

TerveSuomi.fi:n metatietomäärittely

Ver. 1.0 (26.1.2007)

Osma Suominen, Kim Viljanen, Eero Hyvönen, Markus Holi, Petri Lindgren
email: `etunimi.sukunimi@tkk.fi`



SEMANTIC COMPUTING RESEARCH GROUP
FinnONTO project, 2003-2007
<http://www.seco.tkk.fi>

Sisältö

1 Johdanto	1
1.1 Rajaukset	1
2 Kohti yksiselitteisempää metatietoa	2
2.1 Metatiedon esittäminen	3
2.2 Metatietoarvojen tyypit	3
2.3 Metatietokenttien pakollisuus ja toistettavuus	5
2.4 Metatiedon laajennettavuus	5
2.5 Kieli- ja formaattiversiointi	5
3 Metatietokentät	6
3.1 Identifiointitunnus (dc:identifier)	6
3.2 URL-osoite (ts:url)	7
3.3 Nimeke (dc:title) ja tiivistelmä (dcterms:abstract)	7
3.4 Kieli (dc:language)	8
3.5 Julkistamis-, hyväksymis- ja muokkaus aika (dcterms:issued, dcterms:dateAccepted ja dcterms:modified)	8
3.6 Julkaisija ja tekijä (dc:publisher, dc:creator)	8
3.7 Julkaisun laji (ts:genre)	9
3.8 Esitysmuoto (dc:type)	9
3.9 Formaatti (dc:format)	9
3.10 Mediatyyppi (dcterms:medium)	9
3.11 Aihe (dc:subject)	9
3.12 Kohdeyleisö (dcterms:audience)	9
3.13 Alueellinen kattavuus (dcterms:spatial)	10
3.14 Ajallinen kattavuus (dcterms:temporal)	10
3.15 Kuuluu (dcterms:isPartOf)	10
3.16 Oikeudet (dc:rights)	10
3.17 Lähde (dc:source) ja Viittaa (dcterms:references)	11
3.18 Käännöksen lähde (ts:isTranslationOf)	11
3.19 Aikaisempi formaatti (dcterms:isFormatOf)	11
4 Metatiedon esittäminen RDF/XML-muodossa	12
4.1 RDF/XML-dokumentin rakenne	12
4.2 Dokumentin tunniste ja URL-osoite	12
4.3 Vapaatekstikentät	13

4.4	Rakenteiset merkkijonot	13
4.5	Käsiteviittaukset	13
4.6	Yksilöt	14
4.7	Dokumentti- ja yksilöviittaukset	14
4.8	Esimerkki RDF/XML-muotoisesta metatiedosta	14
4.9	Organisaatiokohtaisten yksilöiden esittäminen	15
5	Metatiedon esittäminen osana HTML-sivua	17
5.1	Kenttätunnisteet	17
5.2	Nimiavaruusmäärittelyt	18
5.3	Dokumentin tunniste ja URL-osoite	18
5.4	Vapaatekstikentät	18
5.5	Rakenteiset merkkijonot	18
5.6	Yksilöt ilman viittausta	19
5.7	Käsite-, dokumentti- ja yksilöviittaukset	19
5.8	Esimerkki	20
6	Ontologiat	22

1 Johdanto

TerveSuomi.fi on visio kansallisesta WWW-portaalista, joka toimisi väylänä eri organisaatioiden tuottamaan terveysaiheiseen sisältöön siten, että tiedon löydettävyyden tasolla parantuisi. Nykyisellään tietoa tuotetaan satojen eri organisaatioiden toimesta ja julkaistaan hajautetusti kunkin organisaation verkkosivuilla, jolloin kansalaisen on vaikeaa löytää kaikkea sitä terveyteen liittyvää aineistoa, joka olisi relevanttia hänen kulloisenkin tietotarpeen kannalta. TerveSuomi.fi-portaali tarjoaa keskitetyn väylän tämän sisällön löytämiseen ja käyttämiseen.

FinnONTO-tutkimushankkeen¹ osana kehitettävässä TerveSuomi.fi-prototyypissä pyritään hyödyntämään uusimpia ns. semanttisia viestintäteknikoita (semanttinen web, ontologiat, metatieto) siten, että käyttäjälle tarjotaan mahdollisimman helppokäyttöinen, tehokas ja "älykäs" käyttöliittymä tarjolla olevan tietosisällön käyttämiseen. Portaalin toiminnan edellytyksenä on kuitenkin, että sen sisältämistä dokumenteista on tuotettu laadukasta, standardimuotoista metatietoa. Metatiedoissa kerrotaan mm. dokumenttien julkaisija, tekijät, keskeiset päivämäärät, aiheisällöt ja näiden perusteella portaali osaa ohjata käyttäjän oikeaan tietoon.

Tässä dokumentissa kuvataan TerveSuomi.fi-portaalin metatiedon luonne, määrä ja esitystavat siten, että tiedontuottajat pystyisivät tämän dokumentin perusteella tuottamaan oikeanlaisia ja oikeanlaisessa muodossa tarvittavat metatiedot. Dokumentti perustuu osaltaan aiempaan työhön, erityisesti Kansanterveyslaitoksen vetämän Sanastotyöpajan ehdotukseen metadatakentiksi [Eer06], jota se täsmentää, sekä kansallisiin ja kansainvälisiin metatietostandardeihin, kuten JHS 143 -suositukseen [JHS04] sekä Dublin Core metatietosuositukseen [DCE04, DCT06, DC-02].

Työ on tehty osana FinnONTO-tutkimushanketta, jonka rahoittajina toimivat TEKESin lisäksi yli 30 eri organisaatiota, kuten Kansanterveyslaitos (KTL), STAKES/Luokituskeskus, STAKES/Sosiaaliportti, Valtiovarainministeriö, Elisa Oyj, TietoEnator, Connexor Oy, Siniinen meteoriitti Oy ja Grip Studios Interactive Oy. Tässä kuvattu työ on tehty Teknillisen korkeakoulun Viestintäteknikan laboratoriossa, Semanttisen laskennan tutkimusryhmän SeCo² toimesta.

1.1 Rajaukset

Metatietokentille ei ole määritelty oletusarvoja, koska oletusarvot ovat hyvin tiedontuottajakohtaisia (esim. julkaisija, käyttöoikeudet). Metatiedon tuottamista helpottavat oletusarvot tulisi määritellä järjestelmäkohtaisesti.

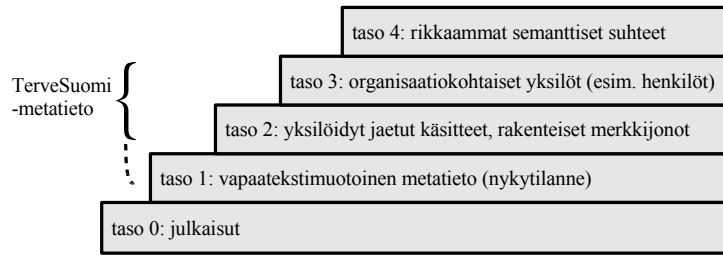
Metatiedolle on määritelty HTML- ja RDF/XML-muotoiset esitysmuodot muttei erillistä, RDF:stä riippumatonta XML-esitysmuotoa. Dublin Core on määritellyt suuntaviivat [DC-03]metatiedon esittämiseen XML-muodossa, joita voidaan tulevaisuudessa käyttää, jos erilliselle XML-esitysmuodolle syntyy tarvetta.

Aihekentän määrittelyssä ei ole huomioitu organisaatiokohtaisten käsitteistöjen (esim. Työterveyslaitoksen sanasto, Stakesin STAMETA) käyttöä. Myöskään tukea vapaalle annotoinnille (uusien käsitteiden lisääminen annotoidessa, kun sopivaa valmista käsitettä ei ole) ei ole tässä dokumentissa määritelty.

Dokumenttien sisältömuotoa, metatiedon tuottamisen tukipalveluita ja metatietojärjestelmien integrointia TerveSuomi-portaaliin käsitellään tarkemmin hajautettua sisällöntuotantoa kuvaavassa dokumentissa [Ter07].

¹<http://www.seco.tkk.fi/projects/finnonto/>

²<http://www.seco.tkk.fi>



Kuva 1: Metatietokerrokset

2 Kohti yksiselitteisempää metatietoa

TerveSuomi.fi-metatietoskeema (ks. luku 3 rakentuu metatiedon tuottajan näkökulmasta kerroksittain (kuva 1) siten, että tiedon tuottamisen vaativuus ja semanttinen tarkkuus kasvavat alhaalta ylöspäin. Jokainen kerros rakentuu alla olevien kerrosten päälle. Tässä jaottelussa ei oteta kantaa dokumenttien tai metatiedon esitysmuotoihin, joita voi olla useita. (Myöhemmissä luvuissa esitetään suositus myös esitysmuotoihin liittyen.)

Taso 0: Julkaisut viittaavat organisaation tuottamiin yksilöitävissä oleviin sisältöihin, joita ovat esimerkiksi tekstidokumentit, kuvat, kalvot, sähköiset palvelut jne. Jokaisella julkaisulla on oltava pysyvä, yksilöivä tunniste kuten arkistointitunnus, ISBN tai URL. Yksilöinti on välttämätöntä, jotta julkaisuun voidaan viitata ja siihen voidaan liittää metatietoa. Tunnisteen pysyvyys on suositeltavaa, jotta eri tahot voisivat viitata julkaisuihin luotettavasti.

Taso 1: Vapaatekstimuotoinen metatieto tarkoittaa metatietoa, joka koostuu ihmisen tulkittavaksi tarkoitetuista merkkijonoista, kuten esimerkiksi otsikosta ja asiasanoista. Monet organisaatiot tuottavat nykyisin vähintään tämän tason metatietoa ja myös useat sovellukset tukevat tällaista metatietoa. Merkkijonomuotoinen metatieto on kuitenkin epätarkkaa, mikäli metatiedossa käytetään vapaasti valittavia, monitulkintaisia asiasanoja tai vapaatekstiä. Tästä seuraa se, että metatietoa on vaikea hyödyntää tehokkaasti.

Taso 2: Yksilöidyt jaetut käsitteet ja rakenteiset merkkijonot antavat mahdollisuuden metatiedon tarkempaan automatisoituun käsittelyyn. Yksilöidyt käsitteistöt ratkaisevat monimerkityksellisten asiasanojen ongelman siten, että metatiedon tuotantoon tarjotaan jaettuja käsitteistöjä. Käsitteet määritellään ns. ontologioiden avulla, joissa kullekin käsitteelle annetaan oma yksilöivä tunnus sekä liitetään se semanttisilla suhteilla muihin käsitteisiin. Yksilöivien tunnisteiden ansiosta on aina selvää mihin käsitteistöön mikäkin käsite liittyy. Käyttämällä tietyn aihealueen eri toimijoiden jaettua käsitteistöä mahdollistetaan se, että sisällöt ovat siirrettävissä aihealueen toimijoiden välillä. Rakenteiset merkkijonot kuten standardimuotoiset päiväykset ovat myös automaattisesti luettavia ja tulkittavia.

Taso 3: Organisaatiokohtaiset yksilöt kuten henkilöt tai organisaation osat ovat kohteita, joihin usein on tarvetta viitata metatiedossa. Mikäli näille annetaan vain merkkijonomuotoinen esitys, kuten henkilön etu- ja sukunimi, ei tämä riitä yksilöimään kohteita tarpeeksi tarkasti, koska organisaatiossa saattaa olla useita samannimisiä henkilöitä. Tämän ongelman ratkaisemiseksi tarvitaan keino tuottaa pysyviä yksilöiviä tunnisteita organisaatiokohtaisille yksilöille.

Taso 4: Rikkaammat semanttiset suhteet viittaa siihen, että dokumentin metatietoa tarkennetaan määrittelemällä käsitteiden (taso 2) ja yksilöiden (taso 3) välisiä suhteita. Esimerkiksi dokumentin kuvaama tapahtuma voi sisältää henkilöitä, organisaatioita, paikkoja, aikoja jne. tietyssä suhteessa toisiinsa. Mitä tarkemmin metatietoon kuvataan dokumentin käsittelemä aihe, sitä tarkemmin dokumentteja voidaan hyödyntää ja hakea.

TerveSuomi.fi-järjestelmän toimivuuden kannalta tärkeää on, että tiedontuottajat pystyvät tuottamaan metatietoa, joka edustaa mahdollisimman korkeaa semanttista tasoa. Mitä korkeammalla tasolla aineisto kuvaillaan, sitä rikkaammat mahdollisuudet tarjoutuvat tiedonhakuun ja eri aineistojen yhteenliittämiseen.

2.1 Metatiedon esittäminen

Metatieto koostuu TerveSuomi.fi:ssä paitsi tekstisisällöstä (esimerkiksi nimeke) myös yhteisistä, yksikäsitteisistä käsiteviittauksista (esimerkiksi dokumentin sisältöluokitus). Yksikäsitteisten käsiteviittausten ansiosta eri tiedontuottajien sisältö voidaan yhdistää toisiinsa. Dokumenttien, käsitteiden, henkilöiden ja organisaatioiden yksilöintiin käytetään URI-tunnuksia, joiden ansiosta kaikki viittaukset voidaan esittää täsmällisessä muodossa. Tavallisimmin URI-tunnukset³ ovat www-selaimesta tuttuja URL-osoitteita, mutta myös URN-tunnusten käyttö on mahdollista (ks. kohta 3.1).

TerveSuomi.fi-metatietoskeema ("metatietokentät") on Dublin Coren metatietoformaatin mukainen. Mukaan on otettu Dublin Coren peruselementtijoukko [DCE04] sekä Dublin Coren määrittelemiä tarkenteita [DCT06]. Joitakin lisätarkenteita on tässä projektissa määritelty: URL-osoite `dc:identifier`-kentän tarkenteena, julkaisutyyppi `dc:type`-kentän tarkenteena sekä viittaus käännetyn dokumentin alkuperäisversioon `dcterms:isVersionOf`-kentän tarkenteena. Kenttäjoukon valinnassa ja kenttien määritelmässä on soveltuvin osin huomioitu myös Dublin Core -metadataformaatin suomalainen versio [DC-02, SFS01] sekä JHS 143-suositus [JHS04].

Tarkenteiden käyttö mahdollistaa kentän täsmällisemmän käyttötarkoituksen ja semantiikan määrittämisen menettämättä kuitenkaan kytkentää alkuperäiseen Dublin Core -elementtijoukkoon. Tarkentamissuhteet on esitetty kuvassa 2.

2.2 Metatietoarvojen tyypit

Metatietokenttien arvot voivat olla joitakin seuraavista tyypeistä:

Vapaatekstin koneellinen tulkinta on pääsääntöisesti hankalaa tai ei onnistu. Tällaisia tietoja ovat mm. dokumentin nimeke ja kuvaus, henkilöiden ja organisaatioiden nimet (standardoinnista huolimatta kirjoitusasu voi vaihdella ja samannimisiä henkilöitä voi olla useita) sekä monimerkitykselliset tai määrittelemättömät asiasanat. Ne edustavat metatiedon tasoa 1.

Rakenteiset merkkijonot, kuten päivämäärät, aikajaksot ja koordinaattimuotoiset paikatiedot, esitetään standardinmukaisessa muodossa ja ovat siis metatiedon tasoa 2.

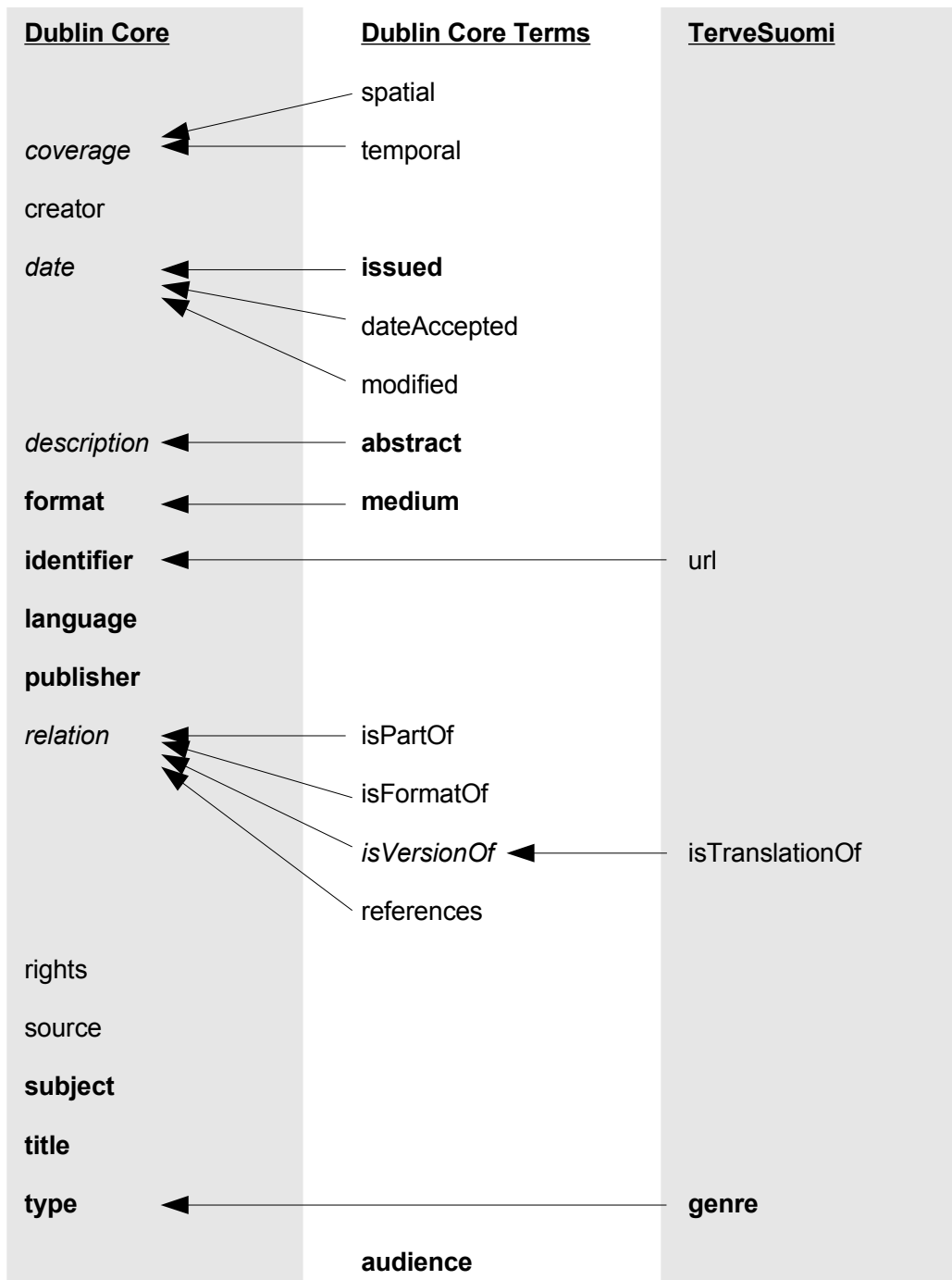
Käsiteviittaukset esitetään viitteinä jaetuissa ontologioissa määriteltyihin käsitteisiin ja ovat siis metatiedon tasoa 2. Viittaus voidaan esittää esimerkiksi URI-tunnisteena. Myös asiasana voi olla yksilöivä, mikäli tiedetään mihin asiasanastoon se kuuluu ja mikäli sama asiasana tarkoittaa vain yhtä asiaa.

Dokumenttiviittaukset esitetään yksikäsitteisinä viitteinä toisiin dokumentteihin ja ne tehdään joko URI-tunnisteiden avulla (suositeltava) tai URL-tunnisteiden avulla. Viittaukset edustavat metatiedon tasoa 2.

Yksilöviittaukset kuten organisaatiot ja henkilöt esitetään viitteinä jaetun toimijatietokannan henkilö- ja organisaatietietueisiin ja ovat metatiedon tasoa 3.

Semanttiset suhteet mahdollistavat yksilöityjen käsitteiden välisten suhteiden määrittelyä eli esimerkiksi tietyn henkilön suhde tiettyyn organisaatioon kuuluvaan työryhmään. Tällaiset suhteet kuuluvat tason 4 metatietoihin eikä niitä ainakaan aluksi oleteta tuotettavan.

³<http://www.ietf.org/rfc/rfc3986.txt>



Kuva 2: Metatietokentät ja niiden tarkennussuhteet. Skeemassa pakolliset kentät on **lihavoitu**. Kentät, joita ei käytetä sellaisinaan vaan ainoastaan tarkennusten kautta on *kursivoitu*.

2.3 Metatietokenttien pakollisuus ja toistettavuus

Kukin kenttä voi olla joko pakollinen tai vapaaehtoinen. Lisäksi se voi olla toistettavissa eli esiintyä joko kerran tai useammin. Jos samasta kentästä, kuten otsikosta, on samaa tarkoitavia eri kielisiä versioita, lasketaan samaa tarkoittavat versiot kentän esiintymiseksi vain kerran.

Dublin Core määrittelee kaikki metatietokentät vapaaehtoisiksi ja toistettaviksi, mutta joidenkin kenttien osalta toisteisuusrajoituksia on tiukennettu metatiedon käsittelyn helpottamiseksi ja epäselvien tilanteiden (esim. mitä tarkoittaa jos dokumentilla on kaksi nimekettä) välttämiseksi.

2.4 Metatiedon laajennettavuus

Tämä dokumentti määrittelee TerveSuomi-portaalin edellyttämän metatiedon määrän ja muodon, mutta ei sulje pois muun kuin tässä esitetyn metatiedon lisäämistä aineistolle. Esimerkiksi organisaatiokohtaiset lisäkentät ovat mahdollisia, vaikka TerveSuomi ei niitä hyödynnä.

Myös organisaatiokohtaisten käsitteistöjen, sanastojen ja luokitusten käyttö metatiedoissa on mahdollista, kunhan nämä metatiedot on esitetty niin, ettei sekaannuksen vaaraa ole. Käsiteviittausten osalta tämä tarkoittaa sitä, että viittauksen kohteena oleva käsitteistö on aina ilmaistu käsitteistötunnuksin.

2.5 Kieli- ja formaattiversiointi

Tyypillisiä tilanteita dokumenttikokoelmissa on se, että samasta dokumentista on useita käännöksiä eri kielillä (kieliversioita) tai että sama dokumentti on esitetty useassa teknisessä formaatissa (esim. HTML ja PDF). Tarkoituksenmukaista on, että eri versiot samasta dokumentista voidaan kuvata metatiedoin niin, että yhteys versioiden välillä säilyy ja näin esimerkiksi TerveSuomi-portaalin käyttöliittymässä pystytään tarjoamaan linkit eri versioiden välille.

Dokumentin eri versioilla voi olla sekä yhteisiä metatietoja (esim. julkaisija, aihe) että toisistaan eriyviä metatietoja (esim. julkaisuajankohta, kieli, formaatti). Julkaisujärjestelmän sisällä kannattaa noudattaa normalisointiperiaatetta eli tallettaa eri versioille yhteiset metatiedot vain kertaalleen. TerveSuomi-metatiedon kannalta dokumentin eri versiot ovat kuitenkin erillisiä ja itsenäisiä dokumentteja.

Yhteys dokumentin eri versioiden välillä ilmaistaan `ts:translation-` ja `dcterms:hasFormat`-kentillä. Näiden kenttien perusteella TerveSuomi-portaali pystyy päättämään, että kysymys on saman dokumentin eri versioista ja vertaamalla kahden eri dokumentin metatietoja päättämään, mitkä metatiedoista ovat yhteisiä. Haittana on tiedon toisteisuus yhteisten metatietojen osalta, mutta varsinaisia ongelmia se ei aiheuta portaalille.

Metatiedon kieliversiointiin liittyy myös kysymys saman tiedon (esim. dokumentin nimi tai julkaisijaorganisaation nimi) esittämisestä eri kielillä. Yksittäisten merkkijonomuotoisten kenttien monikielinen esittäminen on hyvin tuettu sekä RDF- että HTML-kielissä ja näiden esittäminen käsitellään tarkemmin myöhemmin tässä dokumentissa asianomaisten kenttien kohdalla.

3 Metatietokentät

TerveSuomi-portaalin metatietokentät sekä niiden toisto ja pakollisuus esitetään taulukossa 1. Kukin metatietokenttä käsitellään tarkemmin seuraavissa aliluvuissa.

3.1 Identifointitunnus (dc:identifier)

Dokumentin tunniste ilmaistaan URI-muodossa. Tunnisteen tulee yksilöidä dokumentti, ts. sama tunniste vastaa vain yhtä dokumenttia, mutta dokumentilla voi olla useita tunnisteita. Jos tunnisteita on useita, tulee tiedon julkaisijan päättää mikä tunnisteista on se, jolla dokumentti ensisijaisesti yksilöidään.

Taulukko 1: Metatietokentät. Pakolliset kentät on merkitty X:llä P-sarakkeessa ja lisäksi **lihavoitu**. Toisteisuus (useat arvot sallittu) on merkitty X:llä T-sarakkeessa.

	Nimi	Tunniste	P	T	Arvotyyppi	Arvojoukko
perustiedot	Identifointitunnus	dc:identifier	X		URI	
	URL-osoite	ts:url			URL	
	Nimeke	dc:title	X	^a	Vapaateksti	Ei-tyhjät merkkijonot
	Tiivistelmä	dcterms:abstract	X	^a	Vapaateksti	Ei-tyhjät merkkijonot
	Kieli	dc:language	X	X	Rakenteinen merkkijono	RFC 3066
alkamäärät	Julkistamisaika	dcterms:issued	X		Rakenteinen merkkijono	W3CDTF (ISO 8601)
	Hyväksymisaika	dcterms:dateAccepted			Rakenteinen merkkijono	W3CDTF (ISO 8601)
	Muokkusaika	dcterms:modified			Rakenteinen merkkijono	W3CDTF (ISO 8601)
tekijät	Julkaisija Tekijä	dc:publisher dc:creator	X	X X	Yksilö Yksilö	foaf:Organization foaf:Organization, foaf:Person tai foaf:Group
luokittelet	Julkaisun laji	ts:genre	X	X	Käsite	Julkaisun laji -käsitteet
	Esitysmuoto	dc:type	X	X	Käsite	DCMI Type -käsitteet
	Formaatti	dc:format	X		Rakenteinen merkkijono	IANA-rekisteröidyt MIME-tyypit
	Mediatyyppi	dcterms:medium	X		Käsite	Vakioarvo (Mediatyyppi-käsite Verkkoressurssi)
sisällönkuvailu	Aihe	dc:subject	X	X	Käsite	YSO-, MeSH- ja TESA-käsitteet
	Kohdeyleisö	dcterms:audience	X	X	Käsite	Kohdeyleisö-käsitteet
	Alueellinen kattavuus	dcterms:spatial		X	Rakenteinen merkkijono ja/tai käsite	DCMI Point, DCMI Box tai Paikka-käsite
	Ajallinen kattavuus	dcterms:temporal		X	Rakenteinen merkkijono ja/tai käsite	W3CDTF, DCMI Period tai Aika-käsite
suhteet	Kuuluu	dcterms:isPartOf		X	Dokumentti	URI
	Oikeudet	dc:rights		X	Vapaateksti ja/tai dokumentti	URI tai sanallinen kuvaus
	Lähde	dc:source		X	Vapaateksti ja/tai dokumentti	URI (esim. ISBN) tai bibliografinen viite
	Viittaa	dcterms:references		X	Vapaateksti ja/tai dokumentti	URI (esim. ISBN) tai bibliografinen viite
	Käännöksen lähde	ts:isTranslationOf		X	Dokumentti	URI
	Aikaisempi formaatti	dcterms:isFormatOf		X	Dokumentti	URI

^aErikieliset versiot sallittu, kunhan kullakin kielellä on annettu vain yksi arvo.

URI voi perustua esimerkiksi ISBN, ISSN⁴, ISO OID⁵ ja muihin yleisesti käytettyihin tunnistekäytäntöihin, joille on varattu IANA:ssa oma URI⁶- tai URN-skeematunnus. Se voi myös perustua organisaation sisäiseen arkistointitunnukseen tai vastaavaan. Jälkimmäisessä tapauksessa URI tulee muodostaa siten, että se on maailmanlaajuisesti yksilöivä.

Esimerkkejä URI-tunnisteista:

```
http://www.ktl.fi/doc/123456
urn:isbn:951-611-441-5
urn:issn:1234-5678
```

Dokumentin erikieliset versiot sekä eri formaatissa olevat versiot samasta dokumentista (esim. HTML ja PDF) tulkitaan kuitenkin erilliseksi dokumenteiksi, mikäli niistä kaikista tuotetaan metatietoa.

TerveSuomi.fi:n metatiedoissa suositellaan käytettäväksi URI-tunnisteita, jotka toimivat tarvittaessa myös URL-osoitteina. Esimerkiksi

```
http://www.ktl.fi/doc/123456
```

Jos sisällönhallintajärjestelmä on rakennettu niin, että dokumenttien URL-osoitteet ovat luonteeltaan pysyviä (esim. osoitteet eivät muutu vaikka sivuston rakenteeseen tai osiojakoon tehdään muutoksia), voi dokumentin tunnisteena toimia sen URL-osoite. Jos taas dokumentin URL voi vaihtua, dokumentin URI muodostetaan jollakin toisella tavalla (esim. julkaisujärjestelmän sisäisestä dokumenttitunnisteesta) siten, että se pysyy samana dokumentin URL:n muuttuessa. Pysyvien URL-osoitteiden suunnittelusta kertoo enemmän W3C:n dokumentti *Cool URIs don't change*⁷.

3.2 URL-osoite (ts:url)

Kenttä kertoo dokumentin URL-osoitteen, jonka kautta se on ladattavissa esimerkiksi WWW-selaimella. Sitä tarvitaan, jos dokumentin varsinaista tunnistetta (dc:identifier) ei voi käyttää dokumentin lataamiseen.

Suosittelavaa on joko käyttää pysyviä URL-osoitteita kaikkialla (dokumentin tunnisteena ja URL-osoitteena), tai mikäli tunniste poikkeaa sivustolla käytössä olevasta URL-osoitteesta, rakentaa tunniste niin, että sen kautta kuitenkin pääsee käsiksi (esim. HTTP redirect-mekanismilla) dokumenttiin. Näissä tapauksissa tämä kenttä on tarpeeton.

3.3 Nimeke (dc:title) ja tiivistelmä (dcterms:abstract)

Kullakin dokumentilla tulee olla nimeke ("otsikko") ja tiivistelmä, kumpiakin täsmälleen yksi. Nimekkeen tulisi liittyä yksinomaan dokumenttiin eli siihen ei saa sisältyä vakionuotoisia alku- tai loppuliitteitä, esimerkiksi julkaisijaorganisaation nimeä tai sivuston osaa johon se kuuluu. Suositeltavaa on, että nimeke vastaa dokumentissa näkyvää pääotsikkoa.

Tiivistelmän ohjeellinen maksimipituus on 300 merkkiä ja sen tulee kiteyttää dokumentin keskeinen sisältö ja auttaa näin tiedonhakijaa näkemään yhdellä silmäyksellä, onko kyseinen dokumentti hänen kannaltaan kiinnostava. Suositeltavin tapa tuottaa tiivistelmä on tehdä se käsin, jotta tekstin laatu olisi mahdollisimman hyvä ja esittelisi mahdollisimman hyvin dokumentin ydinsisällön. Muita vaihtoehtoja ovat tiivistelmän automaattinen tuottaminen dokumentin ingressin tai leipätekstin alkuosan perusteella.

Jos dokumentin otsikko on käännetty useille kielille (esim. väitöskirjat), voidaan kentässä antaa useita erikielisiä arvoja, kunhan samaa tarkoittavia käännöksiä on vain yksi. Myös tiivistelmästä voidaan vastaavaan tapaan antaa useita erikielisiä versioita.

⁴<http://www.ietf.org/rfc/rfc3044.txt>

⁵<http://www.ietf.org/rfc/rfc3061.txt>

⁶<http://www.iana.org/assignments/uri-schemes.html>

⁷<http://www.w3.org/Provider/Style/URI>

3.4 Kieli (dc:language)

Dokumentin kieli annetaan RFC 3066 -standardin mukaisena kielikoodina [RFC01]. Esimerkiksi `fi` on suomi, `sv` ruotsi ja `en` englanti.

Tarvittaessa kielikoodi voidaan ilmaista kaksiosaisena, esim. `sv-FI` on suomenruotsi ja `en-US` on amerikanenglanti.

Suosittelavaa on kirjoittaa kielikoodin alkuosa pienillä kirjaimilla ja kaksiosaiset koodit kuten edellä, mutta kentän sisältö tulkitaan kuitenkin kirjainkoosta riippumatta.

Mikäli dokumentti sisältää useita erikielisiä osia, voidaan kieli-kenttää toistaa. Suositeltavaa on kuitenkin, että dokumentin eri kieliversiot ovat itsenäisiä dokumentteja, joista kustakin on erilliset metatiedot.

3.5 Julkistamis-, hyväksymis- ja muokkausajankohta (dcterms:issued, dcterms:dateAccepted ja dcterms:modified)

Aikamäärekentät ilmaisevat dokumentin julkistamis-, hyväksymis- ja muokkausajankohdan. Näistä ainoastaan julkistamisaika on pakollinen.

Julkistamisaika kertoo milloin dokumentti ensimmäistä kertaa julkaistiin yleisön saataville. Mikäli ajankohta ei ole tiedossa (esimerkiksi julkaisujärjestelmän vaihtamisen vuoksi), annetaan julkistamisaikaksi varhaisin ajankohta, jolloin dokumentin tiedetään olleen julkaistuna.

Hyväksymisaika kertoo milloin dokumentin asiasisältö on viimeksi käyty läpi ja siinä esitettyjen tietojen ajantasaisuus on varmistettu.

Muokkausajankohta kertoo ajankohdan, jolloin dokumentin asiasisältöä on viimeksi muokattu. Suositus on, että kentän arvoa ei päivitetä sellaisten pienten muutosten (esim. kirjoitusvirheiden korjaus) takia, jotka eivät vaikuta dokumentin asiasisältöön.

Aikamäärät esitetään W3C:n määrittelemässä W3CDTF-muodossa [W3C97], joka määrittelee osajoukon ISO 8601 -standardin sallimista päivämääräformaateista [ISO88]. ISO 8601 mahdollistaa hyvin monta hiukan toisistaan poikkeavaa esitystä päiväyksille, joten rajoittuminen W3CDTF-muotoihin helpottaa päiväysten käsittelyä.

Päiväysten tarkkuus voi vaihdella: ne voidaan ilmaista vuoden, kuukauden, päivän, minuutin, sekunnin tai sekunnin murto-osan tarkkuudella. Silloin kun päiväykseen sisältyy kellonaika, on myös aikavyöhyke ilmaistava W3CDTF:n edellyttämällä tavalla. Suositeltavaa on käyttää päivän tarkkuutta.

Esimerkkejä hyväksyttävistä päiväyksistä:

2006	vuoden tarkkuudella
2006-11	kuukauden tarkkuudella
2006-11-18	päivän tarkkuudella
2006-11-18T16:49+02:00	minuutin tarkkuudella (ilmaistu Suomen talviajassa)
2006-11-18T14:49Z	minuutin tarkkuudella (sama hetki ilmaistu UTC-ajassa)
2006-06-12T12:34:56+03:00	sekunnin tarkkuudella (ilmaistu Suomen kesäajassa)

3.6 Julkaisija ja tekijä (dc:publisher, dc:creator)

Dokumentin julkaisija ja tekijä (organisaatio, henkilö tai ryhmä) voidaan ilmaista joko vapaatekstinä tai semanttisesti rikkaammin yksilöviittausten kautta. Kummassakin tapauksessa yksilöt kuvataan FOAF-ontologian luokkien `foaf:Person`, `foaf:Organization` ja `foaf:Group` avulla, mutta **vapaatekstimuotoiset** yksilömäärittelyt eivät sisällä täsmällistä tietoa siitä, milloin kysymys on samasta yksilöstä, kun taas **yksilöviittaukset** mahdollistavat mm. sen, että saman kirjoittajan artikkelit voidaan automaattisesti liittää toisiinsa, vaikka samassa organisaatiossa työskentelisi esimerkiksi kaksi Matti Virtanen -nimistä henkilöä.

Yksilön nimissä on noudatettava yhdenmukaista kirjoitusasua. Organisaatioiden nimet kirjoitetaan auki käyttämättä lyhenteitä (esim. Kansanterveyslaitos eikä KTL). Henkilöiden nimet annetaan muodossa Sukunimi, Etunimi. Organisaatioiden nimistä voidaan tarjota useita erikielisiä versioita, kunhan ne on asianmukaisesti merkitty kielikoodein.

3.7 Julkaisun laji (ts:genre)

Julkaisun laji kuvaa dokumentin käyttötarkoitusta tai genreä. Mahdolliset arvot on määritetty julkaisun laji -sanastossa. Mahdollisia arvoja ovat mm. *Opas tai ohje*, *Ajankohtaista* ja *Tilasto*.

Kenttä vastaa JHS 143 -suosituksen kenttää *Asiakirjan tyyppi* (DocumentType) [JHS04].

3.8 Esitysmuoto (dc:type)

Esitysmuoto kuvaa kohteen rakennetta tai muotoa. Mahdollisia arvoja ovat DCMI-tyyppisanastossa määritellyt käsitteet kuten *Teksti* ja *Kuva*.

Kenttä vastaa JHS 143 -suosituksen kenttää *Laji* (Type) [JHS04].

3.9 Formaatti (dc:format)

Formaatti-kentässä kerrotaan dokumentin MIME-tyyppi [mim], esimerkiksi `text/html` tai `application/pdf`.

3.10 Mediatyyppi (dcterms:medium)

Mediatyyppi kuvaa kohteen sisällön välittäjänä toimivia entiteettejä, joita on luokiteltu mediatyyppintologiassa. Kentän arvo on toistaiseksi aina *Verkkoresurssi*, koska tässä dokumentissa kuvattu metatietoskeema koskee erityisesti verkossa julkaistuja dokumentteja. Myöhemmin metatietoskeemaa voidaan laajentaa tukemaan muita mediatyyppejä esimerkiksi lisäämällä kentät painotuotteiden sivumäärälle ja painopaikalle, jolloin myös mediatyyppi-kenttä määritellään laajemmalla arvovalikoimalla.

3.11 Aihe (dc:subject)

Aihe-kenttä kertoo dokumentin käsittelemistä aiheista (vrt. asiasanat). Aihe-kentän arvoina käytetään aiheontologioiden (YSO, FinMeSH, TESA) käsitteitä. Suositus on, että jokaiselle dokumentille määritellään useita mahdollisimman kuvaavia aiheita.

3.12 Kohdeyleisö (dcterms:audience)

Kohdeyleisö-kenttä kertoo, kenelle dokumentti on käsittelytavaltaan ja kirjoitustyylyltään suunnattu eli kuka on sen ajateltu lukija. Dokumentilla voi olla useita kohderyhmiä, jotka on määritelty kohderyhmäontologiassa.

Keskeisin erottelu kohderyhmissä on se, onko tieto suunnattu asiantuntijoille (esim. tieteelliset artikkelit, ammattilaisten hoito-oppaat) vai maallikoille (esim. yleistajuiset artikkelit, testit, itsehoito-oppaat).

Kohderyhmäontologia tarjoaa mahdollisuuden määritellä kohderyhmän tarkemmin eri näkökulmista, esimerkiksi iän, sukupuolen, perheaseman, toimintaympäristön, ammatillisen roolin ja potilasryhmän mukaan.

3.13 Alueellinen kattavuus (dcterms:spatial)

Alueellinen kattavuus kertoo mihin maantieteelliseen paikkaan tai alueeseen dokumentti keskittyy. Esimerkiksi tilastotiedot sekä tietyllä paikkakunnalla sijaitsevat palvelupisteet (sairaala, päihdekliniikka jne.) voidaan tämän kentän avulla kytkeä maantieteelliseen sijaintiin.

Kentän arvot ovat joko paikkaontologian käsitteitä (esim. maat ja kaupungit) tai DCSV-syntaksin mukaisia merkkijonoja [DCS06]. Merkkijonot kuvaavat koordinaattimuodossa esitettyä maantieteellistä pistettä (DCMI Point) [DCM06c] tai suorakaiteen muotoista aluetta (DCMI Box) [DCM06a]. Dublin Coren ohjeistuksen mukaisesti koordinaatit ilmaistaan WGS84-järjestelmän mukaisina desimaalilukuina.

Esimerkki koordinaattien ilmaisemisesta DCSV-syntaksilla:

DCMI Point (piste):

```
name=Jorvin sairaala; east=24.688968; north=60.221982;
```

DCMI Box (suorakaiteen muotoinen alue):

```
name=Meilahden sairaala-alue; northlimit=60.191149; eastlimit=24.912598;  
southlimit=60.185824; westlimit=24.900416;
```

3.14 Ajallinen kattavuus (dcterms:temporal)

Ajallinen kattavuus viittaa aikakauteen, jota dokumentin sisällössä käsitellään riippumatta siitä milloin dokumentti on laadittu (esim. tupakoinnin historiasta kertova dokumentti saattaa käsitellä 1700-lukua).

Kentän arvot ovat joko aikaontologian aikaan liittyviä käsitteitä (esim. *syksy, jouluku, toinen maailmansota*), yksittäisiä päivämääriä W3CDTF-muodossa [W3C97] tai DCSV-syntaksin mukaisia merkkijonoja [DCS06]. Merkkijonot kuvaavat aikajaksoa, joka esitetään päivämäärävälinä (DCMI Period) [DCM06b]. Päivämääräväli voi myös olla avoin jommasta kummasta päästä (esim. vuodesta 1960 eteenpäin) ja sille voidaan antaa nimi (esim. Ahtisaaren presidenttikausi).

Esimerkki päivämäärävälin ilmaisemisesta DCSV-syntaksilla:

```
name=Talvisota; start=1939-11-30; end=1940-03-13;
```

3.15 Kuuluu (dcterms:isPartOf)

Kenttä on viite toiseen dokumenttiin (tai useaan), jonka osa kuvailtava dokumentti on. Tyyppillinen osa-kokonaisuussuhde on esimerkiksi kirjan yksittäisen luvun suhde koko kirjaan sekä artikkelin suhde artikkelikokoelmaan. Kokonaisuuden tulisi muodostaa itsenäinen teos; pelkkä sivujen sijainti samassa sivuston osiossa ei vielä tee niistä kokonaisuutta. Sivuston päätaison osioon (esim. uutiset) tai koko sivuston etusivulle ei tule viitata ilman erityisen painavaa syytä.

Kentän arvo on dokumentin URI, ja suositus on, että myös viittauksen kohteena olevasta kokonaisuudesta on olemassa tämän ohjeen mukaiset metatiedot.

3.16 Oikeudet (dc:rights)

Kenttä kertoo dokumenttiin liittyvistä tekijänoikeuksista ja käyttöoikeuksista. Suositeltavaa on antaa kentän arvoksi URI, joka viittaa dokumenttiin, jossa kuvataan käyttöoikeudet (esim. monilta sivustoilta löytyvä sivu Käyttöoikeudet tai Tietoja sivustosta). Myös kolmannen osapuolen julkaisemaan käyttöoikeustietoon, esim. Creative Commons⁸ -lisensseihin, voidaan viitata niiden URIa käyttämällä.

⁸<http://creativecommons.org/>

Kun kentän arvona on dokumentin URI, suositus on, että myös viittauksen kohteena olevasta dokumentista on olemassa tämän ohjeen mukaiset metatiedot.

Vaihtoehtoisesti kentän arvona voidaan antaa sanallinen kuvaus oikeuksista, esimerkiksi Copyright UKK-instituutti. Kaikki oikeudet pidätetty..

3.17 Lähde (dc:source) ja Viittaa (dcterms:references)

Kentissä ilmaistaan toinen dokumentti tai muu lähde, johon kuvattava dokumentti merkittävässä määrin perustuu (Lähde) tai johon se viittaa (Viittaa). Kenttien arvo voidaan antaa joko URI-tunnisteena (esim. URL) tai Vancouver-tyylin mukaisena bibliografisena viitteenä [Van06].

3.18 Käännöksen lähde (ts:isTranslationOf)

Kentässä ilmaistaan dokumentti, joka on sisällöltään sama kuin tämä, mutta on esitetty jollakin toisella (luonnollisella) kielellä. Myöhemmän dokumentin (käännöksen) metatiedoissa viitataan aiempaan, alkuperäiseen dokumenttiin.

3.19 Aikaisempi formaatti (dcterms:isFormatOf)

Kentässä ilmaistaan dokumentti, joka on sisällöltään sama kuin tämä, mutta on esitetty jossakin toisessa teknisessä muodossa (esimerkiksi HTML tai PDF). Myöhemmin luodun dokumentin metatiedoissa viitataan aiempaan, alkuperäiseen dokumenttiin.

4 Metatiedon esittäminen RDF/XML-muodossa

RDF (Resource Description Framework) -tietomalli sopii erittäin hyvin metatietojen esittämiseen [RDF]. Allaoleva ohjeistus RDF/XML-muotoisen metatiedon esittämiseen perustuu Dublin Coren suosituksiin [DCM02, DCQ02].

Tässä esiteltyä RDF-muotoista metatietoa voi tuottaa joko RDF-työkaluin (esimerkiksi Jena-kirjasto⁹ Javalle) tai melko suoraviivaisesti myös XML-työkaluin, jotka eivät ymmärrä RDF:n erityisominaisuuksia.

Oli tiedon tuottamistapa mikä hyvänsä, on syytä kiinnittää erityistä huomiota siihen, että kenttien arvoissa esiintyvät erikoismerkit kuten kulmasulkeet ja lainausmerkit on korvattu XML-standardin mukaisilla koodilla (esim. `"` ja `>`). RDF-työkalut ja useimmat XML-työkalut huolehtivat tästä automaattisesti, mutta mikäli tietoja joudutaan tuottamaan merkkijono-operaatioin on vaara, että lopputulos ei noudata XML-standardia eikä näin ollen ole myöskään portaalin tulkittavissa.

Myös merkistöasiat on huomioitava RDF-muotoisen tiedon tuottamisessa. Yleensä XML-muotoiset dokumentit käyttävät UTF-8-merkistöä, mutta muun merkistön käyttö on mahdollista ilmoittamalla siitä XML-tiedoston otsakkeessa. Metatieto saa sisältää vain ilmoitetun merkistön mukaisia merkkejä. Muut merkit tulkitaan virhetilanteiksi.

4.1 RDF/XML-dokumentin rakenne

Metatietoa sisältävät RDF-dokumentit koostuvat XML-otsakkeesta, `rdf:RDF`-elementistä nimiavaruusesittelyineen sekä yhdestä tai useammasta `ts:Publication`-elementistä. Kukin `ts:Publication`-elementti sisältää yhden dokumentin metatiedot.¹⁰ Kukin metatietokenttä esitetään omalla elementillään, jonka nimi vastaa metatietokentän tunnistetta.

RDF-dokumentin suositeltu rakenne on seuraava:

```
<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:dcterms="http://purl.org/dc/terms/"
  xmlns:ts="http://www.yso.fi/onto/tervesuomi-schema">
  ... ts:Publication-elementtejä ...
</rdf:RDF>
```

4.2 Dokumentin tunnistet ja URL-osoite

Kukin dokumentti kuvataan yhdellä `ts:Publication`-elementillä, jonka attribuutti `rdf:about` kertoo dokumentin identifiointitunnuksen eli URI:n. Sama URI esitetään myös `dc:identifier`-kentässä seuraavasti:

```
<ts:Publication rdf:about="http://www.ktl.fi/portal/8131">
  <dc:identifier rdf:resource="http://www.ktl.fi/portal/8131"/>
  ... muita metatietoja tästä dokumentista ...
</ts:Publication>
```

Mikäli metatietoihin kuuluu dokumentin URL-osoite (`ts:url`), se esitetään `ts:url`-elementillä samaan tapaan kuin `dc:identifier` ylläolevassa esimerkissä.

⁹<http://jena.sourceforge.net>

¹⁰Näin ei välttämättä tarvitse olla; RDF/XML mahdollistaa saman tietosisällön esittämisen hyvin monella eri tavalla. Yksinkertaisuuden vuoksi tämä dokumentti kuvaa sellaisen RDF/XML-dokumentin, joka on helppo tuottaa XML-työkaluin ja myös melko helposti ihmisen tulkittavissa. Muut, RDF-tietosisällöltään yhtäpitävät muodot ovat toki myös sallittuja.

4.3 Vapaatekstikentät

Vapaatekstikentät esitetään elementteinä, joiden tekstisisältönä on kentän sisältö. Mikäli kentän arvon kieli on tiedossa, se merkitään `xml:lang`-attribuutilla.

Esimerkki:

```
<dc:title>Rokottajan käsikirja 2005</dc:title>

<dc:publisher xml:lang="fi">Kansanterveyslaitos</dc:publisher>
<dc:publisher xml:lang="sv">Folkhälsoinstitutet</dc:publisher>
<dc:publisher xml:lang="en">National Public Health Institute</dc:publisher>
```

4.4 Rakenteiset merkkijonot

Rakenteiset merkkijonot tulee esittää luomalla tiedon esitysmuodon kertova RDF-elementti, joka puolestaan sisältää `rdf:value`-elementin, jonka tekstisisältö on kentän arvo DSCV-muodossa esitettynä¹¹. Esimerkiksi päivämäärä esitetään seuraavasti:

```
<dcterms:issued>
  <dcterms:W3CDTF>
    <rdf:value>2004-12-17</rdf:value>
  </dcterms:W3CDTF>
</dcterms:issued>
```

`<<-- esitysmuoto`
`<<-- arvo`

Ajanjakso esitetään puolestaan seuraavasti:

```
<dcterms:temporal>
  <dcterms:Period>
    <rdf:value>name=Talvisota; start=1939-11-30; end=1940-03-13;</rdf:value>
  </dcterms:Period>
</dcterms:temporal>
```

Esitysmuodon kertovat elementit ovat seuraavat:

Tyyppi	Esitysmuoto
Päiväys	dcterms:W3CDTF
Aikajakso	dcterms:Period
Maantieteellinen paikka (pistemäinen)	dcterms:Point
Maantieteellinen paikka (suorakaide)	dcterms:Box
Formaatti (MIME-tyyppi)	dcterms:IMT
Kieli	dcterms:RFC3066

4.5 Käsiteviittaukset

Käsiteviittaukset tehdään viittaamalla käsitteen tunnisteeseen eli URIin. Esimerkki:

```
<dc:subject rdf:resource="http://www.yso.fi/onto/yso/a12345"/>
```

Käsiteviittausten tekeminen siis edellyttää, että käsitteen URI on tiedossa. TerveSuomi-portaalin tukemat käsitteistöt julkaistaan ONKI-palvelimella, jonka kautta käsitteitä voidaan hakea esimerkiksi niihin liittyvien termien perusteella, ja tätä kautta selvittää käsitteen URI. Yhteisten käsitteistöjen ja ONKI-palvelimen käytöstä kertoo enemmän dokumentti Hajautettu sisällöntuotanto TerveSuomi-portaalissa [Ter07].

¹¹Huom. DSCV-syntaksi on käytössä toistaiseksi kunnes Dublin Core -yhteisö julkaisee uuden, RDF-pohjaisen esitysmuodon DSCV-tiedon esittämiseen.

4.6 Yksilöt

Yksilöt (henkilöt, organisaatiot ja ryhmät) voidaan ilmaista vapaatekstyypisesti esittämällä ne RDF-muodossa foaf:Person, foaf:Organization tai foaf:Group-luokkien ilmentyminä seuraavaan tapaan:

```
<dc:creator>
  <foaf:Person>
    <foaf:name>Virtanen, Matti</foaf:name>
  </foaf:Person>
</dc:creator>

<dc:publisher>
  <foaf:Organization>
    <foaf:name xml:lang="fi">Kansanterveyslaitos</foaf:name>
    <foaf:name xml:lang="sv">Folkhälsoinstitutet</foaf:name>
    <foaf:name xml:lang="en">National Public Health Institute</foaf:name>
  </foaf:Organization>
</dc:publisher>
```

Henkilön, organisaation tai ryhmän nimi ilmaistaan foaf:name-ominaisuudella. Useita erikielisiä versioita nimistä voidaan antaa, kunhan ne on merkitty kielikoodiin.

Vaihtoehtoisesti voidaan viitata erikseen määriteltyihin (ks. kohta 4.9) yksilöihin niiden URIn avulla:

4.7 Dokumentti- ja yksilöviittaukset

Dokumentteihin ja yksilöihin viitataan samaan tapaan kuin käsitteisiin viittaamalla niiden URIn. Esimerkki:

```
<dc:rights rdf:resource="http://www.ktl.fi/portal/suomi/palaute/tietoa_sivuista/">
```

4.8 Esimerkki RDF/XML-muotoisesta metatiedosta

Alla on esimerkki RDF/XML-muotoisesta metatietodokumentista, joka on tuotettu metatietotason 2 mukaisesti (eli ei organisaatiokohtaisia yksilömäärittäjiä). Esimerkki perustuu todelliseen dokumenttiin (KTL:n Rokottajan käsikirjan verkkoversion etusivu), jonka metatietoja on kuitenkin täydennetty ja muokattu vastaamaan tämän dokumentin suosituksia. Lisäksi XML-kommentit on kerrottu joidenkin käsitteiden päätermit (esim. "verkkoresursi"), jotka helpottavat esimerkin ymmärtämistä.

```
<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:dcterms="http://purl.org/dc/terms/"
  xmlns:ts="http://www.yso.fi/onto/tervesuomi-schema/">
  <ts:Publication rdf:about="http://www.ktl.fi/portal/8131">

    <dc:identifier rdf:resource="http://www.ktl.fi/portal/8131"/>

    <dc:title>Rokotteiden hävittäminen</dc:title>
    <dcterms:abstract>Vanhentuneet tai väärän käsittelyn takia tuhoutuneet
      rokotteet on hävitettävä. Hävitystapa riippuu siitä, millaisesta
      rokotteesta on kyse.</dcterms:abstract>
    <dc:language>
      <dcterms:RFC3066>
        <rdf:value>fi</rdf:value>
      </dcterms:RFC3066>
    </dc:language>

    <dcterms:issued>
      <dcterms:W3CDTF>
        <rdf:value>2005-01-02</rdf:value>
      </dcterms:W3CDTF>
```

```

</dcterms:issued>

<ts:genre
  rdf:resource="http://www.yso.fi/onto/terveysuomi-genre/guide"/> <!--opas tai ohje-->

<dc:type rdf:resource="http://purl.org/dc/dcmitype/Text"/> <!--Teksti-->
<dc:format>
  <dcterms:IMT>
    <rdf:value>text/html</rdf:value>
  </dcterms:IMT>
</dc:format>
<dcterms:medium rdf:resource="http://www.yso.fi/onto/medium/www"/> <!--verkkoresurssi-->

<dc:subject rdf:resource="http://www.yso.fi/onto/yso/a12345"/> <!--rokotteet-->
<dc:subject rdf:resource="http://www.yso.fi/onto/yso/a12346"/> <!--hävittäminen-->
<dc:subject rdf:resource="http://www.yso.fi/onto/mesh/D015164"/> <!--inaktiivirokotteet-->

<dcterms:audience rdf:resource="http://www.yso.fi/onto/audience/A0076"/> <!--asiantuntijat-->
<dcterms:audience rdf:resource="http://www.yso.fi/onto/audience/A0094"/> <!--aikuiset-->
<dcterms:audience rdf:resource="http://www.yso.fi/onto/audience/A0641"/> <!--sairaanhoidajat-->
<dcterms:audience rdf:resource="http://www.yso.fi/onto/audience/A0636"/> <!--terveydenhoitajat-->
<dcterms:audience rdf:resource="http://www.yso.fi/onto/audience/A0703"/> <!--lääkärit-->
<dcterms:audience rdf:resource="http://www.yso.fi/onto/audience/A0642"/> <!--farmaseutit-->

<dcterms:isPartOf rdf:resource="http://www.ktl.fi/portal/8134"/>

<dc:rights rdf:resource="http://www.ktl.fi/portal/suomi/palaute/tietoa_sivuista"/>

<!-- metatietotason 1 mukaisesti esitetyt organisaatiokohtaiset yksilöt -->

<dc:publisher>
  <foaf:Organization>
    <foaf:name xml:lang="fi">Kansanterveyslaitos</foaf:name>
    <foaf:name xml:lang="sv">Folkhälsoinstitutet</foaf:name>
    <foaf:name xml:lang="en">National Public Health Institute</foaf:name>
  </foaf:Organization>
</dc:publisher>

<dc:creator>
  <foaf:Person>
    <foaf:name>Nohynek, Hanna</foaf:name>
  </foaf:Person>
</dc:creator>

</ts:Publication>
</rdf:RDF>

```

4.9 Organisaatiokohtaisten yksilöiden esittäminen

Organisaatiokohtaiset yksilöt esitetään metatietotasolla 3 siten, että kullekin yksilölle annetaan oma yksilöivä URI-tunniste ja lisäksi yksilöstä kerrotaan metatietoa RDF-muodossa. Organisaatio vastaa siitä, että yksilöiden URI-tunnisteet ovat pysyviä ja yksiselitteisiä.

Yksittäisen henkilön tiedot esitetään FOAF-skeemaa [BM05] hyödyntäen seuraavasti:

```

<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:dcterms="http://purl.org/dc/terms/"
  xmlns:foaf="http://xmlns.com/foaf/0.1/">

  <foaf:Person rdf:about="http://www.ktl.fi/person/1234">
    <foaf:firstName>Hanna</foaf:firstName>
    <foaf:family_name>Nohynek</foaf:family_name>

    ... mahd. lisää tietoa henkilöstä FOAFin avulla ilmaistuna ...

  </foaf:Person>
</rdf:RDF>

```

Yksittäisen organisaation tiedot esitetään FOAF-skeeman [BM05] avulla seuraavasti:

```

<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:dcterms="http://purl.org/dc/terms/"
  xmlns:foaf="http://xmlns.com/foaf/0.1/">

  <foaf:Organization rdf:about="http://www.ktl.fi/organization/1234">
    <foaf:name>Kansanterveyslaitos</foaf:name>

    <!-- Hanna Nohynek kuuluu ao. organisaation resursseihin (henkilökuntaan) -->
    <foaf:member rdf:resource="http://www.ktl.fi/person/1234" />

    ... mahd. lisää tietoa organisaatiosta FOAF:in avulla ilmaistuna ...

  </foaf:Organization>
</rdf:RDF>

```

FOAF on vakiintunut tapa ilmaista henkilötietoja RDF-muodossa, mutta erityisesti ryhmien ja organisaatioiden kuvaamista ajatellen se ei välttämättä sisällä kaikkia tarvittavia tietokenttiä. Tämän vuoksi yksilöiden määrittelytapaan saattaa tulevaisuudessa tulla muutoksia.

5 Metatiedon esittäminen osana HTML-sivua

Mikäli edellä kuvattua RDF-muotoista metatietodokumenttia ei voida tai haluta tuottaa erillisenä dokumenttina, voidaan metatieto myös sisällyttää HTML-sivun `head`-osioon. Tämä ratkaisu noudattaa Dublin Coren ehdotusta[DCQ03] metatiedon esittämiseen HTML-sivun sisällä.

Menetelmän hyvänä puolena on se, että monet julkaisujärjestelmät tukevat Dublin Core -metatiedon sisällyttämistä HTML-sivuihin jo valmiiksi. Menetelmän rajoituksena on, ettei sitä voi soveltaa muihin kuin HTML/XHTML-muotoisiin dokumentteihin. Esimerkiksi PDF-dokumenttien metatiedot on kuvattava muulla tavalla. Menetelmän avulla ei myöskään voida määritellä tietoja yksilöistä kuten henkilöistä ja organisaatioista, vaikka muualla RDF-muodossa (ks. kohta 4.9) määriteltyihin yksilöihin voidaan viitata.

Useimmat ohjeet koskevat sekä HTML- että XHTML-dokumentteja. Mikäli näiden esitysmuoto eroaa, se on mainittu erikseen. Esimerkit ovat XHTML-syntaksin mukaisia. HTML-muoto eroaa näistä siten, että elementin sulkemista merkitsevä kauttaviiva `</>` jää pois.

HTML-tiedostossa tulee esitellä käytetty merkistö (esim. UTF-8 tai ISO-8859-1) käytössä olevan HTML-version mukaisella standarditavalla. HTML-dokumentin sisältämä metatieto ilmaistaan käyttäen samaa merkistöä kuin itse dokumenttikin. Muut merkit tulkitaan virheitilanteiksi.

5.1 Kenttätunnisteet

Kenttätunnisteet HTML-muodossa poikkeavat jonkin verran RDF-muodosta. Tunnistemuoto perustuu Dublin Coren suositukseen [DCQ03]¹². HTML-muodossa käytettävät kenttätunnisteet on kuvattu taulukossa 2.

¹²Tunnisteista on olemassa myös vanhempi muoto aiemmasta Dublin Coren suosituksesta [DCQ00], joka on käytössä myös Dublin Coren suomalaisessa versiossa [SFS01]. Aiemmassa muodossa esimerkiksi Julkistamisaika-kentän tunniste on `DC.Date.issued` eikä `DCTERMS.issued`. Tämä dokumentti ei ota kantaa siihen, hyväksyykö TerveSuomi-portaali muita kenttätunnistemuotoja kuin tässä dokumentissa kuvatut muodot.

Taulukko 2: Metatietokenttien HTML-tunnisteet

Kenttä	Tunniste	HTML-muoto
Identifiointitunnus	<code>dc:identifier</code>	<code>DC.identifier</code>
URL-osoite	<code>ts:url</code>	<code>TS.url</code>
Nimeke	<code>dc:title</code>	<code>DC.title</code>
Kuvaus	<code>dcterms:abstract</code>	<code>DCTERMS.abstract</code>
Kieli	<code>dc:language</code>	<code>DC.language</code>
Julkistamisaika	<code>dcterms:issued</code>	<code>DCTERMS.issued</code>
Hyväksymisaika	<code>dcterms:dateAccepted</code>	<code>DCTERMS.dateAccepted</code>
Muokkusaika	<code>dcterms:modified</code>	<code>DCTERMS.modified</code>
Julkaisija	<code>dc:publisher</code>	<code>DC.publisher</code>
Tekijä	<code>dc:creator</code>	<code>DC.creator</code>
Julkaisun laji	<code>ts:genre</code>	<code>TS.genre</code>
Esitysmuoto	<code>dc:type</code>	<code>DC.type</code>
Formaatti	<code>dc:format</code>	<code>DC.format</code>
Mediatyyppi	<code>dcterms:medium</code>	<code>DCTERMS.medium</code>
Aihe	<code>dc:subject</code>	<code>DC.subject</code>
Kohdeyleisö	<code>dcterms:audience</code>	<code>DCTERMS.audience</code>
Alueellinen kattavuus	<code>dcterms:spatial</code>	<code>DCTERMS.spatial</code>
Ajallinen kattavuus	<code>dcterms:temporal</code>	<code>DCTERMS.temporal</code>
Kuuluu	<code>dcterms:isPartOf</code>	<code>DCTERMS.isPartOf</code>
Oikeudet	<code>dc:rights</code>	<code>DC.rights</code>
Lähde	<code>dc:source</code>	<code>DC.source</code>
Viittaa	<code>dcterms:references</code>	<code>DCTERMS.references</code>
Käännöksen lähde	<code>ts:isTranslationOf</code>	<code>TS.isTranslationOf</code>
Aikaisempi formaatti	<code>dcterms:isFormatOf</code>	<code>DCTERMS.isFormatOf</code>

Dublin Coren suosituksen mukaisesti kenttätunnisteiden tulkinta tehdään yhteensopivuussyistä merkkikoosta riippumatta [DCQ03].

5.2 Nimiavaruusmäärittelyt

Kenttätunnisteissa käytettävät nimiavaruudet esitellään link-elementein seuraavasti:

```
<link rel="schema.DC" href="http://purl.org/dc/elements/1.1/" />
<link rel="schema.DCTERMS" href="http://purl.org/dc/terms/" />
<link rel="schema.TS" href="http://www.yso.fi/onto/tervesuomi-schema/" />
```

Tämä esittely on vakio eli toistuu sellaisenaan kaikissa metatietoa sisältävissä dokumenteissa.

5.3 Dokumentin tunniste ja URL-osoite

Dokumentin tunniste (`dc:identifier`) ja URL-osoite (`ts:url`) esitetään meta-elementillä seuraavasti:

Pysyvä tunniste:

```
<meta name="DC.identifier" scheme="DCTERMS.URI" content="http://www.ktl.fi/portal/8131" />
```

Tämänhetkinen URL-osoite:

```
<meta name="TS.url" scheme="DCTERMS.URI"
content="http://www.ktl.fi/portal/suomi/julkaisut/oppaat_ja_kirjat/rokottajan_kasikirja/" />
```

5.4 Vapaatekstikentät

Vapaatekstikentät esitetään meta-elementillä. Mikäli tekstin kieli on tiedossa, ilmaistaan se erityisellä kielimääreellä. HTML-muodossa kielimääreenä toimii `lang`-attribuutti. XHTML Strict -muodoissa käytetään `xml:lang`-attribuuttia. XHTML Transitional -muodoissa käytetään molempia attribuutteja. Esimerkki:

HTML:

```
<meta name="DC.title" lang="fi" content="Rokotteiden hävittäminen">
```

XHTML Strict:

```
<meta name="DC.title" xml:lang="fi" content="Rokotteiden hävittäminen" />
```

XHTML Transitional:

```
<meta name="DC.title" lang="fi" xml:lang="fi" content="Rokotteiden hävittäminen" />
```

Kielikoodit ovat RFC 3066 -standardin mukaisia, kuten Kieli-kentän (`dc:language`) arvot [RFC01].

5.5 Rakenteiset merkkijonot

Rakenteiset merkkijonot esitetään meta-elementillä. Kentän tunnisteeseen ja arvon lisäksi kerrotaan `scheme`-attribuutilla kentän arvojoukon tunniste.

Esitysmuotojen tunnisteet rakenteisille merkkijonoille ovat seuraavat:

Tyyppi	scheme-tunniste
Päiväys	DCTERMS.W3CDTF
Formaatti (MIME-tyyppi)	DCTERMS.IMT
Kieli	DCTERMS.RFC3066
Aikajakso	DCTERMS.Period
Maantieteellinen paikka (pistemäinen)	DCTERMS.Point
Maantieteellinen paikka (suorakaide)	DCTERMS.Box

Esimerkki rakenteisista merkkijonoista:

```
<meta name="DCTERMS.issued" scheme="DCTERMS.W3CDTF" content="2004-12-17" />
<meta name="DCTERMS.temporal" scheme="DCTERMS.Period"
      content="name=Talvisota; start=1939-11-30; end=1940-03-13;" />
```

5.6 Yksilöt ilman viittausta

Yksilöt (henkilöt, organisaatiot ja ryhmät) voidaan ilmaista vapaatekstityyppisesti *meta*-elementillä, jonka arvona on yksilön nimi. Tällöin *scheme*-attribuutti kertoo sen, onko kysymyksessä henkilö, organisaatio vai ryhmä seuraavien tunnisteiden avulla:

Tyyppi	<i>scheme</i> -tunniste
Henkilö	TS.FOAFPerson
Organisaatio	TS.FOAFOrganization
Ryhmä	TS.FOAFGroup

Useita erikielisiä versioita nimistä voidaan antaa, kunhan ne on merkitty kielikoodein samaan tapaan kuin vapatekstikenttien kohdalla.

Esimerkki yksilöistä:

```
<meta name="DC.creator" scheme="TS.FOAFPerson" content="Virtanen, Matti" />
<meta name="DC.publisher" scheme="TS.FOAFOrganization" xml:lang="fi"
      content="Kansanterveyslaitos" />
<meta name="DC.publisher" scheme="TS.FOAFOrganization" xml:lang="sv"
      content="Folkhälsoinstitutet" />
<meta name="DC.publisher" scheme="TS.FOAFOrganization" xml:lang="en"
      content="National Public Health Institute" />
```

5.7 Käsite-, dokumentti- ja yksilöviittaukset

Käsite-, dokumentti- ja yksilöviittaukset esitetään *link*-elementillä seuraavaan tapaan:

```
<link rel="DC.subject" href="http://www.yso.fi/onto/yso/a12345" />
<link rel="DC.rights" href="http://www.ktl.fi/portal/suomi/palaute/tietoa_sivuista/" />
```

Käsiteviittaukset voidaan vaihtoehtoisesti tai lisäksi esittää *meta*-elementillä. Jokaista käsitettä varten on tällöin oma *meta*-elementtinsä. Tällaisia viittauksia saatetaan hyödyntää myös muissa sovelluksissa kuin TerveSuomi.fi:ssä, joten niiden lisääminen URI-käsiteviittausten rinnalle voi olla järkevää. TerveSuomi-portaali pystyy hyödyntämään molemmiin tavoin esitettyjä viittauksia, eikä ole haitaksi, jos sama tieto esitetään kahteen kertaan.

Kun käsiteviittaukset tehdään *meta*-elementeillä, tulkitaan tämä TerveSuomi.fi:ssä hyödyntäen annetun käsitteistön viimeisintä versiota ja muunnetaan URI-muotoon. Näin tehdyt viittaukset voivat lakata toimimasta, jos esimerkiksi käsitteiden nimet muuttuvat. Tiedontuottaja vastaa siitä, että metatieto hyödyntää aina viimeisintä käsitteistöä ja että käsitteet joihin viitataan ovat sillä hetkellä olemassa. On suositeltavaa käyttää URI-käsitetunnuksia, koska ne ovat aina voimassa vaikka käsitteistöä olisi julkaistu uudempiakin versioita.

Esimerkki *meta*-elementillä esitetyistä käsitteistä:

```
<meta name="DC.subject" scheme="TS.TESA" content="lapset" />
<meta name="DC.subject" scheme="TS.TESA" content="nuoret" />
<meta name="DC.subject" scheme="TS.TESA" content="aikuiset" />
```

Kentän tunnusteen (*name*) ja arvon (*content*) lisäksi kerrotaan *scheme*-attribuutilla sen käsitteistön tunniste, johon kentän arvo kuuluu. Kentän arvo on käsitteeseen liittyvä ensisijainen

termi (preferred term) täsmälleen siinä muodossa, jossa se esiintyy käsitteistössä. Joihinkin kenttiin voi soveltaa useita käsitteistöjä tai samoihin käsitteisiin voi viitata erikielisten termien kautta (englanninkielinen MeSH vs. suomenkielinen FinMeSH).

Käsitteistötunnisteet käsiteviittauksille ovat seuraavat:

Kenttä	Käsitteistö	Käsitteistön tunniste	Huomautukset
Julkaisutyyppi Esitysmuoto	Julkaisutyyppi DCMI Type Vocabulary	TS.GENRE DCTERMS.DCMIType	Suomenkieliset termit Englanninkieliset termit
Mediatyyppi	TerveSuomi- mediatyyppi	TS.MEDIUM	Suomenkieliset termit
Aihe	YSO	TS.YSO	Suomenkieliset termit
Aihe	FinMeSH	TS.FINMESH	Suomenkieliset termit
Aihe	MeSH	DCTERMS.MESH	Englanninkieliset termit
Aihe	TESA	TS.TESA	Suomenkieliset termit
Kohdeyleisö	TerveSuomi- kohderyhmät	TS.AUDIENCE	Suomenkieliset termit
Alueellinen kat- tavuus	Paikkaontologia	TS.SUO	Suomenkieliset termit
Ajallinen katta- vuus	Aikaontologia	TS.AIKA	Suomenkieliset termit

Yksilöviittauksia voi tehdä HTML-sivussa edellä kuvatulla tavalla link-elementeillä, mutta itse yksilöä koskevat tiedot (esim. henkilön tai organisaation nimi ja osoite) tulee määritellä URI-tunnisteen osoittamassa URL-osoitteessa, RDF-muodossa. Tämä johtuu siitä, että HTML:n metatietomäärittelyjen kohde on aina dokumentti itse, joten dokumentista riippumaton yksilöä koskevan metatiedon antaminen ei HTML:ssä onnistu.

5.8 Esimerkki

Alla on esimerkki XHTML Strict -muotoisesta dokumentista, joka sisältää metatietoa. Esimerkki perustuu todelliseen dokumenttiin (KTL:n Rokottajan käsikirjan verkkoversion aliluku Rokotteiden hävittäminen), jonka metatietoja on kuitenkin täydennetty ja muokattu vastaamaan tämän dokumentin suosituksia sekä XHTML Strict -muotoa.

Käsiteviittaukset on esimerkinomaisesti tehty vaihdellen joko meta- tai link-elementillä tai molemmilla. Esimerkiksi dc:type-kenttä on esitetty ainoastaan meta-elementillä kun taas ts:genre-kenttä on esitetty ainoastaan link-elementillä ja dcterms:medium on esitetty molemmilla.

```
<html>
<head>
  <title>KTL | Rokotteiden hävittäminen</title>

  <!-- nimiavaruusmäärittelyt -->

  <link rel="schema.DC" href="http://purl.org/dc/elements/1.1/" />
  <link rel="schema.DCTERMS" href="http://purl.org/dc/terms/" />
  <link rel="schema.TS" href="http://www.yso.fi/onto/tervesuomi-schema/" />

  <!-- metatietokentät -->

  <meta name="DC.identifier" scheme="DCTERMS.URI"
        content="http://www.ktl.fi/portal/8140" />

  <meta name="DC.title" content="Rokotteiden hävittäminen" />
  <meta name="DCTERMS.abstract" content="Vanhentuneet tai väärän käsittelyn
    takia tuhoutuneet rokotteet on hävitettävä. Hävitystapa riippuu siitä,
    millaisesta rokotteesta on kyse." />

  <meta name="DC.language" scheme="DCTERMS.RFC3066" content="fi" />

  <meta name="DCTERMS.issued" scheme="DCTERMS.W3CDTF" content="2005-01-02" />

  <!-- julkaisun laji esitetty pelkästään link-mekanismilla -->
  <link rel="TS.genre"
        href="http://www.yso.fi/onto/tervesuomi-genre/guide" />
```

```

<!-- esitysmuoto ja formaatti esitetty pelkästään meta-mekanismilla -->
<meta name="DC.type" scheme="DCTERMS.DCMIType" content="Text" />
<meta name="DC.format" scheme="DCTERMS.IMT" content="text/html" />

<!-- mediatyyppi esitetty sekä link- että meta-mekanismilla -->
<link rel="DCTERMS.medium" href="http://www.yso.fi/onto/medium/www" />
<meta name="DCTERMS.medium" scheme="TS.MEDIUM" content="Verkkoresurssi" />

<!-- aiheet ja kohdeyleisöt esitetty pelkästään meta-mekanismilla -->
<meta name="DC.subject" scheme="TS.YSO" content="rokotteet" />
<meta name="DC.subject" scheme="TS.YSO" content="hävittäminen" />
<meta name="DC.subject" scheme="TS.FINMESH" content="inaktiivirokotteet" />

<meta name="DCTERMS.audience" scheme="TS.AUDIENCE" content="asiantuntijat" />
<meta name="DCTERMS.audience" scheme="TS.AUDIENCE" content="maallikot" />
<meta name="DCTERMS.audience" scheme="TS.AUDIENCE" content="aikuiset" />
<meta name="DCTERMS.audience" scheme="TS.AUDIENCE" content="sairaanhoidajat" />
<meta name="DCTERMS.audience" scheme="TS.AUDIENCE" content="terveydenhoitajat" />
<meta name="DCTERMS.audience" scheme="TS.AUDIENCE" content="lääkärit" />
<meta name="DCTERMS.audience" scheme="TS.AUDIENCE" content="farmaseutit" />

<!-- viittaus Rokottajan oppaan lukuun Perustietoa rokotteista, jonka
aliluku tämä dokumentti on -->
<link rel="DCTERMS.isPartOf" href="http://www.ktl.fi/portal/8134" />

<!-- viittaus erilliseen sivuun, joka määrittelee käyttöoikeudet -->
<link rel="DC.rights" href="http://www.ktl.fi/portal/suomi/palaute/tietoa_sivuista/" />

<!-- metatietotason 1 mukaisesti esitetyt organisaatiokohtaiset yksilöt -->
<meta name="DC.publisher" scheme="TS.FOAFOrganization" xml:lang="fi"
content="Kansanterveyslaitos" />
<meta name="DC.publisher" scheme="TS.FOAFOrganization" xml:lang="sv"
content="Folkhälsoinstitutet" />
<meta name="DC.publisher" scheme="TS.FOAFOrganization" xml:lang="en"
content="National Public Health Institute" />

<meta name="DC.creator" scheme="TS.FOAFPerson" content="Nohynek, Hanna" />

<!-- metatieto päättyy -->

[... stylesheet, javascript ym. määrittelyksiä ...]
</head>
<body> ... </body>
</html>

```


6 Ontologiat

TerveSuomi.fi:n hajautettu sisällöntuotanto perustuu yhteisten ontologioiden hyödyntämiseen, joissa määritellään sisällön kuvailuun käytettävät käsitteet. Tässä luvussa esitellään kaikki käytettävät ontologiat. Ne löytyvät myös ontologiapalvelimelta www.yso.fi.

Ontologiapalvelin tarjoaa ontoloihin pääsyn sekä selailtavan ONKI-web-käyttöliittymän kautta, Web Service -rajapintojen kautta (kts. web-käyttöliittymän sivu Tietoa Onkista) että koneuettavassa RDF-muodossa.

Kunkin ontologian URL-osoite on muotoa <http://www.yso.fi/onto/tunniste/> missä *tunniste* on annettu taulukossa 3.

Taulukko 3: TerveSuomi-portaalin käyttämät ontologiat

Nimi	Tunniste	Kuvaus
Yleinen suomalainen ontologia (YSO)	yso	Yleinen suomalainen ontologia (YSO) sisältää suuren joukon jokapäiväiseen elämään liittyviä käsitteitä. Se toimii TerveSuomi.fi:ssä ns. yläontologiana, johon muut ontologiat, kuten MeSH ja TESA, kytketään. Yhteinen yläontologia edistää ontologioiden välistä yhteentoimivuutta, mikä edistää mm. loppukäyttäjälle näkyvän käyttöliittymän sisällöllistä yhtenäisyyttä.
MeSH ja FinMeSH	mesh	MeSH ^a (Medical Subject Headings) on kansainvälinen lääketieteellinen käsitteistö ja FinMeSH sen suomenkielinen käännös.
Terveyden edistämisen sanasto	hpmulti	Terveyden edistämisen sanasto ^b on Euroopan komission julkaisema sanasto terveyden edistämiseen liittyvän materiaalin luokitteluun.
Mediaontologia	medium	Mediaontologiassa on luokiteltu käsitteitä, jotka kuvailevat erilaisia sisällön välitystapoja (esim. verkko, kirja, CD-levy). Toistaiseksi se sisältää ainoastaan yhden käsitteen Verkkoresurssi; sitä laajennetaan, jos TerveSuomi-portaaliin otetaan mukaan muun tyyppisiä tietolähteitä.
Kohdeyleisöontologia	audience	Kohdeyleisöontologiassa on luokiteltu käsitteitä, jotka kuvailevat erilaisia ihmisryhmiä, joille viestintää voi suunnata. Käsite ihmisryhmät jakautuu noin 20:een alikäsitteeseen, joita ovat esimerkiksi ammattilaisryhmät, riskiryhmät, ikäryhmät ja sukupuoliryhmät.
Julkaisun laji -sanasto (Genre)	terveysuomi-genre	Julkaisun laji -sanastolla kuvataan julkaisun pääasiallinen sisällöllinen laji tai genre.
DCMI-tyyppisanasto	dcmitype	DCMI-tyyppisanasto (DCMI Type Vocabulary) on Dublin Coren julkaisema yleiskäyttöinen, sovelsalarippumaton luokitus metatiedon kohteiden luokitteluun. TerveSuomi-portaali käyttää DCMI-tyyppisanastoa sellaisenaan. Ontologiapalvelimella olevaan versioon on lisätty suomenkieliset käännökset termeistä ja määritelmistä.
Paikkaontologia (SUO)	suo	Paikkaontologia sisältää Suomen paikat, kuten Helsinki ja Länsi-Suomen lääni.
Aikaontologia (Aika)	aika	Aikaontologia sisältää ajan ilmaisemiseen liittyviä käsitteitä.
TerveSuomi-metatietoskeema	terveysuomi-schema	Metatietoskeema määrittelee TerveSuomi.fi:n metatietokentät formaalissa muodossa.

^a<http://www.nlm.nih.gov/mesh/>

^b<http://www.hpmulti.net>

Viitteet

- [BM05] Dan Brickley and Libby Miller. Foaf vocabulary specification, June 3 2005. <http://xmlns.com/foaf/0.1/>.
- [DC-02] Dublin Core metadataformaatin suomalainen versio, 2002. http://www.lib.helsinki.fi/dublin_core/dc-sfs.html.
- [DC-03] Guidelines for implementing Dublin Core in XML, 2003. <http://dublincore.org/documents/dc-xml-guidelines/>.
- [DCE04] Dublin Core metadata element set, version 1.1, 2004. <http://dublincore.org/documents/dces/>.
- [DCM02] Expressing simple Dublin Core in RDF/XML, 2002. <http://dublincore.org/documents/dcmes-xml/>.
- [DCM06a] DCMI Box encoding scheme: specification of the spatial limits of a place, and methods for encoding this in a text string, 2006. <http://dublincore.org/documents/dcmi-box/>.
- [DCM06b] DCMI Period encoding scheme: specification of the limits of a time interval, and methods for encoding this in a text string, 2006. <http://dublincore.org/documents/dcmi-period/>.
- [DCM06c] DCMI Point encoding scheme: a point location in space, and methods for encoding this in a text string, 2006. <http://dublincore.org/documents/dcmi-point/>.
- [DCQ00] Recording qualified Dublin Core metadata in HTML meta elements, 2000. <http://dublincore.org/documents/2000/08/15/dcq-html/>.
- [DCQ02] Expressing qualified Dublin Core in RDF/XML, 2002. <http://dublincore.org/documents/dcq-rdf-xml/>.
- [DCQ03] Expressing Dublin Core in HTML/XHTML meta and link elements, 2003. <http://dublincore.org/documents/dcq-html/>.
- [DCS06] DCMI DCSV: A syntax for representing simple structured data in a text string, 2006. <http://dublincore.org/documents/dcmi-dcsv/>.
- [DCT06] DCMI metadata terms, 2006. <http://dublincore.org/documents/dcmi-terms/>.
- [Eer06] Johanna Eerola. Ehdotus metadatakentiksi. Hyväksytty KTL:n Sanastotyöpajan kokouksessa 24.8.2006, Sep 15 2006.
- [ISO88] ISO 8601: Data elements and interchange formats - information interchange - representation of dates and times, 1988.
- [JHS04] JHS 143: Asiakirjojen kuvailun ja hallinnan metatiedot, 2004. <http://www.jhs-suositukset.fi/suomi/jhs143>.
- [mim] Mime media types. <http://www.iana.org/assignments/media-types/>.
- [RDF] Resource description framework (RDF). <http://www.w3.org/RDF>.
- [RFC01] RFC 3066: Tags for the identification of languages, 2001. <http://www.ietf.org/rfc/rfc3066.txt>.
- [SFS01] SFS 5895: Dublin Core -metadataformaatin suomalainen versio, 2001. Suomen Standardoimisliitto.
- [Ter07] Hajautettu sisällöntuotanto TerveSuomi.fi-portaalissa, 2007.
- [Van06] Uniform requirements for manuscripts submitted to biomedical journals: Writing and editing for biomedical publication, 2006. <http://www.icmje.org/>.
- [W3C97] Date and time formats, 1997. <http://www.w3.org/TR/NOTE-datetime>.