

Metatiedon rakennemuutos: avaamista, linkitystä ja rikastamista

Kuvailun opintopiiri 3.5.2011: RDA

Marja-Liisa Seppälä
Kansalliskirjasto

Avoin linkitetty data (Linked Open Data)

- Avoimuus = Kirjaston rikkaan ja luotettavan metatiedon avaaminen verkkoon koneluettavaan muotoon kirjaston vaikuttavuuden lisäämiseksi verkkotiedonhaussa
- Linkittäminen = Metatietojen linkittäminen yli kansallis- ja organisaatorajojen
- Globaalit/kansalliset metatietovarannot verkossa: kontrolloidut sanastot, tunnistejärjestelmät, metatietorekisterit, resoluutiopalvelut, tietomallit jne.

Vaikutus metatiedon tuotantoon

- Suuri muutos kirjastoissa: Suljetusta metatiedon tuotantoympäristöstä avoimempaan verkkoympäristöön
- Metatiedon viitekehyksen ja jäsentämisen muutos
- Metatiedon globaali yhteiskäyttöisyys:
 - metatietotuotannon tehostuminen ja tasalaatuistuminen
 - yhteistyön moninaistuminen muistiorganisaatioiden, julkaisualan ja tiedonhakijoiden kanssa

RDA ja linkitetty data

- RDF (Resource Description Framework) linkitetyn datan tietomallina
- Triple datan jäsenyyksen perusrakenteena:
 - RDAn entiteetti = RDFn luokka (class) = triplen subjekti : esim. teos
 - RDAn elementti = RDFn ominaisuus (property) = triplen predikaatti : esim. kirjoittaja
 - RDAn elementin arvo = RDFn arvo (value) = triplen objekti : esim. Thomas Mann
- Triple selväkielisenä: Teoksen kirjoittaja on Thomas Mann

Linkitetty data

- Datan konelukuisuus vaatii URI-tunnisteet (Uniform Resource Identifier) ja XML-skeeman

- Triple konelukuisena:

- Teos:

```
<rdf:Description rdf:about=" http://istc-international/0F9/2010/123456789 ">
```

```
<rdf:type rdf:resource=" http://RDVocab.info/uri/schema/FRBRentitiesRDA/Work "/>
```

- kirjoittaja ja Thomas Mann (ominaisuus-arvo-pari):

```
<rdarole:authorWork rdf:resource=" http://isni-international/1234432156788765 "/>
```

Metatiedon rakennemuutos

- Tietueiden muodossa oleva metatieto pilkotaan osiin elementeiksi, elementtien kootaan entiteettikohtaisiksi seteiksi ja entiteettien välille muodostetaan suhteet/linkit
- Pois nykyisestä tietuejaosta (bibliografinen, auktoriteetti- ja varastotietue) entiteettikohtaisiin kuvailuihin
- Ekonomisuus: useammalle entiteeteille yhteisten elementtien käytön maksimointi

RDF-skeema

```
<?xml version="1.0" encoding="UTF-8"?>
```

```
<rdf:RDF xmlns=" http://www.w3.org/1999/02/22-rdf-syntax-ns# "  
  xmlns:rdf=" http://www.w3.org/1999/02/22-rdf-syntax-ns# "  
  xmlns:rda=" http://RDVoc.info/Elements/ "  
  xmlns:rdagr2=" http://RDVoc.info/ElementsGr2/ "  
  xmlns:rdarole=" http://RDVoc.info/roles/ "  
  xmlns:rdaRelWEMI=" http://RDVocab.info/RDARelationshipsWEMI/ ">
```

XML-
nimialueet

```
<!-- Teos -->
```

```
+ <rdf:Description rdf:about=" http://istc-international/0F9/2010/123456789 ">
```

```
<!-- Teoksen ekspressio -->
```

```
+ <rdf:Description rdf:about=" http://istc-international/0F9/2011/341278956 ">
```

```
<!-- Ekspression manifestaatio -->
```

```
+ <rdf:Description rdf:about=" http://urn.fi/urn:isbn:951-612-448-3 ">
```

```
<!-- manifestaation kappale -->
```

```
+ <rdf:Description rdf:about=" http://fennica.linnea.net.fi/hold/58930 ">
```

```
<!-- Henkilö entiteettinä -->
```

```
+ <rdf:Description rdf:about=" http://isni-international/1234432156788765 ">
```

entiteetit

Kuvailutietoja RDFS-muodossa

<!--Teos-->

```
- <rdf:Description rdf:about=" http://istc-international/0F9/2010/123456789 ">  
  <rdf:type rdf:resource=" http://RDVocab.info/uri/schema/FRBReentitiesRDA/Work "/>  
  <rda:preferredTitleOfWork xml:lang="de">Der Tod in Venedig</  
  rda:preferredTitleOfWork>  
  <rdaRelWEMI:expressionOfWork rdf:resource=" http://istc-international/0F9/2011/341278956 "/>  
  <rda:authorWork rdf:resource=" http://isni-international/1234432156788765 "/>  
  <rda:identifierForTheWork>(ISTC)0F9 2010 12345678 9</rda:identifierForTheWork>  
</rdf:Description>
```

Thomas Mann

<!--Teoksen ekspressio-->

```
- <rdf:Description rdf:about=" http://istc-international/0F9/2011/341278956 "/>  
  <rdf:type rdf:resource=" http://RDVocab.info/uri/schema/FRBReentitiesRDA/Expression "/>  
  <rda:contentType rdf:resource=" http://RDVocab.info/termList/RDAContentType/1020 "/>  
  <rda:languageOfExpression>fin</rda:languageOfExpression>  
  <rdaRelWEMI:workExpressed rdf:resource=" http://istc-international/0F9/2010/123456789 ">  
  <rdaRelWEMI:expressionManifestated rdf:resource=" http://urn.fi/urn:isbn:951-612-448-3 ">  
</rdf:Description>
```

Tekstiä

<!--Ekspression manifestaatio--> ...

Henkilötiedot RDFS-muodossa

<!--Henkilö-->

```
- <rdf:Description rdf:about=" http://isni-international/1234432156788765 ">  
  <rdf:type rdf:resource=" http://RDVocab.info/uri/schema/FRBEntitiesRDA/Person "/>  
  <rdagr2:preferredNameForThePerson>Thomas Mann  
</rdagr2:preferredNameForThePerson>  
  <rdagr2:identifierForThePerson>(ISNI) 1234 4321 5678 8765  
</rdagr2:identifierForThePerson>  
  <rdagr2:identifierForThePerson>(VIAF) 54151065</rdagr2:identifierForThePerson>  
  <rdarole:authorWork rdf:resource=" http://istc-international/0F9/2010/123456789 ">  
</rdf:Description>
```

Tunnisteet ja linkitykset

- URI-muotoiset tunnisteet ovat keskeisiä entiteettien identifiointissa, entiteettien välisissä linkityksissä (suhteissa) ja entiteetteihin liittyvien kuvailuelementtien kokoajina
- Globaalien ja standardoitujen tunnisteiden (kuten ISTC, ISNI ja ISBN) käyttö URI-tunnisteissa parantavat metatietojen linkittymistä ja luotettavuutta
- Tunnisteet vähentävät perinteisten hakutietojen merkitystä bibliografisessa kuvailussa

Tunnisteet ja linkitykset (2)

- URN-tunnuksen käyttö manifestaation tunnisteiden yhteydessä parantaa itse julkaisun löytyvyyttä
- Kuvailuelementit ja niiden arvot ilmaistaan vakioidussa muodossa johonkin rekisteröityyn nimialueeseen liittyen
- Nimialue kertoo viitekehyksen (esim. RDA-standardi), joka määrittelee, miten ko. elementtiä käytetään kuvailussa ja mihin arvoon koodi liittyy

Globaalit metatietovarannot verkossa

- Metadatarokisterit sisältävät kuvailuelementtien määritelmät ja URIt, esim. Open metadata registry <http://metadatarokistry.org/rdabrowse.htm>
- Tunnistojärjestelmien tietokantojen metatiedot, esim. ISTC-, ISNI- ja ISSN-tietokannat
- VIAF (Virtual International Authority File) sisältää eri maiden tietokannoista haravoituja henkilöiden ja yhteisöjen nimiä: <http://viaf.org/>
- Ideaalissa tilanteessa kunkin maan avoimet metatietovarannot loisivat aukottoman toisiinsa linkittyvien metatietojen verkoston

Shared, re-used and valued metadata

- Muutama poiminto UKSG-konferenssista (4.4.-6.4.2011, Harrogate, UK)
- Linkitetty data vasta kehityksensä alussa
- Pyrittävä prosessin avoimuuteen läpi informaatioketjun (kustantajat/kirjastot/tieteentekijät)
- Avoimuus lisää innovointia
- Kirjastojen löydettävä uusia avoimuutta hyödyntäviä toimintamalleja (business models)
- Metadata on luuranko, johon kaikki web-palvelut (esim. graafiset kartta- ja aikajanapalvelut) kiinnittyvät
- Tiedonhakijat haluavat tietää metatiedon provenienssin ja luotettavuuden asteen

Linkitetyn ja avoimen metatiedon mahdollisuudet

- Linkitetyn metatiedon joustavuus
- Tiedonhakijoiden osallistuttaminen metatiedon rikastamiseen
- Hakutuloksen jäseneltävyys sekä metatiedon kontekstualisointi, klusterointi ja skaalattavuus
- Uusien ja yllättävien näkökulmien löytäminen metatietoja yhdistelemällä

