilla www-
muodossa <<http://www.finlex.fi/>>.

Swartout W. & Tate A. 1999. Ontologies. IEEE Intelligent Systems 14 (1),18-19.

Swartout B., Patil R., Knight K. & Russ T. 1996. Toward Distributed Use of
Large-Scale Ontologies. Proceedings of the 10th Knowledge Acquisition
for Knowledge-Based Systems Workshop (KAW'96). Banff, Canada.

Tekes 2003. Suomalaiset semanttisen webin ontologiat [online]. Helsinki:
Teknologian kehittämiskeskus [viitattu 07.10.2003]. Saatavilla www-
muodossa <<http://akseli.tekes.fi/Resource.phx/tivi/vaui/z-07684633.htx>>.

Tiitinen P., Salminen A. & Lyytikäinen V. 1997. EU-lainsäädäntöasiakirjat
Suomessa. RASKE-projektin raportti. Eduskunnan kanslian julkaisu
1/1997. Helsinki: Edita.

Torniainen A. & Valtari E. 2001. Valtioneuvoston johtamista tukevien
tietojärjestelmien arviointi. Helsinki: Valtiovarainministeriö. Saatavilla
pdf-muodossa <<http://www.vm.fi/tiedostot/pdf/fi/17002.pdf>>.

Uschold M. & Grüninger M. 1996. Ontologies: Principles, Methods and
Applications. The Knowledge Engineering Review 11(2), 93-155.

Uschold M. & King M. 1995. Towards a Methodology for Building Ontologies. Proceedings of the Workshop on Basic Ontological Issues in Knowledge Sharing (IJCAI'95). Montreal, Canada.

Valente A. 1995. Legal Knowledge Engineering; A Modelling Approach. University of Amsterdam, The Netherlands. Amsterdam: IOS Press.

Valtion tietotekniikan rajapintasuosituksia 2001. Valtiovarainministeriön työryhmämuistioita 27/2001. Helsinki: Valtiovarainministeriö, Hallinnon kehittämisosasto.

Valtioneuvoston ohjesääntö 2003. Helsinki: Valtioneuvosto. Saatavilla pdf-muodossa <<http://www.vnk.fi/tiedostot/pdf/fi/35626.pdf>>.

Valtioneuvoston tietohallintostrategia 2003-2007 2003. Tyvestä puuhun: tuottavaa – toimivaa – turvallista. Helsinki: Valtiovarainministeriö. Saatavilla pdf-muodossa <<http://www.vm.fi/tiedostot/pdf/fi/39132.pdf>>.

Valtioneuvoston yhteisten asianhallintaprosessien kuvaukset ja suositukset 2000. Valtiovarainministeriön työryhmämuistioita 7/2000. Helsinki: Valtiovarainministeriö. Saatavilla pdf-muodossa <<http://www.vn.fi/vm/julkaisut/tyoryhmuistioid/pdf/tr700.pdf>>.

Van Kralingen R.W. 1995. Frame-Based Conceptual Models of Statute Law. Computer / Law Series 16. The Hague, the Netherlands: Kluwer Law International.

- Visser P.R.S. 1995. Knowledge Specification for Multiple Legal Tasks; A Case Study of the Interaction Problem in the Legal Domain. Computer / Law Series 17. The Hague, the Netherlands: Kluwer Law International.
- Wache H., Vögele T., Visser U., Stuckenschmidt H., Schuster G., Neumann H. & Hübner S. 2001. Ontology-Based Integration of Information - A Survey of Existing Approaches. Proceedings of the Workshop on Ontologies and Information Sharing (IJCAI'01). Seattle, Washington, USA, 108-117.
- Wang N. & Xu X. 2000. A Method to Build Ontology. Proceedings of the Fourth International Conference/Exhibition on High Performance Computing in the Asia-Pacific Region, 672-673.
- Wiederhold G. (toim.) 1996. Intelligent Integration of Information. Boston, MA: Kluwer Academic Publishers.
- YSA 2003. Yleinen suomalainen asiasanasto [online]. Helsinki: Helsingin yliopiston kirjasto, Kansalliskirjasto [viitattu 25.09.2003]. Saatavilla [www-muodossa <http://www.lib.helsinki.fi/kirjastoala/asiasanastot/ysa.htm>](http://www.lib.helsinki.fi/kirjastoala/asiasanastot/ysa.htm).