

Informaatiotutkimuksen päivät 2010
21. - 22. lokakuuta, Tampere

ABSTRAKTI

Marjatta Autio-Tuuli & Marja-Liisa Seppälä

Kirjaston vastaus informaatioympäristön muutokseen : Semanttinen web, ontologiat ja standardit osana kirjastojen tiedonhallinnan infrastruktuuria

Marjatta Autio-Tuuli, Kansalliskirjasto, marjatta.autio-tuuli@helsinki.fi
Marja-Liisa Seppälä, Kansalliskirjasto, marja-liisa.seppala@helsinki.fi

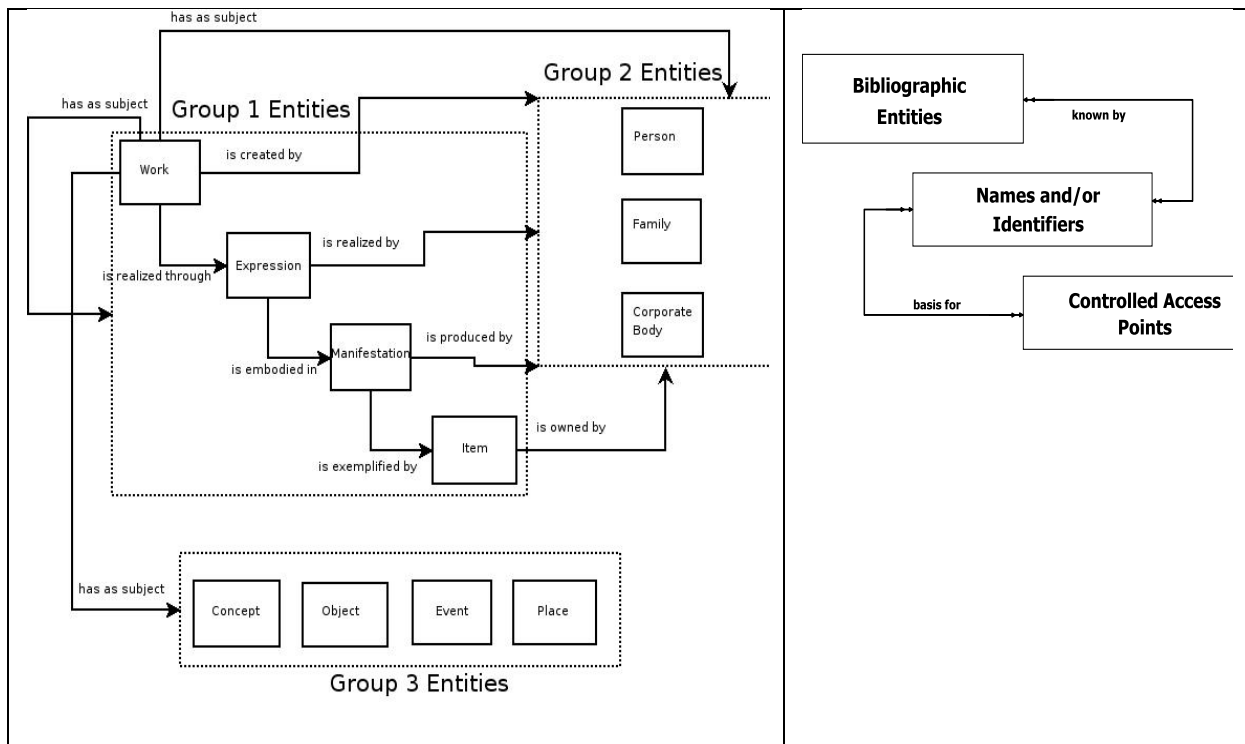
Informaatioympäristön muuttuminen

Informaatioympäristön keskeisiä muutoksia aineiston kuvailutietojen näkökulmasta ovat informaation ja julkaisemisen uudet muodot. Sisältöjen muuttuessa yhä enemmän digitaaliseen muotoon informaation tallentaminen ja välittäminen on moninaistunut. Samalla julkaisemisen määrä ja tiedonsaannin mahdollisuudet ovat lisääntyneet. Tiedonsaannin helpottuessa käyttäjä etsii tietoa omatoimisesti ilman tiedonhaun ohjausta.

Digitaalinen toimintaympäristö merkitsee uusia vaatimuksia ja mahdollisuuksia metatietojen tallentamiselle, esittämiselle ja käytettävyydelle. Tiedonhakijalle voidaan tarjota mielekästä ja tarkoituksenmukaista linkittämiseen perustuvaa navigointiympäristöä. Erityyppisten metatiedontuottajien yhteistyöllä ja yhteensopivilla standardeilla edistetään metatiedon välittämistä ja rikastamista.

Uusi kuvailustandardi - uusi ajattelutapa

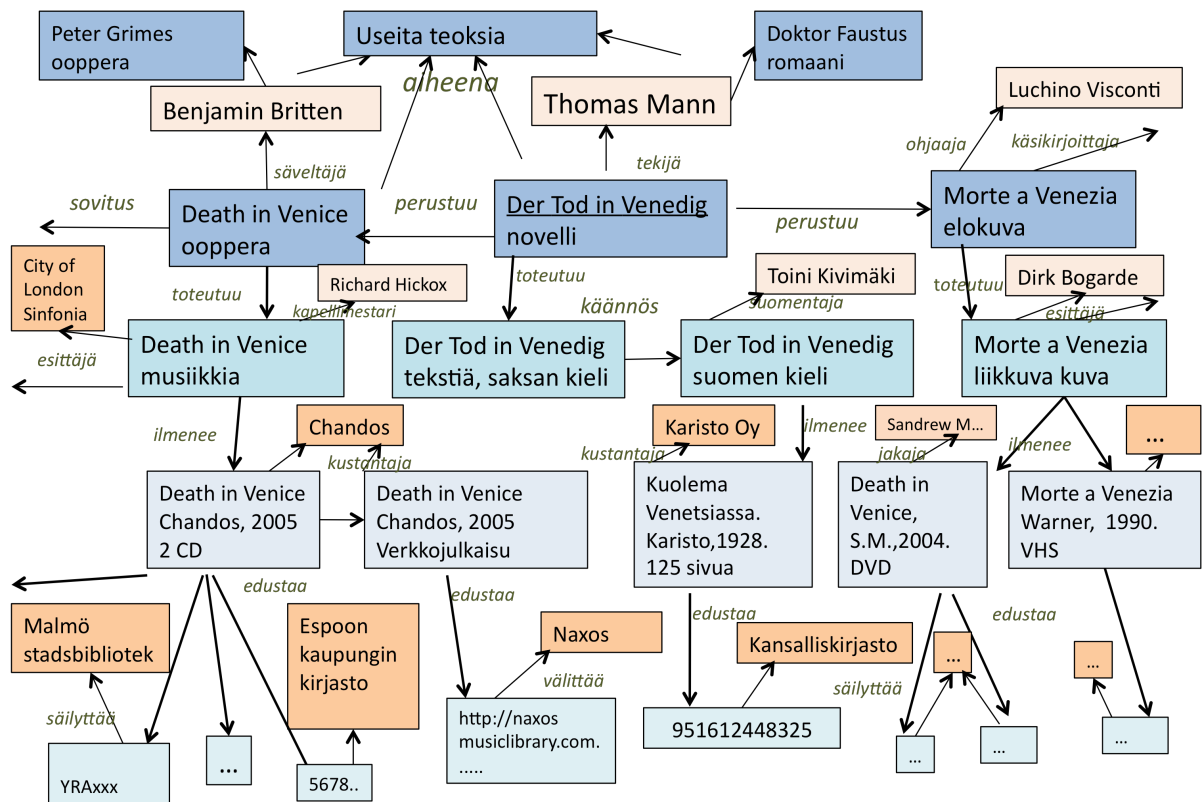
Uusi luettelointistandardi RDA (RDA 2008) on vastaus informaatioympäristön muutokseen. Anglo-amerikkalaisiin luettelointisääntöihin AACR2:een (Anglo-American Cataloguing Rules) pohjautuva RDA ilmestyi kesäkuussa 2010. Keskeinen uusi asia luettelointisäännöissä on kuvailu- ja hakutieto-ohjeiden perustuminen tiedonmallinnuksen teoriaan. FRBR (IFLA Study Group 1997) ja sen johdannaiset FRAD (IFLA Working Group on Functional Requirements and Numbering 2009) ja FRSAD (IFLA Working Group on Functional Requirements for Subject 2009) muodostavat RDA:n teoreettisen viitekehyksen. Ko. käsittemallit määrittelevät tiedonhakijan mielenkiinnon kohteet, entiteetit, ja näitä kuvailevat tiedot, attribuutit, sekä entiteettien keskinäiset suhteet.



Kuva 1
FRBR-käsitelmä ja FRAD-käsitelmä

FRBR-mallissa määritellyt tiedonhakijoiden tarpeet ohjaavat metatietojen muodostamista. Funktionaalisen kuvailun tulee auttaa tiedonhakijaa löytämään, tunnistamaan, valitsemaan ja saamaan käyttötarkoitukseensa sopivaa tietoa sekä navigoimaan entiteettien välille luotujen suhteiden avulla. Lisäksi FRAD-malli tukee tiedon löytymistä yhdistämällä entiteetin ensisijaisen nimimuodon ja sen varianttimuodot. Ensisijaisesta nimimuodosta ja muista kuvailutiedoista muodostettu hakutieto kokoaa toisiinsa funktionaalisessa suhteessa olevia entiteettejä. (Ks. kuva 1)

RDA-ohjeita noudattava metatieto tallennetaan vastaamaan FRBR-mallissa määriteltyjä tiedonhakijan tarpeita. Keskeinen tavoite kuvailussa on sisällön nostaminen esiin teos/ekspressio-tason kautta. Tämä tukee sisältölähtöistä tiedonhakua moniulotteisessa digitaalisessa tietomassassa. Entiteettien linkittäminen toisiinsa tarjoaa tiedonhakijalle metatietokontekstin ja joustavan liikkumisen metatiedossa (ks. kuva 2). Linkittäminen yhdessä funktionaalisen metatiedon palastelun (rakenteistamisen) kanssa mahdollistaa ekonomisen metatiedon tuottamisen: jokainen tietoelementti tallennetaan vain kerran.



Kuva 2

Esimerkki RDA-kuvailusta

Entiteettiryhmä 1:

■ = teos ■ = ekpressio □ = manifestaatio □ = kappale

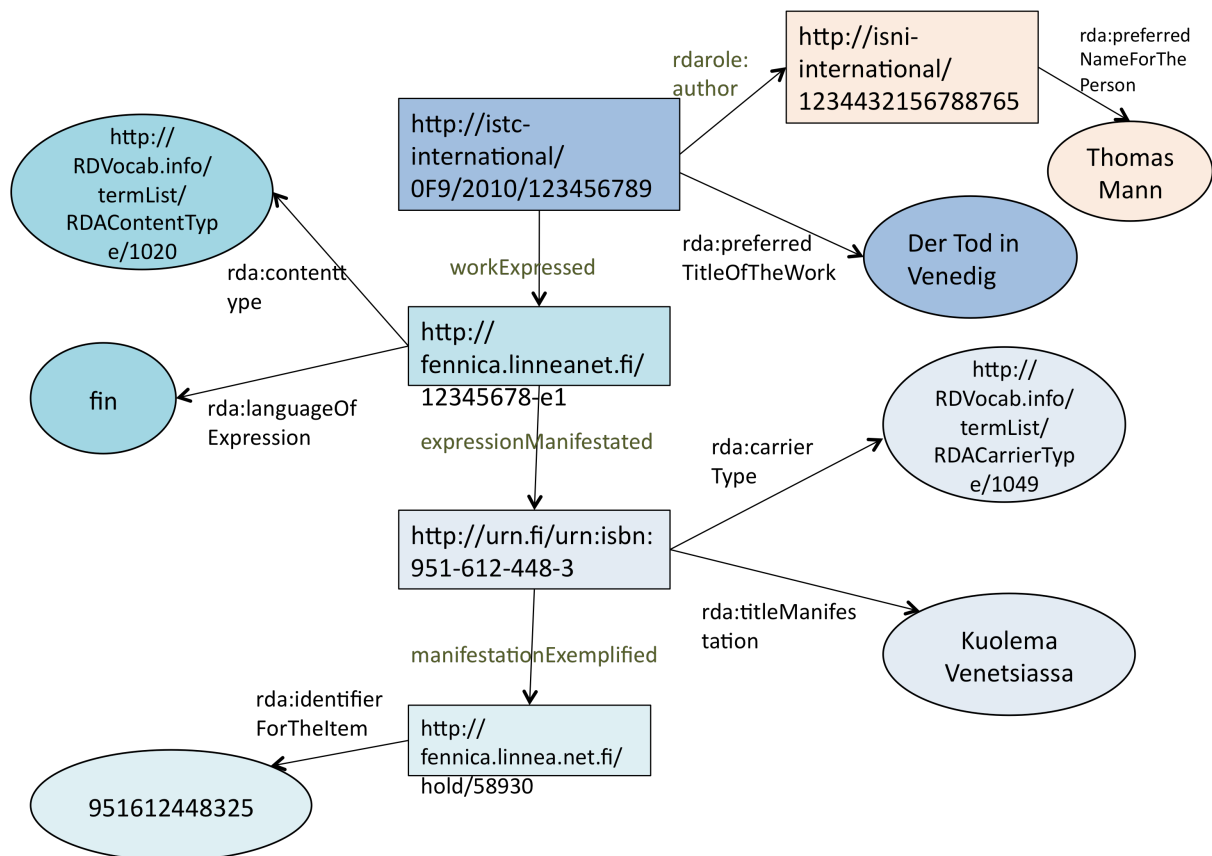
Entiteettiryhmä 2:

□ = henkilöt ■ = yhteisöt

Entiteettien väliset suhteet (vihreällä)

Metatieto avoimeen verkkoon

Funktionaaliselle tiedonmallinnukselle pohjautuva RDA muodostaa ontologiarakenteisen, merkityksellisten (semanttisten) suhteiden verkoston. Koska RDA ei määrittele, miten metatiedot esitetään, mallinnetaan ne semanttiseen webiin esim. DCMI- (Powell 2007), Indecs Metadata Framework- (Rust 2000) ja RDF-viitekehysten (Klyne 2004) avulla. RDA:n entiteetit, elementit ja arvot vastaavat RDF-mallin luokkia (class) ja ominaisuus/arvo-pareja (property/value-pair) (Kiorgaard 2008).



Kuva 3
Esimerkki RDF-mallinnuksesta
(Osa URI-tunnisteista kuvitteellisia)

Avoin linkitetty data (Berners-Lee 2009) rakentuu kielineutraaleille luokkien, ominaisuuksien ja arvojen URI-tunnisteille, jotka mahdollistavat tietojen automaattisen linkittymisen yli organisaatorajojen. Globaalit metadatarokisterit, kuten RDA-sanasto (NSDL 2010), määrittelevät arvojen ja ominaisuuksien termejä metatietojen laadun ja yhdenmukaisuuden parantamiseksi. Tietojärjestelmä, johon globaali metadatarokisteri on integroitu, auttaa tiedonhakijaa valitsemaan täsmällisempiä hakutermejä ja näin parantamaan hakutuloksia.

Kohti rikkaampaa metatietoa

Kirjastojen metatiedon avaaminen verkkoon korostaa luotettavan metatiedon keskeisyyttä. URI-tunnisteiden rekisteröityjen nimialueiden (namespace) avulla voidaan osoittaa metatiedon proveniensi. Entiteetti-relaatio-mallin (FRBR) eksplisiittisesti esitetyt suhteet antavat tiedonhakijalle jäsenneilyn hakutuloksen ja mahdollisuuden tiedon kontekstualisointiin (esim. teosten suhteet), klusterointiin (esim. teos/manifestaatiot-suhde) ja skaalattavuuteen (esim. kokonaisuus/osa-suhde). Tiedonhakijat voivat osallistua tiedon organisointiin luomalla uusia näkökulmia käsitteiden välille esim. assosiatiivisin suhtein. RDA:n pohjautuminen teoreettiseen tietomalliin tekee kuvailusta joustavan: jokainen tietomalliin linkitetty uusi entiteetti perii ympärillä olevien entiteettien ominaisuudet.

Uusi metatiedon viitekehys osallistaa kirjaston semanttisen verkon kudontaan.

Lähteet

Berners-Lee, Tim. Linked Data (27.7.2006. Päivitetty 18.6.2009). [katsottu 23.6.2010]
<http://www.w3.org/DesignIssues/LinkedData.html>

IFLA Study Group on the Functional Requirements for Bibliographic Records. Functional Requirements for Bibliographic Records (1997). [katsottu 23.6.2010]
http://www.ifla.org/files/cataloguing/frbr/frbr_2008.pdf

IFLA Working Group on Functional Requirements and Numbering of Authority Records (FRANAR). Functional Requirements for Authority Data. München: K.G. Saur, 2009.

IFLA Working Group on Functional Requirements for Subject Authority Records (FRSAR). Functional Requirements for Subject Authority Data (FRSAD) (2009). [Draft]. [katsottu 23.6.2010]
<http://nkos.slis.kent.edu/FRSAR/report090623.pdf>

Kiorgaard, Deirdre. RDA Element analysis (26.10.2008). [katsottu 23.6.2010]
<http://www.rda-jsc.org/docs/5rda-elementanalysisrev2.pdf>

Klyne, Graham and Jeremy Carroll, editors. Resource Description Framework: Concepts and Abstract Syntax. W3C Recommendation (10.2.2004). [katsottu 23.6.2010]
<http://www.w3.org/TR/2004/REC-rdf-concepts-20040210/>

NSDL Registry. [katsottu 23.6.2010]
<http://metadataregistry.org/rdabrowse.htm>

Powell, Andy, Mikael Nilsson, Ambjörn Naeve, Pete Johnston, and Thomas Baker. DCMI Abstract Model (4.6.2007). [katsottu 23.6.2010].
<http://dublincore.org/documents/abstract-model/>

RDA Toolkit Resource Description & Access. [Full Draft] (2008). [katsottu 23.6.2010].
<http://www.rdatoolkit.org/constituencyreview>

Rust, Godfrey and Mark Bide, The <indec> Metadata Framework: Principles, Model and Data Dictionary (2000). [katsottu 23.6.2010].
http://www.doi.org/topics/indec/indec_framework_2000.pdf