

*Mikko Lappalainen, Susanna Nykyri,  
Tuomas Palonen*

# Semanttisesta ja funktionaalista tiedonhaun ja –tallennuksen ontologiasta : suomalaisen yleisontologian laadinnan haasteita

This article discusses theoretical and pragmatical challenges concerning the Finnish general ontology (Yleinen suomalainen ontologia, YSO). The central challenges are due to e.g. operating on a more conceptual level than thesauri traditionally do, and that the meanings of the words are bound to their cultural and historical context.

*Address: Susanna Nykyri, Mikko Lappalainen, Tuomas Palonen, Kansalliskirjasto – Tutkimuskirjasto PL 26 (Teollisuuskatu 23) 00014 Helsingin yliopisto  
E-mail: susanna.nykyri@helsinki.fi, mikko.lappalainen@helsinki.fi, tuomas.palonen@helsinki.fi*

## Johdanto

Ontologioista on tullut 2000-luvun alun trendi verkkomaailmassa. Visiossa uudesta älykkäästä internetistä eli semanttisesta webistä, ontologiat on ajateltu eräänlaisiksi linkitetyn ja koneluettavan (meta)tiedon selkärangaksi. Kirjastojen ja muiden muistiorganisaatioiden piirissä on viime vuosina herätty huomaamaan perinteisten resurssien kuten asiasanastojen potentiaali tässä uuden semanttisen tiedon maailmassa.

Suomessa on lähdetty suhteellisen varhain mukaan kehittämään semanttisen webin ideologian mukaisia teknologisia ratkaisuja. Aalto-yliopistossa vuosina 2003-2012 toimineessa FinnONTO-projektissa luotiin suomalaisia semanttisen webin ontologioita avoimien standardien ja sisältöjen pohjalta. Projektin yksi keskeisiä tavoitteita oli luoda Yleinen suomalainen ontologia (YSO), joka perustui Kansalliskirjaston kehittämään ja ylläpitämään Yleiseen suomalaiseen asiasanastoon (YSA). FinnONTO-projektissa tehtyjä ontologioita

on hyödynnetty mm. joidenkin muistiorganisaatioiden portaaleissa, kuten Kulttuuri - ja Kirjasammossa sekä Yritys-Suomessa, mutta niiden tehokkaan sovellettavuuden on katsottu vaativan vielä osin radikaaliakin kehitystyötä.

Kansalliskirjastossa maaliskuussa 2013 käynnistynyt ONKI-projekti jatkaa FinnONTO:n työtä ja vastaa YSON sekä kansallisen ontologiapalvelun, Finton<sup>1</sup>, nykyisestä kehityksestä ja ylläpidosta Finto-palveluun sisältyy KOKO, jossa eri alojen erikoisontologiat on sillattu YSOon. ONKI-projektin myötä erityistä huomiota kiinnitetään ontologioiden käytettävyyteen ja sovellettavuuteen erilaisissa yhteyksissä. Kehitystyössä otetaan huomioon myös muut kuin perinteiset kirjastoalan tarpeet ja ulotetaan kehitystyö koskemaan laajemminkin muistiorganisaatioiden (kuten museoiden) ja median tarpeita. Kansalliskirjasto-vetoisessa kehitystyössä fokus siirtyy kohti tiedonhakijan tarpeita, kun aiemmin on keskitytty erityisesti ontologian kehittäjiä ja ammattimaisten käyttäjiä kuten sisällönkuvailijoiden näkökulmiin.

Fokuksen muuttuminen tuo uusia haasteita Yleisen suomalaisen ontologian kehittäjille. Tässä artikkelissa tarkastellaan ontologiatyön ja erityisesti YSON rakentamisen teoreettisia ja pragmaattisia haasteita suhteessa kirjasto-maailman perinteisiin vastuualueisiin kuten asiasanapohjaiseen sanastotyöhön. Pääpaino artikkelissa on ihmiskäyttäjän näkökulmassa ja ihmistieteiden erityishaasteissa. Esimerkeissä painotetaan kirjoittajien havaintoja YSON nykytilasta.

## Sanastoista ja ontologioista

Sisällönkuvailu voidaan toteuttaa automaattisesti tai intellektuaalisesti. Automaattisen kuvailumenetelmien etuja ovat tyhjäntävyys (kaikki dokumentin yksityiskohdat ovat haettavissa), nopeus ja pienet kustannukset. Ongelmia tyypillisesti aiheuttavat mekaanisuus, jolloin tekstin ja ilmausten merkityksiin liittyvät tulkinnat eivät tule mukaan kuvailuun. Intellektuaalinen kuvailu sen sijaan perustuu ihmisen tekemään dokumentin sisällön analyysiin ja kuvailutietojen valintaan. Siitä erotetaan perinteisesti kaksi tasoa: luettelointi (dokumenttien ulkoisten piirteiden kuvailu) ja sisällönkuvailu. Viimeksi mainittu kohdistuu dokumenttien sisältöön ja se tapahtuu joko luonnollista kieltä tai dokumentaatiokieltä käyttäen. (Järvelin & Sormunen 1999, 131)

Kun sisällönkuvailussa käytetään luonnollista kieltä, ovat sisältöä kuvaavat sanat tai –fraasit avainsanoja. Tällaisen kuvailun edut ovat vähäiset, koska sitä rajoittavat samat luonnolliseen kieleen liittyvät monitulkintaisuusongelmat kuin automaattista indeksointia. Dokumentaatiokieltä käyttäessä annetaan asiasanoja, ja tavoitteena on tällöin olla mahdollisimman vapaa monitulkintaisuudesta. Asiasanastoissa ja tesauksissa, jotka siis sisältävät tiettyjen periaatteiden mukaisesti laadittuja asiasanoja, sanat ja ilmaisut on pyritty saamaan mahdollisimman yksiselitteiseksi dokumentaatiokielen mukaan. (Emt., 131-132)

Automaattista kuvailua puoltavia seikkoja ovat intellektuaalisen kuvailun vaatimat riittämättömät resurssit ja välineet, tesaurus ei sisällä tietyn tyyppisiä asiasanoja (esimerkiksi nimiin kohdistuvia hakuja varten), tesauksien sanasto on liian yleisellä tasolla (spesifiset kyselyt) ja että tesauksien sisältö on osin vanhentunutta (uusia ilmiöitä koskevat haut). (Järvelin &

Sormunen 1999, 133) Ontologian voidaan ajatella olevan myös automaattisen sisällönkuvailun taustalla, mutta niin sitä kuin intellektuaalista sisällönkuvailua palvellessaan sen tulisi kyetä ylittämään em. puutteet ja rajoitteet. Teknisen kehityksen ja myös toimintatapojen muutosten myötä voidaankin ajatella, että ontologia onnistuu lisäämään niin ilmaisun rikkautta (spesifisyys), ajantasaisuutta kuin tehostamaan sisällönkuvailua intellektuaalisena prosessina. Hyödyntäkövät sekä sisällönkuvailijalle että tiedonhakijalle.

Dokumenttien sisällönkuvailua ja tiedonhakua varten luotujen kontrolloitujen sanastojen rakentamisella ja käytöllä on pitkät perinteet kirjastokentällä. Kongressin kirjaston Subject Headings (LCSH) juhli satavuotista historiaansa vuonna 2000 ja YSAnkin käyttöön otosta on vierähtänyt kohta kolmekymmentä vuotta.

YSA on ollut varhaisesta vaiheesta lähtien tesaurusmallinen sanasto, eli siinä on esitetty varsinaisten asiasanojen lisäksi myös termien suhteita toisiinsa (laajemmat, suppeammat ja rinnakkaistermit). YSA on laadittu noudattaen tesaurus-standardeja (ks. standardi SFS 5471, myös Järvelin & Sormunen 1999, 132). Ontologiat voidaan nähdä eräänlaisina verkkoajan päivitettyinä tesauksina (Hyvönen 2005).

Keskeisin ero perinteisiin tesauksiin nähden ontologioissa on termeistä niiden taustalla olevalle käsitetasolle siirtyminen sekä koneluettavuuden vaatimus. Ideana on saada kone ”ymmärtämään” käsitteiden merkityksiä ja niiden välisiä suhteita samaan tapaan kuin ihmiset. Tiedonhakutilanteessa kone voisi luoda haetulle tiedolle kontekstia ja auttaa tiedonhakijaa eteenpäin tämän tiedontarpeen määrittelyssä. Lisäksi käsitetasolle siirryttäessä pyritään irtautumaan termitason kieliriippuvuudesta ja näin mahdollistamaan ontologialle laajemmat sovellettavuusmahdollisuudet.

Ontologisoimisen mahdollistamia keskeisiä edistyskäsiä ovat siis teknisten ratkaisujen mahdollistama käsitetason suurempi tarkkuus ja toisaalta kielellisten variaatioiden suurempi rikkaus. Edistyksellistä on myös teknologian kehittymisen mahdollistama konepäättelyn suurempi osuus ja se, että ontologiat on rakennettu nimenomaan verkkomaailman tarpeisiin ja sen ehdoilla. Semanttisessa webissä on mahdollista tukea erilaisten diskursien kirjoja mm. erilaisten ja moninaisempien ohjaustermien ja assosiativisten suhteiden lisäämisellä ja esittämisellä. Semanttisessa webissä toimiminen tarjoaa myös mahdollisuuksia dynaamisempaan sisältöön kuin

perinteisessä kirjasto- ja tesaurusmaailmassa toimiminen. Voidaankin ajatella, että ontologian tuominen lähemmäs käyttäjien kirjoja mahdollistaa perinteisiä asiasanastoja merkittävästi laajemman käytettävyyden. - Ontologiat ovat dynaamisempia muuttumaan käyttötarpeiden mukaan ja kykenevämpiä vastaamaan moninaisempiin odotuksiin kuin tesauukset, joiden päivitystä edelleen leimaa ajallinen jähmeys ja tiukemmin tietyn näkökulman dominoivuus. Silti ontologiat ovat sisällönkuvailun ja tiedonhaun apuvälineinä edelleen nimenomaan erikoiskieltä ja aivan kuten tesauukset ja asiasanastotkin, dokumentaatiokieltä. Tiedonhaun ja sisällönkuvailun apuvälineeksi laadittavan ontologian perusfunktio on identtinen dokumentaatiokielen perusfunktion kanssa. Semanttisen webin mahdollistamat ratkaisut tehostavat sisällönkuvailun funktion mahdollisimman täyttä toteutumista.

Tietokonesemantiikan vaatima tiukka käsitehierarkia luo kuitenkin uusia haasteita sanaston rakentajalle. Ontologioita on kritisoitu muun muassa siitä, että käsitteet voidaan sijoittaa vain yhteen paikkaan hierarkkisessa kokonaisuudessa ja ontologian käyttäjän on hyväksyttävä tämä päätös. Onkin todettu, että tietyn ontologian käyttäminen edellyttää myös kyseisen ontologian sanaston ja maailmankatsomuksen hyväksymistä. (Shirky 2003, Wyssusek 2004) Näin ollen voidaan kysyä, kenellä on oikeus kategorisoida maailma.

Tämä kysymys on hankala erityisesti YSON kaltaisia yleisontologioita rakennettaessa, kun niiden määrittelyalue on periaatteessa koko maailma. Maailmaa voidaan kuitenkin aina tarkastella eri näkökulmista ja jokapäiväisen elämän merkitystä luova todellisuus pakenee usein loogisuutta. Tämä pitää ottaa huomioon myös ontologioita rakennettaessa ja erityisesti hyödynnettäessä niitä erilaisissa tiedonhaku järjestelmissä. YSON kohdalla on kuitenkin muistettava, että se ei ole – eikä myöskään pyri olemaan – maailmanselitys, vaan sen sisällöllisen kehitystyön käsitekorpus on peräisin YSAsta eli se pohjautuu sisällönkuvailukäsitteistöön jo todettuun tarpeeseen.

YSON kaltaisen ontologian rakentamisessa tasapainoillaan toisaalta normatiivisen sanastotyön – jossa pyritään käsitteiden ja termien sopimuksenvaraiseen yksiselitteisyyteen ja siten takaamaan tiedonkulun häiriöttömyys – ja toisaalta deskriptiivisen, eli käsitteen käytön merkityseroja kuvaavan sanastotyön välillä. Käytännön ontologiatyössä yksi tärkeistä työtä

ohjaavista ja ratkaistavista kysymyksistä onkin se, ovatko tiedonhaun ja -tallennuksen ontologiat myös luonteeltaan kieltä standardoivia ja miten pitkälle niiden tarkoitus on myötäillä luonnollisen kielen ja sen käytön ominaisuuksia.

Sisällönkuvailuun ja tiedonhakuun tarkoitettujen ontologioiden uudenlainen ja perinteisiä asiasanastoja ja tesauksia moninaisempi käyttöympäristö, semanttinen web, on kuitenkin kehittynyt myös osin spontaanisti ja perinteisistä kirjastoalan traditioista ja standardeista piittaamatta. Mm. media-alalla on tavallista, että sisällönkuvailua tuottavat myös muut kuin kirjastoalan ammattilaiset, ja tämä tuo sisällönkuvailun tukemisen haasteet myös lähemmäs ns. tavallisten tiedonhakijoiden haasteita. Ontologiaa rakennettaessa onkin mietittävä ratkaisuja, jotka tukevat perinteistä intuitiivisempaa ja myös näkökulmiltaan moninaisempaa ontologian käyttöä. Kirjastomaailman trendit noudattelevat semanttisen webin trendejä mm. siltä osin, että palveluiden tuottaminen on aiempaa yhteisöllisempää – ei niin paikallista, eikä aina kirjastoammattilaisen tekemää (kirjastoalan trendeistä ks. Almgren & Jokitalo, 2010, erit. s. 14-16). Paitsi että sisällönkuvailijana toimii aiempaa useammin ei-ammattilainen, niin myös verkossa kuvailtava aineisto on varsin heterogeenista. Esimerkiksi blogien tagittaminen on yleistä, mutta on huomattava, että niissä metadatan tuottaminen ei useinkaan tähtää informaation löydettävyyteen tai sitä ei osata toteuttaa tiedonhakua palvelevin keinoin (Nykyri & Niemi-Pynttari 2012).

Kun puhutaan asiasanoittamisesta verkkoproosan kohdalla, viitataan usein virheellisesti vain rajattuun osaan metadattaa. - Asiasanoitus (indeksointi, dokumentaatio) ymmärretään alan standardeissa (ISO 5963-1985) tekona, jossa kuvaillaan tai identifioidaan dokumentti sen asiasisällön perusteella. Asiasanat voidaan kuitenkin nähdä myös dialogisuuden muotona, intertekstuaalisuutena (vrt. Fairclough 1992), tai palvelevan muuta kuin asiasisällön kuvausta. Sosiaalisessa mediassa metadata voi välittää viestiä myös emotiivisesta tai poeettisesta kokemuksesta, ne ovat diskursiivisia kielen käytön tapoja. Asiasanojen käyttö niistä edellyttää reflektiivistä etäisyyttä, jota sosiaalisen median tuottajalla ei asiasanoittaessaan useinkaan ole, vaan hänelle sisällönkuvailu on osa juuri siinä hetkessä syntyvää ja itse kirjoitustilanteeseen liittyvää sisällönluomista. (Nykyri & Niemi-Pynttari 2012)

## Kielestä ja merkityksistä

Tiedonhakutilanteessa tiedonhakijan on pyrittävä muotoilemaan haluamansa informaatio niin, että se vastaa käytettävän tietojärjestelmän kuvausta. Sekä tiedonhakupyynnöt, että järjestelmien sisällönkuvailut perustuvat kieleen, merkityksiin ja ymmärtämiseen. (Blair 2006, 2) Ongelmallista on, että kielenkäytössä tapojen ja näkökulmien kirjo on moninainen. Wittgenstein rinnastaa kielen ”polkujen labyrinttiin”: yhdestä suunnasta lähestyttäessä reitti on tuttu, mutta toisesta suunnasta tullessa outo. (sit. tässä emt., 28-29)

Tiedonhakuympäristössä kieli ja kielenkäyttö, ja sitä kautta ymmärrys käsitteiden sisällöstä, vaihtelee yksilöiden välillä ja myös ajassa. Ns. valtakulttuurien välistä eroavaisuutta suurempi vaikutus voi olla alakulttuureista aiheutuvilla eroavaisuuksilla. Suuremmat eroavaisuudet esiintyvätkin todennäköisesti esimerkiksi suomalaisen sisällönkuvailijan ja suomalaisen tiedonhakijan kuin vaikkapa suomalaisen ja brittiläisen sisällönkuvailijan lähestymistavoissa. (Nykyri 2010)

Mitä eri merkityksiä asiasanaston sanalla tai termillä voi sitten olla? Vehmas-Lehto (1999, 74) jakaa sanan merkitykset neljään päätyyppiin Svejcerin (1973) esittämää jakoa noudattaen: denotatiiviseen, konnotatiiviseen, pragmaattiseen ja kielensisäiseen. Jako denotatiiviseen ja konnotatiiviseen on yleinen perusjako – denotatiivisen ilmaistessa sanan päämerkitystä ja konnotatiivisen sivumerkitystä.

Hörmannin (1986) mukaan sanan denotatiivinen merkitys viittaa sanan tarkoittamaan objektiin, ja konnotatiivinen merkitys siihen mitä sana assosiatiiivisesti ehdottaa. Esimerkiksi kuun denotatiivinen merkitys voidaan määritellä olevan ”maan luonnollinen kiertolainen” ja konnotatiivinen merkitys olevan esim. kylmä, kaukainen, yksinäinen. (emt., 142) Eri kielissä, kulttuureissa ja diskursseissa sanan konnotatiivinen merkitys voi olla erilainen, vaikka sanat denotaatiotasolla viittaisivat samaan tarkoitteeseen (ks. Nykyri 2010, 31). Esimerkiksi etanan ranskankielinen vastine escargot assosioituu herkullisena, mutta saksan kielen Schnecke inhottavana ja limaisena. Suomen kielen etana on lähinnä hidas (Vehmas-Lehto 1999, 75). Vastaavasti sama konnotaatio voidaan eri kielissä ilmaista denotaatioltaan erimerkityksisin sanoin tai fraasein. Kun Suomessa vihainen vaimo voidaan stereotyyppisesti esittää

naisena joka pitelee kaulinta, niin Venäjällä se on vastaavasti nainen silitysraudan kanssa. Suomen keppiä ja porkkanaa olisi Venäjällä ruoskaa ja piparkakkua. (Emt.)

Pragmaattinen merkitys määräytyy puolestaan kielenkäyttötilanteesta - ilmauksen ja osallistujien välisessä suhteessa viestintätalanteesta. Suhde vaihtelee kahdesta syystä. Ensinnäkin kielenkäyttäjät eroavat toisistaan mm. syn-nyinperänsä, ikänsä, koulutustasonsa ja ammattinsa puolesta. (Esimerkiksi sivistyssanojen käyttöaste korreloi usein koulutustason kanssa.) Toiseksi kielenkäyttötalanteet vaihtelevat. Yleisö voi olla pieni tai suuri, tuttu tai tuntematon, viestintämuoto voi olla suullinen tai kirjallinen, viestinnän intiimiyysaste voi vaihdella hyvinkin tuttavallisesta erittäin muodolliseen, käytetty kieli voi olla luonteeltaan murretta tai yleiskieltä, jne. Pragmaattinen merkitys on sukua tylille – sana voi olla esimerkiksi vanhahtava, muodollinen, tai kuulua vaikka mainosdiskurssiin. (Vehmas-Lehto 1999, 75-76)

Kielensisäistä merkitystä kutsutaan myös kieliopilliseksi merkitykseksi. Sillä viitataan merkitykseen, joka liittyy kieliyksiköiden väliin suhteisiin. Kielensisäinen merkitys ilmaistaan mm. etu- tai loppuliittein, prepositioin, sanajärjestyksin jne. (Vehmas-Lehto 1999, 74-76)

Merkitysten ja näkökulmien moninaisuus on haaste sanastojen laadinnassa ja sisällönkuvailussa. Rowleyn (1988) mukaan sisällönkuvailun tulee kattaa erilaiset mahdolliset lähestymistavat kuvailtua aihetta kohtaan. Tästä aiheutuu kahdenlaisia ongelmia. Ensinnäkin sisällönkuvailijan tulee nimetä aiheet tehokkaasti kiinnittäen huomiota esimerkiksi synonymiaan (esim. salt vs. sodium chloride), lähisynonymiaan (esim. salaries vs. income), sanaliittoihin (esim. merchant ships) ja erilaisiin kieliopillisiin versioihin (esim. an encyclopaedia of bibliography vs. a bibliography of encyclopaedias). Toiseksi sisällönkuvailijan tulee löytää keino ilmaista aiheiden väliset suhteet eri hakustrategioiden kannalta. (Emt., 49-50)

Tyypillisiä ongelmia erityisesti suomen kielessä ovat monitulkintaisuus, taiputusmuodot, yhdyssanat ja johdetut sanat (Järvelin 1995, 165-169). Kontrolloidun asiasanaston käyttö sisällönkuvailussa voi säästää tiedonhakijan vaivaa tämänkaltaisissa luonnollisesta kielestä juontuvissa ongelmatilanteissa, sillä hänen ei tarvitse ottaa haussa huomioon käsitteen erilaisia ilmaisumuotoja. Lisäksi mahdollista käsittehierarkiaa voidaan käyttää laajentamaan tai supistamaan hakua. (Emt., 180-184)

Kontrolloitujen sanastojen käyttö voi tuoda mukanaan myös ongelmia kuten joustamattomuuden (asiasanat eivät kehity samassa tahdissa kuin se kieli jota dokumenteissa käytetään) ja asiasanojen keinotekoisien ja vieraan suhteen luonnolliseen kieleen. Pyrittäessä kontrolloituun sanastoon myös monien sivumerkitysten ilmaisu voi olla hankalaa. Luonnollisesti ongelmia esiintyy myös silloin, kun yhteisymmärrystä ei synny käytettävästä sanastosta ja eri tietokannoissa päädytään käyttämään eri sanastoja. (Järvelin 1995, 181, ks. myös Cleveland & Cleveland 1983, 46-52)

## Termeistä käsitteisiin

Erilaisia merkityksiä kantava sana on symboli, sovittu merkki, joka viittaa tiettyyn merkitysyksikköön eli käsitteeseen. Ogdenin ja Richardsin (1923) esittämä semanttinen kolmio havainnollistaa symbolin, käsitteen (thought or reference) ja tarkoitteen (referent) suhdetta (kuva 1).

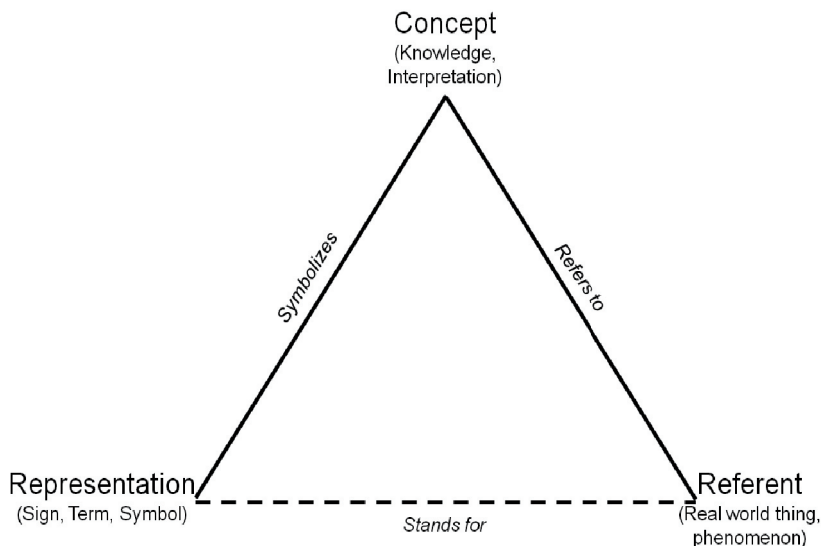
Karlsson (2004) pelkistää tämän puhumalla kielestä, mielestä ja maailmasta. Hänen mukaansa kielellä on fyysinen olomuoto (puhunnokset, viitonnokset, tekstit), mentaalinen olomuoto (representaatiot ym.), ja sosiaalinen olomuoto.

TSK (2006, 10) määrittelee käsitteen (concept) tiedon yksiköksi, joka muodostuu käsittepiirteiden ainutkertaisesta yhdistelmästä:

”Käsitteitä käytetään tiedon jäsentämiseen. Käsitteet eivät välttämättä ole kielisidonnaisia, mutta yhteiskunta- ja kulttuurisidonnaiset tekijät vaikuttavat usein erilaisen käsitejaon syntymiseen eri kielissä. Kaikille käsitteille ei välttämättä ole vakiintunut niitä vastaavaa kielellistä ilmausta.” (emt.)

Tarkoite voi olla konkreettinen (esine, ihminen tms.) tai abstrakti (ominaisuus, tapahtuma tms.). Ihminen muodostaa tarkoiteista käsitteitä eli siis ajatuskokonaisuuksia. Käsitteet voidaan edelleen jakaa yksilökäsitteisiin ja yleiskäsitteisiin. Yksilökäsite vastaa vain yhtä tiettyä tarkoitetta. Yleiskäsite syntyy yhdistämällä yksilötarkoitteiden yhteiset ominaisuudet yhteiseksi käsitteeksi. Viestintää varten ihminen muodostaa käsitteestä termin, nimen tai symbolin. YSO koostuu pääosin yleiskäsitteistä (esim. skandaalit), mutta mukana on myös yksilökäsitteitä (esim. Watergate).

Perinteiset tesaurukset toimivat enemmän kielellisellä symbolitasolla (termi, representation), kun taas ontologiat käsitetasolla (käsite, merkitys, mieli). Voidaan ajatella, että termitasolta käsitetasolle siirryttäessä tietyt kielisidonnaiset



Kuva 1: Semanttinen kolmio (Ogden ja Richards 1923)

ongelmat siirtyvät sivummalle ja osittain korvautuvat käsiteanalyysillä. Ontologian sisällön suunnittelussa tämä merkitsee perinteistä tesaurusajattelua tarkempaa semanttista tarkkuutta.

Jakoa käsitteen ja tarkoitteen välillä voidaan valaista esimerkin kautta. Teoksensa Kauneuden historia johdannossa Umberto Eco (2008, 14) toteaa, että ko. kirja

--- ”lähtee liikkeelle periaatteesta, että kauneus ei koskaan ole ollut absoluuttista ja muuttumatonta, vaan se on omaksunut erilaiset kasvot eri aikakausina ja eri maissa. Tämä ei pidä paikkaansa vain fyysiseen (miehen, naisen tai maiseman) kauneuteen nähden, vaan myös siihen, mitä tulee

jumalan, pyhimysten tai aatteiden kauneuteen ja moneen muuhun niiden lisäksi.” (emt.)

Econ (2008) mukaan se, mitä pidetään kauniina, riippuu aikakaudesta ja kulttuurista. Kauniin ja hyvän välillä on kuitenkin ollut läheinen yhteys monina historian aikakausina. (Emt. s. 8, 10)

YSAssa (VESA 10.6.2013) asiasana kauneus (ruotsinkielinen vastine skönhet) on asiasanan arvot suppeampi käsite, rinnakkaistermeinään esteettisyys ja rumuus, ja se kuuluu ryhmään [02] Filosofia. Absoluuttisen ”oikeaa” paikkaa ei tällaisille käsitteille ole – ’kauneus’ voitaisiin myös sijoittaa esimerkiksi ominaisuuksiin.

Myös YSOssa kauneus on esitetty arvona:

▼ **kauneus** ★

---

**Hierarkia:** [yso-käsitteet](#)

- └ [pysyvä](#)
- └ [ilmiöt](#)
  - └ [mentaaliset ilmiöt](#)
    - └ [arvot \(ilmiöt\)](#)
      - └ **kauneus** 12 vieruskäsitettä
      - └ [altruismi](#)
      - └ [epäoikeudenmukaisuus](#)
      - └ [hyöty](#)
      - └ [ihmisarvo](#)
      - └ [itseisarvo](#)
      - └ [luontoarvot](#)
      - └ [oikeudenmukaisuus](#)
      - └ [rehellisyys](#)
      - └ [sankaruus](#)
      - └ [tasa-arvo](#)
      - └ [viisaus](#)
      - └ [ympäristöarvot](#)

---

**Lähikäsite:** [esteettisyys](#)  
[rumuus](#)

---

**Tyyppi:** [YSO-käsite](#)

---

**URI:** <http://www.yso.fi/onto/yso/p18118>

---

**Nimikkeet ja vastaavat käsitteet:** [skönhet \(sv\)](#)  
[beauty \(en\)](#)

[skönhet \(vastaava käsite\)](#)  
[Y95995\\* \(vastaava käsite\)](#)

Kuva 2: 'kauneus' YSOssa (onki.fi-versio 11.6.2013)

Voidaankin ajatella, että käsitteenä (concept, mieli) ‘kauneus’ on staattisempi, kuin sen ilmentymät (referent, maailma). Kauneuteen voidaan melko pysyvästi liittää sellaisia arvoja kuin hyvyys ja miellyttävyys, mutta niiden ilmenemismuodot (referent, maailma) muuttuvat taustalla olevia arvoja nopeammin (ks. kuva 1, semanttinen kolmio –malli).

Käsitetasolla toimiessa ontologiassa on tiettyjä etuja suhteessa asiansanastoihin. Esimerkiksi kielen termin tason muutoksiin voidaan mahdollisesti reagoida nopeammin ja muutokset voivat olla vähempitöisiä. Kun esimerkiksi YSA:n päätetään ottaa uusi termi tietokoneohjelmat vastaamaan käsitettä johon aiemmin on viitattu termillä atk-ohjelmat, täytyy sanastoon luoda kokonaan uusi tietotyypin ja viittaukset vanhoihin termeihin. Ontologiassa kyseinen operaatio hoituisi ainoastaan suomenkielisen termin muuttamisella käsitteen pysyessä samana. Lisäksi ontologian monikielisyys ja tietokoneystävällisyys voivat osaltaan madaltaa erilaisten toimijoiden kynnystä päästä yhteisymmärrykseen saman sanaston käytöstä.

## Ontologian funktionaalisuudesta

Sisällönkuvailun ja tiedon tallennuksen nähdään yleensä palvelevan tulevaa tiedon hakua. Asiaa voidaan lähestyä myös muista näkökulmista, kuten Suominen (2004) tekee erottaessaan käsitteellisesti myöhempää hakua varten tehtävän sisällön kuvauksen ja metadokumentaarisen sisältörekonstruktion. Hänen mukaansa jälkimmäistä määrittävät ensisijaisesti olemassa olevat dokumentaariset sisällöt, edellistä ennakoitavat haut ja tarpeet. (emt., 17)

Sisällönkuvailun tutkimuksessa ei ole useinkaan kiinnitetty huomiota käsitetason ilmiöihin. Iivonen ja Kivimäki (1998, 91) toteavat, että valitettavasti useimmat tutkimukset keskittyvät termeihin eivätkä käsitteisiin, ja niin tekevät myös useimmat sisällönkuvailusäännöt. Voidaankin ajatella, että kun ontologioihin siirryttäessä sisällönkuvailuvälineiden kehittämisessä siirrytään käsitetasolle, myös itse sisällönkuvailuprosessissa ja sen ohjeistuksessa käsitetaso tulee korostumaan.

Tesaurusia on perinteisesti laadittu ihmisten lähestymistavoista lähtien. Tällöin termien semantiikan kuvaus on voitu jättää osittain ihmisen tulkinnan varaan. Yhä enenevässä määrin tesaurusia käyttävät kuitenkin tieto-

konesovellukset, ja mm. semanttinen web ja web-palvelut. Tietokoneelta puuttuu ihmisen piilotietämys ja tulkintakyky, joten termien merkitys tulee määritellä esim. semanttisen webin sanastoissa aiempaa täsmällisemmin ontologioina. (Hyvönen 2005)

Ontologioiden rakentaminen perustuu pitkälti käsitteiden sisällön määrittelyyn ja käsitteiden ryhmittelyyn. Ryhmittelyssä pätevät saman keskeiset periaatteet kuin luokituksessa. Perusajatuksena luokituksessa on, että luokitaja on tiedon välittäjä dokumentin ja tietoa tarvitsevan välillä (Meriläinen 1984, 3). Luokitusjärjestelmän pohjana on fasettianalyysi, jossa termit ryhmitellään sisäisesti homogeenisiin ja toisensa poissulkeviin ryhmiin yhden erottavan piirteen perusteella (emt., 46-48). Luokitusta käytetään paitsi organisoimiseen ja tiedonhaun apuvälineenä, myös sivistyksellisiin tarkoituksiin (ks. esim. Suominen 1984, 8-15).

Luokitusjärjestelmät ja sisällönkuvailuun tarkoitetut sanastotyökalut perustuvat yleensä tiettyyn aineistoon. Niiden korpus on se aineisto, jota on niiden luokituksessa ja asiansanoittamisessa käytetty. Näin on myös Yleinen suomalainen asiansanasto rakennettu, ja sen päivitys perustuu etupäässä siihen, millaisia termiehdotuksia sisällönkuvailijat tekevät uusien aineistotarpeiden myötä.

Sisällönkuvailun sanastovälineen tulisi yleisten periaatteiden mukaan täyttää ainakin kaksi pääasiallista takuuta. Nämä takuut on NISO (1994, sit. tässä Hjørland 2007) määritellyt seuraavasti:

- Kirjallisuustakuu (literary warrant): asiansanasto perustuu alan kirjallisuudessa esiintyviin sanoihin ja ilmauksiin. Kun kahdella tai useammalla vaihtoehdoisella tavalla ilmaista käsite on kummallakin kirjallisuustakuu, tulisi näistä useammin esiintyvä valita asiansanaksi.

- Käyttäjätakuu (user warrant): se miten käsitteen edustamaa informaatiota voidaan ajatella tarvittavan ja haettavan tulisi määrätä mikä käsitettä edustava esitystapa valitaan suosituksi termiksi.

Takuuta – tai perustoja, kuten Suominen (2004) niitä kutsuu – on muitakin, ja niitä ovat koonneet mm. Beghtol (1986) ja Suominen (2004). Pääasiallisia ovat tieteellis-filosofinen perusta (scientific-philosophical warrant), kasvatuksellinen perusta (educational warrant), ja kulttuurillinen perusta (cultural warrant), jonka alla on hakuperusta (enquiry warrant). (ks. Suominen 2004, 17-19, Hjørland 2007)

Tiedon haun ja tallennuksen tarpeisiin laadittu ontologia ei siis ole kuva ”tosiasiallisesta kokonaisuudesta”, vaan sen sisältö määräytyy luonnollisemmin kahden edellä kuvatun pääasiallisen takuun kautta – kirjallisuus- ja käyttäjätakuun. Sen vuoksi voidaan ajatella, että on luontevaa, jos ontologiassa ei vastoin yleistä tietoa ole viikonpäivien alla lueteltuina seitsemää viikonpäivää, vaan pelkästään esim. lauantai ja sunnuntai – mikäli erityisesti niihin kohdistuvaa kirjallisuutta on olemassa ja tiedonhakijoiden voidaan olettaa em. käsitteillä informaatiota hakevan, kun taas torstai olisi irrelevantti.

Ontologian laatija on kuitenkin haastavan kysymyksen edessä – miten paljon ontologiassa pitäisi tarjota myös kontekstuaalista tietoa ja millaisella koherenssilla? Ratkaistava dilemma on ontologioissa näkyvämpi kuin tesauksissa, sillä ontologiassa kaikki käsitteet ovat suhteessa toisiinsa, ja kontekstuaaliset puutteet voivat aiheuttaa enemmän vahinkoa kuin esim. perinteisessä tesaurusympäristössä.

Tiedon hakua ja tallennusta varten laadittavan tesauksien ja ontologian tulee olla paitsi semanttisesti looginen, myös pragmaattisesti – sekä kontekstuaalisesti että funktionaalisesti tarkasteltuna. Esimerkiksi YSON nykyversiossa (12.7.2013) tutkimusmenetelmät ja tutkimus (esim. narratiivinen tutkimus) ovat eri puolilla ontologiaa ja sen hierarkiaa. Logiikkana em. sijoittelussa on se, että koska sana on muodossa ”-tutkimus” eli on johdettu verbistä, se on laitettu toiminnan alle. Samaten esimerkiksi estetiikka (filosofian alakäsite) on ilmiö ja poetiikka (kirjallisuudentutkimuksen alakäsite) on prosessi. Tämä voi olla sisällönkuvailu- ja tiedonhakutilanteessa hämmentävää - etenkin, mikäli assosiaatiosuhteisiin ei ole panostettu kyllin.

Ontologian funktion hämärtyminen voi johtaa tiedonhaussa toisenkinlaisiin toimimattomiin ratkaisuihin, kun ajatellaan tiedonhaun laajentamista tai supistamista ontologian hierarkiarakenteeseen perustuen. Tällaista tapausta edustaa YSON nykyversiossa (11.6.2013) esimerkiksi käsite siirto. Se on ns. monimerkityksinen käsite, ja se on YSOssa kolmessa eri merkityksessä – täsmenteinään kooste, liikuttaminen ja viestintä. Liikuttamismerkityksisen alta sen alakäsitteinä ovat alkionsiirto, elimenssiirto, energiansiirto, hiustenssiirto, ksenotransplantaatio, kudoksenssiirto, maanssiirto, teleportaatio, tiedonsiirto, verenssiirto

ja voimansiirto. Hiustenssiirrosta kiinnostuneen voidaan kuitenkin olettaa olevan maanssiirtoa kiinnostuneempi vaikkapa kosmeettisesta kirurgiasta.

Tiedonhakuun ja -tallennukseen suunniteltavan ontologian tulisikin onnistua ratkaisemaan em. kaltaiset semanttisesti toimivat, mutta ontologian tarkoitustiedon kannalta toimimattomat ratkaisut.

## Ontologioiden käytön erityishaasteista kirjasto- ja muistiorganisaatioissa

Semanttisen webin kehitys on tähän mennessä edennyt paljolti tietoteknisten ambitioiden ja innovaatiopaineiden tahdittamana, kun taas teknisen keskustelun ulkopuolella on pohdittu olemassa olevan teknologian hyötyjä ja sovellusmahdollisuuksia. Kirjastoalalla sovellusmahdollisuudet ovatkin hieman vaikeammin määriteltävissä kuin esimerkiksi yrityssectorilla.

Semanttisen webin osalta kritiikkiä ovat saaneet ontologioihin melko yleisesti liitettävät näkemykset siitä, että ontologia on a) objektiivinen, b) staattinen ja c) kieliriippumaton reaali maailmaa kuvaava käsitteistö. Edellä on jo puhuttu semantiikan ja luokittelun haasteista, joiden valossa kukin esitetyistä ontologian määreistä voidaan todeta paikkansapitämättömäksi: ontologia hahmottaa tietyn hetken maailmaa senhetkisillä käsitteillä jonkin kielen, kielialueen ja kulttuurin kautta. (Cramer 2007, Pietarinen 2003, Wyssusek 2004).

Toisaalta on huomioitu, että ihmistieteissä merkitys niin itsessään kuin yksittäisten käsitteidenkin tasolla määrittyy paljon monimutkaisemmin kuin monissa muissa semanttisen webin sovellusaloissa kuten verkkokaupankäynnissä tai lääketieteessä. (Veltman 2004, Weller 2007). Yksi semanttisen webin ideologian kulmakivistä on looginen päättely: tässä oletuksena on, että kone pystyy käsittelemään ontologian ja siihen kytkeytyvän järjestelmän sisältämää dataa ja tarjoamaan päättelyn kautta uutta tietoa käyttäjälle. Päättelystä on kuitenkin enimmäkseen hyötyä joko ontologian ylläpitovaiheessa, jolloin ontologiasta voidaan etsiä loogisia virheitä tai tilanteissa, joissa ontologia ja mahdollisesti siihen liitetty tietokanta sisältävät runsaasti päättelyllisesti hyödynnettävää tietoa (ikä, paikkatieto, hinta



tms), johon voidaan esim. kohdistaa erilaisia hakutoimintoja (Shirky 2003, Zöllner-Weber 2009, Oberle 2009). Kontrolloidulla sanastolla tehty sisällönkuvailu ei kuitenkaan tuota samalla tavalla laskennallisesti hyödynnettävää dataa kuin esim. paikkatieto tai hintatieto. Kirjastoalalla haasteena onkin selvittää, miten metatieto voitaisiin saattaa paremmin konepäättelyn piiriin. Keskeisiä ratkaistavia kysymyksiä ovat käsitteen sijainti hierarkiassa ja käsitteen assosiativiset suhteet. Geneerisen hierarkiaan sijoittamisessa on päätettävä mikä valitaan käsitteen sisällön keskeisimmäksi merkittäväksi semanttiseksi komponentiksi. Keskeinen komponentti määrää käsitteen paikan ontologian hierarkiassa. Valinnan tulee tukea mahdollisimman laajasti käyttäjiensä intuitiivista päättelyä ja heijastaa ontologialla kuvattua aineistoa. Toinen sisällöllisesti keskeinen haaste on miten tuotetaan relevantteja - ja siten myös ajassa eläviä - assosiaatiosuhteita.

#### Johtopäätöksiä

Perinteinen ontologia-ajattelu on selkeästi geneeriseen käsittehierarkiaan perustuvaa. Yksinomaan semanttisiin ominaisuuksiin pohjautuvassa ajattelussa on kuitenkin riskinsä, kun ajatellaan ontologian toimivuutta tiedonhaun ja -tallennuksen apuvälineenä.

Tiedonhaun ja -tallennuksen tarpeisiin tehtävä ontologia ei tekovaiheessaan ja ylläpidossa säästy luonnollisen kielen ominaispiirteiden haasteilta. Ontologioiden vahvuutena pidetään semanttista tarkkuutta, mutta asialla on myös toinen puoli. Mikäli kulttuurialalla - ja erityisesti kirjastoalalla - siirrytään asiasanastoista ontologioihin, on otettava huomioon, että ontologioilla saattaa olla toisenlaisia tarpeita kuin muilla sovellusaloilla, joissa ontologioiden käytöstä on enemmän kokemusta. Jos ontologiaa aiotaan käyttää pääasiassa dokumenttien sisällönkuvailuun ja tiedonhaakuun, sen rakentamisen pitäisi tapahtua nimenomaan näiden käyttötärpeiden ehdoilla. Ontologiamallille voidaan määrittää toisenlaisia vaateita kuin perinteisesti semanttisessa webissä korostetut looginen täsmällisyys, käsitteiden eksplisiittisyys tai periytyvien ominaisuuksien runsaus. Ihmistieteissä käsitteet ovat merkitykseltään, käytöltään ja konteksteiltaan laajalaisia ja dynaamisia: ontologian pitäisi tukea tätä, ei olla sen esteenä. Ontologian kehitystyö edellyttää siis paneutumista niin laadinnassa kuin päivittämisessä, jopa tesauksia enemmän, mutta sen ansiot ovat vastaavasti suuremmat tiedonhaussa ja -tallennuksessa.

Miten edistää ontologiatyössä semanttisuuden loogisuutta käsitetasolla niin, että se palvelee sisällönkuvailua ja tiedonhakuja? Koska ontologiat perustuvat tesaurus- ja asiasanastomaailmaa korostetummin nimenomaan käsitetasoon, ja ominaisuudet voivat periytyä erittäinkin hierarkkisesti tarkasteltuna usean tason takaa, tulisi käsiteanalyysin olla erityisen huolellista. Oleellista onkin pohtia, mitkä ovat ne keskeiset semanttiset piirteet, joiden katsotaan olevan hierarkiaketjussa oleellisia, ja tiedonhaussa merkittäviä.

Käytännön ontologiatyö on informaatiotutkimuksen alalla vielä kovin nuorta, ja siksi se nostaakin esiin enemmän kysymyksiä kuin vastauksia.

- Miten tehdä käsiteanalyysi ja millainen hierarkiamalli tulisi valita?

- Missä määrin ontologiatyö lähtee terminologisista periaatteista ja miten paljon pragmaattisista käsin? Miten taataan käyttäjä- ja kirjallisuustakuu?

- Miten käsiteperustaiseen sanastoon siirtyinen vaikuttaa sisällönkuvailuun?

- Missä määrin merkitysten kontekstisidonnaisuus voidaan ottaa huomioon? Mikä on eksaktissa toimintaympäristössä suhtautuminen relativismiin?

Tiedonhaakuun ja -tallennukseen laadittavan ontologian rakentamisen haasteet ovat kahdella, osin ristiriitaisilla ja jo yksinään varsin vaativilla tasoilla – ontologiassa tulee päästä semanttiseen tarkkuuteen ja sen on oltava toimiva tiedonhaun ja -tallennuksen apuvälineenä. Suurimpia haasteita ihmistieteiden osalta on, miten saada ontologia vastaamaan ihmisen intuitiivista tapaa lähestyä asioita ja tukea sanaston temaattista käyttöä. Lisäksi voidaan kysyä, että mitä webin diskursseja ja metadatan moninaistuvia muotoja ontologioiden voidaan ajatella tulevaisuudessa palvelevan. - Mitkä erittäin heterogeenisen web-materiaalin sisällönkuvailun ja tiedonhaun haasteista tulisi ratkaista ontologian puolella ja mitkä hakujärjestelmissä esimerkiksi aineistotyypin valinnan ja kohdentamisen kautta.

Hyväksytty julkaistavaksi 4.12.2013

## Viite:

<sup>1</sup>Kansalliskirjaston ONKI-projektin alkuvaiheessa, 03/2013-10/2013, ontologia-palvelusta käytettiin Aalto-yliopiston projektin aikaista nimitystä ONKI. Finto nimitys, Finnish thesaurus and ontology service, on otettu käyttöön 11/2013.

## Lähteet

- Almgren, Päivi & Jokitalo, Päivi 2010: Kirjasto 2011 – Vaikutteita maailmalta. BTJ Finland Oy, Vantaa.
- Beghtol, C. 1986: Semantic validity: Concepts of warrant in bibliographic classification systems. *Library Resources & Technical Services*, 30(2), 109-125. URL: <http://downloads.alcts.ala.org/lrts/lrtsv30no2.pdf> (7.6.2013)
- Blair, David 2006: Wittgenstein, Language and Information : back to the rough ground. Dordrecht : Springer.
- Cleveland, Donald & Cleveland, Ana 1983: Introduction to Indexing and Abstracting. Libraries Unlimited, Inc, Littleton, Colorado.
- Cramer, Florian 2007: Animals that Belong to the Emperor – Failing universal classification schemes from Aristotle to the Semantic Web. Luentokäsikirjoitus, Quaero Forum, Jan van Eyck Academy, Maastricht. URL: <http://www.nettime.org/Lists-Archives/nettime-l-0712/msg00043.html> (12.8.2013)
- Eco, Umberto 2008: Kauneuden historia. Helsinki: WSOY.
- Fairclough, Norman 1992: Discourse and Social Change. Cambridge, England: Polity Press.
- Hjørland 2007: Literary warrant (and other kinds of warrant). Core Concepts in Library and Information Science (LIS). URL: [http://www.iva.dk/bh/lifeboat\\_ko/concepts/literary\\_warrant.htm](http://www.iva.dk/bh/lifeboat_ko/concepts/literary_warrant.htm) (6.6.2013)
- Hörmann, H. 1986: Meaning And Context. An Introduction to the Psychology of Language. (Toim. ja johdanto Robert E. Innis.) New York: Plenum.
- Hyvönen, Eero 2005: Miksi asiasanastot eivät riitä vaan tarvitaan ontologioita? URL: <http://www.seco.tkk.fi/publications/2005/hyvonen-miksi-asiasanastot-eivat-riita-2005.pdf> (12.8.2013)
- Iivonen, Mirja & Kivimäki, Katja 1998: Common Entities and Missing Properties: Similarities and Differences in the Indexing of Concepts. *Knowledge Organization* 25(3), pp. 90-102.
- Järvelin, Kalervo 1995: Tekstiedonhaku tietokannoista : Johdatus periaatteisiin ja menetelmiin. Suomen ATK-kustannus, Espoo.
- Karlsson, Fred 2004: Yleinen kielitiede. Yliopistopaino, Helsinki.
- Meriläinen, Outi 1984: Perustiedot kirjastoluokitusten ymmärtämiseksi. Kirjastopalvelu Oy, Helsinki.
- NISO 1994: National Information Standards Organisation (1994). ANSI/NISO Z39.19-1993 Guidelines for the construction, format and management of monolingual thesauri. Bethesda, MD: NISO Press.
- Nykyri Susanna 2010: Equivalence and Translation Strategies In Multilingual Thesaurus Construction. Åbo Akademis förlag, Åbo.
- Nykyri Susanna & Niemi-Pynttari Risto 2012: Päiväkirjablogien metadatan funktioista. *Informaatiotutkimus*. Vol 31, Nro 3 (2012).
- Oberle, Daniel 2009: How Ontologies Benefit Enterprise Applications. *Semantic Web Journal*. URL: [http://www.semantic-web-journal.net/system/files/swj212\\_2.pdf](http://www.semantic-web-journal.net/system/files/swj212_2.pdf) (12.8.2013)
- Ogden, C. K. and I. A. Richards, I. A. (1923). *The Meaning of Meaning : A Study of the Influence of Language Upon Thought and of the Science of Symbolism*. London: Routledge & Kegan Paul.
- Ooi, K. & Liew, C. L. 2011. Selecting fiction as part of everyday life information seeking. *Journal of Documentation* 67 (5), 748–772.
- Pietarinen, Ahti-Veikko 2003: The Semantic + Pragmatic Web = The Semiotic Web. Proceedings of the IADIS International Conference WWW/Internet 2003, ICWI 2003, Algarve, Portugal, Marraskuu, 2003. 981-984. URL: <http://www.helsinki.fi/~pietarin/publications/The%20Semiotic%20Web-Pietarinen.pdf> (12.8.2013)
- SFS 5471 (1988). Suomenkielisen tesaauruksen laatimis- ja ylläpito-ohjeet. Suomen Standardisointiliitto SFS ry.
- Shirky, Clay 2003: The Semantic Web, Syllogism, and Worldview. Blogi-kirjoitus. [[http://www.shirky.com/writings/herecomeseverybody/semantic\\_syllogism.html](http://www.shirky.com/writings/herecomeseverybody/semantic_syllogism.html)]
- Suominen, Vesa 1984: Luokitusopas ja vaihtoehtoiset luokitusperusteet. Tampereen yliopisto, Tampere.
- Suominen, Vesa 2004: Grammaattis-luokituksellinen johdatus dokumentaaristen sisältöjen rekonstruktioon ja tuntemukseen: Metodologia, metodiikka ja illustraatio. - *Finnish Information Studies* 22. Oulu, Oulun yliopisto. 145
- TSK 2006 = Sanastokeskus TSK ry 2006: Termino-

- logian sanasto (TSK 36). URL: <http://www.tsk.fi/tiedostot/pdf/TerminologianSanasto.pdf> (3.6.2013)
- Vehmas-Lehto, Inkeri 1999: Kopiointia vai kommunikointia : Johdatus käännteoriaan. Oy Finn Lectura Ab, Helsinki.
- Veltman, Kim 2004: Towards a Semantic Web of Culture. *Journal of Digital Information* . Vol 4, No 4. URL: <http://journals.tdl.org/jodi/index.php/jodi/article/view/113/112> (12.8.2013)
- Weller, Katrin 2007: Folksonomies and Ontologies: Two New Players in Indexing and Knowledge Representation. Teoksessa H. Jeppard (toim.), *Applying Web 2.0: Innovation, Impact and Implementation. Proceedings of the Online Information Conference, London, Great Britain* (pp. 108–115). London: Learned Information Europe Ltd.
- Wyssusek, Boris 2004: Ontology and Ontologies in Information Systems Analysis and Design: A Critique. *Proceedings of the Tenth Americas Conference on Information Systems, New York, USA*. Elokuu, 2004. 4303-4307.
- Zöllner-Weber 2009: Ontologies and Logic Reasoning as Tools in Humanities? *Digital Humanities Quarterly*. Vol 3, No 3. URL: <http://www.digitalhumanities.org/dhq/vol/3/4/000068/000068.html> (12.8.2013)