

Marianne Hacklin

**MARC-LUETTELOINTIFORMAATIN JA BIBFRAME-LUETTELOINTIMALLIN
VERTAILEVA ANALYYSI**

**MARC-LUETTELOINTIFORMAATIN JA BIBFRAME-LUETTELOINTIMALLIN
VERTAILEVA ANALYYSI**

Marianne Hacklin
Opinnäytetyö
Syksy 2014
Kirjasto- ja tietopalvelun
koulutusohjelma
Oulun ammattikorkeakoulu

TIIVISTELMÄ

Oulun ammattikorkeakoulu
Kirjasto- ja tietopalvelun koulutusohjelma

Tekijä: Marianne Hacklin

Opinnäytetyön nimi: MARC-luettelointiformaatin ja BIBFRAME-luettelointimallin vertaileva analyysi

Työn ohjaaja: Pirjo Tuomi

Työn valmistuslukukausi ja -vuosi: Syksy 2014

Sivumäärä: 56

Opinnäytetyön tarkoituksena on kartoittaa MARC-luettelointiformaatin ongelmia ja syitä siihen, miksi MARCista halutaan luopua ja sen tilalle kehittää uusi luettelointimalli, sekä selvittää millainen on kehitteillä oleva BIBFRAME-luettelointimalli ja miten se eroaa MARC-formaatista. Työn tilaajana toimii Kansalliskirjasto, joka seuraa BIBFRAMEn kehittämistä, muttei itse toistaiseksi ole osallistunut kehitystyöhön.

MARCin ongelmien kartoituksen lähdemateriaalina on käytetty kirjastoalan asiantuntijoiden kansainvälistä verkkokeskustelua. BIBFRAME-mallin selvityksen aineistona ovat Yhdysvaltain kongressin kirjaston julkaisemat raportit, selvitykset ja muut siitä jo julkaistut dokumentit. BIBFRAMEn ja MARCin vertailussa keskitytään niiden rakenteisiin ja skeemoihin. Näkökulma luettelointiin on kirjastolähtöinen ja muiden bibliografista luettelointia tekevien organisaatioiden tarpeita ei opinnäytetyössä erikseen tarkastella.

MARC-formaatin ongelmien kartoituksen perusteella formaatti koetaan kirjastoalalla varsin ongelmalliseksi ja puutteelliseksi kirjastoluetteloinnin nykyisiin tarpeisiin. Selvitys osoittaa, että BIBFRAME-malli on vielä keskeneräinen eikä toistaiseksi ole varmaa, mitkä luetteloinnin osa-alueet mallin on tarkoitus kattaa itsenäisesti ja kuinka paljon BIBFRAMEn käyttö vaatii muiden standardien tai formaattien tukea. Jos kuitenkin BIBFRAMEn kehittämisen tavoitteet toteutuvat, siitä voi tulla MARCia paremmin bibliografisen luetteloinnin tarpeisiin vastaava, monipuolisempi ja bibliografisesti syvällisempi luettelointimalli.

Keskeisimpänä MARCin ja BIBFRAMEn välisenä erona voidaan pitää sitä, että MARC on formaatti yksittäisten resurssien bibliografiseen kuvailuun kun BIBFRAME taas keskittyy kuvaamaan resurssien välisiä suhteita. MARC-formaatista luopuminen tulevaisuudessa vaikuttaa todennäköiseltä, jos BIBFRAME saadaan kehitettyä valmiiksi ja Yhdysvaltain kongressin kirjasto siirtyy sitä käyttämään. Tällöin MARCin päivittäminen mahdollisesti jopa lopetetaan ja muutkin luettelointiorganisaatiot siirtyvät MARCista joko BIBFRAMEn tai johonkin muuhun luettelointiformaattiin tai -malliin. BIBFRAMEn valmistumisen ja mahdollisen siirtymän aikataulua on kuitenkin vaikea vielä tässä vaiheessa arvioida.

Asiasanat: bibliografinen valvonta, luettelointi, formaatit, MARC, BIBFRAME

ABSTRACT

Oulu University of Applied Sciences
Degree Programme in Library and Information Services

Author: Marianne Hacklin

Title of thesis: Comparative analysis of MARC format and BIBFRAME model

Supervisor: Pirjo Tuomi

Term and year when the thesis was submitted: Autumn 2014

Number of pages: 56

The aim of the thesis is to look into the issues of the MARC format and clarify reasons why there is a need to give up MARC and to develop a new cataloging model, to describe the model under development aimed to replace MARC, and the differences between MARC and the new model, BIBFRAME. This thesis is commissioned by The National Library of Finland, which has been following the development of the BIBFRAME model, but has not taken an active role in the process.

The part of the thesis dealing with the problems of MARC format is based on the international online discussion of the library field. The study of the BIBFRAME model is based on reports and other documents published by The Library of Congress. The comparison of MARC and BIBFRAME focuses on their schemes and structures. The thesis point of view to cataloging is derived from the library field. The requirements of other cataloging organizations are not looked into.

The analysis of the problems of MARC indicates that the format is found widely inconvenient and inadequate considering the present requirements of library cataloging. The study of the BIBFRAME model shows that the model is still unfinished, and it is not certain which sections of cataloging the model will cover independently and how much support is required from other formats and standards. However, if the goals of the development of the new cataloging model are fulfilled, BIBFRAME can better respond to the needs of library cataloging. Furthermore BIBFRAME can be more diverse and bibliographically more profound and in-depth than the MARC format.

The key difference between MARC and BIBFRAME appears to be the fact that MARC is a format for describing single bibliographic resources and BIBFRAME focuses on describing the relationships between the resources. Abandoning the MARC format in the future seems probable if the BIBFRAME model will be completed and The Library of Congress will change over from MARC to BIBFRAME. In that case updating of MARC may cease, and other cataloging organizations will move from MARC to BIBFRAME or to some other cataloging format or model. The schedule of the completion of BIBFRAME and the possible transfer is, however, difficult to estimate at the present stage.

Keywords: bibliographic control, cataloging, formats, MARC, BIBFRAME

SISÄLLYS

1 JOHDANTO	6
2 FORMAATTEJA JA STANDARDEJA	8
2.1 Luettelointiformaatti	8
2.2 Metadata	8
2.3 Tiedonsiirtoformaatti	9
2.4 FR-tietomallit	10
2.5 RDA	12
2.6 RDF	13
3 MARC	14
3.1 MARCin rakenne	14
3.2 Formaatin ongelmat	16
3.3 MARCiin kohdistunutta kritiikkiä	17
3.3.1 Käytössä vain kirjastoissa	17
3.3.2 Historiallinen painolasti	18
3.3.3 Yhteentoimivuus	19
3.3.4 Kenttäongelmat	21
3.3.5 Tehottomuus	22
4 BIBFRAME	24
4.1 BIBFRAME-malli	27
4.2 Entiteetit	28
4.3 Suhteet ja ominaisuudet	31
4.4 Skeema ja sanasto	32
4.5 BIBFRAME-profiilit	36
4.6 Audiovisuaalisen aineisto kuvailu	37
5 VERTAILU	39
6 JOHTOPÄÄTÖKSET	45
7 POHDINTA	49
LÄHTEET	51

1 JOHDANTO

Kirjastoalalla luettelointityötä on jo vuosikymmeniä hallinnut MARC (Machine Readable Cataloging Record) -formaatti. MARC-formaattia on päivitetty vastaamaan luetteloinnin muuttuvia käytäntöjä ja tarpeita, mutta kirjastoluetteloiden käyttöympäristön painottuessa yhä enemmän internetiin, on MARC-formaatti havaittu yhä sopimattommaksi luettelointityöhön ja siihen kohdistunut kritiikki lisääntyneet. MARC-formaatin päivityksetkään eivät tunnu ratkaiseen kaikkia keskeisiksi koettuja ongelmia.

Opinnäytetyön tarkoituksena oli selvittää, miksi MARC-luettelointiformaatin käytöstä halutaan kirjastoalalla luopua, millainen on MARCin seuraajaksi kehitteillä oleva BIBFRAME (Bibliographic Framework) -malli sekä verrata MARCia ja BIBFRAMEa toisiinsa niiltä osin kuin BIBFRAME-mallista on tällä hetkellä tietoa saatavilla. MARCin ongelmia käsittelevässä osuudessa käydään läpi formaattiin kohdistunutta kritiikkiä. Aineistona on lähinnä kirjastoalan asiantuntijoiden blogikirjoituksia ja blogeissa käytyjä keskusteluja sekä joitakin ammattilehdissä julkaistuja artikkeleja. Ongelmien kartoituksessa on keskitytty kirjastoalalla käytyyn kansainväliseen keskusteluun, sillä varsinaisia tutkimuksia aiheesta on vähän ja tietämys MARCin ongelmista on pitkälti sen parissa työskentelevien asiantuntijoiden kokemuksen varassa. BIBFRAME-mallia koskeva aineisto on etupäässä mallin kehittämistä vastaavan Yhdysvaltain kongressin kirjaston julkaisemia keskustelupöytäkirjoja, analyseja ja raportteja, sekä niiltä osin kuin niitä on saatu valmiiksi, myös luonnoksia tarkemmista mallia koskevista kuvauksista, määrittelyistä ja ohjeista. Vertailu painottuu MARCin ja BIBFRAME:n rakenteisiin ja skeemoihin. Koska BIBFRAME on keskeneräinen, ei BIBFRAME- ja MARC-luetteloinnin vertaaminen käytännön luettelointityön kannalta ole vielä mielekäs. Opinnäytetyön näkökulma luettelointiin on kirjastolähtöinen eikä työssä juurikaan tarkastella MARCin ja BIBFRAME:n soveltuvuutta muiden luetteloivien kulttuuriorganisaatioiden tarpeisiin.

Opinnäytetyön toimeksiantaja on Kansalliskirjasto. Kansalliskirjasto on seurannut BIBFRAME-projektin etenemistä ja BIBFRAME-mallin kehittämistä, raportoinut niistä ja alustavasti esitellyt niitä suomeksi. Mallin kehittämiseen ei kuitenkaan ole Kansalliskirjaston toimesta osallistuttu eikä kattavaa suomenkielistä selvitystä BIBFRAMEsta ole tehty. Myöskään MARC-formaatista luopumisen syistä ei ole tehty suomenkielisiä selvityksiä, vaikka Suomessakin on kirjastoalalla MARCin ongelmista keskusteltu. Opinnäytetyön tavoitteena on siis myös olla suomenkielinen selvitys

BIBFRAMEsta ja osaltaan edesauttaa suomalaista kirjastoalan keskustelua bibliografisen luetteloinnin tulevaisuudesta.

Opinnäytetyössä on BIBFRAME-mallin suomenkielisten termien kohdalla käytetty Kansalliskirjaston alustavia suomennoksia niiltä osin kuin niitä on tehty. Muut termit on käännetty MARC-formaatin suomenkielisten termien mukaan, jos vastaava termi löytyy. Jos MARCissa ei vastaavaa termiä ole, on käytetty pääsääntöisesti BIBFRAMEn alkuperäistä englanninkielistä ilmausta. BIBFRAMEn englanninkieliset termit on joka tapauksessa mainittu suomennoksen yhteydessä.

MARCin keskeisimmät ongelmat pohjautuvat sen rakenteeseen eivätkä siis ole täysin ratkaistavissa formaattia päivittämällä. MARCin ongelmallisimmiksi puoliksi on koettu esimerkiksi sen soveltumattomuus internetiin toimintaympäristönä ja entiteetti-relaatiomallien ilmaisemiseen, sekä hankaluudet tiedon linkittämisessä ja MARCia tukevien sovellusten kehittämisessä. BIBFRAMEn on tarkoitus olla käsitteellinen ja käytännöllinen malli bibliografiseen kuvailuun ja kuvailutietojen tallentamiseen. Keskeisiä tavoitteita, joihin pyritään, ovat riippumattomuus luettelointisäännöistä, tiedon linkittyminen ja tunnisteiden käyttäminen. BIBFRAMEssa on tarkoitus myös tehdä selkeä ero luetteloitujen resurssien käsitteellisten sisältöjen ja niiden erilaisten manifestaatioiden välillä sekä tunnistaa ja nimetä informaatioentiteettejä.

MARCia ja BIBFRAMEa yhdistää niiden perimmäinen tarkoitus, joka on mahdollistaa vastaaminen luettelon käyttäjän tarpeisiin löytää, tunnistaa, valita ja saada käyttöön haluttuja aineistoja. Keskeisimpänä MARCin ja BIBFRAMEn välisenä erona voidaan pitää sitä, että MARC on formaatti yksittäisten resurssien bibliografiseen kuvailuun kun BIBFRAME taas keskittyy kuvaamaan resurssien välisiä suhteita. MARCissa tiedot luetteloidusta resurssista tallennetaan tietueena. BIBFRAMEssa tietoa ei esitetä yksittäisinä dokumentteina vaan luetteloitua resurssia koskevana hajautettuina lauseina.

BIBFRAME on toistaiseksi vielä hyvin keskeneräinen ja jopa periaatteellisia kysymyksiä on ratkaisematta. Toistaiseksi on vielä epäselvää esimerkiksi se, pyrkikö BIBFRAME kattamaan kaikenlaisen luetteloinnin kaikki tarpeet vai olemaan yksinkertainen ja joustava pohja, jota voi soveltaa erilaisten luettelointitapojen tarpeisiin. Epäselvää on myös se, kuinka laajasti BIBFRAME tulee itsenäisesti kattamaan hallinnollista metadataa ja kuinka paljon se tulee tarvitsemaan tukea erillisiltä metadastandardeilta.

2 FORMAAATTEJA JA STANDARDEJA

2.1 Luettelointiformaatti

Luettelointiformaatteja käytetään bibliografisten tietueiden luomiseen ja jakamiseen. Kirjassa *Bibliografinen valvonta* luettelointiformaatin sanotaan määrittävän ne muodot, joissa luetteloitavaa dokumenttia koskevat tiedot esitetään. Dokumenttia koskevat tiedot, eli dokumentin metadata, voidaan esittää luettelointiformaatin avulla koneluettavassa muodossa. (Suominen, Saarti & Tuomi 2009, 49.)

Luettelointityön kannalta käytetyllä luettelointiformaatilla on suuri merkitys, sillä se on yksi luettelointityön keskeisimmistä työvälineistä. Luetteloitavaa dokumenttia koskevien tietojen esitysmuodon lisäksi käytetty formaatti vaikuttaa osaltaan myös luettelon sisältöön. Luettelointiformaatti määrittää luetteloitavien tietojen välisiä suhteita ja asettaa rajoja sille, mitä tietoja luetteloitavasta dokumentista on mahdollista luettelon tallentaa. Käytetty formaatti vaikuttaa myös luettelon kohdistuvaan tiedonhakuun. Luettelon mahdollisesti tallennettavien tietojen muodon ja sisällön rajaamisen lisäksi formaatti voi vaikuttaa myös siihen, mihin tallennetuista tiedoista tiedonhaun voi kohdistaa. Hakua ei siis voi välttämättä kohdistaa kaikkeen luettelon tallennettuun tietoon ja osa tiedosta voi jäädä haun ulottumattomiin. Sisällönkuvailun kannalta luettelointiformaatti vaikuttaa keskeisesti esimerkiksi mahdollisuuteen ilmaista luetteloitujen resurssien, nimien ja dokumentin sisältöä kuvaavien asiasanojen välisiä suhteita ja muotoon, jossa nämä suhteet voidaan ilmaista. (Ks. Chowdhury 2010, 27–33.)

2.2 Metadata

Luetteloitavaa dokumenttia koskeva metadata on tietoa, joka kuvailee, paikallistaa tai muutoin edesauttaa dokumentin löytämistä tai käyttöä. Dokumenttia kuvailevan tiedon jakaminen yksittäisiin tietokokonaisuuksiin on tärkeää, sillä samaa tietoa voidaan käyttää useisiin eri tarkoituksiin. Metadataa voidaan hahmottaa useiden erilaisten jaottelujen avulla. Sen merkitystä voidaan kuvata esimerkiksi käyttötarkoituksen mukaisen jaon kautta. Metadataa tallennetaan viiteen käyttötarkoitukseen, jotka ovat: resurssin kuvailu, tiedonhaku, informaatioresurssien hallinnointi, digitaalisten resurssien omistajuuden ja autenttisuuden dokumentointi, ja järjestelmien välisen yhteentoimivuuden tukeminen (Rowley & Hartley 2008, 43–44).

Järjestelmien yhteentoimivuutta taas voidaan hahmottaa kolmella tasolla. Semanttinen yhteentoimivuus saavutetaan kuvailusääntöjen kautta, rakenteellinen yhteentoimivuus tiedonsiirtomallien ja syntaktinen yhteentoimivuus merkintäkielien avulla (sama, 47–48). Metadata voidaan myös jaotella sen sisällön perusteella kolmeen ryhmään, jotka ovat kuvaileva, rakenteellinen ja hallinnollinen metatieto. Kuvaileva metadata kuvaa kohteen bibliografisia ominaisuuksia. Kuvailevaa metadataa ovat esimerkiksi teoksen nimeke, tekijä ja asiasanat. Kuvailevan metadatan tarkoituksena on muun muassa edesauttaa dokumentin löytämistä ja tunnistamista. Rakenteellinen metadata määrittää kohteen osien väliset loogiset tai fyysiset suhteet. Hallinnollista metadataa ovat kyseessä olevan resurssin hallintaan liittyvät tiedot, esimerkiksi tiedostomuoto ja muu resurssin teknisiä ominaisuuksia kuvaava metadata, resurssin oikeuksien hallintaan liittyvät tiedot, esimerkiksi tiedot dokumentin teollisoikeuksista, ja tietojen säilyttämiseen tarvittavat tiedot. (National Information Standards Organization 2004, 1.)

Tiedonhaun kannalta luetteloidun metadatan yhteentoimivuus on tärkeää, ja yhteisiä periaatteita, standardeja ja sääntöjä noudattava luetteloidun aineiston kuvailu on yhteentoimivuuden edellytys (Autio-Tuuli & Seppälä 2012, viitattu 6.6.2014). Metadatakeemojen avulla standardisoidaan ja hallinnoidaan metadataa. Metadatakeema koostuu metadataelementeistä eli yksittäisistä kategorioista, joissa kuvataan resurssia koskevaa tietoa. Metadataelementtejä ovat esimerkiksi nimeke ja asiasanat. Useat luettelointiformaatit, esimerkiksi MARC, ovat sekä metadatakeemoja että tiedonsiirtoformaatteja.

2.3 Tiedonsiirtoformaatti

Luettelointiformaateista esimerkiksi MARC on sekä tiedontallennus- että tiedonsiirtoformaatti. Pelkkä tiedontallennusformaatti mahdollistaa vain tiedon tallentamisen ja datan siirtämisen saman järjestelmän sisällä. Bibliografisen datan siirtämiseen järjestelmien välillä tarvitaan lisäksi tiedonsiirtoformaatti. (Chowdhury 2010, 28.)

G. G. Chowdhury määrittää kolme ominaisuutta, joista bibliografinen tiedonsiirtoformaatti koostuu. Nämä ovat fyysinen rakenne (physical structure), sisällön merkitsijät (content designators) ja sisältö (content). Fyysisellä rakenteella Chowdhury tarkoittaa sääntöjä, joiden mukaan tallennettavan tiedon tallennusväline, eli formaatti, rakentuu. MARCissakin käytetyn ISO 2709-formaatin

käyttäminen tiedonsiirtoformaattien pohjana mahdollistaa eri tiedonsiirtoformaattien yhteentoimivuuden. Sisällön merkitsijät ovat tunnisteita, joilla tiedoston eri elementit ovat tunnistettavissa. Eri luettelointiformaateissa metadataskeemat ja niiden merkitsemiseen käytetyt tunnisteet voivat olla keskenään erilaisia jopa saman formaatin eri versioiden välillä. Usein nämä erot ovat kuitenkin pinnallisia ja automaattinen konvertointi formaattien välillä on tavallisesti mahdollista. Sisältö tarkoittaa tiedoston sisältöä, tietoa, jota formaatin avulla kuvataan ja hallinnoidaan. Luettelointiformaatin ominaisuuksien lisäksi sisältö on riippuvainen käytetyistä luettelointisäännöistä. (Sama, 28–29.)

2.4 FR-tietomallit

FR-malleista keskeisin on FRBR, The Functional Requirements for Bibliographic Records. FRBR on bibliografisen universumin käsitelmä, jonka avulla voidaan jäsentellä bibliografista tietoa. Se pyrkii täyttämään kirjastoluettelon käyttäjän tarpeet löytää, identifioida ja valita erilaisia bibliografisia entiteettejä ja päästä käyttämään niitä. FRBR on entiteetti-relaatio-malli, eli se määrittää tietyt entiteetit, joiden välisiä suhteita sen avulla voidaan jäsentää. Sen tarkoituksena on erityisesti myös, kuten Tillett (2005, 2) *Mikä on FRBR?* -raportissa toteaa, ”olla yleinen, luettelointisäännöistä sekä käytännön toteutuksista riippumaton”.

FRBR-käsitelmä perustuu IFLAn (International Federation of Library Associations and Institutions) asettaman työryhmän selvitykseen kirjastojen tuottaman bibliografisen tiedon käytettävyydestä. FRBR kuvaa sitä, minkälaisesta metadatasta käyttäjät tavallisesti ovat kiinnostuneita aineistoja etsiessään, ja mitä aineistoja kuvailevia tietoja käyttäjät tarvitsevat voidakseen löytää, tunnistaa, valita ja saada käyttöönsä tarvittua aineiston. Selvityksen perusteella käyttäjät ovat kiinnostuneita tieteellisestä tai taiteellisesta sisällöstä eli teoksesta, siitä miten sisältö on esitetty eli ekspressiosta, siitä missä muodossa sisällön voi saada käyttöönsä eli manifestaatiosta ja yksittäisestä kappaleesta. Käyttäjien kiinnostus kohdistuu myös aineistoihin liittyviin henkilöihin ja yhteisöihin eli muun muassa tekijöihin, esittäjiin, julkaisijoihin ja kappaleiden omistajiin. Lisäksi käyttäjät ovat kiinnostuneita teoksen aiheista. FRBR:ssä näitä käyttäjiä kiinnostavia asioita kutsutaan entiteeteiksi. (Autio-Tuuli & Seppälä 2012, viitattu 6.6.2014.)

Mikä on FRBR? -raportissa FRBR:n terminologia jakautuu kolmeen entiteettien ryhmään. Ryhmän 1 entiteettejä ovat teos, ekspressio, manifestaatio ja kappale. Ryhmän 2 entiteetit, henkilö ja

yhteisö, ovat suhteessa ryhmän 1 entiteetteihin. Ryhmien 1 ja 2 entiteettien väliset suhteet ilmoittavat henkilön tai yhteisön roolin tai tehtävän, joka koskee jotakin ryhmän 1 entiteettiä. Esimerkiksi tietty henkilö on teoksen tekijä tai yhteisö manifestaation julkaisija. Ryhmän 3 entiteetit ovat teosten aiheita, ne voivat olla edellisten ryhmien entiteettien lisäksi käsitteitä, kohteita, tapahtumia tai paikkoja. (Tillett 2005, 3.)

Ryhmän 1 entiteettien sisäisten ja kaikkien ryhmien entiteettien välisten suhteiden lisäksi FRBR:n mukaisesti voidaan ajatella muitakin bibliografisia suhteita. Teosperheellä eli teoksella, sen ekspressioilla, manifestaatioilla, siitä johdetuilla uusilla teoksilla ja sitä kuvailevilla teoksilla on myös sisältösuhteita, jotka FRBR-mallin avulla on mahdollista hahmottaa toisiinsa liittyvänä kokonaisuutena. FRBR auttaa erottamaan tästä kokonaisuudesta käyttäjän kannalta olennaiset suhteet. FRBR:n mukaan voidaan määrittää myös bibliografiselta kannalta olennaisia kokonaisuus-osa ja osa-osa -suhteita. Kokonaisuus-osa-suhde vallitsee esimerkiksi kokonaisen verkkosivuston ja sen komponenttien välillä. Osa-osa-suhteet taas ovat kokonaisuuden eri osien suhteita toisiinsa. (Tillett 2005, 4–5.) Artikkelissaan *Laadukas metatieto kirjastojen kansallisen metatietovarannon pohjana* Marjatta Autio-Tuuli ja Marja-Liisa Seppälä antavat esimerkeiksi entiteettien välisistä eritasoisista FRBR:n määrittelemistä suhteista elokuvasta tehdyn näytelmän, tietokirjan uuden version ja laulajan, joka on myös yhtyeen jäsen. Näiden suhteiden ilmaiseminen näkyvästi ja hierarkkisesti luo käyttäjille tiedonhakuja tukevan navigointiympäristön. (Autio-Tuuli & Seppälä 2012, viitattu 6.6.2014.)

FRAD (Functional Requirements for Authority Data) ja FRSAD (Functional Requirements for Subject Authority Data) pohjautuvat FRBR:n kanssa samoihin periaatteisiin ja määrittävät bibliografisen kuvailun kannalta varsinaiseen bibliografiseen entiteettiin liittyviä muita entiteettejä. FRAD on entiteetti-relaatio-malli nimiä koskevien auktoriteettitietojen ja niiden suhteiden esittämiseen. FRAD määrittää siis tarkemmin ryhmän 2 entiteetit eli henkilön, perheen tai suvun ja yhteisön, sekä niihin liittyviä suhteita. FRAD:n mukaan bibliografinen entiteetti yhdistetään esimerkiksi henkilön nimeen, joka toimii perustana auktorisoidulle hakutiedolle. FRSAD on entiteetti-relaatio-malli, joka käsittelee teosten aiheita eli ryhmän 3 entiteettejä, jotka ovat käsitteitä, tapahtumia, paikkoja tai materiaalisia objekteja. Se määrittää teoksen ja esimerkiksi siihen liittyvien asiasanojen suhteita. (IFLA Working Group on Functional Requirements and Numbering of Authority Records (FRANAR) 2008, 4, 7; Zeng, Žumer & Salaba 2010, 15.)

2.5 RDA

Marjatta Autio-Tuuli esittelee RDA-luettelointistandardin *Tietolinja*-lehden artikkelissa *RDA – uusi luettelointistandardi*. RDA (Resource Description and Access) on luettelointistandardi, joka pohjautuu angloamerikkalaisiin luettelointisääntöihin (AARC2). RDA on kehitetty laajapohjaisessa yhteistyössä myös kirjastoalan ulkopuolisten tahojen kanssa. Uuden standardin kehittämisen tavoitteena oli saada yleisiin periaatteisiin pohjautuvat säännöt, joita voi soveltaa kaikenlaisien aineistojen, myös uusien vielä ennakoimattomissa olevien julkaisumuotojen, kuvailuun. Vanhojen sääntöjen ongelmana koettiin soveltumattomuus etenkin digitaalisten aineistojen kuvailuun. (Autio-Tuuli 2009, viitattu 3.6.2014.)

RDA:n lähtökohtana on antaa sovellettavissa olevia ohjeita yksityiskohtaisten sääntöjen sijasta. Sen perustana ovat FR-tietomallit. RDA:n tavoitteena on vastata käyttäjän tarpeisiin löytää, tunnistaa, valita ja saada käyttöön tietoa ja aineistoja, sekä tarjota käyttäjälle mahdollisuus navigoida tietokannoissa entiteettien välille luotujen suhteiden avulla. Keskeistä on tuoda luettelosta esiin sisältöjä käyttäjää varten. RDA on siis sisältöstandardi, joka antaa ohjeita siitä, mitä kuvailutietoja käyttäjä todennäköisimmin tarvitsee. (Sama, viitattu 1.9.2014.)

RDA määrittelee mitä teosten, ekspressioiden, manifestaatioiden, kappaleiden, henkilöiden ja yhteisöjen välisiä suhteita kuvataan, mutta se ei aseta sääntöjä sille, miten tiedot esitetään tai tallennetaan. RDA:ssa metatiedot ovat itsenäisiä tietoelementtejä, jotka jäsentyvät dataseteiksi eli kuvailukokonaisuuksiksi (Autio-Tuuli & Seppälä 2012, viitattu 6.6.2014). RDA:n ominaisuuksia Autio-Tuuli kuvaa seuraavasti:

Elementit tallennetaan siten, että hakutulokset ryhmittyvät mielekkäällä tavalla. Entiteettien välisten suhteiden luonnetta on selkeytetty. Hakukoneet voivat ryhmitellä hakutuloksia näiden mukaisesti ja esittää ne näytöissä niin, että käyttäjä voi helposti valita ja edelleen rajata hakutulospöytäkirjasta aineistot, joista on kiinnostunut. (Autio-Tuuli 2009, viitattu 3.6.2014.)

RDA:n sovellettavuus eri metadataformaatteihin mahdollistaa sen käyttämisen muuallakin kuin kirjastoissa ja entiteetti-relaatio-mallin vuoksi se sopii verkkoympäristöön ja linkitetyn datan hyödyntämiseen.

2.6 RDF

RDF (Resource Description Framework) on Wide Web Consortiumin (W3C) kehittämä standardi tiedon siirtoon verkossa. Se ei ole riippuvainen mistään tietystä formaatista ja se soveltuu käytettäväksi erilaisilla merkintäkielillä, esimerkiksi XML (Extensible Markup Language) -kielellä. RDF on kehitetty erityisesti tiedon linkittämiseen URI (Uniform Resource Identifier) -tunnisteiden avulla. RDF:n avulla ilmaistaan jäseneltyä tietoa ja asioiden välisiä suhteita koneluettavassa muodossa. RDF:n mukaan mikä tahansa tietoa sisältävä lause jaetaan kieliopin mukaisiin lauseenjäseniin: subjekti, predikaatti ja objekti. Subjekti nimeää kuvatun resurssin, objekti on ominaisuus tai asia, johon subjekti on jonkinlaisessa suhteessa ja predikaatti määrittää subjektin ja objektin välisen suhteen. Kolmikon subjektille, predikaatille ja objektille annetaan omat tunnisteet. Tunnisteita käyttämällä subjekti-predikaatti-objekti-kolmikot muodostavat verkon, jossa tiedot kohteista linkittyvät. (Tauberer 2008, viitattu 12.6.2014.) Marjatta Autio-Tuuli ja Marja-Liisa Seppälä painottavat luettelointiformaatin RDF:n kanssa yhteentoimivuuden tärkeyttä artikkelissaan *Laadukas metatieto kirjastojen kansallisen metatietovarannon pohjana* kirjoittamalla:

Yhteentoimivuus RDF:n kanssa on oleellista, sillä RDF semanttisen webin kielioppina ja globaalina metatiedon alustana mahdollistaa metatiedon avaamisen verkkoon muodossa, jota eri sovellukset pystyvät hyödyntämään ja jonka myötä kirjaston metatieto on haettavissa verkon hakupalveluissa. (Autio-Tuuli & Seppälä 2012, viitattu 12.6.2014.)

3 MARC

MARC (MACHine-Readable Cataloguing) on luettelointi-, tiedontallennus- ja tiedonsiirtoformaatti. Suomessa MARCin versioista käytössä ovat FINMARC ja MARC 21. FINMARC on kehitetty vastaamaan Suomalaisia luettelointisääntöjä ja se on vielä käytössä suuressa osassa yleisiä kirjastoja. FINMARCin ylläpito ja kehittäminen ovat päättyneet vuonna 2000 ja tarkoituksena on vuoden 2014 loppuun mennessä saada FINMARC-tietokannat muutettua MARC 21-formaattiin. (Kansalliskirjasto 2014a, viitattu 3.6.2014; Kansalliskirjasto 2014b, viitattu 3.6.2014.)

MARC-formaatti kehitettiin Yhdysvaltain kongressin kirjaston toimesta 1960-luvulla korttiluetteloiden pohjalta. Tarkoituksena oli kehittää fyysisen kortiston koneluettava vastine. Alkuvaiheessa tärkeää oli formaatin koneluettavuuden ja sähköisen tiedonsiirron lisäksi myös tietueiden tulostettavuus fyysiseksi kortistoksi. Korttiluettelopohjaisuuden vuoksi MARCissa on yhä kortistoihin ja niiden logiikkaan liittyviä jäänteitä. Toisaalta myös elektronisissa tietokannoissa on korttiluetteloille ominaisilla funktioilla merkitystä etenkin selailuhakuun liittyen. (Suominen ym. 2009, 49.)

3.1 MARCin rakenne

MARC 21 koostuu kolmesta perusosasta: tiedonsiirtoformaatin rakenteesta, tietokenttien joukosta ja datasta, joka jäsentyy sisältöstandardien mukaisesti. Formaatin rakenne on suunniteltu tiedon siirtämiseen järjestelmien välillä, jotta samaa dataa voitaisiin käyttää eri ympäristöissä. Formaatin rakenne on NISON (Z39.2) sekä ISON (2709) standardiksi hyväksymä. MARC 21 käyttää ISO 2709 -standardin mukaista rakennetta tiedonsiirtoon. (Library of Congress 2011, viitattu 26.6.2014.)

MARC-tietueet noudattavat rakenteeltaan vaihtomuotostandardia ISO 2709. ISO 2709 on kansainvälinen standardi, joka määrittää elektronisten bibliografisten luetteloiden yleisrakenteen. MARC-tietue rakentuu nimiöstä, hakemistosta sekä kiinteä- ja vaihtuvamittaisista tietokentistä. Nimiö on kiinteä kenttä, joka käsittää tietueen 24 ensimmäistä merkkipaikkaa. Se sisältää tietoa kyseessä olevasta tietueesta, muun muassa sen pituudesta ja tilasta eli siitä onko tietue uusi vai muokattu, ja järjestelmä tuottaa sen tietueen alkuun. Hakemisto on myös järjestelmän tuottama osa tietuetta. Se on hakemisto vaihtuvamittaisten kenttien sijainnista ja se muodostuu kahden-toista merkin pituisista jaksoista, joissa kuvataan kenttäkoodi, kentän pituus ja kentän aloituskoh-

ta. Tietueen jokaisella tietokentällä on kolminumeroinen kenttäkoodi. Kiinteämittaisten kenttien kenttäkoodi on muotoa 00X eikä kiinteämittaisissa kentissä ole indikaattoreita tai osakenttäkoodeja ja ne sisältävät esimerkiksi kontrollinumeroita ja tietueiden käsittelyssä käytettävää kooditietoa. Vaihtuvamittaisten kenttien koodit vaihtelevat välillä 01X–8XX ja niissä käytetään sekä indikaattoreita että osakenttäkoodeja. Vaihtuvamittaiset kentät sisältävät pääasiallisen bibliografisen tiedon, kuten pääkirjauksen ja lisäkirjaukset sekä lisäksi luokituksen ja asiasanat. (Kansalliskirjasto 2008a, viitattu 3.6.2014.)

MARCIa on päivitetty jatkuvasti vastaamaan ajan luettelointitarpeita. Sitä on muokattu uudenlaisien materiaalien, esimerkiksi äänitteiden ja tietokoneohjelmien kuvailua varten ja käytettäväksi erilaisten luettelointisääntöjen kanssa. Myös elektronisten aineistojen kuvailu ja niiden linkittäminen tietueeseen on tehty mahdolliseksi. MARCIin sisältyy kuvailevan metadatan lisäksi myös muunlaista metadataa, joka mahdollistaa MARCin käytön luetteloinnin lisäksi muissakin kirjastojen toiminnoissa kuten hankinnassa ja aineiston kierron hallinnoinnissa. (Library of Congress 2011, viitattu 26.6.2014.)

MARC 21 -yhtenäisformaateista kirjastot käyttävät tavallisesti bibliografisten tietojen formaattia, auktoriteettitietojen formaattia ja varastotietojen formaattia. Bibliografisten tietojen formaattia käytetään aineistojen kuvailuun. Auktoriteettitietojen formaattia käytetään MARC-tietueissa käytettävien nimien ja asiasanojen ohjeellisten muotojen, niiden lisämääremuotojen ja muotojen välisiä suhteita kuvaavien tietojen tallentamiseen. Nimi viittaa auktoriteettitietojen formaatissa henkilön-, yhteisön-, kokousten tai hallintoalueiden nimiin, yhtenäistettyihin nimekkeisiin tai tekijä/nimekeyhdistelmiin. Asiasana viittaa aikaa tai aihetta ilmaiseviin termeihin, maantieteellisiin nimiin, nimiin, joihin liittyy asialisämääreitä, tai termeihin ja nimiin, joita käytetään asialisämääreinä. Varastotietojen formaattia käytetään varastotietojen merkitsemiseen yksi- tai moniosaisille monografioille ja kausijulkaisuille. Varastotiedot voivat sisältää kappalekohtaisia tietoja luetteloidusta kohteesta, omistavaan organisaatioon liittyviä ominaistietoja, kohteen paikalliseen käsittelyyn, ylläpitoon tai säilytykseen liittyviä tietoja tai versiotietoja. (Kansalliskirjasto 2007a; 2007b; 2008b, viitattu 8.7.2014.)

3.2 Formaatin ongelmat

MARC on kirjastoalalla kansainvälisesti merkittävin ja tunnetuin luettelointiformaatti. Formaattilla on pitkä historia, siitä on tehty useita versioita ja sitä on jatkuvasti päivitetty vastaamaan luetteloinnin uusia haasteita ja monimuotoistuvan aineiston luetteloinnin vaatimuksia. 2000-luvun alusta lähtien MARC-formaattiin on kuitenkin kohdistunut yhä enenevässä määrin kritiikkiä ja useat luetteloinnin asiantuntijat katsovat sen olevan liian jäykkä etenkin soveltuakseen verkko- ja muun elektronisen aineiston yksityiskohtaiseen ja käyttäjälle mahdollisimman hyödylliseen luettelointiin. MARCiin kohdistuu myös paljon kritiikkiä liittyen sen käytettävyyteen verkkoympäristössä. Yhtenä MARCin ongelmista voidaan pitää sitä, että vain osa MARC-tietueisiin kirjatusta tiedosta on hakutietoa. Tiedonhaku voi siis kohdistua vain tiettyihin MARC-kenttiin, vaikka haun ulkopuolelle jäävissäkin kentissä voi olla tiedonhaun kannalta relevanttia tietoa. Suurena ongelmana koetaan lisäksi se, ettei MARC toteuta FRBR-mallia. MARC-tietueita on kuitenkin mahdollista käyttää kirjastojärjestelmän ja sen lisäosien puitteissa järjestää ilmaisemaan joitakin FRBR:n mukaisia suhteita. Osaan MARC-kritiikistä onkin vastattu painottamalla, etteivät ongelmat olisi MARCissa sinänsä vaan ne olisivat pikemmin käytetyn kirjastojärjestelmän puutteita (ks. Coyle 2000, viitattu 2.9.2014).

Luettelointiformaatin tarkoituksenmukaisuutta ja käytettävyyttä voidaan tarkastella useista erilaisista näkökulmista. Formaatin soveltuvuutta tiedonhaakuun voidaan arvioida esimerkiksi tietyn, tunnetun aineiston tai entuudestaan hakijalle tuntemattoman halutunlaisen aineiston löytämisen kannalta. Lisäksi voidaan arvioida sen soveltuvuutta ilmaisemaan aineiston sisältöä, jotta aineiston voi tunnistaa halutunlaiseksi ja myös formaatin tarjoamia mahdollisuuksia rajata hakutulosta tai esittää löydetty aineisto eri näkökulmista. Luettelointiformaatin pitäisi mahdollistaa vastaamisen käyttäjän tarpeisiin löytää, tunnistaa, valita ja saada käyttöön aineistoja. Käyttäjän tarpeisiin vastaamisen lisäksi voidaan arvioida formaatin soveltuvuutta luettelointitietueiden luomisen ja niiden säilyttämisen kannalta. Formaattia voidaan arvioida koneluettavuuden ja yhteentoimivuuden kannalta, mutta sen luettavuutta voidaan arvioida myös ihmiskäyttäjän näkökulmasta. (Ks. Chowdhury 2010, 27–29.)

Vielä 2000-luvun alussa MARCiin kohdistuneen kritiikin osoittamien ongelmien nähtiin usein olevan ratkaistavissa soveltamalla, formaattia päivittämällä ja muuta toimintaympäristöä, kuten kirjastojärjestelmiä ja luettelointisääntöjä kehittämällä. 2000-luvun ensimmäisen vuosikymmenen

loppua kohden tultaessa luetteloinnin, etenkin MARCiin liittyvän ohjelmoinnin ja systeemin suunnittelun parissa työskentelevät asiantuntijat yhä laajemmin tulivat johtopäätökseen, että hyvinkin yksityiskohtaisten ja keskenään erilaisten ongelmien korjaaminen ei pidemmällä aikavälillä olisi tyydyttävä ratkaisu. Tähän perimmäisenä syynä on se, ettei MARC-formaatin perusrakennetta ei kuitenkaan saataisi vastaamaan vallalla olevaa bibliografisen ajattelun teoriaa, FR-tietomalleja. Mielekkäämmäksi koettiin kokonaan uuden, FRBR-malliin todella pohjautuvan ja RDA-standardia tukevan formaatin tai mahdollisesti laajemman luettelointityön välineenä toimivan ja joustavasti sovellettavissa olevan bibliografisen datamallin kehittäminen.

3.3 MARCiin kohdistunutta kritiikkiä

3.3.1 Käytössä vain kirjastoissa

Vuonna 2006 Yhdysvaltain kongressin kirjasto perusti työryhmän tarkastelemaan bibliografisen valvonnan tulevaisuutta. Työryhmän tarkoituksena oli selvittää kuinka bibliografinen valvonta ja luettelointikäytännöt voisivat mahdollisimman tehokkaasti tukea kirjastomateriaalien hyödyntämistä kehittyvässä tieto- ja teknologiaympäristössä sekä antaa suosituksia toimenpiteistä, joihin tavoitteen saavuttamiseksi olisi ryhdyttävä. Pohtiessaan teknologiaa ja internetiä toimintaympäristönä työryhmä kiinnitti erityisesti huomiota MARC-formaattiin. Ongelmalliseksi koettiin etenkin se, että MARC on käytössä ainoastaan kirjastoissa eivätkä kirjastoalan ulkopuolella kehitetyt ja käytetyt formaatit välttämättä pysty hyödyntämään MARC-muotoista tietoa. MARC-formaatissa pysyminen voi tulevaisuudessa johtaa siihen, ettei kirjastojen tuottamaa bibliografista dataa hyödynnettäisi kirjastoalan ulkopuolella. (The Library of Congress Working Group on the Future of Bibliographic Control 2008, 1, 24–26.)

Työryhmän mukaan uudet ja toivotut tavat bibliografisen tiedon käyttämiseen ja tuottamiseen vaativat formaatin, jonka avulla metadatan voi tuottaa niin asiantuntijat, koneet kuin käyttäjätkin. Formaatin tulisi olla myös joustava ja sovellettavissa erilaisten sitä käyttävien yhteisöjen tarpeisiin. Työryhmä huomioi, että tekstimuotoisten tunnisteiden käyttö sitoo formaatin käyttäjän tiettyyn kieleen ja vaikeuttaa sekä auktoriteetti- että bibliografisten tietueiden kansainvälistä jakamista. Uudessa formaatissa tunnisteiden olisi siis tärkeää olla kielestä riippumattomia. Työryhmä myös selkeästi toteaa, ettei MARCiin perustuminen ole hyvä lähtökohta uuden bibliografisen tiedonsiirtomenetelmän kehittämiseksi. Työryhmä suositteli uuden joustavamman ja laajennettavissa ole-

van metadatansiirtomenetelmän kehittämistä, kirjastostandardien integroimista verkkoympäristöön ja standardoitujen tunnisteiden laajempialaista käyttämistä. Tuloksena toivotaan olevan kirjastojen tuottaman bibliografisen datan siirtyminen eristyneistä järjestelmistä verkkoperustaiseen malliin, jossa kirjastojen tuottama tieto olisi myös koneiden käytettävissä ja kirjastot itse voisivat hyödyntää muita verkossa olevia tietoresursseja. (Sama, 24–26.) MARCin rakenteen ongelmallisuuden tiedon automaattisen prosessoinnin ja käyttämisen kannalta kiinnitetään huomiota myös Online Computer Library Centerin (OCLC) raportissa *Implications of MARC Tag Usage on Library Metadata Practices* (Smith-Yoshimura 2010, 27).

3.3.2 Historiallinen painolasti

Roy Tennant käsittelee artikkelissa *A Bibliographic Metadata Infrastructure for the 21st Century* MARCin ongelmia ja kirjastojen metadainfrastruktuurin tulevaisuutta. Kun MARC kehitettiin, tietokoneet olivat huomattavasti tehottomampia kuin nykyään. Tallennustilaa oli vähän ja tietokoneet vain harvojen käytössä. (Tennant 2004, 1.) Nykyään MARC-tietueen tallennustilan rajoittuminen 99 999:än ja kunkin kentän rajoittuminen 9999:än tavuun aiheuttaa ongelmia kun tietueeseen haluttaisiin lisätä sisältöä kuten tiivistelmiä ja arvosteluja (Karen 2010, viitattu 4.6.2014).

Artikkelissaan Tennant huomauttaa myös, että nykyään kirjastoissa käytetään kirjastoluetteloita useisiin eri tarkoituksiin, joihin MARCia ei ole alun perin suunniteltu, esimerkiksi kaukolainaamiseen tai verkkomateriaalin hakemistona. Lisäksi luettelointikäytännöt ovat muodostuneet fyysisten kappaleiden luetteloinnin pohjalta, minkä seurauksena luodaan useita tietueita, joissa toistetaan samaa tietoa. Koska MARCin mukaan luetteloidaan vain manifestaatioita, joudutaan teoksen kuvailu toistamaan jokaisen luetteloidun manifestaation yhteydessä. Tiedon toistaminen eri aikoina ja eri luettelointijärjestelmien toimesta voi johtaa keskenään hyvinkin erilaisiin kuvauksiin saman teoksen sisällöstä. Käyttäjän kannalta tavallisesti kuitenkin aineiston sisältö on olennainen, eikä esimerkiksi sillä ole väliä, mikä kirjan painoksista on kyseessä. Keskeisenä ongelmana Tennantin mukaan on se, ettei MARC rakennu FRBR:n varaan. (Tennant 2004, 2.)

Karen Coyle kirjoittaa blogissaan otsikolla *MARC vs RDA* MARCin yhteensopimattomuudesta RDA-standardin kanssa. MARCin metadataskeema ei tue RDA:ta, sillä toisin kuin MARC, RDA toteuttaa FRBR-mallia sekä mahdollistaa tunnisteiden käytön entiteettien ja termien ilmaisemiseen. Suurinpana hankaluutena MARCissa Coyle pitää FRBR:n mukaisten entiteettien välisten

suhteiden ilmaisemista järjestelmän sisällä. Coylen johtopäätös on, että olisi parempi keskittyä kehittämään kokonaan uutta FRBR-malliin ja RDA-standardiin pohjautuvaa formaattia kuin koettaa soveltaa MARCia RDA:n kanssa yhteensopivaksi. (Coyle 2011, viitattu 5.6.2014.) MARCia on kuitenkin päivitetty RDA:n mukaiseksi. Siihen on esimerkiksi lisätty uusia kenttiä, ja vuonna 2012 Yhdysvaltain kongressin kirjasto vahvisti ohjeet RDA:n käytöstä MARC 21:ssä (McCallum 2012, viitattu 5.6.2014).

Se, että MARCin ylläpidossa on huomioitu vanhemmat tietueet, jotka on luotu nykyisten luettelointikäytäntöjen mukaan puutteellisesti tai jopa nykyisten sääntöjen kanssa ristiriidassa olevien, päällekkäisyyttä aiheuttavien sääntöjen mukaan, aiheuttaa ongelmia yritettäessä yksinkertaistaa formaatin kenttien ja kenttäkoodien rakennetta. Myös uudet tavat datan järjestämiseen ovat MARCissa hankalia toteuttaa, sillä uudella tavalla järjestetty data olisi yleensä yhteensopimatonta vanhan datan kanssa. Lisäksi yksinkertaisemmat tiedonsiirtomallit ovat usein paremmin koneiden ymmärrettävissä ja ohjelmistokehityksen myötä myös monipuolisemmin koneiden hyödynnettävissä. MARCin kaltaiselle yksityiskohtaiselle datan jaottelulle koodien avulla ei ole enää yhtä suurta tarvetta. Toisaalta taas nykyään on kasvavassa määrin tarvetta datan yksityiskohtaiselle jaottelulle sellaisella tavalla, joka edesauttaisi sen hyödyntämistä linkitetyn tiedon periaatteiden mukaisesti. Siihen MARCin indikaattoreista, kentistä ja osakentistä muodostuvaa semanttisesti monimutkaista rakennetta ei pidetä sopivana. (Library of Congress 2011, viitattu 26.6.2014.)

3.3.3 Yhteentoimivuus

Kirjoituksessaan *Is MARC Dead?* Karen Coyle jaottelee MARC-formaattiin kohdistuneen kritiikin MARCin ongelmiin tietokantaformaattina, sen joustamattomuuteen, MARCin näyttöön kohdistuvaan kritiikkiin ja huomioihin MARCin liiasta monimutkaisuudesta sekä yksinkertaisuudesta. Yhtenä perustavanlaatuisimmista ongelmista Coyle näkee sen, että MARC on lähinnä ainoastaan kirjastojen käytössä. Coyle ei kuitenkaan pidä mitään MARCin yksittäistä vikaa ylitsepääsemättömänä vaan painottaa kirjastoluetteloiden sisällön olevan kunnossa ja näkee ratkaisuna konvertoinnin ja formaattien yhteentoimivuuden kehittämisen. Coylen mukaan ongelmat ovat pikemmin käytetyissä kirjastojärjestelmissä kuin luettelointiformaatissa. Hän hahmottelee kirjastojärjestelmän voivan olla se väline, jonka avulla muualla verkossa olevaa haluttua tietoa voitaisiin kirjaston oman luettelon ohessa tarjota kirjaston käyttäjille. (Coyle 2000, viitattu 5.6.2014.)

Lähes kymmenen vuotta myöhemmin Coyle sanoo artikkelissaan *Metadata mix and match*, että luettelointityön tulevaisuutta ajatellen tärkeintä ei ole keskittyä bibliografiseen luetteloon ja sen ominaisuuksiin vaan kaikkein tärkeimmät asiat ovat luettelon sisältämä tieto sekä se, että käytettyjen tietokantasovellusten tulisi olla yhteensopivia useiden eri standardien kanssa. Kirjastot eivät käytä ainoastaan itse tuottamaansa metadataa vaan suuri osa tietokantoihin tallennetusta datasta tulee muista lähteistä, jotka käyttävät omia luettelointiformaattejaan. Kirjastot hankkivat valmista dataa esimerkiksi Amazon.comista ja kustantamoilta, joilla usein on käytössään ONIX-formaatti. Vaikka esimerkiksi ONIX-muotoinen metadata voidaan konvertoida MARC-muotoon, muunnos ei ole täysin häviötön. Kirjastojen bibliografiset tietueet eivät Coylen mukaan ole muuttumattomia vaan ne laajentuvat eri lähteistä lisätyn tiedon myötä. Myös luetteloiden välillä tapahtuvan tiedon linkittämisen merkitys korostuu tulevaisuudessa yhä enemmän. (Coyle 2009a, 9.)

Jason Thomale määrittää artikkelissaan *Interpreting MARC: Where's the Bibliographic Data?* MARCin perimmäiseksi ongelmaksi yhteentoimivuuden kannalta sen, että MARC perustuu luettelointitietueisiin eikä itse bibliografiseen dataan. Luetteloitavat kohteet ilmaistaan standardisoidun luettelointitietueen kautta eikä suoraan jäsennehtynä datana. MARCin ongelmana on myös luettelointitietojen esittämiseen painottuminen kun nykyisessä toimintaympäristössä korostuu erilaisten järjestelmien välinen yhteentoimivuus. MARC vaatii avukseen monimutkaisia ja yksityiskohtaisia ohjelmia tiedon siirtämiseen muihin formaatteihin. Thomalen mukaan kirjastoissa käytettävien järjestelmien kehittämisessä suurimmat haasteet ovat järjestelmien yhteensopivuudessa ja niiden käyttäjien odotuksiin vastaamisessa. Bibliografisen datan erittely MARCista sen järjestelmien välistä siirtämistä varten vaatii enemmän kuin vain MARCin varsinaisen rakenteen tulkitsemista, sillä MARCiin sisältyy yhtä aikaa suorja ja epäsuoria rakenteita. MARCin kentät ja osakentät muodostavat sen varsinaisen, selkeän rakenteen, mutta luettelointisäännöt määrittävät rakennetta epäsuoralla tavalla. Ne limittyvät yhteen ja tuottavat semanttisia merkityksiä, jotka katoavat konversiossa, jos molempia rakenteita ei oteta huomioon. Thomalen mukaan formaattien välisen konvertoinnin suurin tekninen haaste voi olla luettelointisääntöihin pohjautuvan datan tulkitseminen, sillä MARCin kanssa on käytetty ja käytetään yhä useita erilaisia luettelointisääntöjä ja niiden eri versioita. (Thomale 2010, viitattu 8.6.2014.)

3.3.4 Kenttäongelmat

Blogikirjoituksessaan *Who needs MARC?* Lukas Koster nimeää MARCin ongelmaksi sen, että MARC-tietueet ovat ainoastaan digitoituja luettelointikortteja. MARC noudattaa korttiluettelon loogikkaa siinä, että se pyrkii sekä säilyttämään että esittämään tietoa samassa yhtenäisessä muodossa. Esimerkkeinä Koster käyttää MARC 21:n kenttiä 100 (Pääkirjaus – henkilönimi) ja 773 (Linkkikenttä – emokohde). Kentän 100 vikana Koster näkee sen, että henkilönimi kirjataan joko kaiseen tietueeseen erikseen sen sijaan, että luotaisiin tekijästä auktoriteettitietue, johon tekijän teoksia koskevat tietueet linkitettäisiin. Kentän 773 osakentässä ꞡg (yhteys, sijainti) Koster pitää ongelmana sitä, että osakenttään voi kirjata tietoa monessa eri muodossa. Jos osakenttä olisi normalisoitu eli jaoteltu valmiisiin osiin, joista kuhunkin kirjattaisiin tietty osa osakohteen sijaintitiedoista, esimerkiksi vain lehden, jossa artikkeli on julkaistu, vuosikerta, numero tai sivu, josta artikkeli alkaa tai johon se päättyy, olisi mahdollista tuottaa tietueen näyttöä varten automaattisesti tietueeseen kirjattua metadataa selittävää tekstiä. (Koster 2009, viitattu 5.6.2014.) Kosterin blogikirjoituksen kommentteissa Karen Coyle nostaa esiin sen, että tulevaisuudessa tietoelementtien standardisoiminen on tärkeää, mutta sitä ei pitäisi sitoa mihinkään tiettyyn standardisoituun formaattiin. MARCissa tietoelementit ovat standardisoituja, mutta ne kytkeytyvät erottamattomasti formaattiin, mikä aiheuttaa hankaluuksia tiedon siirtämisessä formaattien välillä. (Coyle 2009b, viitattu 5.6.2014.)

Blogikirjoituksessaan *MARC: from mark-up to data* Karen Coyle käy läpi MARC 21 -formaatin kenttien ja osakenttien ongelmia. Ongelmiksi Coyle määrittää indikaattorit, päällekkäisyyden, epäjohdonmukaisuuden ja datan, joka on hautautunut tekstiin. Yksi indikaattoreiden ongelmista on se, että osa niistä sisältää itsessään metadataa. Käyttämällä kentässä tiettyä indikaattoria ilmaistaan samalla tietoa luetteloidusta kohteesta. Toisena indikaattoreiden ongelmana on se, että kentän sisällön merkitys voi muuttua ratkaisevasti kentän eri indikaattoreita käytettäessä. Päällekkäisyydellä Coyle tarkoittaa teksti- ja ei-tekstimuotoisen datan päällekkäisyyttä. MARC 21 -tietueessa samaa tietoa toistetaan sekä kiinteissä kentissä, joissa se on helposti koneluettavassa koodimuodossa, että vaihtuvamittaisissa kentissä, joissa tieto esitetään tekstimuodossa. Samaa tietoa ilmaistaan tietueessa kontrollikentissä (00X), numero- ja koodikentissä ja luokituskentissä (0XX) sekä pääsääntöisesti tekstimuotoisissa kuvailevissa kentissä (1XX–9XX). Päällekkäisyys voi johtaa ristiriitaisiin merkintöihin saman tietueen sisällä. Epäjohdonmukaisuudesta Coyle käyttää esimerkkinä kenttiä 100 (Pääkirjaus – henkilönimi) ja 760 (Linkkikenttä – pääsarja), joista

kentän 100 osakenttiin Ꞥa, Ꞥb, Ꞥc, Ꞥd, Ꞥe, Ꞥj, Ꞥq ja Ꞥu kirjatut tiedot kirjataan kentässä 760 kaikki osakenttään Ꞥa. (Coyle 2010, viitattu 5.6.2014.) Tällä hetkellä hankalimpana ongelmana MARC 21:ssä nimitietojen osalta etenkin Suomessa pidetään sitä, että henkilön etu- ja sukunimi kirjaan samaan osakenttään.

Tekstiin hautautuneesta datasta Coyle taas käyttää esimerkkinä ISBNää. MARC 21 -tietueessa ei ole paikkaa, johon ISBN kirjattaisiin yksiselitteisenä dataelementtinä vaan kentän 020 (ISBN-tunnus, International Standard Book Number) osakenttään Ꞥa voidaan pelkän ISBN-tunnuksen lisäksi kirjata selittävää tietoa esimerkiksi kustantajasta tai sidosasusta. Tämän vuoksi järjestelmän, joka prosessoi MARC 21 -tietueita, täytyy olla ohjelmoitu erottelemaan kirjauksesta pelkkä ISBN-tunnus, jos sitä halutaan hyödyntää automaattisesti. Tämä ongelma ISBN-tunnuksen kohdalla ei ole kuitenkaan koskenut MARCin kaikkia variantteja ja on MARC 21:n viimeisimmässä versiossa ratkaistu. (Sama.) Kommentissaan Karen Coylen *MARC as Data: A Start* -artikkeliin Eric Nord nostaa esiin MARCin ongelmana myös sen suurpiirteisyyden. Nord mainitsee kontrollikentän 007 (Ulkoasua koskevat informaatiokoodit), johon on hankala luetteloita uusia aineistomuotoja esimerkiksi videotallenteiden formaattien nopean kehityksen myötä. (Eric Nord 2012, viitattu 8.6.2014.)

MARCin joustamattoman, mutta hyvin yksityiskohtaisen rakenteen hankaluutena on kenttien ja osakenttien suuri määrä ja jatkuva tarve lisätä niitä kuvaamaan uudenlaista aineistoa tai tarkentamaan aineiston kuvailua. MARC-kenttien käyttöä tutkittaessa on havaittu, että MARCissa on paljon kenttiä ja osakenttiä, joita käytännössä tarvitaan vain harvoin, sillä niitä käytetään ainoastaan tietynlaisen aineiston, esimerkiksi karttojen, kuvailuun (Moen & Bernardino 2003, 7–9; Smith-Yoshimura, Rowilson de Ortiz & Dickey 2010, 11). Mahdollisuus tarvittaessa kuvata monenlaisia aineistoja yksityiskohtaisesti voidaan toisaalta nähdä MARCin vahvuutena. Toisaalta hyvin harvoinkin käytetyt kentät ja osakentät täytyy ottaa huomioon systeemin suunnittelussa, ja koska vähän käytettyjä kenttiä on paljon, se lisää työmäärää huomattavasti, mikä taas voi vaikuttaa kustannuksiin. (Moen & Bernardino 2003, 7.)

3.3.5 Tehottomuus

Laura Krier nostaa esiin blogikirjoituksessaan *What's wrong with MARC? (NISO BIBFRAME roadmap meeting, part 2)* sen seikan, etteivät kirjastot tule luopumaan MARC-formaatista, ellei

sen vaihtoehtona tarjottu formaatti ole selkeästi tehokkaampi ja edullisempi. Kirjoituksessaan Krier pohtii syitä, jotka hänen käsityksensä mukaan tekevät MARCista tehottoman ja kalliin käyttää. Hänen mukaansa ohjelmoijat eivät ole kiinnostuneita työskentelemään MARC-datan kanssa, koska MARC ei ole kovinkaan yhteentoimiva sellaisten järjestelmien kanssa, joita muut tahot kuin kirjastot käyttävät. Lisäksi, koska MARC on käytössä lähinnä vain kirjastoissa, MARCin kanssa paremmin yhteentoimivien järjestelmien kehittämiseen ei ole suurtakaan kiinnostusta. Krierin mukaan luettelointityö kirjastoissa on kallista, koska sekä MARCin käytön opetteleminen että sen käyttäminen on vaikeaa ja monimutkaista. (Laura K 2013, viitattu 6.6.2014.)

Krier uskoo, että helppokäyttöisemmän formaatin käyttöön ottaminen toisi kirjastoille huomattavia kustannussäästöjä. MARCin vaikeakäyttöisyys johtaa lisäksi virheisiin luetteloidussa metadatas-
sa. Usein painotetaan kirjastojen luoman metadatan arvoa, sillä sitä pidetään laadukkaana, mutta MARCia käytettäessä mahdollisuus virheiden tekemiseen on suuri. Yksinkertaistamalla meta-
dataa siitä saataisiin yhteentoimivampaa ja joustavammin käytettävää, halvempaa tuottaa ja sen
käytettävyys lisääntyisi. (Laura K 2013, viitattu 6.6.2014.)

Toisaalta luettelointityön kalleuteen vaikuttavat käytetyn formaatin vaikeuden lisäksi myös luette-
lointisääntöjen käyttämisen vaatima asiantuntemus ja julkaisusta tallennettavan metadatan laa-
juus, joten helppokäyttöisempi formaatti ei sinänsä välttämättä muuttaisi tilannetta ellei myös luet-
telointisääntöjä yksinkertaistettaisi. Lisäksi mitä tahansa formaattia käytettäessä virheet ovat
mahdollisia ja käytännössä oleellista on, että luetteloiija on ammattitaitoinen ja tuntee luettelointi-
säännöt sekä käyttämänsä luettelointivälineet. Tällä hetkellä luettelointityössä säästöjä on saatu
aikaan poimintaluetteloinnilla, jota formaattien parempi yhteentoimivuus tai yhteinen formaatti
myös tukisivat.

4 BIBFRAME

Vuonna 2011 Yhdysvaltain kongressin kirjasto käynnisti Bibliographic Framework Initiative -projektin. Sen tavoitteena on luoda uusi bibliografisen kuvailun ympäristö, jossa luetteloidun datan verkottuminen ja yhteen kytkeytyminen toteutuvat. Uusien luettelointinormien, esimerkiksi RDA:n, tukemisella on myös merkittävä osa projektissa. Kongressin kirjasto, joka on päävastuussa MARC-formaatin ylläpidosta ja kehittämisestä, perustelee MARCista luopumista ja uuden luettelointimallin kehittämistä sen tilalle etupäässä MARC-muotoisen datan jaettavuuden ongelmilla ja sillä, ettei MARC vastaa luetteloinnin tarpeita nykyisessä toimintaympäristössä. (Library of Congress 2012, 3.) MARCin puutteita tarkastelevista raporteista Kongressin kirjasto nojaa pitkälti asettamansa työryhmän *On the Record: Report of the Library of Congress Working Group on the Future of Bibliographic Control* -raporttiin ja National Agricultural Libraryn, National Library of Medicine ja Kongressin kirjaston itsensä yhdessä johtamaan RDA-testaukseen (Marcum 2011, viitattu 25.6.2014). Taustana sille, että MARC korvattaisiin uudella luettelointimallilla, on vahvasti Kongressin kirjaston asettaman työryhmän *On the record* -raportissa antama suositus ja kirjastoalan vuosia jatkunut keskustelu MARC-formaatin ongelmista.

Kongressin kirjasto määritteli vaatimuksia, jotka uuden luettelointikehyksen tulisi täyttää. Sen tulisi olla riippumaton käytetyistä luettelointisäännöistä, sillä eri luetteloivat organisaatiot käyttävät erilaisia sääntöjä ja keskittyvät luetteloinnissa eri puoliin. Riippumattomuus luettelointisäännöistä myös tukisi jo olemassa olevan datan säilyttämistä ja hyödyntämistä, sillä ajan myötä samatkin luettelointisäännöt tai niiden käyttö on voinut muuttua ja osa metadatasta ei perustu mihinkään luettelointisääntöihin. Bibliografista kuvailua tukevan datan lisääminen esimerkiksi linkittämällä sekä tekstimuotoisen datan korvaaminen tunnisteilla katsottiin myös tärkeäksi. Tunnisteiden käyttö tekstin sijasta olisi huomioitava, koska käytössä olevat kirjastojärjestelmät ja muu luettelointiympäristö on moninainen ja ohjelmistoilla on keskenään hyvinkin eriasteisia valmiuksia siirtää ja tulkita tekstimuotoista tietoa. Uuden luettelointimallin kehittämisessä olisi huomioitava myös tiedonsiirtoformaattien, tiedonsyötön, -siirron, -tallentamisen ja -käsittelyn käytännöt, suositukset ja niiden väliset suhteet. Nykyisessä teknologiassa nämä sekoittuvat yhteen, mutta tulevaisuudessa on odotettavissa, että tallennetun datan esittämisellä on vähenevässä määrin yhteyttä tiedonsiirtoformaatteihin. Kaikenkokoisten ja -tyyppisten kirjastojen huomioiminen mallin käyttäjinä on myös tärkeää. Saman mallin tulisi palvella niin suurta tieteellistä kirjastoa kuin pientä, yksinkertaisempaa luettelointia tarvitsevaa kirjastoakin. Kirjastojen lisäksi uuden mallin täytyisi sopia myös

muille kulttuurimateriaalia luetteloiville muistiorganisaatioille. MARCin korvaavan mallin pitäisi erityisen hyvin soveltua toteuttamaan funktiot, joihin MARC on tarkoitettu. Näitä ovat tiedon esittäminen ja välittäminen. (Library of Congress 2011, viitattu 26.6.2014.)

Kongressin kirjasto huomioi MARC 21:n olevan todennäköisesti käytössä vielä vuosia ja sen, että kirjastoille järjestelmämuutokset ovat hankalia ja kalliita. MARC-formaatin tuesta ja ylläpidosta täytyisi huolehtia siihen asti kunnes se ei enää ole tarpeellista. MARCia voidaan vielä joutua päivittämään RDA:han mahdollisesti tehtävien muutosten takia, mutta muutoin MARCin päivitykset aiotaan toistaiseksi jättää mahdollisimman vähiin. Uuden mallin tulisi olla molempiin suuntiin yhteensopiva MARCin kanssa. Ihanteellista olisi, jos kaikki MARC-tietueissa oleva data pystyttäisiin säilyttämään muutoksesta huolimatta. Konvertoitumisen edestakaisuus ei ole keskeistä ainoastaan luetteloivien organisaatioiden kannalta, vaan se on välttämätöntä myös uuden mallin kehittämisessä ja testaamisessa. (Sama.)

Linkitetyn tiedon periaatteet ovat olleet uuden mallin kehittämisessä keskeisessä osassa. Linkitetystä tiedossa on pohjimmiltaan kyse jäsenneilyn datan julkaisemisesta internetin protokollien mukaisessa muodossa ja datan linkittämisestä yhteen muun datan kanssa, jotta tiedon löydettävyyden paransi. Datamalleista tähän sopivimpana on uuden luettelointimallin kannalta nähty RDF. RDF:n ominaisuuksien lisäksi sen käyttöä tukee se seikka, että sitä julkaisevat ja ylläpitävät vakiintuneet ja asiantuntevat organisaatiot, etupäässä Wide Web Consortium (W3C). (Ford 2012, 47, 49.)

Linkitetyn tiedon periaatteiden nähdään sopivan kirjastojen käytäntöihin. Keskeistä molemmissa on tiedon tai datan jakaminen ja julkaiseminen. Linkitetyn tiedon periaatteet ovat laajalti ohjelmistojen kehittäjien tuntemia ja linkitettyyn tietoon liittyy joukko HTTP-protokollaan pohjautuvia yleisesti käytettyjä menettelytapoja, joiden avulla RDF-muotoista dataa voidaan hyödyntää toimittavissa internetissä. Näiden metodien käyttäminen toisi kirjastojen datan paremmin esiin ja tekisi siitä käytettävämpää, minkä lisäksi ne tarjoaisivat myös hyvän ja tarkkaan määritellyn keinon ilmaista dataa. Linkitetyn tiedon malli toisi tiedon verkottumisen keskeiseksi osaksi tiedon jakamista ja käyttämistä sen sijaan, että kirjastojen data olisi pakattuna tietueina tietokantoihin, joista se joudutaan poimimaan ulos ja saattamaan käyttöön tiedonsiirtoprotokollan avulla. (Sama, 49.)

W3C suositti BIBFRAMEssa käytettäväksi RDF:ää, sillä se ei ole riippuvainen metadataskeemoista ja sitä voidaan käyttää niiden muuttuessa, minkä lisäksi RDF on suunniteltu erityisesti verkossa linkittyvän tiedon ilmaisemiseen ja hyödyntämiseen. RDF:ään ja muihin W3C:n malleihin tukeutumisen odotetaan mahdollistavan kirjastojen ja muiden kulttuuriorganisaatioiden datan integroitumisen internetiin ja olevan näin laajemmin käytettävissä. Yleisesti tunnetun ja käytetyn, asemansa vakiinnuttaneen standardin käyttäminen myös tekee kirjastoluetteloinnin dataympäristöstä ymmärrettävämmän esimerkiksi ohjelmistojen kehittäjille ja tekijöille, mikä mahdollisesti lisää kiinnostusta kirjastojen käyttämien ohjelmistojen kehittämiseen. Tämä mahdollistaa kilpailun lisääntymisen alalla, minkä toivotaan lisäävän kirjastojen mahdollisuuksia hyödyntää tietotekniikan ammattilaisten osaamista sekä tuovan kirjastojen ulottuville laajemman valikoiman palveluja ja teknisiä ratkaisuja. Sen toivotaan myös johtavan kirjastoille taloudellisesti edullisempiin käytäntöihin ja ratkaisuihin. (Library of Congress 2011, viitattu 26.6.2014.) BIBFRAMEssa RDF:n täsmällistä noudattamista ei kuitenkaan katsottu edulliseksi, sillä avoimen linkitetyn datan periaatteiden painotus on linkkienvälisyydessä eikä minkään tietyn mallin noudattamisessa. Kuitenkin, koska RDF koostuu pohjimmiltaan linkeistä, tunnisteista ja niiden nimiöistä, siihen pohjautuvat järjestelmät ovat sovitettavissa yhteen minkä tahansa muun vastaavat osat sisältävän järjestelmän kanssa. (Miller, Mueller, Ogbuji & MacDougall 2013, viitattu 27.6.2014.)

Koska BIBFRAMEn on tarkoitus olla yhteensopiva erilaisten kirjastojärjestelmien ja linkitetyn tiedon mallien kanssa, jopa pohjana kirjastojen väliselle luettelointitietojen siirrolle, on siinä huomioitava erilaisia tiedonsiirtomalleja. BIBFRAMEssa tiedonsiirron pohjana on tiedon sarjallistaminen RDF/XML-muotoon. BIBFRAMEn tiedonsiirtomalli voidaan purkaa URI-tunnisteisiin, linkkeihin ja tekstimuotoisiin nimiöihin, joten tiedonsiirtomalli, joka koostuu samoista perusosista, on BIBFRAMEn kanssa teoriassa jo yhteensopiva. Tiedonsiirron tukena voidaan käyttää myös muita RDF:ään perustuvia malleja kuten Turtlea tai RDF/JSON:ää. Pääasiallisena kyselykielenä BIBFRAMEssa on SPARQL, mutta muitakin voidaan käyttää. Uudessa tiedonsiirtomallissa on tärkeää huomioida linkitetyn tiedon käyttöön saamisen lisäksi myös vanhemmat ja nyt käytössä olevat tiedonhakuun tarkoitetut standardit kuten SRU ja Z39.50. BIBFRAMEssa tullaan mahdollisesti hyödyntämään jaettujen luettelointitietueiden päivityksistä tai uusien tietueiden lisäyksistä ilmoittamiseen web triggereitä, käyttäjäkohtaisia, automaattisia huomautuksia verkkosivujen välillä. Kun kirjasto esimerkiksi lisää luetteloonsa uuden painoksen jostain teoksesta ja luetteloisi siis teokselle uuden manifestaation ja linkittää sen yhteen teoksen kanssa, lähtisi teoksen uudesta

manifestaatiosta ilmoitus kirjastolle, joka teoksen on luetteloinut. Teoksen luetteloinut kirjasto voi näin hyväksyä uuden teosta koskevan manifestaation. (Sama.)

Yhdysvaltain kongressin kirjaston ja projektissa konsultoidun W3C:n lisäksi BIBFRAMEn kehittämiseen osallistuu semanttisen webin standardeihin ja teknologioihin erikoistunut, kirjastojen käyttämien ohjelmien ja sovellusten kanssa aiemminkin työskennellyt Zepheira-yhtiö. BIBFRAME-projektia on toteutettu varhaisesta vaiheesta lähtien käytännössä testaamisen ja kokeilun kautta. Vuoden ajan BIBFRAMEn skeemaa kehitettiin sitä testanneiden kirjastojen avulla. Varhaisiin testaajiin kuuluivat British Library, George Washington University, Princeton University, Deutsche National Bibliothek, National Library of Medicine, OCLC (Online Computer Library Center) ja Kongressin kirjasto. Vuosina 2014–2015 toivotaan ilmoittautuvan lisää instituutioita koekäyttämään BIBFRAMEa. MARCin nykyinen merkittävä rooli on huomioitu kehittämässä tekemällä yhteistyötä keskeisten MARCia käyttävien instituutioiden, Libraries and Archive Canadian, British Library ja Deutsche National Bibliothekin kanssa. Lisäksi pyydetään arvioita ja palautetta neuvoo-antavalta MARC-ryhmältä, johon kuuluvat American Library Association's Machine-Readable Bibliographic Information Committee ja Canadian Committee on MARC. Projektissa konsultoidaan OCLC:tä, ohjelmistojen myyjäyhteisöä, metadataa tuottavia ja käyttäviä tahoja, standardien asiantuntijoita kirjastoissa ja niiden ulkopuolella, sekä kerätään palautetta ja kommentteja ja yritetään herättää keskustelua asiasta kiinnostuneiden keskuudessa.

4.1 BIBFRAME-malli

Uusi luettelointimalli BIBFRAME, Bibliographic Framework, ei ole pelkkä luettelointiformaatti vaan se on käsitteellinen ja käytännöllinen malli tarkkaan bibliografiseen sekä kirjasto- että muunkinlaisen kulttuurimateriaalin kuvailuun ja kuvailutietojen tallentamiseen. BIBFRAME on verkkoteknologiaa käyttävä kehys tai metamalli kirjastojen ja muiden muistiorganisaatioiden tuottaman ja tallentaman tiedon hakemiseen ja siirtämiseen. Sen on tarkoitus olla joustava ja sovellettavissa eri organisaatioiden tarpeisiin mahdollistaen kaikille yhteisen luettelointimallin käyttämisen. (Miller, Mueller, Ogbuji & Baker 2014, viitattu 19.6.2014.)

BIBFRAMEn tavoitteena on tehdä selkeä ero luetteloitujen resurssien käsitteellisten sisältöjen ja niiden fyysisten tai digitaalisten manifestaatioiden välille sekä tunnistaa ja nimetä informaatioentiteettejä kuten auktoriteetteja. Olennaista on myös luetteloitujen resurssien välisten suhteiden esit-

täminen ja niiden hyödyntäminen. BIBFRAMEn on tarkoitus olla mukautettavissa erilaisiin luettelointisääntöihin ja sisältömalleihin. (Library of Congress 2014a, viitattu 24.6.2014) Tarkoituksena on myös toteuttaa parhaalla mahdollisella tavalla käytäntöjä datan ja tiedon näkyville tuomiseen, jakamiseen ja yhdistämiseen verkossa. Välineenä tähän käytetään tunnisteita, esimerkiksi URI-tunnisteita (Uniform Resource Identifier) ja IRI-tunnisteita (Internationalized Resource Identifier), joiden avulla tietty resurssi tunnistetaan ja nimetään. (Miller ym. 2013, viitattu 27.6.2014.)

4.2 Entiteetit

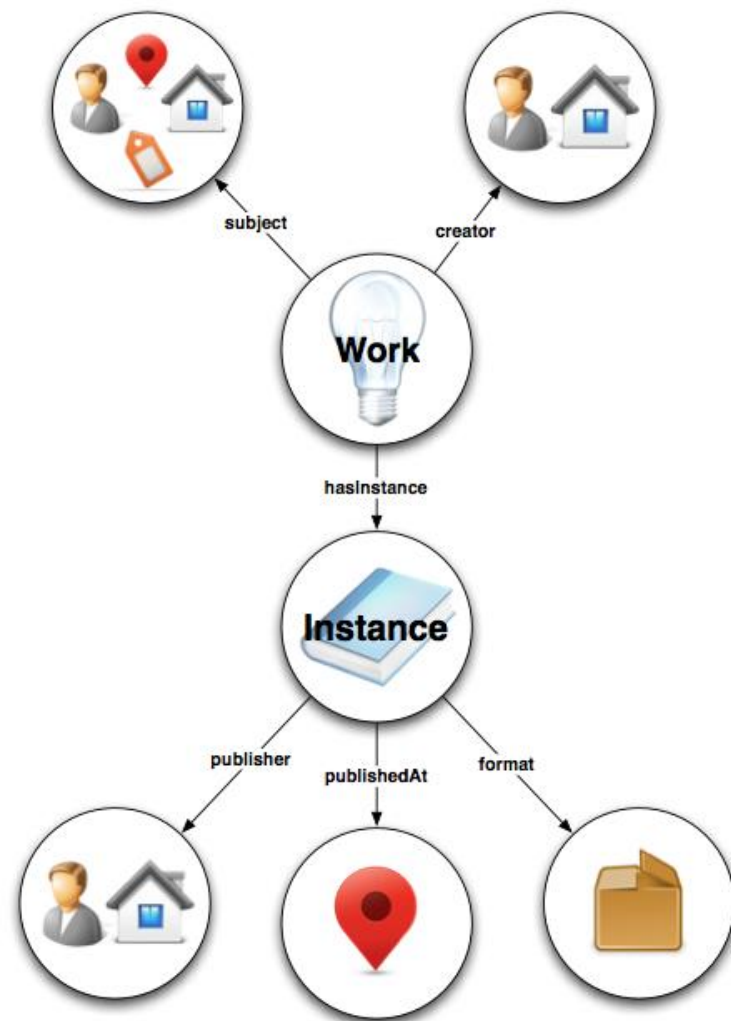
Vaikka BIBFRAME nojaa FRBR:ään, se on itsenäinen entiteetti-relaatio-malli. BIBFRAMEssa entiteettien väliset suhteet eivät ole yhtä vahvasti hierarkkisia kuin FRBR:ssä, mikä tekee mallista joustavamman. BIBFRAMEssa on neljä pääentiteettiä: teos (work), manifestaatio (instance), auktoriteetti (authority) ja liitännäinen (annotation). BIBFRAMEssa tehdään selkeä ero luetteloitavan kohteen sisällön ja sen fyysisten ilmentymien välille määrittämällä ja luetteloimalla erikseen entiteetit teos ja manifestaatio.

BIBFRAME-teos on abstrakti entiteetti, joka ei viittaa mihinkään fyysiseen kohteeseen. Teosentiteetin kautta saadaan erilaiset samansisältöiset manifestaatiot yhdistettyä toisiinsa ja sen kautta voidaan myös viitata toisiin teoksiin, jotka ovat esimerkiksi kokonaisuus-osa-suhteissa toisiinsa. FRBR:stä poiketen BIBFRAMEssa ei ole ekspression käsitettä vaan teoksen FRBR:n mukaiset ekspressiot tulkitaan pääsääntöisesti teoksiksi. Teoksille määritellään auktoriteetti-entiteettejä, jotka liittyvät teoksen aiheisiin sekä sen luomiseen. Teokseen liittyvä metadata on yhdistelmä osasta siitä datasta, joka sisältyy yhtenäistettyyn nimekkeeseen ja osasta MARCin bibliografiseen tietueeseen sisältyvästä kuvailutiedosta. Jokaiselle luetteloidulle resurssille kuvailaan ja tallennetaan sitä vastaava teos ja, jos resurssi on olemassa fyysisessä tai sähköisessä muodossa, myös yksi tai useampi manifestaatio.

BIBFRAME-manifestaatio on tietyn teoksen tietty fyysinen tai digitaalinen ilmentymä. Manifestaatioon kuuluu sen aineellisten ominaisuuksien kuvailu ja teoksen lisäksi siihen yhdistetään auktoriteetti-entiteettejä, jotka liittyvät kyseisen aineellisen ilmentymän tuotantoon, julkaisuun ja jakeluun.

BIBFRAME-auktoriteetti on BIBFRAME-resurssi, joka kuvaa esimerkiksi henkilöä, organisaatiota, paikkaa tai aihetta, joka puolestaan liittyy johonkin muista BIBFRAME-resursseista, siis teokseen, manifestaatioon tai liitännäiseen. BIBFRAME-auktoriteetti ei ole yhteneväinen yleisen auktoriteetti-käsitteen kanssa vaan BIBFRAME-auktoriteettien on tarkoitus tarjota erilaisille auktoriteettimalleille yhteinen kehys. Luetteloinnin kannalta auktoriteetit ovat tapa yhdenmukaistaa tietoa ja käyttäjälle ne toimivat kiintopisteinä luetteloa selatessa.

BIBFRAME-liitännäinen on resurssi, joka luetteloituun kohteeseen liitettynä lisää tietoa kyseisestä luetteloidusta kohteesta. Liitännäisen voi lisätä niin teokselle, manifestaatiolle, auktoriteetille kuin toiselle liitännäisellekin. Liitännäisen avulla voi ilmaista mielipiteen luetteloidusta resurssista, kertoa instituutiokohtaista tietoa siihen liittyen tai syventää sen kuvausta. BIBFRAME-liitännäisten yhtenä tarkoituksena on saada kirjastojen ja muiden muistiorganisaatioiden tuottama laadukas data yhdistettyä muuhun verkossa olevaan tietoon ja samalla hyödyntää luettelossa muiden tahojen tuottamaa tietoa. Liitännäisten avulla myös ilmaistaan esimerkiksi varasto- ja muita kirjasto-kohtaisia tietoja. Liitännäiset mahdollistavat lisäksi sen, että kirjaston ulkopuolinen taho, kuten kustantaja, voisi liittää teokseen tai manifestaatioon siihen liittyvää tietoa. (Library of Congress 2012, 8–12.)



Kuvio 1. BIBFRAME-malli (Library of Congress 2014b, viitattu 19.6.2014)

Kongressin kirjaston *Bibliographic Framework as a Web of Data* -julkaisussa painotetaan, että verkkoympäristössä, joka kirjastojen kannalta on erittäin olennainen toimintaympäristö, on tärkeää pystyä erottamaan teos ja sen jokainen eri manifestaatio toisistaan. Tärkeää on myös pystyä osoittamaan selvästi luetteloidun resurssin luomiseen liittyvät entiteetit, kuten tekijä ja julkaisija, ja resurssin sisältöön liittyvät käsitteet (kuvio 1). Luettelointimallin yhteentoimivuus verkkojen hakukoneiden kanssa on olennaista, sillä käyttäjät usein aloittavat tiedon etsimisen hakukoneella. Jotta käyttäjät löytäisivät kirjastojen ja muiden kulttuuriresursseja luetteloivien instituutioiden tallentaman tiedon mahdollisimman helposti, on tiedon oltava hakukoneiden löydettävissä. Tarvitaan siis myös koneille ymmärrettävää kuvailutietoa luetteloiduista resursseista. Antamalla entiteeteille tunnisteet mahdollistetaan se, että koneet pystyvät tulkitsemaan helposti entiteettien välisiä suhteita. Entiteettien erotteluisella toisistaan itsenäisiksi osiksi ja niiden yhteen linkittämisellä on

asioiden välisten suhteiden esiintuomisen lisäksi etuna se, että päivitetessä esimerkiksi henkilöön liittyviä tietoja tarvitsee muokata vain kyseistä henkilöä koskevaa tietuetta. BIBFRAMEssa käytetään RDF:ään pohjautuvaa BIBFRAME RDF -skeemaa kaikkien entiteettien ja niiden välisten suhteiden esittämiseen verkkoresursseina. RDF-yhteensopivassa muodossa esitettyinä tiedot ovat saumattomammin käytettävissä verkkoympäristössä. (Library of Congress 2012, 4.)

4.3 Suhteet ja ominaisuudet

BIBFRAME-suhde on suhde BIBFRAME-teoksen tai -manifestaation ja toisen BIBFRAME-teoksen tai -manifestaation välillä. Suhteita on siis pääasiassa neljäntyyppisiä: teoksen suhde toiseen teokseen, teoksen suhde manifestaatioon, manifestaation suhde teokseen ja manifestaation suhde toiseen manifestaatioon. Suhde ilmaistaan RDF-kolmikon avulla. Esimerkiksi haluttaessa ilmaista, että teos A edeltää teosta B, suhde ilmaistaan pelkistetysti muodossa subjekti-ominaisuus-objekti, siis teosA bf:edeltää teosB.

BIBFRAME-resursseille, luetteloiduille BIBFRAME-entiteeteille, määritellään ominaisuuksia (property), joista osa kuvaa resurssien keskinäisiä suhteita ja osa kuvailee yksittäistä resurssia. Ominaisuuksille on sekä yleisluontoisempia luokkia (class) että tarkempia alaluokkia. Resurssien suhteet ja ominaisuudet jakautuvat seuraavassa taulukossa (taulukko 1) esitettyihin tyypeihin, ja ominaisuudet vielä alaluokkiin, joilla voidaan kuvata tarkemmin resurssien välistä suhdetta. (Library of Congress 2014c, viitattu 19.6.2014.)

Taulukko 1. BIBFRAME-suhteet ja -ominaisuudet (Library of Congress 2014c, viitattu 19.6.2014)

BIBFRAME Relationship Type	Represented by Property	Domain	Range
Work to Work	bf:relatedWork	bf:Work	bf:Work
Work to Instance	bf:hasInstance	bf:Work	bf:Instance
Instance to Work	bf:instanceOf	bf:Instance	bf:Work
Instance to Instance	bf:relatedInstance	bf:Instance	bf:Instance
unspecified	bf:relatedTo	unspecified	unspecified

Ominaisuuksia määritetään teosten ja manifestaatioiden lisäksi myös auktoriteeteille ja liitännäisille. Kaikille resursseille, teoksille, manifestaatioille, auktoriteeteille ja liitännäisille, kirjataan auktorisoitu hakutieto (authorized access point), tunniste (identifier) ja nimiö (label) -ominaisuudet, joiden avulla tietty resurssi voidaan nimetä ja tunnistaa. Näiden lisäksi, resurssin tyypistä riippuen, resurssille voidaan kirjata sitä tarkemmin määrittäviä ja kuvailevia ominaisuuksia. (Library of Congress 2014d, viitattu 18.6.2014.)

4.4 Skeema ja sanasto

BIBFRAMElle on määritelty oma sanastonsa tai ontologiansa eli sen skeema. Sanasto on joukko termejä, jotka kuvaavat erityyppisiä resursseja, resurssien välisiä suhteita ja resursseihin liittyviä ominaisuuksia. Jokaisella sanaston ennalta määritellyllä termillä on URI, mutta myös sanastonulkoisia suhteita voidaan BIBFRAMEssa ilmaista käyttämällä tekstimuotoista tunnistetta. Koska käsitteiden ja termien sisällöt muuttuvat ja uusia syntyy, on katsottu tarpeelliseksi, että BIBFRAMEn sanasto on joustava ja muokattavissa. Lisäksi sanaston on tarkoitus olla sovitettavissa lueteloivien yhteisöjen mahdollisiin omiin yksityiskohtaisempiin sanastoihin. (Miller ym. 2013, viitattu 27.6.2014.)

BIBFRAME-sanasto (vocabulary) on joukko ennalta määriteltyjä luokkia ja ominaisuuksia (property). Luokka nimeää resurssista kuvailtavan puolen ja ominaisuudet tarkentavat kuvausta. Sanasto määrittää BIBFRAMEssa käytettyjä RDF:n mukaisten kolmikoiden predikaatteja ja kuvaa niiden välisiä suhteita ominaisuuksien tai tarkempien alaominaisuuksien avulla. Alustavasti BIBFRAME-sanastoon määriteltiin *Description of the Category View of the BIBFRAME Vocabulary* -dokumentissa (Library of Congress 2014e, viitattu 11.7.2014) kategoriat: General Information, Cataloging Resource Category Information, Title Information, Instance Description Information, Identifier Information, Note Information, Subject Information, Subject Classification Information, Subject Term Information, Relationship Information, General Relationship Information, Agent Relationships, Cataloging Resource Relationships, Annotation Information, Holdings Annotation Information ja Administrative Information.

Yleinen informaatio (General Information) sisältää ominaisuuksia, jotka periytyvät suurimpaan osaan muista sanaston luokista ja alaluokista. Yleistä informaatiota ovat esimerkiksi nimiö, tun-

niste ja auktorisoitu hakutieto. Myös teokseen tai manifestaatioon liittyvien auktoriteettien (hasAuthority) ja liitännäisten (hasAnnotation) merkinnät ovat yleistä informaatiota.

Luetteloidun resurssin kategoria (Cataloging Resource Category Information) on teoksen tai manifestaation alaluokka, joka kuvaa luetteloidun resurssin tyyppiä ja sisältöä. Teoksille kategorioita, eli teoksen alaluokkia, ovat esimerkiksi teksti, audio ja liikkuva kuva. Manifestaatioiden alaluokkia ovat muun muassa painettu julkaisu, sähköinen julkaisu ja käsikirjoitus. Luetteloidun resurssin kategoriaan sisältyvät myös RDA:n sisältö-, media- ja tallennetyyppejä vastaavat ominaisuudet contentCategory, mediaCategory ja carrierCategory.

Nimeketiedoilla (Title Information) on erikseen alakategoriat teoksen nimekkeelle (workTitle) ja manifestaation nimekkeelle (instanceTitle). Manifestaation nimeketiedot koostuvat pääasiassa neljästä osasta, jotka ovat nimeke (titleValue), alanimeke (subtitle), osan numero (partNumber) ja osan nimi (partName). Jos manifestaation nimekkeeseen tarvitaan muita osia, ne voidaan lisätä nimekkeen ominaisuus -ominaisuuden (titleAttribute) avulla. Teoksen nimeke vastaa jotakuinkin yhtenäistettyä nimekettä ja se koostuu samoista osista kuin manifestaation nimeke. Nimeketietoihin kuuluu myös muita ominaisuuksia, kuten nimekemerkinä (titleStatement), johon voidaan kirjata nimekkeen tiedonlähde, ja muu nimeke (titleVariation), johon voidaan kirjata nimekkeen vaihtoehtoisia muotoja.

Manifestaation kuvailutiedot (Instance Description Information) sisältävät tietoa teoksen tietystä manifestaatiosta. Manifestaation kuvailutietoihin kuuluvat esimerkiksi tiedot kyseisestä painoksesta, ulkoasusta, kestosta ja sarjan ensimmäisestä ja viimeisimmästä osasta. Lisäksi voidaan kirjata merkintä-ominaisuuksia (statement) kuten vastuullisuusmerkintä tai merkintä julkaisutiedoista. Yleensä manifestaation kuvailutiedot ovat tekstimuotoisia, mutta joitakin julkaisutietoja voidaan merkitä myös kontrolloidussa muodossa.

Tunnistetietoihin (Identifier Information) voidaan kirjata standardoituja tunnisteita sekä teoksille (ISTC, ISWC) ja manifestaatioille (ISBN, ISSN). Huomautuksiin (Note Information) kirjataan luetteloitua resurssia koskevia sekalaisia tietoja esimerkiksi palkinnoista tai kohderyhmästä. Aihetta koskevat tiedot (Subject Information) jakautuvat kahteen pääryhmään, luokitukseen ja asiasanoihin. MARC 21 -formaatin mallin mukaisesti vain muutama luokitusjärjestelmä on määritelty omak-

si ominaisuudekseen ja muut luokituskeemat merkitään yleisemmällä luokitus-ominaisuudella (classification), jonka voi nimetä käytetyn skeeman mukaan.

Suhteita koskevista tiedoista (Relationship Information) kaksi tärkeintä tyyppiä ovat resurssin ja toimijan välisten sekä resurssin ja resurssin välisten suhteiden tiedot. Resurssien väliset suhteet yhdistävät kaksi resurssia toisiinsa. Toistaiseksi erilaisia resurssien välisiä suhteita on määritelty noin viisikymmentä. Resurssin ja toimijan välistä suhdetta kutsutaan rooliksi (role). Rooleja ovat esimerkiksi toimija, tekijä ja julkaisija. Roolit yhdistävät teoksen tai manifestaation BIBFRAME-auktoriteettiin.

Auktoriteetteja koskevat tiedot (Authority Information) liittyvät toimijoihin, paikkoihin ja aiheisiin. BIBFRAMEssa oletuksena on auktorisoidun datan käyttäminen ja auktoriteettitiedostojen yhdistäminen luetteloituun resurssiin myös joidenkin sellaisten kuvailuelementtien tapauksessa, joiden kohdalla näin ei aiemmin ole tehty.

Liitännäisiä koskevia tietoja (Annotation Information) on määritetty viidentyyppisiä: kansikuva, tiivistelmä, arvostelu, sisällysluettelo ja varastotieto (Holdings Annotation Information). Muunkinlaisia liitännäisiä on mahdollista käyttää annotates-ominaisuuden avulla. BIBFRAME-liitännäinen lisää tietoa BIBFRAME-resurssista. Tämä "aboutness" tavallisesti ilmaistaan liitännäisen luokan avulla (Denenberg 2013, viitattu 29.6.2014). Luokat ovat kansikuva, tiivistelmä, arvostelu, sisällysluettelo ja varastotieto. Lisättäessä liitännäisiä resurssiin on mahdollista samalla tallentaa tiedot siitä milloin kirjaus on tehty, kuka sen on tehnyt ja mistä kirjauksen sisältämä tieto on peräisin. Kansikuva, tiivistelmä, arvostelu ja sisällysluettelo voidaan URIn avulla yhdistää luettelon ulkopuoliseen lähteeseen.

HeldMaterial on yleistä resurssia koskevaa sijainti- ja säilytystietoa. Sillä on tarkempi alaluokka kappale (HeldItem). HeldMaterialiin kuuluvia ominaisuuksia ovat esimerkiksi electronicLocator, joka osoittaa sähköisessä muodossa olevan resurssin sijainnin, heldBy, joka kertoo kenen tai minkä instituution kokoelmaan kyseinen resurssi kuuluu, ja lendingPolicy, johon kirjataan resursseja koskevat lainauskäytännöt. HeldItem sisältää muun muassa hyllypaikan, viivakoodin ja kappaleen kiertoon liittyvän tilan (circulation status).

BIBFRAMEssa ilmaistua hallinnollista metatietoa (Administrative Information) on kuvailutiedon luomiseen liittyvä tieto. Tällaista tietoa ovat esimerkiksi creationDate, johon kirjataan metadatan luomisajankohta, descriptionLanguage, johon merkitään kieli, jolla metadata on esitetty, ja descriptionSource, johon kirjataan metadatan luonut taho. Muutakin hallinnollista tietoa BIBFRAMEssa voidaan esittää muun muassa nidetiedoissa ja manifestaation kuvailutiedoissa.

Bibframe.org-sivustolla, jossa BIBFRAME:n skeema on varsinaisesti esitetty, BIBFRAME:n sanasto on kuitenkin hieman edellä esitetystä poikkeava. Bibframe.org-sivustoa päivitetään melko aktiivisesti kulloistakin sanaston kehittämisen vaihetta vastaavaksi ja se kuvaa BIBFRAME:n skeemaa suunnitelmaa huomattavasti yksityiskohtaisemmin. Sivustolla (Library of Congress 2014f, viitattu 11.7.2014) BIBFRAME:n sanaston kerrotaan koostuvan RDF-ominaisuuksista, luokista, ja niiden välisistä suhteista. BIBFRAME-sanaston ominaisuudet kuuluvat tietyille BIBFRAME-sanaston luokalle. Luokat voivat jakaantua tarkempiin alaluokkiin ja ominaisuudet tarkempiin alaominaisuuksiin. Bibframe.org-sivuston [Model View](#) esittää sen, mitkä ominaisuuksista kuuluvat millekin BIBFRAME:n pääluokista, jotka ovat BIBFRAME-entiteettien mukaiset. [Category View](#) taas esittää ominaisuudet aihepiireittäin ja selventää ominaisuuksien ja alaominaisuuksien suhteita. Bibframe.org sivustolla on lisäksi [List View](#) -osuus, jossa kaikki toistaiseksi määritellyistä luokista ja alaluokista sekä ominaisuuksista ja alaominaisuuksista on listattu kahteen aakkosjärjestyksenmukaiseen listaan.

Sivuston [Model View](#) -osassa sanaston pääluokiksi määritellään BIBFRAME-mallin entiteetit, eli teos, manifestaatio, auktoriteetti ja liitännäinen (Library of Congress 2014g, viitattu 11.7.2014). Näistä jokaiselle kirjataan yleiset ominaisuudet eli hakutieto, tunniste, nimiö ja relatedTo, jonka avulla määritetään mihin muihin resursseihin kyseinen luetteloitu resurssiin liittyy, sekä kyseiseen resurssiin liittyviä muita ominaisuuksia. [Model View](#)'ssa ominaisuudet on siis jaoteltu sen mukaan, mille entiteetille kyseinen ominaisuus voidaan kirjata. Jokaisella entiteettityypillä on myös tarkemmin määriteltyjä tyyppejä tai alaluokkia, teoksella esimerkiksi teksti, audio tai multimedia, liitännäisellä kansikuva, varastotieto ja niin edelleen.

Sivuston [Category View](#) -osiossa (Library of Congress 2014h, viitattu 11.7.2014) sanastossa määritellyt ominaisuudet esitetään jaoteltuna kategorioihin, eli luokkiin. [Category View](#)'ssa on siis listattu samat ominaisuudet kuin [Model View](#)'ssakin, mutta ne on esitetty tavalla, joka ilmaisee ominaisuuksien välisiä suhteita myös entiteeteistä riippumattomalla tavalla. Toistaiseksi esitetyt

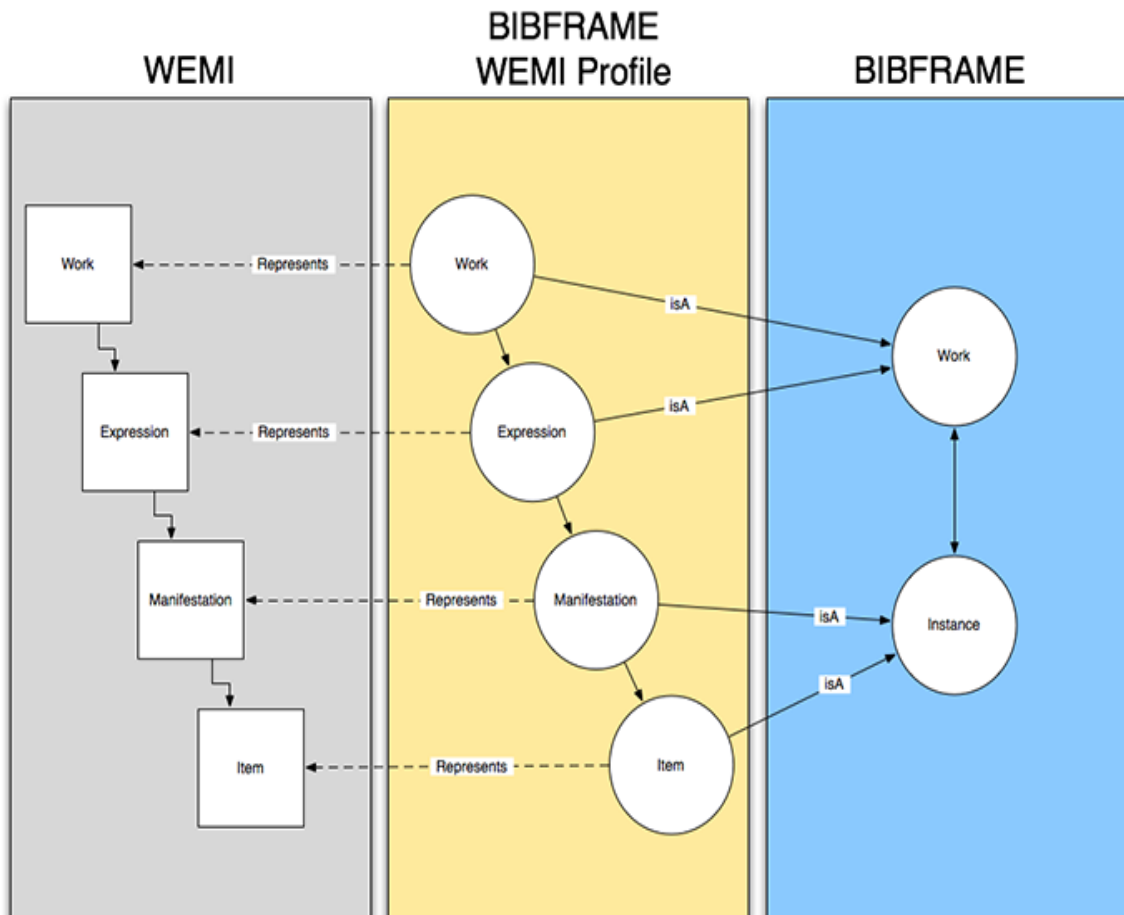
luokat ovat: General information, johon kuuluvat hakutieto, tunniste ja nimiö; Category information, johon kuuluvat ominaisuudet sisältö-, media-, ja tallennekategoria ja genre; Title information, johon kuuluu nimekkeeseen liittyviä ominaisuuksia; Instance information, johon sisältyy manifestaatioon liittyviä ominaisuuksia; Identifier information, johon kuuluvat standardisoidut tunnisteen, kuten ISBN; Note information, joka sisältää ominaisuuksia erilaisille huomautuksille; general Resource Relationship information, johon kuuluvat luokitukseen ja asiasanoihin liittyvät ominaisuudet; Very general related work or instance relationship information, johon kuuluu erilaisia yleisluontoisia teosten ja manifestaatioiden suhteita kuvaavia ominaisuuksia; Holdings Annotation information, joka sisältää liitännäisten varastotietoihin liittyvät ominaisuudet ja Administrative information, johon kuuluu hallinnollista tietoa esimerkiksi kuvailukäytännöistä ja tiedoston muokaamisesta.

4.5 BIBFRAME-profiilit

Koska BIBFRAME ei ole sidottu mihinkään tiettyyn luettelointistandardiin tai -sääntöihin, se ei itsessään tarjoa luottelijalle ohjeita luettelointityöhön. BIBFRAMEa käyttävät luetteloivat organisaatiot voivat tehdä itselleen BIBFRAME-profiilin, jonka avulla organisaation käyttämät luettelointia ohjaavat standardit, säännöt ja sanastot voidaan määritellä. BIBFRAME-profiili on dokumentti tai dokumenttiryhmä, joka määrittää paikalliset luettelointikäytännöt ja siten ohjaa luottelijaa käytäntöjen noudattamisessa. BIBFRAME-profiili voi olla myös jonkin yksittäisen kuvailustandardin tai -mallin, esimerkiksi RDA:n tai FRBR:n esittämisen ohje. BIBFRAME-profiilin perusrakenne muodostuu resource templateista, joita tarkentavat property templatet. Ominaisuuksille voidaan lisätä ominaisuuden arvon määrittämiseen kohdistuvia rajoitteita (value constraint). (Miller ym. 2014, viitattu 19.6.2014.)

FRBR:n tai RDA:n mukaisten teoksen, ekspression ja manifestaation suhteiden esittäminen BIBFRAMEssa on mahdollista. Teoksen ja ekspression suhde ilmaistaan kahden BIBFRAME-teoksen välisinä hasExpression ja expressionOf suhteina. Ekspression ja manifestaation suhde voidaan esittää BIBFRAME-teoksen ja -manifestaation välisinä hasInstance ja instanceOf suhteina. (Library of Congress 2014i, viitattu 22.6.2014.) Kahden FRBR:n tai RDA:n mukaisen teoksen, kahden ekspression tai kahden manifestaation välisen suhteen ilmaiseminen BIBFRAMEssa taas riippuu suhteen luonteesta. Suhde voidaan esimerkiksi määritellä yleisellä tasolla kahden teoksen väliseksi käyttämällä suhdetta kuvaavaa luokkaa relatedWork tai kirjasarjan osien välinen järjes-

tys käyttämällä luokkia continues tai continuedBy. Teoksen, ekspression, manifestaation ja kappaleen ilmaisemista varten voidaan myös tehdä BIBFRAME-profiili (kuvio 2), jolloin luetteloiva organisaatio saa yhtenäisen pohjan esimerkiksi RDA:n mukaiseen luettelointiin.



Kuvio 2. Teos, ekspressio, manifestaatio ja kappale -BIBFRAME-profiili (Miller, Mueller, Ogbuji & Baker 2014, viitattu 19.6.2014)

4.6 Audiovisuaalisen aineisto kuvailu

BIBFRAME AV Modeling Study -raportissa on käsitelty audiovisuaalisen aineiston luetteloinnin tarpeita ja annettu ehdotuksia siitä, kuinka tarpeet voitaisiin BIBFRAMEssa huomioida. Raportissa on tarkastelu erilaisia, nykyisin käytössä olevia datamalleja AV-aineiston luetteloinnin näkökulmasta. Yleisen, kattavan datamallin, joka tukisi myös audiovisuaalisen aineiston luettelointia, täytyisi kattaa niin aineiston sisällön kuin sen tallennusvälineenkin (carrier) kuvailu käyttäjän tar-

peita vastaavalla tavalla. Nykyiset datamallit ovat erikoistuneita tietynlaisen aineiston kuvailuun eikä mikään niistä kata kaikkia AV-aineiston kuvailun tarpeita.

Raportissa BIBFRAME nähdään mahdollisuutena saada aikaan yleinen, AV-aineiston kannalta riittävän kattava datamalli. BIBFRAMEn avulla on mahdollista luetteloida muun muassa tapahtumia (event), mutta AV-aineiston kannalta ongelmallista on se, että toistaiseksi BIBFRAMEssa tapahtuma ei voi olla irrallinen teoksesta. AV-aineiston kohdalla joissain tapauksissa, esimerkiksi jos kuvaillaan äänitettä joka sisältää linnun laulua, varsinaista teosta ei ole vaan manifestaatio on tallenne tapahtumasta. BIBFRAMEn tietomalliin ehdotetaan lisättäväksi hierarkkisesti teoksen yläpuolelle entiteettiä sisältö (content), jonka alapuolella teos ja tapahtuma olisivat tasaveroisina. Näin saataisiin mahdollisuus selvästi erottaa luova sisältö, eli teos, ja tapahtuman dokumentointi. Raportissa teoksen ja tapahtuman erottamista toisistaan omiksi entiteeteikseen perustellaan muun muassa teollisoikeudellisella näkökulmalla. Teokseen lähtökohtaisesti liittyy esimerkiksi tekijänoikeuksia, kun taas tapahtumaan, esimerkiksi mielenosoitukseen, itseensä ei liity. Tapahtumaa koskeva dokumentti, esimerkiksi puhtaasti dokumentaarinen video mielenosoituksesta, on raportin mukaan manifestaatio, ei itsenäinen teos. RDA:n tavoin BIBFRAME vaatii jokaiselle manifestaatiolle luetteloitavaksi myös siihen liittyvän teoksen, ja koska raportin mukaan teosta ei aina ole, tarvitaan BIBFRAMEen teoksen kanssa hierarkkisesti rinnakkainen tapahtuma-entiteetti. (Van Malssen 2014, 41–43.)

Raportissa ehdotetaan, että jotkut entiteettien ominaisuudet, kuten kesto (duration) tai kuvasuhde, pitäisi laajentaa manifestaation lisäksi koskemaan myös teos- ja sisältö-entiteettejä. Suosituksena annetaan myös sen varmistaminen, että yksi manifestaatio voi edustaa yhtä aikaa useita manifestaatiotyyppisiä, olla yhtä aikaa esimerkiksi sähköinen julkaisu, kokoelma ja monografia tai moniosainen monografia. Koska audiovisuaalisen aineiston varastotietojen ja aineiston käyttöominaisuuksien kuvailu on osin toisistaan erottamatonta, ehdotetaan raportissa, että BIBFRAMEn olisi mahdollisesti hyvä kattaa laajemmin teknistä metadataa, joka mahdollistaisi myös AV-aineiston teknisten ominaisuuksien kuvailun. Raportissa suositellaan myös liitännäisiin lisättäväksi luokkia, jotka kuvaavat FRBR:n mukaisiin kappaleisiin liittyviä ominaisuuksia ja ehdotetaan tutkittavaksi mahdollisuutta esittää tietoa esimerkiksi elokuvaan sisältyvistä kohtauksista ja niiden välisistä suhteista, sekä yhdistää tietty sisältö manifestaation tiettyyn ajalliseen kohtaan. (Sama.)

5 VERTAILU

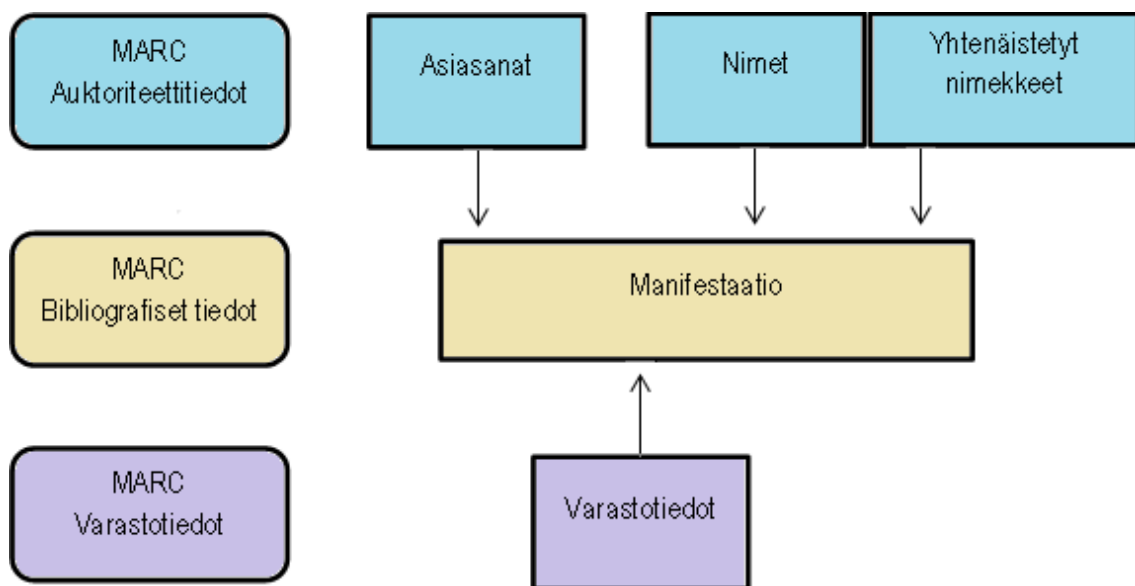
Keskeisin MARCin ja BIBFRAME:n välinen ero on se, että MARC on formaatti yksittäisten resurssien bibliografiseen kuvailuun kun BIBFRAME taas keskittyy kuvaamaan resurssien välisiä suhteita. Kaikissa MARC-formaatin varianteissa luodaan tietueita, jotka ovat toisistaan riippumattomia ja yksittäisinä ymmärrettäviä, sekä kerätään yhteen tietoa sekä teoksesta että sen tietystä fyysisestä ilmentymästä. MARCissa tunnisteet ovat yleensä merkkijonoja, joilla on sama sisältö myös tietueen ulkopuolella, esimerkiksi henkilön nimiä tai asiasanoja. Käytössä on kuitenkin joitakin koodimuotoisia tunnisteita kuten kielikoodi tai maantieteellisen alueen koodi. BIBFRAMEssa tunnisteet, joilla kuvataan asioita esimerkiksi henkilöitä, paikkoja tai kieliä, pyritään säännönmukaisesti kontrolloimaan. (Library of Congress 2014a, viitattu 3.7.2014.)

Rurik Greenall nostaa esiin *Brinxmat's blog* -blogissaan kirjoituksessa *What does it mean to "provide support for BIBFRAME"?* MARCin ja BIBFRAME:n keskeisimpänä erona sen, että MARC keskittyy dokumentteihin kun BIBFRAME taas keskittyy lauseisiin (statement). BIBFRAMEssa tietoa ei esitetä yksittäisenä dokumenttina vaan luetteloitua resurssia koskevana hajautettuina lauseina. MARCissa koodataan tietoa yksittäisiin dokumentteihin, tietueisiin, joita voidaan tulkita sopivien ohjelmistovälineiden avulla. MARCin tavoin myös BIBFRAME on tiedonkantaja (carrier) ja tiedonsiirtoformaatti, mutta lisäksi se on metadataskeema. Toisin kuin MARC, BIBFRAME siis tarjoaa ohjeita myös datan rakenteeseen. (Greenall 2014, viitattu 9.7.2014.)

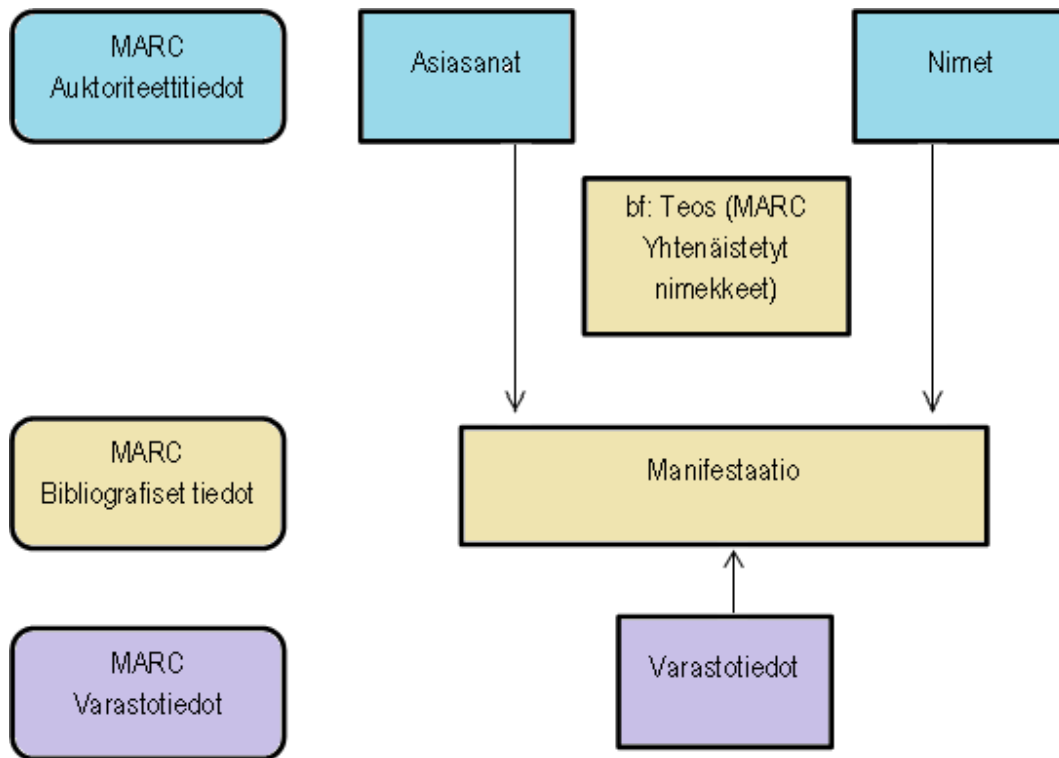
BIBFRAMEa ja MARCia yhdistää niiden käyttötarkoitus. Molempien tavoitteena on mahdollistaa vastaaminen luettelon käyttäjien tarpeisiin löytää, tunnistaa, valita ja saada käyttöön haluttuja aineistoja. Samankaltaisuutta on myös jaossa bibliografisiin ja auktoriteettitietoihin, vaikka suurikin eroavaisuuksia BIBFRAME:n ja MARCin välillä niissä on. Bibliografiset tiedot MARCissa ilmaistaan manifestaatioiden kautta, kun BIBFRAMEssa taas tehdään selvä ero teoksen ja manifestaation välille. BIBFRAMEssa on myös mahdollista ilmaista muita entiteetti-relaatio-malleja kuin vain teoksen ja manifestaation sekä niihin liittyvien auktoriteettien suhteita. Auktoriteettitiedot BIBFRAMEssa ovat jäsennellympiä ja itsenäisempiä kuin MARCissa ja entiteettien välisten suhteiden ilmaiseminen on syvällisempää ja monipuolisempaa.

Sally McCallum (2014, 7–12) kuvaa *BIBFRAME and Cataloging: changing landscapes* -esityksessään MARC-formaatin ja BIBFRAME:n eroja MARCista BIBFRAMEen siirtymisen

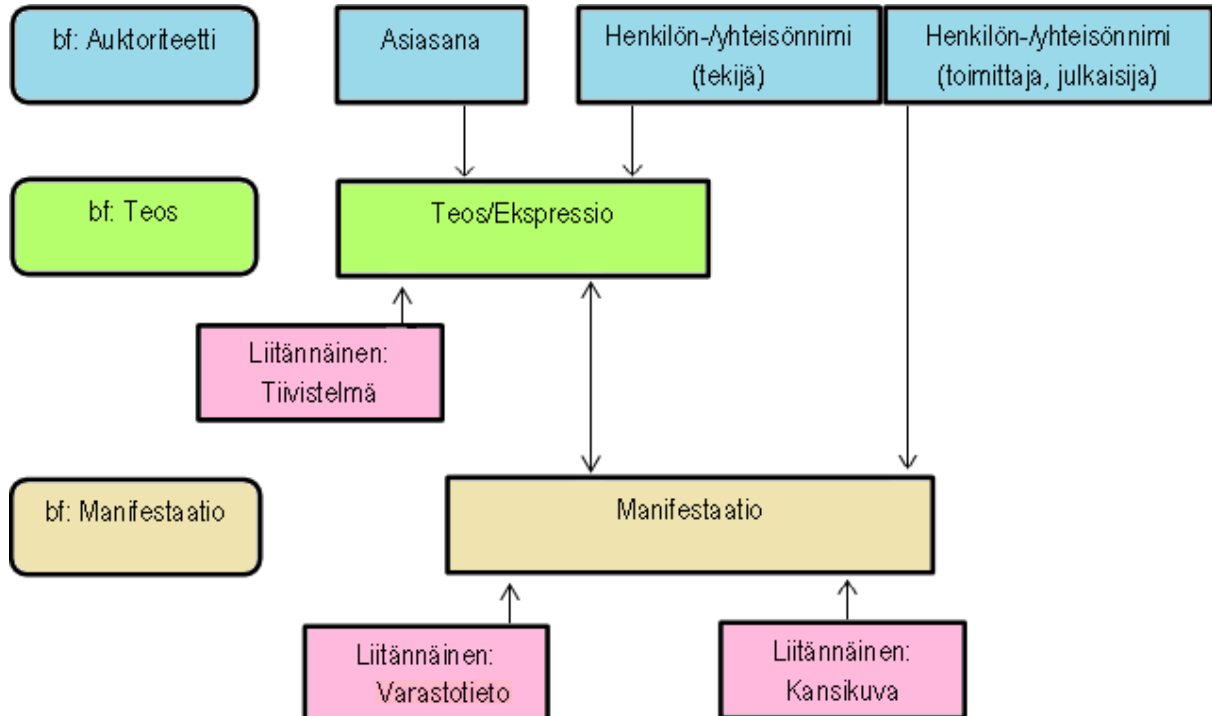
näkökulmasta. McCallumin mukaan MARCista BIBFRAMEen siirryttäessä MARCin auktoriteettitietojen nimistä ja asiasanoista tulee BIBFRAME-auktoiteetteja ja yhtenäistetyistä nimekkeistä BIBFRAME-teoksia. Kaikkia yhtenäistettyjä nimekkeitä vastaa siis BIBFRAME-teos ja saman julkaisun eri manifestaatioita BIBFRAME-manifestaatio. Vaikka MARC-tietueessa ei olisi julkaisun yhtenäistettyä nimekettä, luodaan kuitenkin BIBFRAME-teos, sillä jokaisella BIBFRAME-manifestaatiolla täytyy olla sitä vastaava BIBFRAME-teos. MARC-tietueen teosta tai ekspressiota koskevat kuvailutiedot siirretään BIBFRAME-teoksen kuvailutiedoiksi. BIBFRAME-manifestaation kuvailutiedoiksi siirretään vain MARC-tietueen manifestaatiota koskevat tiedot. Teosta koskevia tietoja ei siis toisteta manifestaation kuvailussa. McCallum kuvaa MARCin ja BIBFRAMEen eroa ja konversiota MARCista BIBFRAMEen seuraavanlaisina kuvioina (kuvio 3, 4, 5).



Kuvio 3. MARC-malli Sally McCallumin mukaan (McCallum 2014,7)



Kuvio 4. MARCin yhtenäistetty nimeke BIBFRAMEn teokseksi (McCallum 2014, 10)



Kuvio 5. BIBFRAME-malli Sally McCallumin mukaan (McCallum 2014, 12)

MARC ja BIBFRAME poikkeavat toisistaan bibliografisen kuvailun toteutustavoiltaan. Kortistopohjainen MARC tuottaa luettelointitietueita, yksittäisiä kiinteitä kokonaisuuksia, joissa pyritään ilmaisemaan kaikki tarvittava tieto resurssista. MARCissa kuvaillaan manifestaatioita, joista jokaisen yhteydessä kuvaillaan myös manifestaatioon liittyvää teosta. BIBFRAMEssa taas teos kuvaillaan kokonaisuutena ja linkitetään manifestaatioihin, joista kuvaillaan vain teoksen kyseiselle manifestaatiolle ominaiset piirteet. MARCissa siis toistetaan jokaisen manifestaation yhteydessä teoksen kuvailu, kun BIBFRAMEssa manifestaation kuvailussa ei tallenneta lainkaan tietoa teoksesta, mutta kattava kuvailu voidaan luoda yhdistämällä manifestaation ja teoksen kuvailut toisiinsa.

MARC-luetteloinnissa data tallennetaan tietueisiin. BIBFRAMEssa tietueita ei ole. Tietuetta vastaavan yksittäisen luetteloidun entiteetin kuvailun pohjana ovat entiteetin nimeävät nimiö, tunniste ja auktorisoitu hakutieto. Tähän pohjaan yhdistetään BIBFRAME-sanaston kyseiselle entiteetille määrittämistä kuvailuelementeistä ne, jotka luetteloivan tahon käyttämien kuvailusääntöjen ja muiden käytäntöjen mukaan ovat tarpeellisia. Toisin kuin MARCissa BIBFRAMEssa tiedon linkittyminen on siis väistämättömästi sidoksissa sen rakenteeseen.

Erona MARCin ja BIBFRAMEn välillä on lisäksi se, että BIBFRAMEn kehittämisessä on laajasti huomioitu myös ohjelmistosuunnittelun näkökulma. Siitä on pyritty tekemään ohjelmoijille ymmärrettävämpi, jotta, toisin kuin MARCissa, ohjelmoijan ei tarvitse ymmärtää luettelon rakenteen vaikutusta sisältöön tai luettelointikäytäntöjä kovinkaan syvällisesti voidakseen esimerkiksi tehdä päivityksiä ohjelmistoon. Lähtökohtaisena ohjelmistollisena erona MARCin ja BIBFRAMEn välillä on myös näkökulma yhteentoimivuuteen muiden järjestelmien kanssa. Ollakseen yhteentoimiva MARC vaatii useita tiedonsiirtoon liittyviä standardeja ja yksityiskohtaisia konversio-ohjelmia, koska MARCin lähtökohtana on tiedonsiirto saman formaatin sisällä. BIBFRAMEssa tiedonsiirron lähtökohdaksi on otettu järjestelmien välinen yhteentoimivuus.

Verrattaessa MARCin ja BIBFRAMEn niitä puolia, jotka MARCissa koetaan erityisen ongelmalliseksi, vaikuttaa BIBFRAME niiltä osin selkeästi ongelmattomammalta. MARCin tehottomuuden, sen vaikeaselkoisuuden ja tiedon päällekkäisyyden ongelmat BIBFRAMEssa on ratkaistu tekemällä rakenteesta koneluettavan lisäksi myös ihmiskäyttäjälle helpommin luettavissa oleva ja jakamalla tallennettava tieto yhdestä itsenäisestä kokonaisuudesta, tietueesta, useisiin osiin, joista yhdistämällä muodostetaan luetteloitua resurssia kuvaava kokonaisuus. MARCista tekee erityisen vaikeaselkoisen ja inhimillisille virheille alttiin se, että luettelointitietueen rakenne, kentät,

osakentät ja indikaattorit, on ilmaistu kokonaan koodimuotoisena ja näiden numero- ja kirjainkoodien tulkinnan avuksi tarvitaan erilliset ohjeet. Vastaavasti taas suuri osa kuvailutiedosta kirjataan tietueeseen tekstimuodossa, mikä lisää virheiden mahdollisuutta ja erilaisten tulkinnanvaraisten ilmaisutapojen, esimerkiksi lyhenteiden, käyttämistä. BIBFRAMEssa luokat ja ominaisuudet ilmaistaan luetteloijalle tekstinä, mikä vähentää tarvetta erillisille luettelointiohjeille. Tallennettavan tiedon jakaminen osiin ja mahdollisuus liittää esimerkiksi valmiiksi olemassa olevat tiedot tekijästä uuteen teokseen vähentää sekä virheiden mahdollisuutta että tehtävän työn määrää.

BIBFRAMEssa MARCin rajoittuneisuus lähinnä kirjastojen käyttöön ja hankaluudet yhteentoimivuudessa on pyritty ratkaisemaan käyttämällä yleisesti tunnettua ja käytettyä RDF-standardia. BIBFRAME ei kuitenkaan noudata yleisesti tunnettuja RDF-skeemoja, vaan sille määritellään oma skeema, jota muut kuin BIBFRAME-sovellukset eivät välttämättä ymmärrä. Tällaisia tapauksia varten esimerkiksi nimitiedot joudutaan muuttamaan FOAF (Friend of a Friend) -muotoon tai jonkin muun yleisesti tunnetun skeeman mukaisiksi.

Kun kirjastojen tallentaman tiedon rakenne voidaan helposti muuttaa yhteismitalliseksi muiden toimijoiden tuottaman tiedon kanssa, voidaan muiden tuottamaa tietoa hyödyntää kirjastoissa ja kirjastojen tuottamaa tietoa muualla huomattavasti sujuvammin. MARC-formaatin rakenne vaikuttaa tallennetun tiedon sisältöön ja MARC-luettelointi on kiinteästi sidoksissa useisiin standardeihin, kuten luettelointisääntöihin, jotka osaltaan vaikuttavat tiedonsiirtoon ja konvertointiin tarvittaviin teknisiin ratkaisuihin. BIBFRAMEssa pyritään siihen, ettei tiedonsiirtoa ja konversiota varten tarvitsisi jokaisen tallennetun tiedon kohdalla erikseen ottaa huomioon tiedon rakenteen vaikutusta tallennetun tiedon sisältöön. Toisin kuin siis MARCissa, joka sisältää sekä suoria että epäsuoria rakenteita, BIBFRAMEssa kaikki tallennettu tieto koostuu RDF-kolmikoista ja jäsentyy kauttaaltaan yhtenäisesti. Jos yhteentoimivuusongelma saadaan ratkaistuksi, ei sekään, että BIBFRAME jäisi vain kirjastojen käyttämäksi malliksi, olisi todennäköisesti enää yhtä ongelmallinen tilanne.

BIBFRAME on toistaiseksi malli pääasiassa vain aineiston kuvailuun, ei niinkään aineiston hallintoihin tarvittavan tiedon tallentamiseen, vaikka esimerkiksi elektronisen aineiston pitkäaikais säilytykseen tarvittava metadata on toisaalta myös aineiston kuvailutietoa. Tällä hetkellä vaikuttaa siltä, ettei BIBFRAME välttämättä pyri kattamaan kaikkea e-aineiston hallinnollista metadataa. Samoin kuin MARCin myös BIBFRAMEn käyttäminen toistaiseksi vaatisi e-aineiston pitkäaikais-

säilytykseen ja käyttöön, käsittelyyn ja käyttöoikeuksiin liittyen erillisiä metadastandardeja. BIBFRAME tulisi siis olemaan riippuvainen siitä erillisistä metadastandardeista tai datamalleista esimerkiksi METSistä (Metadata Encoding and Transfer Standard) tai PREMISistä (Preservation Metadata Implementation Strategies). MARCIin verrattuna BIBFRAMEssa on toistaiseksi vielä vähemmän mahdollisuuksia tallentaa e-aineistojen hallinnollista metadataa, ja hallinnollinen metadata mahdollisesti jää BIBFRAMEn kanssa yhteentoimivien, sen käsittelyyn erityisesti kehitettyjen formaattien tehtäväksi. On kuitenkin tehty selvitys erilaisista datamalleista audiovisuaalisen aineiston luetteloinnin näkökulmasta ja selvityksen raportissa esitetään ehdotuksia, kuinka audiovisuaalisen aineiston luetteloinnin tarpeet voitaisiin BIBFRAMEssa huomioida (ks. s. 37–38). Elektronisen ja audiovisuaalisen aineiston luetteloinnin tarpeet hallinnollisen metadatan, muun muassa oikeudellisten ja formaattia koskevien tietojen osalta, ovat pitkälti yhteneväiset, ja jos ehdotukset toteutetaan BIBFRAME voi tulla kattamaan hyvinkin laajasti myös hallinnollisen metadatan tallentamisen.

6 JOHTOPÄÄTÖKSET

MARC-formaatti koetaan kirjastoalalla nykyisiin luetteloinnin tarpeisiin puutteelliseksi ja monilta osin ongelmalliseksi. Keskeisimpiä ongelmia MARCissa ovat tiedon linkittämisen rajalliset toteutusmahdollisuudet ja formaatin sopimattomuus internetiin toimintaympäristönä, sekä sopimattomuus ilmaisemaan entiteetti-relaatio-malleja. MARC-formaatti on ongelmallinen myös ohjelmistojen kannalta. MARC-formaattia tukevan sovelluksen kehittäminen ja ylläpito vaatii ohjelmointitaidon lisäksi myös syvällistä ymmärrystä bibliografisesta luetteloinnista. Koska MARCin ongelmat pohjautuvat sen perustavanlaatuisiin ominaisuuksiin kuten rakenteeseen, ei ongelmia voida täydellisesti ratkaista formaattia päivittämällä. Kokonaan uuden, luetteloinnin ja luettelon käytön nykyisten tarpeiden ja toimintaympäristön lähtökohdista toteutetun luettelointiformaatin kehittäminen on katsottu tarpeelliseksi, jotta MARC-formaatin historiallisesta painolastista päästäisiin eroon.

BIBFRAME-mallin teoreettinen pohja on jo sinänsä valmis ja tavoitteet, joihin pyritään, ovat selkeitä. Malli on kuitenkin toistaiseksi toteutukseltaan vielä niin keskeneräinen, että siitä on vaikea esittää täsmällisiä johtopäätöksiä, sillä tekninen toteutus voi vaikuttaa ratkaisevasti mallin lopulliseen muotoon. BIBFRAMEsta käydään vielä aktiivisesti keskustelua ja jo alustavasti valmiiksi saatuja osiakin muokataan ja täsmennetään. Alkuperäisen aikataulun mukaan BIBFRAMEn pitäisi olla tätä kirjoitettaessa jo valmis, mutta aikataulua on venytetty eikä mitään sitovaa takarajaa ole asetettu. Tällä hetkellä on vaikeaa arvioida, kuinka paljon työtä on vielä tekemättä, ja kuinka kauan aikaa vie ennen kuin BIBFRAMEa tukevia sovelluksia voidaan ryhtyä rakentamaan. Todennäköistä kuitenkin on, että siirtymäkausi MARC-formaateista BIBFRAMEen tulee kestämaan pitkään.

BIBFRAME-keskustelua on käyty monista seikoista, jotka voisivat vaikuttaa mallin lopulliseen muotoon ratkaisevasti, esimerkiksi luettelointisääntöjen mukaisten välimerkkien upottamisesta dataan sekä tietoelementtien URI-pohjaisesta linkittämisestä ja sen vaikutuksesta tiedonhaun tekniseen toteuttamiseen (esim. Sanderson, Coyle & Ford sähköpostiviestiketju 28.7.2014; Berger, Soroka, Coyle, Thompson, Yeates, Sanders & Ford sähköpostiviestiketju 27.7.2014). Pääasiallisena linjauksena toistaiseksi on, ettei välimerkkejä BIBFRAMEssa käytetä, mikä voidaan nähdä selkeänä parannuksena MARC 21:een verrattuna, ja mikä vastaisi esimerkiksi UKMARC- ja FINMARC-formaattien käytänteitä. Jos kuitenkin välimerkit jostakin syystä päätettäisiin MARC

21:n mukaisesti jättää dataan, voisi siitä seurata suuria muutoksia nykyiseen BIBFRAME-versioon.

BIBFRAMEn ja MARCin vertailu jää BIBFRAMEn keskeneräisyyden vuoksi melko pinnalliseksi. Jos esimerkiksi haluttaisiin selvittää MARCin ja BIBFRAMEn käytettävyyttä luotteloijan näkökulmasta tai luettelon käytettävyyttä asiakkaan näkökulmasta, olisi huomioitava myös luettelointityöhön käytettyjen välineiden muodostama kokonaisuus, esimerkiksi käytetty kirjastojärjestelmä, luettelointisääntöjen ja MARCia tai BIBFRAMEa täydentävien formaattien vaikutukset. Tuloksellisempi vertailu vaatisi rajatumpia näkökulmia tai laajemman työn, jossa yksityiskohtiin voisi syventyä tiiviimmin.

BIBFRAMEsta voidaan kuitenkin jo tehdä yleisluontoisempia johtopäätöksiä ja yksittäisiä havaintoja. Se vaikuttaisi sopivan erityisen hyvin luettelointitiedon jakamiseen sekä laajoihin yhteisluetteluihin ja keskitettyyn luettelointiin. Luetteloidun datan hyödyntämisen on tarkoitus BIBFRAMEssa toimia kuitenkin myös alhaalta ylöspäin niin luutteloivan organisaation sisäisesti kuin sen ulkopuolistenkin tahojen kanssa. Keskitetyn luettelon ylläpitäjä voi halutessaan helposti linkittää luetteloonsa tietoa, jonka on tuottanut jokin muu taho ja mahdollisesti hyväksyä luettelon osaksi muun tahon tuottamaa ei-bibliografista tietoa. BIBFRAME tarjoaa myös mahdollisuuden linkittää tietoja keskenään hyvinkin erilaista aineistoa sisältävistä luetteloista; esimerkiksi kirjastoluettelosta voitaisiin viitata taidemuseon luetteloimaan maalaukseen. Tiedonhakijalle tällainen olisi eduksi, ja internetistä tietoa hakemaan tottunut voi jopa pitää lähtökohtaisena oletuksena eri tahojen tuottaman tiedon linkittymistä. Tiedonhakijan kannalta BIBFRAMEn linkittyvyys on hyödyllistä niin laajempien asiakokonaisuuksien ja resurssien välisten suhteiden hahmottamisen kuin luettelon selailtavuudenkin kannalta.

Vaikka BIBFRAMEn kehittämisessä on koettu lähteä luetteloinnin nykytilanteen tarpeista tulevaisuuden mahdolliset muutostarpeet huomioiden, voi osa MARCin historiallisesta painolastista siirtyä BIBFRAMEen. Luettelointimallina BIBFRAMEa ei rakenneta MARCin varaan, mutta jos konversiota MARCista BIBFRAMEen ei onnistuta toteuttamaan esimerkiksi tiedon linkittymistä toteuttavalla tavalla, voi hyvin suuri määrä tietueista jäädä BIBFRAME-muotoisinakin käytettävyydeltään edelleen osin samalle tasolle kuin MARC-muotoisina. Jotta MARCin historiallinen painolasti saataisiin kokonaan karistettua, tulisi konversion olla perinpohjainen.

MARC-formaattia on koetettu muokata vastaamaan lähes kaikkia tiedontallennuksen osa-alueita ja sen toiminnoista monet, esimerkiksi e-aineiston hallinnollisen metadatan tallentamiseen liittyen, ovat puutteellisia. E-aineistojen kattava kuvailu onkin mahdollista vain jos MARC-pohjaista kuvailua täydennetään hallinnollisen metadatan formaattien avulla. BIBFRAME on näiltä osin vielä keskeneräinen eikä selkeää päätöstä hallinnollisen metadatan kattavuuden suhteen ole tehty. On mahdollista, että BIBFRAME ei tule kattamaan juuri lainkaan hallinnollista metadataa ja BIBFRAMEn tueksi tarvitaan MARC-formaatin tavoin hallinnollisen metatiedon formaatteja. BIBFRAME todennäköisesti tulee kuitenkin MARCin tavoin kattamaan ainakin sen osan hallinnollisesta metadatasta, joka on tarpeellista myös AV- ja e-aineiston kuvailussa. Sekin on mahdollista, että BIBFRAMEssa päätetään pyrkiä kattamaan kaikki tarpeellinen myös hallinnollisen metadatan osalta, sillä *BIBFRAME AV Modeling Study* -raportin mukaan vaikuttaa siltä, ettei mikään yksittäinen olemassa oleva datamalli ole riittävä kaikenlaisen aineiston kuvailuun. Tällainen päätös tukisi BIBFRAMEn tavoitetta olla yhteinen luettelointiväline kaikille aineistoa luetteloiville kulttuuri- ja muistiorganisaatioille.

BIBFRAMEn kehittämisen heikkoutena on periaatteellinen epäselvyys siitä, pyrkiikö BIBFRAME mallina kattamaan kaikenlaisen luetteloinnin kaikki tarpeet vai olemaan yksinkertainen ja joustava pohja, jota voi soveltaa erilaisten luettelointitapojen tarpeisiin. Molemmille vaihtoehdoille on kannatusta ja kumpi tahansa suunta tai niiden jonkinlainen välimuoto olisi tässä vaiheessa mahdollinen. BIBFRAMEssa voidaan pyrkiä joko niin laajaan kattavuuteen, ettei sen lisäksi minkään ulkopuolisen standardin tai formaatin tuki luetteloinnissa olisi välttämätön, tai siitä voidaan pyrkiä kehittämään joustava pohja, johon muita tarvittavia standardeja ja formaatteja voitaisiin yhdistää, jolloin se toimisi myös erilaisten luettelointivälineiden ja tapojen yhteisenä alustana. Selkeää linjausta siitä, kumpaan pyritään, ei toistaiseksi ole tehty.

Tulevaisuudessa MARC-formaatista luopuminen vaikuttaa todennäköiseltä, vaikkei BIBFRAME nousisikaan yhtä laajasti käytetyksi vaihtoehdoksi kuin MARC tällä hetkellä. Jos Yhdysvaltain kongressin kirjasto ja muut MARCin ylläpitoon osallistuvat tahot siirtyvät käyttämään BIBFRAMEa, on luultavasti edessä pitkä siirtymäkausi, jonka aikana kirjastot vähitellen siirtyvät MARCista BIBFRAMEen tai johonkin muuhun luettelointiformaattiin tai -malliin. Muutosta voi hidastaa esimerkiksi nykyisten kirjastojärjestelmien modernisointi BIBFRAME-pohjaisiksi. Todennäköisesti senkin jälkeen kun Kongressin kirjasto ja muut keskeiset toimijat luetteloivat BIBFRAMElla, tiedu- eet ovat vielä pitkään saatavilla MARC-muodossa.

Ei ole kuitenkaan selvää, että BIBFRAME tulisi MARCin tavoin käyttöön useimmissa kirjastoissa. Monet kirjastot kehittävät luettelointiaan osallistumatta BIBFRAME-projektiin. Nämäkin kirjastot tuntuvat pitävän RDF-pohjaista luettelointia hyvänä ratkaisuna tiedon linkittymisen ja entiteettien välisten suhteiden ilmaisemisen kannalta. Esimerkiksi Ruotsin ja Norjan kirjastot ovat siirtymässä luettelointiin ja datan ilmaisemiseen RDF:n avulla (Kadesjö 2014, viitattu 24.7.2014; Westrum 2014, 2). Jos käytetyt luettelointisäännöt ovat yhteismitallisia, RDF-pohjaisuus tarjoaa mahdollisuuden metadatan edestakaisin toimivaan konvertointiin tällaisten paikallisten formaattien ja BIBFRAME:n, sekä myös MARC 21 -formaatin välillä. Siirtymä nykyisestä luetteloinnista mahdollisesti tapahtuu siis ainakin RDF:n jonkinlaiseen soveltamiseen datan ilmaisemisessa.

7 POHDINTA

Yhtenä opinnäytetyön tavoitteista oli kartoittaa MARC-luettelointiformaatin ongelmia kirjastoalan kirjoitusten ja keskustelujen pohjalta. Tältä osin aineisto rajautui lähinnä blogikirjoituksiin ja blogeissa käytyihin keskusteluihin, sillä tutkimuksia aiheesta ei juurikaan ole tehty ja tietämys formaatin ongelmista on pitkälti MARCin parissa työtä tekevien henkilöiden kokemuksen varassa. Kartoituksen tarkoituksena ei ole ollut olla kattava selvitys formaatin ongelmista. Kattava selvitys ongelmista vaatisi oman työnsä, tässä keskityttiin hahmottelemaan ongelmien laajuutta ja sitä, mitä formaatin puolia ne koskevat. Yksittäisiä ongelmia on nostettu esiin esimerkinomaisesti. MARC-formaatin ongelmien kartoituksen merkitys opinnäytetyön kannalta oli siis selvittää syitä MARCista luopumiseen ja uuden luettelointimallin kehittämiseen.

Toinen tavoitteista oli selvittää, millainen on MARCin seuraajaksi kehitteillä oleva BIBFRAME-malli. BIBFRAME-mallin esittely ei sen keskeneräisyyden vuoksi ole kattava eikä opinnäytetyön puitteissa ollut mahdollisuutta syventyä mallin kaikkiin valmiiksi hahmoteltuihinkin puoliin syvällisesti. Opinnäytetyön lähtökohtana oli MARCin ja BIBFRAMEn vertaaminen toisiinsa. Tämä tavoite odotetusti toteutui BIBFRAMEn keskeneräisyyden vuoksi vain MARCin ja BIBFRAMEn rakenteiden ja skeemojen osalta. Esimerkiksi BIBFRAMEn sanasto tulee todennäköisesti vielä muuttamaan, laajenemaan ja täsmentymään huomattavasti.

MARC-formaatin ongelmien kartoitus oli haastavaa, sillä ongelmista käyty keskustelu ulottuu pitkälle aikavälille ja on laajaa ja hajanaista. MARCiin kohdistunut kritiikki koskee usein vain formaatin jotakin tiettyä versiota eivätkä samat ongelmat esiinny kaikissa MARCin varianteissa. Useat ongelmista on myös päivitysten yhteydessä korjattu. Vertailussa ja johtopäätösten tekemisessä haasteena olivat BIBFRAMEn keskeneräisyys ja jo valmiisiin osiin mahdollisesti tehtävät muutokset. MARCin ongelmien hahmottamisessa ja johtopäätösten tekemisessä käytännön kokemuksesta luettelointityön parissa olisi ollut huomattavaa etua ja opinnäytetyön tekemisen aikana saatu etenkin edellä mainittuja osa-alueita koskeva palaute toimksiantajan taholta oli hyödyllistä.

Kaikkia opinnäytetyössä käsiteltyjä aiheita olisi mahdollista selvittää vielä kattavammin ja syvällisemmin. Tässä työssä näkökulma on rajattu yleisluontoiseksi, ja esimerkiksi kirjastoalan ulkopuolisten luotteloivien organisaatioiden tai erityisesti suomalaisten kirjastojen näkökulmia voisi olla myös mielekästä selvittää. BIBFRAMEa olisi lisäksi kiinnostavaa tarkastella myös UKJ (Uusi kir-

jastojärjestelmä) -hankkeen näkökulmasta. Myös BIBFRAMEn suhdetta muihin jo olemassa ole-
viin luettelointiformaatteihin ja sitä, millaisia muita formaatteja tällä hetkellä on suunnitteilla ja ke-
hitteillä, olisi kiinnostavaa kartoittaa.

LÄHTEET

Autio-Tuuli, M. 2009. RDA – uusi luettelointistandardi. Kansalliskirjasto. Tietolinja 1/2009. Viitattu 3.6.2014 ja 1.9.2014, <http://www.kansalliskirjasto.fi/kirjastoala/tietolinja/0109/rda.html>.

Autio-Tuuli, M. & Seppälä, M. 2012. Laadukas metatieto kirjastojen kansallisen metatietovarannon pohjana. Tietolinja 1/2012. Viitattu 6.6.2014 ja 12.6.2014, http://www.kansalliskirjasto.fi/kirjastoala/tietolinja/0112/laadukas_metatieto.html.

Berger, T., Soroka, A., Coyle, K., Thompson, T., Yeates, S., Sanders, S. & Ford, K. 2014. [BIBFRAME] Our darkest alley [Was: Re: [BIBFRAME] bf:Title]. Sähköpostiviestiketju. 27.7.2014.

Chowdhury, G. 2010. Introduction to modern information retrieval. London: Facet publishing.

Coyle, K. 2000. Is MARC Dead?. Viitattu 5.6.2014 ja 2.9.2014, <http://www.kcoyle.net/marcdead.html>.

Coyle, K. 2009a. Metadata mix and match. Information Standards Quarterly 21 (1), 8–11. <http://www.kcoyle.net/isqv21no1.pdf>.

Coyle, K. 2009b. May 21st, 2009 at 01:30. Kommentti blogikirjoitukseen Who needs MARC?. Viitattu 5.6.2014, <http://commonplace.net/2009/05/who-needs-marc/#comment-2199>.

Coyle, K. 2010. MARC: from mark-up to data. Viitattu 5.6.2014, <http://kcoyle.blogspot.fi/2010/03/marc-from-mark-up-to-data.html>.

Coyle, K. 2011. MARC vs RDA. Viitattu 5.6.2014, <http://kcoyle.blogspot.fi/2011/09/marc-vs-rda.html>.

Denenberg, R. 2013. BIBFRAME Annotation Model. BIBFRAME Community Draft, 26 August 2013. Viitattu 29.6.2014, <http://bibframe.org/documentation/annotations/>.

Eric Nord. 2012. 7. Kommentti artikkeliin MARC as Data: A Start. Viitattu 8.6.2014, <http://journal.code4lib.org/articles/5468#comment-5980>.

Ford, K. 2012. LC's Bibliographic Framework initiative and the Attractiveness of Linked data. *Information Standards Quarterly* 24 (2/3), 46–50. http://www.niso.org/apps/group_public/download.php/9405/SP_Ford_LC_isqv24no2-3.pdf.

Greenall, R. 2014. What does it mean to “provide support for BIBFRAME”? Viitattu 9.7.2014, <http://brinxmat.wordpress.com/2014/04/28/what-does-it-mean-to-provide-support-for-bibframe/>.

IFLA Working Group on Functional Requirements and Numbering of Authority Records (FRAN-AR). 2008. Tarkistettu ja korjattu 2013. Functional Requirements for Authority Data: A Conceptual Model. Viitattu 1.8.2014, http://www.ifla.org/files/assets/cataloguing/frad/frad_2013.pdf.

Kadesjö, M. 2014. Libris roadmap. Viitattu 24.7.2014, <http://www.kb.se/libris/Om-LIBRIS/Libris-roadmap/>.

Kansalliskirjasto. 2007a. Johdanto. MARC 21 -formaatti: auktoriteettitiedot. Viitattu 8.7.2014, <http://www.kansalliskirjasto.fi/extra/marc21/aukt/johdanto.htm>.

Kansalliskirjasto. 2007b. Johdanto. MARC 21 -formaatti: varastotiedot. Viitattu 8.7.2014, <http://www.kansalliskirjasto.fi/extra/marc21/hold/johdanto.htm>.

Kansalliskirjasto. 2008a. Yleistä tietoa yhtenäisformaateista. MARC 21 yhtenäisformaatit. Viitattu 3.6.2014, <http://www.kansalliskirjasto.fi/extra/marc21/yleista.htm>.

Kansalliskirjasto. 2008b. Johdanto. MARC 21 -formaatti: bibliografiset tiedot. Viitattu 8.7.2014, <http://www.kansalliskirjasto.fi/extra/marc21/bib/johdanto.htm>.

Kansalliskirjasto. 2014a. MARC-formaatit Suomessa. Viitattu 3.6.2014, <http://www.kansalliskirjasto.fi/kirjastoala/formaatit/marcit.html>.

Kansalliskirjasto. 2014b. MARC 21 -formaattiin siirtyminen. Viitattu 3.6.2014, <http://www.kansalliskirjasto.fi/kirjastoala/formaatit/aikataulu.html>.

Karen. 2010. On the brokenness of MARC. Viitattu 4.6.2014, <http://www.alatechsource.org/blog/2010/10/continuing-the-conversation-new-models-of-metadata.html#comment-2803>.

Koster, L. 2009. Who needs MARC?. Viitattu 5.6.2014, <http://commonplace.net/2009/05/who-needs-marc/>.

Laura K. 2013. What's wrong with MARC? (NISO BIBFRAME roadmap meeting, part 2). Viitattu 6.6.2014, <http://www.lauraek.net/2013/04/18/whats-wrong-with-marc-niso-bibframe-roadmap-meeting-part-2/>.

Library of Congress. 2011. A Bibliographic Framework for the Digital Age (October 31, 2011). Viitattu 26.6.2014, <http://www.loc.gov/bibframe/news/framework-103111.html>.

Library of Congress. 2012. Bibliographic Framework as a Web of Data: Linked Data Model and Supporting Services. Viitattu 15.6.2014, <http://www.loc.gov/bibframe/pdf/marclid-report-11-21-2012.pdf>.

Library of Congress. 2014a. BIBFRAME Frequently Asked Questions. Viitattu 24.6.2014, 3.7.2014, <http://www.loc.gov/bibframe/faqs/>.

Library of Congress. 2014b. Overview of the BIBFRAME Model. Viitattu 19.6.2014, <http://www.loc.gov/bibframe/docs/model.html>.

Library of Congress. 2014c. BIBFRAME Relationships. Viitattu 19.6.2014, <http://www.loc.gov/bibframe/docs/bibframe-relationships.html>.

Library of Congress. 2014d. Category View BIBFRAME Vocabulary. Viitattu 18.6.2014, <http://bibframe.org/vocab-category/>.

Library of Congress. 2014e. Description of the Category View of the BIBFRAME Vocabulary. Viitattu 11.7.2014, <http://www.loc.gov/bibframe/docs/vocab-category.html>.

Library of Congress. 2014f. Vocabulary. Viitattu 11.7.2014, <http://bibframe.org/vocab/>.

Library of Congress. 2014g. Model View. Viitattu 11.7.2014, <http://bibframe.org/vocab-model/>.

Library of Congress. 2014h. Category View. Viitattu 11.7.2014, <http://bibframe.org/vocab-category/>.

Library of Congress. 2014i. Category View of the BIBFRAME Vocabulary. Viitattu 22.6.2014, <http://www.loc.gov/bibframe/docs/vocab-category.html>.

The Library of Congress Working Group on the Future of Bibliographic Control. 2008. On the Record. Viitattu 12.6.2014, <http://www.loc.gov/bibliographic-future/news/lcwg-ontherecord-jan08-final.pdf>.

Marcum, D. 2011. A Bibliographic Framework for the Digital Age (October 31, 2011). Viitattu 25.6.2014, <http://www.loc.gov/bibframe/news/framework-103111.html#ftn2>.

McCallum, S. 2012. RDA in MARC. Viitattu 5.6.2014, <http://www.loc.gov/marc/RDAinMARC.html>.

McCallum, S. 2014. BIBFRAME and Cataloging: changing landscapes. Viitattu 9.7.2014, http://www.jbi.hio.no/bibin/korg_dagene/korg2014/20140519_mccallum_sally.pdf.

Miller, E., Mueller, V., Ogbuji, U. & Baker, M. 2014. BIBFRAME Profiles: Introduction and Specification. Additional editing by Kevin Ford, Library of Congress. Viitattu 19.6.2014, <http://www.loc.gov/bibframe/docs/bibframe-profiles.html>.

Miller, E., Mueller, V., Ogbuji, U. & MacDougall, K. 2013. BIBFRAME Use Cases and Requirements. Viitattu 27.6.2014, <http://bibframe.org/documentation/bibframe-usecases/>.

Moen, W. & Benardino, P. 2003. Assessing Metadata Utilization: An Analysis of MARC Content Designation Use. Viitattu 9.6.2014, http://www.unt.edu/wmoen/publications/MARCPaper_Final2003.pdf.

National Information Standards Organization. 2004. Understanding metadata. <http://www.niso.org/publications/press/UnderstandingMetadata.pdf>.

Rowley, J. & Hartley, R. 2008. Organizing Knowledge. Aldershot: Ashgate.

Sanderson, R., Coyle, K. & Ford, K. 2014. [BIBFRAME] bf:Title queries from use cases. Sähköpostiviestiketju. 28.7.2014.

Smith-Yoshimura, K. 2010. MARC Tag Usage in WorldCat. Teoksessa Karen Smith-Yoshimura et al. Implications of MARC Tag Usage on Library Metadata Practices. Viitattu 13.6.2014, <http://www.oclc.org/content/dam/research/publications/library/2010/2010-06.pdf?urlm=162940>.

Smith-Yoshimura, K., Rowlison de Ortiz, L. & Dickey, T. 2010. Executive Summary: Implications of MARC Tag Usage on Library Metadata Practices. Teoksessa Karen Smith-Yoshimura et al. Implications of MARC Tag Usage on Library Metadata Practices. Viitattu 9.6.2014, <http://www.oclc.org/content/dam/research/publications/library/2010/2010-06.pdf?urlm=162940>.

Suominen, V., Saarti, J. & Tuomi, P. 2009. Bibliografinen valvonta. Helsinki: BTJ Kustannus.

Tauberer, J. 2008. What is RDF and what is it good for?. Viitattu 12.6.2014, <https://github.com/JoshData/rdfabout/blob/gh-pages/intro-to-rdf.md>.

Tennant, R. 2004. A Bibliographic Metadata Infrastructure for the 21st Century. Author's version, published in Library Hi Tech, vol 22 (2) 2004, pp.175-181. Viitattu 5.6.2014, <http://roytennant.com/metadata.pdf>.

Thomale, J. 2010. Interpreting MARC: Where's the Bibliographic Data?. Code{4}lib 2014 (11). Viitattu 8.6.2014, <http://journal.code4lib.org/articles/3832>.

Tillett, B. 2005. Mikä on FRBR?. Viitattu 3.6.2014, <http://www.loc.gov/catdir/cpsol/Mika-on-FRBR.pdf>.

Van Malssen, K. 2014. BIBFRAME AV Modeling Study: Defining a Flexible Model for Description of Audiovisual Resources. Viitattu 28.7.2014, <http://www.loc.gov/bibframe/pdf/bibframe-avmodelingstudy-may15-2014.pdf>.

Westrum, A. 2014. The library catalogue of the future. *Scandinavian Library Quarterly* 47 (2), <http://slq.nu/?article=volume-47-no-2-2014-6>.

Zeng, M., Žumer, M. & Salaba, A. (toim.). 2010. Functional Requirements for Subject Authority Data (FRSAD): A Conceptual Model. Viitattu 1.8.2014, <http://www.ifla.org/files/assets/classification-and-indexing/functional-requirements-for-subject-authority-data/frsad-final-report.pdf>.