

Metatiedot organisaatioiden sisällönhallinnassa

Airi Salminen

Tallennetun tiedon hallinnassa tarvitaan aina metatietoa. Metatietoa on esimerkiksi tieto siitä mille tietovälineille tiedot on tallennettu, miten ne ovat saatavilla ja miten ne on organisoitu. Ilman systemaattisesti koottuja ja tallennettuja metatietoja tietovarannot ränsistyvät ja tulevat vähitellen käyttökelvottomiksi tai ainakin vaikeasti käytettäviksi. Omaan kirjahyllyymme ja työpöydällemme kertyvistä kirjoista, monisteista ja paperilappusista saatamme löytää kokemuksemme ja asiantuntemuksemme perusteella etsimämme tiedot vain omia aistejamme käyttäen. Henkilökohtaistenkin tietovarantojen hallinnassa toisenlaisia toimintatapoja tarvitaan, mikäli tiedot on tallennettu digitaaliseen muotoon. Laajan yhteisön käyttöön tarkoitettun digitaalisen tiedon hallintaan tarvitaan ihmisten aistien, kokemuksen ja asiantuntemuksen lisäksi monipuolista metatietoa ja ohjelmistoja, jotka osaa- vat käyttää tuota metatietoa.

Metatieto

Metatieto on tietoa tiedosta. Mihin tahansa tallennettuun tietokokonaisuuteen voidaan liittää metatietoa. Metatietojen systemaattinen tuottaminen ja käyttö liittyy aina johonkin *tietovarantoon*. Esimerkkejä tietovarannoista, joihin liitetään metatietoa ovat:

- kirjaston julkaisukokoelma
- Internetissä osoitettavissa olevat resurssit
- Internetin HTML-sivut
- dokumenttien hallintajärjestelmän dokumentit
- suomalaisen julkishallinnon lomakkeet
- yrityksen tietovarasto (data warehouse)
- yrityksen tietokanta (database)

Koska tietovarannot ovat hyvin erilaisia, on myöskin niihin liittyvissä metatiedoissa suuria eroja. Nykyisissä monimutkaisissa, verkottuneissa toimintaympäristöissä on aina monentyyppisiä tietovarantoja. Siitä syystä

myöskin metatietoja on hyvin erilaisia, eri tahoilla luotuja ja eri tavoin käytettävissä olevia. Puhuttaessa metatiedosta tietyssä ympäristössä onkin aina syytä tarkentaa, minkä tietovarannon metatiedoista puhutaan.

Tietovarantoon liittyvä metatieto kuvaa aina jotain kohdetta. *Kohde* voi olla koko tietovaranto, sen osa tai jokin asia joka liittyy tietovarantoon tai sen hallintaan. Yhteen kohteeseen saattaa liittyä monenlaista, eri tahoilla luotua metatietoa. Metatieto voi olla olemassa myös sellaisesta kohteesta, joka on ollut olemassa, mutta jota ei enää ole olemassa. Esimerkiksi viraston kirjaimossa voi olla tietoja asiakirjoista, jotka ovat hävinneet tai jotka on tuhottu. Taulukossa 1 on esitetty esimerkkejä metatiedon kohteista, kun tietovarantona ovat Internetin tietoresurssit.

Taulukko 1. Metatietoja ja metatiedon kohteita Internetin tietoresursseihin liittyen.

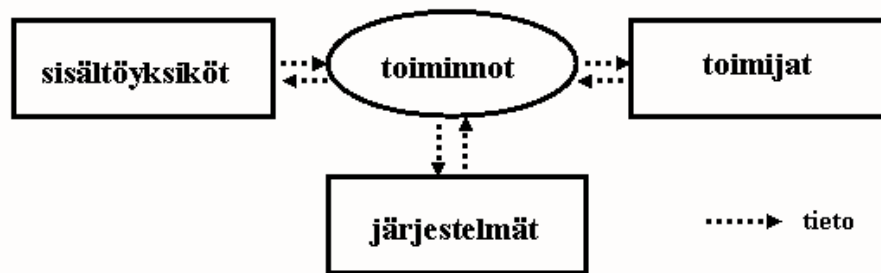
Kohde	Esimerkkejä metatiedosta
Internetin kaikki tallennetut tietoresurssit	- Google-hakukoneen indeksi
Internetin HTML-sivut	- HTML-kielen spesifikaatio
HTML-sivu	- Sivuuun sisältyvä META-elementti - Sivuuun sisältyvä META-elementtien joukko - Sivun kuvaus digitaalisen kirjaston hakemis- tossa
Suomalaisen julkishallinnon Internetissä tarjoamat palvelut	- Suomi.fi-portaalin tiedot
Internetin teknologia	- http-yhteyksikäytännön kuvaus - XML-teknologian kuvaus
Internetin kehittämistä koordinoivat tahot	- Kuvaukset W3C-organisaation toiminnasta - Kuvaukset Internet Society:n toiminnasta

Metatietoja tallennetaan metatietokuvauksiin. *Metatietokuvaus* kuvaa jotain esitystapaa käyttäen yhden tai useamman kohteen ominaisuuksia. Laajuudeltaan kuvaus voi vaihdella suuresti. Kuvaus voi koskea esimerkiksi yhden kohteen yhtä ominaisuutta (HTML-sivun tekijä), yhden kohteen useita ominaisuuksia yhdessä (HTML-sivun tekijä ja luomispäivä) tai sitten yhteen kuvaukseen on koottu monien kohteiden ominaisuuksia (esimerkiksi hakemisto, johon on koottu julkaisukokoelman julkaisutiedot).

Sisällönhallinta

Kun tiedonhallintaa organisaatioissa tarkastellaan tallennetun tiedon käytettävyyden näkökulmasta, voidaan puhua sisällönhallinnasta. Jotta tieto olisi organisoitavissa, manipuloitavissa ja käytettävissä täytyy se ryhmitellä tunnistettavissa oleviksi sisältöyksiköiksi. Kuva 1 esittää sisällönhallinnan

peruskomponentteja ja tiedonkulkua niiden välillä (Salminen, 2005). Soikio esittää kohdeympäristön toimintaprosessien *toimintoja*, katkoviivanuolet esittävät sitä tietoa, joka syntyy tai jota käytetään toiminnoissa. Suorakaiteet kuvaavat toimintoihin liittyviä kolmentyyppisiä tietoresursseja. *Toimijat* ovat toimintoja suorittavia organisaatioita ja ihmisiä. Internet-ympäristössä toimijat saattavat olla myös ihmisiä tai organisaatioita edustavia agenttiohjelmaa. *Järjestelmät* ovat laitteistoja, ohjelmistoja, standardeja ja sopimuksia, joita käytetään toiminnoissa. Nykyisissä verkottuneissa ympäristöissä järjestelmiin kuuluvat Internet, WWW, niihin liittyvät standardit, erilaiset intranet- ja ekstranet-järjestelmät sekä näiden verkostojen piirissä käytettävät ohjelmistot, laitteet ja standardit. *Sisältöyksiköt* ovat tallennettuja tietokokonaisuuksia, joita syntyy toiminnoissa ja joita käytetään toiminnoissa. Ne voivat olla esimerkiksi tietokantatietoja, tekstinkäsittelyohjelmilla syntyneitä dokumentteja, rakenteisia dokumentteja ja niiden osia, digitaalisia valokuvia tai erilaisia paperitallenteita.



Kuva 1. Sisällönhallinnan osa-alueet

Jotta sisältöyksiköt olisivat löydettävissä ja käytettävissä täytyy niihin liittää tavalla tai toisella metatietoa. Koska verkottuneissa toimintaympäristöissä on monenlaisia ja eri tahoilla luotuja sisältöyksiköitä, on siellä myöskin hyvin monentyyppistä metatietoa. Kaikki tietovarantoja käsittelevät ohjelmistot tallentavat metatietoja, joiden avulla ne pystyvät käsittelemään ja käyttämään omia tietovarantojaan. Ohjelmistokohtaiset metatiedot ovat usein sellaisessa muodossa, ettei niitä ole mahdollista hyödyntää ilman kyseistä ohjelmaa. **Kun organisaatioissa etsitään pitkäikäisiä, järjestelmäriippumattomia ratkaisuja, metatietoja täytyy tarkastella toimijoiden, toimintojen ja ohjelmistojen välisen yhteistyön näkökulmasta, ei yksittäisen ohjelmiston näkökulmasta. Toimintaympäristössä luotavat metatietokuvaukset pitää tallentaa sellaisiksi sisältöyksiköiksi, joita on mahdollista käyttää erilaisilla ohjel-**

milla, eri osallistujaorganisaatioissa ja käsitellä tietoresursseina niin kuin muutakin tallennettua tietoa.

Metatietojen käyttötapoja toimintaprosesseissa

Kun metatiedot tallennetaan digitaaliseen muotoon, ovat metatietojen käyttäjinä tietokoneet ja ihmiset. Tarkasteltaessa tiettyä organisaatioympäristöä ja sen toimintaprosesseja, voidaan metatietojen käyttö jakaa kahdelle tasolle. Toisaalta metatietoja tarvitaan ja käytetään sisältöyksiköiden, niiden kielen ja esitystavan hallintaan, toisaalta varsinaiseen asianhallintaan.

Sisältöyksiköiden hallinta. Metatietoa tarvitaan sisältöyksiköiden, esimerkiksi dokumenttien, organisoimiseen, tuottamiseen, ylläpitämiseen, säilyttämiseen, hakemiseen, tarjoamiseen, jakelemiseen ja tuhoamiseen. Sisältöyksiköiden hallintaan kuuluu myöskin sisällön kielen, rakenteen ja esitystavan hallinta. Sisältöyksiköiden hallinnassa tarvitaan yleensä myös metatietoa liittyen sisältöyksiköiden sisältämään tietoon, mutta itse toimenpiteet ovat sisältöyksiköiden hallintatoimenpiteitä. Näitä toimenpiteitä voivat ainakin osittain suorittaa ihmiset, jotka eivät ole varsinaisen asian käsittelyn toimijoita.

Asianhallinta. Näissä toimenpiteissä ensisijaisena tavoitteena on jotain muuta kuin sisältöyksiköiden käsittely. Sisältöyksiköt ja niihin liittyvä metatieto ovat apuvälineitä jonkin asian hallinnassa (esimerkiksi lainsäädäntöasian). Toimijat ovat asiankäsittelyn toimijoita. Metatietoa käytetään esimerkiksi tiedonhakuun, todistamaan tehtyjä toimenpiteitä tai tukemaan kommunikointia.

Näiden kahden tason erottaminen ei suinkaan ole kaikissa tilanteissa selvä, mutta esimerkiksi lainsäädäntöprosessin yhteydessä nämä kaksi tasoa erottuvat monissa toimenpiteissä. Esimerkiksi asiakirjojen kirjaaminen, Webjulkaiseminen ja painaminen ovat asiakirjahallintatoimenpiteitä kun taas lakitiedonhaku, asiantuntijatietojen haku, päätöspäivien haku ovat esimerkkejä asiankäsittelyyn liittyvistä toimenpiteistä. Tietokoneiden välistä kommunikointia ja metatiedon hyödyntämistä siinä voi tapahtua kummallakin tasolla.

Metatietoluokituksia

Metatietoja luokitellaan mitä erilaisimmilla tavoilla, esimerkiksi käyttötarkoituksen mukaan, jonkin metatiedon ominaisuuden mukaan tai kohteen ja sen ominaisuuden mukaan. Metatietoluokitteluja on vertailtu esimerkiksi Virpi Lyytikäisen väitöskirjassa (2004). Taulukkoon 2 on koottu metatiedon ominaisuuksia. Taulukko on koottu pitkälti artikkelista Gilliland-Swetland (2000). Kukin ominaisuus vaihtoehtoisine arvoineen on mahdollinen luokitteluperuste. Esimerkiksi pysyvyyden perusteella metatieto voidaan luokitella staattiseksi tai dynaamiseksi, tallennustavan mukaan voidaan metatiedot luokitella upotetuiksi tai ulkoisiksi ja organisoinnin perusteella hajautetuiksi tai keskitetyiksi.

Taulukko 2. Metatiedon ominaisuuksia.

Ominaisuus	Vaihtoehdot	Esimerkkejä
tuottaja	tietokone/ihminen	tekstisisällöstä muodostettu indeksi/asiakirjan kuvailutieto
tuottamisajankohta	yhdessä sisällön tuottamisen yhteydessä /jälkikäteen	kirjan julkaisija/kirjaston hakemisto
käyttäjä	tietokone/ihminen	XML-muotoinen metatieto/dokumenttiin kirjoitettu tiivistelmä
formaaliusaste	formaali / vapaamuotoinen	asiakirjan RDF/XML-kuvaus / asiakirjaan liitetyt vapaamuotoiset kommentit
pysyvyys	staattinen/dynaaminen	asiakirjan luomis-päivä/asiakirjan päivityspäivä
tallennus suhteessa kohteeseen	upotettu/ulkoinen	asiakirjaan kirjoitetut metatiedot/asiakirjaan liitetty erillinen metatietokuvaus
organisointi	hajautettu/keskitetty	kokoelman asiakirjoihin sisältyvät metatiedot/metatietotietokanta
kohteen taso	luokka/ilmentymä	dokumenttityypin kuvaus/asiakirjan kuvailutiedot

Metatiedon käyttötapojen tai taulukossa 2 esitettyjen ominaisuuksien perusteella tapahtuvan luokittelun lisäksi metatietoja voidaan luokitella sen mukaan, mitä metatiedolla kuvataan. Tällaisessa luokittelussa metatiedot voidaan kaikkein ylimmällä tasolla jakaa kahteen luokkaan: metatiedoilla kuvataan sisältöyksiköiden ominaisuuksia tai sitten niiden kontekstin ominaisuuksia.

sia. Tämän kahtiajaon ovat esittäneet esimerkiksi Hill ym. (1999). Sisältöyksiköiden ominaisuudet puolestaan jaetaan useissakin lähteissä kahteen luokkaan: sisällön merkitystä kuvaava semanttinen tieto ja sisällön rakennetta kuvaava tieto. Tällä tavalla päästään metatietojen kolmeen luokkaan, joita seuraavassa esitellään sisällönhallintaympäristössä. Kussakin näistä luokista metatiedolla voi olla edelleen taulukon 2 mukaisia ominaisuuksia.

Semanttinen metatieto on sisällön merkitystä kuvaavaa tietoa. Luokkatasolla esimerkkinä semanttisesta metatiedosta on tietovarannon sisältöjen käsitteistöä kuvaava ontologia. Ilmentymätasolla ontologiaa voidaan käyttää esimerkiksi asiakirjan asiasanoina. Myöskin asiakirjan nimeke, aihe ja tiivistelmä ovat esimerkkejä ilmentymätason semanttisista metatiedoista.

Rakennemetatieto on sisältöyksikön fyysistä tai loogista rakennetta tai sisällön kieltä kuvaavaa tietoa. Luokkatasolla tyypillisiä esimerkkejä rakennemetatiedosta ovat rakenteisten dokumenttien (esimerkiksi XML-muotoisten) dokumenttityyppimäärittelyt. Ilmentymätasolla viittaus rakennekuvauksen sisältävään dokumenttityyppiin, rakenteen merkkkaus jollain merkkaukielellä, tiedoston pituus, sisällön kieli ja tiedostoformaatti ovat esimerkkejä rakennemetatiedosta.

Kontekstuaalinen metatieto kuvaa sisältöyksiköiden ympäristöä jossain tietyssä tilanteessa. ISO:n kehitteillä olevassa asiakirjojen metatietostandardissa (ISO, 2003) erotetaan asiakirjojen tuottamisympäristökonteksti ja käyttöympäristökonteksti. Ympäristöjen toimintaprosesseissa puolestaan erotetaan asiakirjojen käsittelyprosessit ja liiketoimintaprosessit. Kontekstuaalisilla metatiedoilla voidaan kuvata näihin prosesseihin, niiden toimijoihin ja niissä käytettäviin järjestelmiin liittyviä tietoja. Myös tapahtumatiedot (esimerkiksi asiakirjan luomisaika), sisältöyksiköiden tunnisteet tietyssä ympäristössä ja sisältöyksiköiden suhteet muihin sisältöyksiköihin voidaan lukea kontekstuaaliseen metatietoon. Lisäksi kontekstuaaliseen metatietoon voidaan lukea tietosuojan, julkisuuteen ja käyttöön liittyvät tiedot, joille joissakin luokituksissa on varattu oma kontrollimetatiedon luokka (esimerkiksi Sami Jokelan väitöskirjassa, 2001).

Tietty metatieto ei suinkaan aina asetu vain yhteen edellä kuvatuista luokista. Esimerkiksi XML-dokumenttien rakennekuvauksien elementtimissä on ihmisen ymmärrettävissä olevaa semanttista tietoa dokumenttien sisällöstä.

Samoin dokumenttien tuottamisympäristön toimintaprosessin kuvaus saattaa olla edellytys prosessissa syntyneiden dokumenttien sisällön ymmärtämiseen.

Metatietojen standardointi

Metatietoja on standardoitu pitkään esimerkiksi kirjastoissa. Metatietojen laaja-alainen standardointi tuli välttämättömäksi siinä vaiheessa, kun metatietojen tallentaminen, käsittely ja käyttö siirrettiin tietokoneille. Erityisesti Internetin huimaa vauhtia kasvavat tietosisällöt ovat vaatineet sellaisten metatietostandardien kehittämistä, että tietokoneohjelmat pystyvät hyödyntämään metatietoa. Tietokoneiden tulkittavissa oleva metatieto on noussut tärkeään asemaan pyrittäessä rakentamaan nykyisestä webistä *semanttinen web*. Semanttinen web on visio tulevasta webistä, mikä entistä webiä huomattavasti paremmin palvelee yhteisöjen ja niiden käyttämien ohjelmien välistä yhteistyötä. Semanttisella webillä pyritään edistämään webin tietosisältöjen ja palvelujen käytettävyyttä kokoamalla erityisesti semanttista ja Internetin käytön turvallisuuteen liittyvää, standardoitua metatietoa.

Metatietojen standardointia varten on kehitetty lukuisia kansainvälisiä ja kansallisia suosituksia (joita kutsutaan usein myös standardeiksi). Niitä on mahdollista hyödyntää kehitettäessä metatietostandardia tietyn organisaatioympäristön sisällönhallinnan tarpeisiin. Internet-ympäristössä tärkeimpiä metatietojen standardoinnissa käytettäviä kansainvälistä pohjasuosituksia ovat XML ja RDF, jotka tarjoavat mallin ja syntaksin metatietojen esittämiseen sekä Dublin Core, joka kehitettiin metatietosuosituksiksi verkkojulkaisuille ja jota on laajalti käytetty perustana organisaatio- ja sovellusaluekohtaisten metatietoskeemojen kehittämisessä. Nämä suositukset esitellään lyhyesti seuraavassa. Niiden jälkeen esitellään suomalainen julkishallinnolle kehitetty suositus JHS 143.

XML

Internetin standardoinnin perustana on XML-kieli (Extensible Markup Language)¹. Sitä voidaan käyttää monella eri tavalla tietosisältöjen ja niihin liittyvien metatietojen esittämisessä. XML-kieli on suunniteltu erityisesti rakennemetatiedon ja sen mukaisten rakenteiden esittämiseen. XML:n dokumenttityyppimäärittelyllä (DTD) on mahdollista kuvata dokumenttiluo-

¹ XML-kielen versio 1.1: <http://www.w3.org/TR/2004/REC-xml-20040204/>.

kan rakenne. XML-kielellä voidaan myös kuvata muuntyyppisiä metatietoja. Semanttisen webin teknologiat² perustuvat pitkälle XML:n käyttöön metatietojen esitystapana.

RDF

RDF (Resource Description Framework)³ on W3C:n malli metatietojen liittämiseksi resursseihin. RDF:ssä resurssi on mikä tahansa Internetissä identifioitavissa oleva asia. Resurssien identifiointimekanismi on eräs RDF-mallin keskeisimpiä komponentteja. RDF:ssä identifiointi tapahtuu URI-tunnisteella (URI = Uniform Resource Identifier), joka voi olla joko WWW:stä tuttu, resurssin sijaintiin viittaava URL (Uniform Resource Locator) tai sitten yleisempi identifiointitunniste URN (Uniform Resource Name). Erityisesti pitkäaikaisen resurssien löydettävyyden varmistamiseksi yksikäsitteisestä, sijainnista riippumattomasta identifioinnista on tärkeää sopia samalla kun metatietoratkaisuja suunnitellaan. RDF-metatietokuvaukset koostuvat väittämistä, missä kohteeseen liitetään ominaisuus ja sen arvo. Samasta kohteesta voidaan esittää useita väittämiä ja ominaisuuden arvot voivat olla resursseja, joista esitetään uusia väittämiä. Näin RDF-metatietokuvauksista voi syntyä monimutkaisia resurssien ja niiden välisten suhteiden kuvauksia. Graafisesti RDF-kuvauksia esitetään solmuina ja nuolina. Kuvausten koneille ymmärrettävässä muodossa esittämiseen RDF-spesifikaatiossa on määritelty XML-syntaksi.

Dublin Core

Erityisesti Internetin rerurssien kuvaamiseen on kehitetty kansainvälinen metatietosuositus nimeltään Dublin Core⁴. Tavoitteena kehittämisessä oli parantaa Internetin tietoresurssien löydettävyyttä. Suosituksessa määritellään 15 metatietoelementtiä liitettäväksi Internetin tietoresursseihin. Elementteillä kuvataan esimerkiksi tietoresurssin nimeke (title), tekijä (creator), julkaisija

² Semanttisen webin tekniikoita on esitelty esimerkiksi Airi Salmisen luennossa *Web ja semanttinen web organisaatioissa*, <http://www.cs.jyu.fi/~airi/presentations/SemanttinenWeb-18112003.ppt>.

³ <http://www.w3.org/RDF/>

⁴ <http://dublincore.org/>

(publisher), identifiointitunnus (identifier) ja oikeustiedot (rights). Elementteihin on mahdollista liittää myös tarkenteita. Esimerkiksi elementtiä, jolla annetaan päivämäärä (date) on mahdollista tarkentaa niin, että päivämäärän merkitys täsmentyy tietoresurssin luomispäiväksi (date.created) tai julkistamisajankohdaksi (date.issued). Dublin Core -suositusta on käytetty pohjana erittäin monilla tahoilla julkishallinnon metatietostandardeja (metatietoskeemoja) kehitettäessä (European Committee for Standardization, 2003). Niihin on valittu kyseiseen ympäristöön sopivat Dublin Core -elementit ja tarkenteet sekä määritelty, mitkä elementeistä ovat pakollisia ja mitä elementtejä on mahdollista toistaa useaan kertaan. Lisäksi paikallisissa metatietoskeemoissa on mahdollista määritellä sallittu tai suositeltu tapa antaa elementtien arvot. Esimerkiksi on mahdollista määritellä, missä muodossa päiväys täytyy antaa. Dublin Coreen pohjautuva metatietostandardi on mahdollista määritellä RDF-mallin mukaisesti ja XML-syntaksia käyttäen.

JHS 143

Julkisen hallinnon neuvottelukunta JUHTA⁵ on hyväksynyt suomalainen julkishallinnon metatietosuosituksen JHS 143⁶, joka perustuu Dublin Core -suositukseen. Sen ensimmäinen versio nimeltään *Asiakirjojen kuvailuformaatti JHS 143* julkaistiin vuonna 1999. Vuonna 2004 suosituksesta kehitettiin uusi laajennettu versio, joka on tarkoitettu tukemaan sähköisten asiakirjojen elinkaaren hallintaa. Uutta suositusta kutsutaan nimellä *JHS 143 Asiakirjojen kuvailun ja hallinnan metatiedot*.

Yllä mainittuja suosituksia on esitelty suomeksi esimerkiksi Nuutti Laitalan pro gradu -tutkielmassa (2004). RDF:ää on käsitelty laajasti Maiju Virtasen pro gradu tutkielmassa (2004). RDF:ää ja erityisesti metatietoja tietovarastoympäristöissä on kuvattu Ville Niemijärven pro gradu -tutkielmassa (2002).

⁵ Sisäasiainministeriön asettama julkisen hallinnon tietohallinnon neuvottelukunta JUHTA on valtion ja kuntien yhteinen tietohallinnon kehittämisfoorumi.

⁶ <http://www.intermin.fi/intermin/juhta>

Lähteet

European Committee for Standardization. (2003). CWA 14859 Guidance on the use of Metadata in eGovernment.

<http://www.cenorm.be/cenorm/businessdomains/businessdomains/iss/cwa/cwa14859.asp>.

Gilliland-Swetland, A.J. (2000). Introduction to metadata: Setting the stage.

<http://www.getty.edu/research/institute/standards/intrometadata/>

Hill, L.L., Janée, G., Dolin, R., Frew, J., & Larsgaard, M. (1999). Collection metadata solutions for digital library applications. *Journal of the American Society for Information Science* 50 (13), 1169-1181.

Jokela, S. (2001). Metadata enhanced content management in media companies. PhD Thesis, Helsinki University of Technology.

Laitala, N. (2004). Metatiedot tietojenkäsittelyn integraation tukena. Tapauksena suomalainen lainsäädäntöprosessi. Pro gradu -tutkielma, Jyväskylän yliopisto, Tietojenkäsittelytieteiden laitos.

Lyytikäinen, V. (2004). Contextual and structural metadata in enterprise document management. PhD Thesis, University of Jyväskylä.

Niemijärvi, V. (2002). Metatieto tietovarastoymäristössä. Pro gradu -tutkielma, Jyväskylän yliopisto, Tietojenkäsittelytieteiden laitos.

ISO (2003). ISO/PDTS 23081, Information and documentation – Records management processes – Metadata for records. Part 1: Principles. International standard / Proposed draft technical specification.

Salminen, A. (2005). Building digital government by XML. In R.H. Sprague, Jr. (Ed.), *Proceedings of the Thirty-Eighth Hawaii International Conference on System Sciences*. Los Alamitos, CA: IEEE Computer Society.

Virtanen, M. (2004). RDF-tietomalli toimintaprosessin tiedonhallinnan tukena. Esimerkkinä suomalainen lainsäädäntöprosessi. Pro gradu -tutkielma, Jyväskylän yliopisto, Tietojenkäsittelytieteiden laitos.