

Opinnäytetyö (AMK)

Kirjasto- ja tietopalvelun koulutusohjelma

2015

Elina Meriö

OIKEAT SANAT KUNNALLE KIRJOJA

– Turun kaupunginkirjaston vanhojen digitoitujen
kirjojen asiansanoittamisprojekti



TURUN AMMATTIKORKEAKOULU
TURKU UNIVERSITY OF APPLIED SCIENCES

OPINNÄYTETYÖ (AMK) | TIIVISTELMÄ

TURUN AMMATTIKORKEAKOULU

Kirjasto- ja tietopalvelun koulutusohjelma

Toukokuu 2015 | 61 + 9 sivua

Ohjaaja Olli Mäkinen

Elina Meriö

OIKEAT SANAT KUVAILEMAAN KIRJOJA – TURUN KAUPUNGINKIRJASTON VANHOJEN DIGITOITUJEN KIRJOJEN ASIASANOITTAMISPROJEKTI

Tämä opinnäytetyö kertoo niiden 1800-luvulla painettujen teosten asiasanoittamisesta, jotka Turun kaupunginkirjasto on asettanut saataville yleisten kirjastojen digitoidun aineiston kokotekstitietokantaan Digiin. Opinnäytetyön tekijä valitsi asiasanat 63 teokselle ja raportoi tätä työtään. Opinnäytetyöprojektissa asiasanoitetuista teoksista jokainen sai keskimäärin 8-13 asiasanaa.

Kirjastoaineiston sisällönkuvailuun kuuluvassa asiasanojen valinnassa on kyse tekstien sisällön tiivistämisestä termeihin. Termit ovat asiasanoituksessa käytettäviksi vakiintuneissa muodoissa olevia, viralliseen käyttöön valittuja sanoja tai usean sanan muodostamia termejä.

Asiasanat valitaan asiasanastoista. Tässä opinnäytetyössä asiasanoituksessa on käytetty Yleistä suomalaista asiasanastoa suomenkielisten tietokirjojen asiasanoittamisessa ja Verkko-Kaunokkia suomenkielisen kaunokirjallisuuden asiasanoittamisessa. Nämä asiasanastot näyttävät myös ruotsinkielisiin vastineisiinsa Allärsiin ja Bellaan kuuluvat termit, jotka ovat tarpeellisia antaessa ruotsinkielisille kirjoille asiasanat molemmilla kotimaisilla kielillä. Suurin osa asiasanoitettavista teoksista on suomenkielisiä.

Asiasanoituksen tekemisessä lähtökohtana tulee olla kuvailtavan aineiston käyttäjien huomioon ottaminen. Tietokannan tiedonhakutoimintojen käyttäjät saavat hakutuloksiksi dokumentteja, joiden sisältöä indeksoija on kuvailut samoilla asiasanoilla joita tiedonhakuun sisältyy. Asiasanat tarjoavat hakutuloksia silmäileville tiedonhakijoille mahdollisuuden saada nopeasti käsitys teosten olennaisimmista sisällöistä.

Asiasanoittamisen peruseriaatteita ovat kokoelman sisäinen yhtenäisyys asiasanoituksessa ja termien käytön johdonmukaisuus. Tavallisimpia pohdittavia kysymyksiä asiasanoja valitessa ovat, onko asiasanoja liikaa yhdelle kirjalle ja onko sisällönkuvailun spesifiyden tai pelkistämisen aste sopiva. Vanhan aineiston ominaispiirteisiin kuuluvat takakansitekstien ja toisinaan sisällysluettelojenkin puuttuminen, kirjoissa käytetyt vanhat sanat ja suomennetut nimet sekä tieto- ja kaunokirjallisuuden rajan epäselvyys joissakin kirjoissa.

ASIASANAT:

Asiasanoitus, sisällönkuvailu, vanhat kirjat, kirjastot

BACHELOR'S THESIS | ABSTRACT

TURKU UNIVERSITY OF APPLIED SCIENCES

Degree Programme in Library and Information Services

2015 | 61 + 9 pages

Instructor Olli Mäkinen

Elina Meriö

RIGHT WORDS FOR DESCRIBING BOOKS – THE INDEXING PROJECT OF THE OLD DIGITALIZED BOOKS AT TURKU CITY LIBRARY

The topic of my thesis is indexing of digitalized books which Turku City Library has set to a full-text database Digi where the books are available for customers. The author of the thesis selected indexing terms for 63 books and reported the process. Every book got average 8-13 indexing terms.

Indexing belongs to describing of library material. Indexing means the content of books have to condense into a few terms. Terms are official words or phrases which have accepted by the authors of thesauri and are used in certain forms.

Controlled vocabularies are tools for selecting indexing terms. Vocabularies used in this indexing project are General Finnish Controlled Vocabulary for non-fiction and Kaunokki for fiction. Most of the books in the digitalized collection are in Finnish, but some are in Swedish language. Swedish controlled vocabularies are Allärs for non-fiction and Bella for fiction. Books in Swedish got indexing terms in both languages.

The starting point of indexing is to think the information needs of users. Indexing is useful because it offers possibilities to do keyword searches with a database and also to browse which themes are included in a book. The main principles in indexing are an internal cohesion in the indexing terms of a collection and consistency how to use terms. Most common issues are if there are too many indexing terms for a book and if describing is enough detailed or simplified.

In old books there are not texts in the covers nor tables of contents in every non-fiction book. Authors in 19th century have used different terms than at present. The Finnish translators of old books are translated some foreigner names to Finnish. In some cases one has to draw a conclusion if an old book is fiction or non-fiction and it is not absolutely clear when one reads the first page of a book.

KEYWORDS:

Indexing, keywords, old books, libraries

SISÄLTÖ

SANASTO	6
1. JOHDANTO	7
1.1 Opinnäytetyöprojekti ja sen tavoitteet	7
1.2 Sisällönkuvailu viitekehyksenä	8
2. TURUN KAUPUNGINKIRJASTON VANHAT KIRJAKOKOELMAT	11
2.1 Turun kansankirjaston kokoelman historiaa	11
2.2 Turun kaupunginkirjaston vanhan kokoelman tutkimushanke	12
3. KIRJOJEN SISÄLLÖNKUVAILUN TEORIAA	14
3.1 Kuvailu asiasanoilla ja luokituksilla	14
3.2 Tiedonhaun tulosten arviointi	16
3.3 Asiasanastojen historiaa	18
3.4 Asiasanoittaminen valintoina ja vallankäyttönä	19
3.5 Hyvän indeksoinnin ominaisuuksia	21
3.6. Asiasanojen valitseminen	23
3.7 Asiasanojen tyypit ja muoto	24
3.8 Termien suhteet toisiinsa	27
4 DIGITOIDUN KOKOELMAN ASIASANOITUSTYÖ	31
4.1 Asiasanoitettava digitaalinen kokoelma	31
4.2 Työskentelyperiaatteet asiasanoittamisessa	33
4.2.1 Kaunokirjallisuuden asiasanoittaminen	35
4.3 Ongelmia ja ratkaisuja	35
4.3.1 Sisällön pelkistäminen	35

4.3.2 Sanojen merkitykset	37
4.3.3 Väliotsikot ja vanhat sanat	39
4.3.4 Henkilöiden nimet asiasanoina	42
4.3.5 Kauno- ja tietokirjallisuuden rajanveto	44
4.4 Esimerkkejä kirjojen asiasanoittamisesta	44
4.4.1. Pitkä matkakertomus	44
4.4.2. Monta ihmiskohtaloa lyhyessä romaanissa	48
5 LOPUKSI	51
5.1 Asiasanaprojektin lopputulos	51
5.2 Asiasanastojen käyttö vanhojen kirjojen kuvailuun	53
5.3 Oppimista ja haasteita	54
LÄHTEET	56

LIITTEET

Liite 1. Lista digitoiduista ja asiasanoitetuista Turun kansankirjaston kokoelman kirjoista

KUVAT

Kuva 1. Ensimmäinen tekstisivu kirjasta *Kongo – Uusi vapaavaltio mustien maanosassa*. 2. osa. (Stanley 1887, 1).

Kuva 2. Henkilöiden ja aiheiden nopeaa vaihtelua teoksessa *Koivuluoto ja Pohjakoski*. (Koivuluoto ja Pohjakoski 1861, 8-9).

SANASTO

termi / lyhenne

termin / lyhenteen selitys (Lähdeviite)

asiasanasto

Asiasanastoon kuuluu virallisina asiasanoina käytettäviä siihen tarkoitukseen hyväksytyjä termejä. Ilmaisujen käyttökelpoisuus asiasanoina tarkistetaan asiasanastoista. (Alaterä & Halttunen 2002, 77.)

sisällönkuvailu

Sisällönkuvailu esittää tiivistetyn kuvauksen dokumentin asiasisällöstä. Keinoja tähän ovat asiasanat, luokituskaavan koodit sekä tiivistelmät. (Alaterä & Halttunen 2002, 68–69, 78–79; Suominen, Saarti & Tuomi 2009, 189–190.)

tesaurus

Tesaurus on kontrolloidun indeksointikielen sanasto, joka tekee termien väliset suhteet näkyviksi ja jota käytetään tiedonhaussa ja indeksoinnissa. (Aitchison, Gilchrist ja Bawden 2004, 1.)

YSA

Yleinen suomalainen asiasanasto (Kansalliskirjasto 2015a. Finto - Suomalainen asiasanasto- ja ontologiapalvelu.)

1 JOHDANTO

1.1 Opinnäytetyöprojekti ja sen tavoitteet

Turun kaupunginkirjaston kokoelmiin kuuluu omaksi kokoelmakseen erotettu Turun kansankirjaston kokoelma, joka koostuu 1800-luvulla painetuista, kansankirjaston ajasta asti kokoelmissa säilyneistä kirjoista. Tuon kokoelman teoksista on digitoitu yleisön saataville 63 kappaletta, jotka ovat internetissä yleisten kirjastojen yhteisessä Digi-kokotekstitietokannassa (Kirjastot.fi 2015a. Digi – yleisten kirjastojen digitoitut aineistot). Tämän opinnäytteen aiheena on Turun kansankirjaston kokoelman digitoitujen teosten asiasanoittamisprojekti.

Turun kaupunginkirjasto on toimeksiantajana tässä opinnäytteessä, jonka sisältönä on teoreettista taustaa asiasanoittamisesta ja asiasanoittamisprojektin raportointi. Kolmas luku sisältää teoriataustaa ja neljäs luku raportointia. Työn empiiriseen osaan kuuluu indeksoinnin lisäksi vanhan kirjallisuuden tietoja tallentaneen projektitutkijan haastattelu, jonka tuloksia on luvussa 2.2. Projektitutkija on työskennellyt kaupunginkirjaston vanhan kirjakokoelman tutkimusprojektissa, johon sisältyy yliopiston jatko-opiskelijoiden tekemää kirjaston vanhimman aineiston kuvailutietojen tallentamista tietokantaan.

Tätä opinnäytetyötä tarkastellaan projektityön kaavan mukaisesti. Kaupunginkirjaston ammattilaisten ja opinnäytetyön tekijän kirjaston tarpeista käymien keskustelujen tuloksena valittiin aihe laajuudeltaan rajatulle projektille, jonka lopputulokseksi tulee lista asiasanoja jokaiselle digitoitulle kirjalle. Vakiintuneet projektityön vaiheet ovat toimeksianto, suunnittelu, resursointi, toteutus, raportointi ja arviointi. Suunnittelu ja resursointi tarkoittavat tässä projektissa ajankäytön suunnittelua ja ajan varaamista työskentelylle. Suunnitteluun kuuluu myös perehtyminen asiasanoittamista koskevaan teoriaan ennen toteutusvaihetta. Toteutus ja raportointi toteutuvat osittain limittäisinä työskentelyvaiheina, mutta raportti saa jäsennellyn ja viimeistellyn muotonsa vasta toteutuksen ollessa niin pitkällä, että opinnäytteen tekijä voi jäsentää asiasanoja antaessa tekemiään havaintoja ja pohdintoja paremmin kuin alkuvaiheessa.

Opinnäytteessä raportoidun työn tavoitteena on asiasanojen avulla helpottaa digitoitujen kirjojen löytyvyyttä tiedonhakijoille sekä selventää teosten sisältöä digitaalisten kirjojen tietokannan selaajille. Jos tiedonhakija kirjoittaa Digi-tietokannan hakuruutuun sellaisen hakusanan, joka kuuluu viralliseen asiasanastoon ja jota on käytetty tietokannassa olevien kokotekstien asiasanoittamiseen, tietokanta antaa hakutuloksiksi ne dokumentit, joita asiasanoittaja on kuvaillut kyseisellä sanalla. Kun tietokannan käyttäjä avaa sivun, jolla on teoksen metatietoja, hän voi asiasanoja lukemalla saada käsitystä teoksen sisällöstä ja päätellä, kiinnostako häntä lukea teosta ja kertooko se sellaisista aiheista, joista hän haluaa lukea. Tavoitteiden toteutumisen ensimmäinen vaihe on, että kirjakokoelma tulee asiasanoitetuksi. Toinen vaihe on, että Turun kaupunginkirjasto on tyytyväinen projektin tuloksiin. Asiasanoituksia on tarvittaessa muokattava vastaamaan paremmin toimeksiantajan toiveita. Lopullisesti projektin tavoitteet toteutuvat asiasanoituksen hyödyttäessä Digi-tietokannan käyttäjiä.

1.2 Sisällönkuvailu viitekehyksenä

Asiasanoittaminen eli indeksointi kuuluu kirjastoaineiston sisällönkuvailuun, jonka tarkoituksena on tiivistetyn kuvauksen esittäminen dokumentin asiasisällöstä. Tätä kuvailua voidaan tehdä dokumentaatiokielten eli luokitusten ja indeksoinnin avulla ja kirjoittamalla tiivistelmiä luonnollisella kielellä. Sisällönkuvailussa voidaan käyttää rinnakkain luokitusta, asiasanoja ja tiivistelmiä. Asiasanoittamista muistuttaa luonnollisen kielen menetelmä avainsanoittaminen, mutta avainsanoja ei valita asiasanaston kaltaisesta listasta, vaan ne poimitaan dokumentista. (Alaterä & Halttunen 2002, 68–69, 78–79; Vakkari 2006, 22; Suominen, Saarti & Tuomi 2009, 189–190.) Avainsanat voivat olla vakiintumattomia eikä niiden joukossa ole sellaisia sanoja, jotka eivät esiinny suoraan tekstissä. Näiden puutteiden takia varsinainen indeksointi on hyödyllistä.

Sisällönkuvailu palvelee tiedonhakijoita, jotka etsivät aineistoa jostakin aiheesta tietämättä hakuunsa sopivien kirjoittajien ja teosten nimiä. Kuvailutiedot myös

auttavat tietokannan käyttäjiä arvioimaan, mitkä hakutulosten listassa olevat dokumentit kiinnostavat heitä. Aineiston sijaintipaikkoja näyttävien viitetietokantojen lisäksi sisällönkuvailua käytetään erittäin usein myös kokotekstitietokannoissa, jollainen myös Digi on. Kokotekstitietokannoista hakija saa luettavakseen dokumentteja, joiden tekstimassaan on useissa tapauksissa mahdollista tehdä tiedonhakuja. Silloinkin asiasanat ja luokitus tukevat tiedonhakijoita ja mahdollistavat sisällön nopeaa arvioimista. Sisällönkuvailu kokotekstitietokannoissa vähentää kielen monimuotoisuuden ja liian laajojen, epätarkkojen hakutulosten aiheuttamia vaikeuksia. (Alaterä & Halttunen 2002, 68.) Tiedonhaku sisältää sekä hakukomentoja että selailua eivätkä ne ole toistensa vastakohtat (Suominen, Saarti & Tuomi 2009, 187).

Tietokantojen sisältämät tietueet ovat dokumenttien korvikkeita, jotka sisältävät informaatiota dokumenteista (Olson 2002, 2–3). Kokotekstitietokannoissa koko dokumentti on helposti saatavilla, mutta sen esittelysivulla on kuvailevia tietoja, joita voi silmäillä helpommin kuin lukea koko dokumentin. Dokumenttien sisällönkuvailu voi parhaimmillaan lisätä käyttäjän kiinnostusta niiden lukemiseen.

Tässä opinnäytteessä on lähteinä sisällönkuvailua käsittelevää ammattikirjallisuutta. Aiheesta on kirjoitettu niukasti uusia suomenkielisiä kirjoja. Hyödyllinen perusteos on Vesa Suomisen, Jarmo Saartin ja Pirjo Tuomin kirjoittama *Bibliografinen valvonta. Johdatus luetteloinnin ja sisällönkuvailun menetelmiin*. Muita lähteitä ovat esimerkiksi Anu Alaterän ja Kai Halttusen *Tiedonhaun perusteet – osa lukutaitoa*, Heidi Suonuutin *Sanastotyön opas* ja G. G. Chowdhuryn *Introduction to Modern Information Retrieval*. Monet sisällönkuvailusta kirjoitetut teokset on suunnattu asiasanastojen tekijöille eikä niiden käyttäjille. Useissa teoksissa kerrotaan suunnilleen samoja asioita muun muassa siitä, miksi tesauksia tarvitaan ja millaisia asiasanojen pitää olla.

Kirjaston kokoelmiin kuuluvia dokumentteja ei kuvailla millä tahansa sanoilla, vaan hyväksyttävien termien pitää kuulua asiasanoittavan organisaation käyttöönsä valitsemiin asiasanastoihin. Suomessa yleiset kirjastot ja suuri osa tieteel-

lisistä kirjastoista käyttää dokumenttiansa asiasanojen valinnassa Yleistä suomalaista asiasanastoa eli YSA:a (Alaterä & Halttunen 2002, 77). YSA on olennainen väline myös Turun kaupunginkirjaston digitoitujen teosten asiasanoittamisessa.

Yleinen suomalainen asiasanasto on yksikielinen, kaikki tieteen- ja tiedonalat kattava yleissanasto, jonka Kansalliskirjasto on julkaissut käytettäväksi useiden aineistotyyppien asiasanoittamiseen ja tiedonhakuun. Lisäksi sen tarkoituksena on toimia perussanastona eri alojen erikoissanastoja laadittaessa. YSA:sta voi etsiä johonkin tarkoitukseen sopivia sanoja kolmella tavalla. Käyttäjä voi hakea haluaansa sanaa hakukentän avulla tai selata aakkosjärjestyksessä olevaa termien listaa. Kolmas vaihtoehto on käyttää aiheenmukaista hakemistoa, joka sisältää kirjastoluokituksessa olevista aiheista termilistat, joissa ovat valitun aihepiirin termit aakkosjärjestyksessä. YSA:an kuuluu runsaat 32 600 päätermiä. (Kansalliskirjasto 2015a. Finto – Suomalainen asiasanasto- ja ontologiapalvelu.)

Turun kaupunginkirjasto valitsi projektissa käytettävät asiasanastot. Suomenkielisen tietoaineiston asiasanoittamiseen käytettiin Yleistä suomalaista asiasanastoa ja suomenkielisen kaunokirjallisen aineiston asiasanoittamiseen Verkko-Kaunokkia. Ruotsinkielisen tietoaineiston asiasanat valittiin Allärs-asiasanastosta ja ruotsinkielisen kaunokirjallisuuden Bellasta. YSA:ssa näkyvät Allärsiin kuuluvat vastineet YSA:ssa oleville asiasanoille ja Verkko-Kaunokissa myös Bellaan kuuluvat sanat.

2 TURUN KAUPUNGINKIRJASTON VANHAT KIRJAKOKOELMAT

2.1 Turun kansankirjaston kokoelman historiaa

Turun kansankirjasto, jonne digitoitujen kirjojen alkuperäiskappaleet hankittiin, aloitti toimintansa vuoden 1863 alussa Alfthánin talossa Jokikadun varrella. Kansankirjasto perustettiin kansanvalistuksen ihanteiden mukaan rahvaan sivistämiseksi. (Turun kaupunki 2013.)

1800-luvun jälkimmäinen puolisko oli tärkeä kirjastojen kehittymiselle Suomessa, ja varsinaisten kansankirjastojen perustaminen alkoi maassamme vuoden 1845 tienoilla. Turkuun vuonna 1863 perustettu kansankirjasto oli avoinna aluksi vain kaksi tuntia viikossa, ja kirjaston sijaintipaikka vaihteli tiheästi; se oli milloin keuhkin henkilön talossa tai koulun huoneistossa. Lainat olivat maksuttomia tai panttia vastaan saatavia. Köyhyys ja nälkävuodet haittasivat kirjaston hankintojen tekemistä ja toimintaa, mikä johti kansankirjaston johtokunnan vuonna 1870 tekemään päätökseen lopettaa kirjaston toiminta muutamaksi vuodeksi. (Seppälä 1963, 57–74.) Vuonna 1878 Turun kaupunki päätti kunnallistaa kansankirjaston (Seppälä 1963, 80).

Turun kaupungin nykyisen pääkirjaston vanhassa osassa alkoi kirjastotoiminta vuonna 1903. Ensimmäisessä kerroksessa oli rahvaalle suunnattu kansankirjasto ja ylemmässä kerroksessa tieteellinen kaupunginkirjasto sivistyneistöä varten. Vuonna 1912 Turun kaupunginvaltuusto teki päätöksen kansankirjaston ja Turun kaupunginkirjaston yhdistämisestä Turun kaupunginkirjastoksi, minkä jälkeen kirjastorakennusta uudistettiin ennen kuin se avattiin yleisölle uudestaan vuonna 1914. Silti kesti kauan ennen kuin kansankirjasto ja tieteellinen kirjasto todellisuudessa yhdistyivät, eikä vanhan kaupunginkirjaston kokoelmaa koskaan liitetty osaksi uuden kaupunginkirjaston kokoelmaa. (Näättsaari 2012, 75–79.)

Turun kansankirjaston luetteloita on säilynyt 1870-luvulta asti. Niiden perusteella suurin osa kokoelmasta oli käännöksiä ja uskonnollista kirjallisuutta. Yleisiä kirjojen aihepiirejä olivat myös historia, maantiede, matkakertomukset, elämäkerrat, luonnontiede ja terveydenhoito. Myös kaunokirjallisuutta ja lastenkirjoja kuului kokoelmaan, joka noudatti pääpiirteittäin Viipurin Suomalaisen Kirjastoseuran tekemää suosituslistaa kirjastoille. (Seppälä 1963, 77.)

2.2 Turun kaupunginkirjaston Vanhan kokoelman tutkimushanke

Turun kaupunginkirjastolla on kolme erillistä vanhojen kirjojen kokoelmaa. Nämä ovat Turun kansankirjaston kokoelma, ruotsinkielinen näytelmäkokoelma *Svensk dramatik* ja niin sanottu Vanha kokoelma eli kansankirjaston kanssa samaan aikaan toimineen, aikoinaan sivistyneistölle tarkoitettun kaupunginkirjaston kokoelma. Kansankirjaston ja vanhan kaupunginkirjaston kokoelmat eivät nykyäänkään ole samoissa tiloissa niin kuin eivät olleet hankkimisaikanaankaan. (Yli-Seppälä 2015.)

Turun kaupunginkirjasto ja Turun yliopiston Suomen historian ja yleisen historian oppiaineet ovat tehneet yhteistyötä Turun kaupunginkirjaston vanhan kokoelman tutkimushankkeessa. Tämä Varsinais-Suomen kulttuurirahaston rahoittama ja merkittävänä pitämä projekti alkoi vuonna 2013 ja päättyi helmikuussa 2015. Hankkeessa on tutkittu Vanhan kokoelman perustamista, sisältöä, järjestämistapaa ja hankintaperusteita. Vanhan kokoelman tunnettuutta on lisätty luentosarjalla ja Opintokokoelma-blogilla sekä tekemällä kokoelmasta kertova julkaisu, joka ilmestyy joulukuussa 2015. (Yli-Seppälä 2015.)

Tutkimushankkeessa on luetteloitu 9 036 Vanhaan kokoelmaan ja *Svensk dramatik* –kokoelmaan kuuluvan teoksen viitetiedot erilliseen tietokantaan (joka on vanhakokoelma.kirjastot.fi), muttei asiasanoitettu. Luettelointia ovat tehneet kaksi pitkällä opinnoissaan olevaan historian opiskelijaa yliopistosta, kumpikin eri aikaan kahdeksan kuukautta. (Yli-Seppälä 2015.) Tässä opinnäytteessä tutkimushankkeesta kertovana lähteenä on vanhaa kokoelmaa luetteloineen projektitutkijan, Suomen historian jatko-opiskelija Laura Yli-Seppälän haastattelu.

Luetteloinnissa merkittiin selvästi kirjojen perustiedot eli nimeke, tekijä, julkaisu-vuosi, kustantaja, kustannuspaikka, kieli sekä mistä kirja on saatu tai hankittu kirjastoon. Kaupunginkirjaston Vanha kokoelma sisältää kirjoja monella kielellä, joista tavallisimmat ovat ruotsi, suomi, saksa ja ranska, mutta joukossa on teoksia myös englanniksi, tanskaksi, norjaksi, viroksi, kreikaksi, venäjäksi, italiaksi ja espanjaksi. Tutkimushankkeessa luetteloitujen teosten yleisin kieli on ruotsi. Sak-sankielistä aineistoakin kuuluu luetteloitujen joukkoon enemmän kuin suomen-kielistä. (Yli-Seppälä 2015.)

Hankkeessa tutkitun kokoelman kirjoista suurin osa on 1800-luvulta tai 1900-lu-vun alusta, mutta vanhimmat lahjoitukset ovat 1500-luvulta ja kokoelman uusim-mat niteet 1930-luvulta. Luetteloiduista teoksista noin 39 % on painettu vuosina 1900–1930, noin 44 % 1800-luvulla, noin 12 % 1700-luvulla, noin 0,5 % 1600-luvulla ja noin 0,07 % 1500-luvulla. Laskelmista puuttuvat ne teokset, joihin ei ole merkitty lainkaan painovuotta. (Yli-Seppälä 2015.)

Vanhan kokoelman tutkimushankkeen tärkein kohderyhmä ovat eri alojen tutkijat, sillä Vanhan kokoelman teoksista kiinnostuneet saavat lukea teoksia ainoastaan lukusalilainoina kirjaston kellarivarastossa. Hankkeen tuottamat materiaalit kuten blogi ja kirja sopivat myös kirjojen ja kirjastojen historiasta kiinnostuneille ihmi-sille. (Yli-Seppälä 2015.)

Yli-Seppälän mukaan haastavinta projektiin osallistumisessa ovat olleet vaikea-selkoisten käsialojen lukeminen ja puutteelliset tiedot kirjojen alkuperästä. Koko-elman teosten provenienssia eli alkuperää ja omistushistoriaa on pyritty projek-tissa selvittämään kirjoihin käsin kirjoitetuista omistajamerkinnoista ja omistuskir-joituksista sekä lahjoitusasiakirjoista, joita tehtiin kun ihmiset lahjoittivat kirjastolle kirjoja. (Yli-Seppälä 2015.)

Historiantutkijan taidot ovat hyödyksi Vanhan kokoelman tutkimisessa ja luette-loinnissa, jossa on tarpeellista osata lukea nopeasti frakturekstiä ja vanhoja käsialoja. Taustatietojen selvittämiseen tarvitaan tiedonhakutaitoja, sillä projek-tissa on tutkittu muun muassa sitä, keitä kirjojen entisiksi omistajiksi ja lahjoitta-jiksi merkityt henkilöt olivat. (Yli-Seppälä 2015.)

3 KIRJOJEN SISÄLLÖNKUVAILUN TEORIAA

3.1 Kuvailu asiasanoilla ja luokituksilla

Tiedonhaussa hyödynnetään luetteloinnin ja sisällönkuvailun tuottamia viitetietoja. Luetteloinnissa merkittävää metadataa ovat muun muassa dokumentin nimi, tekijä sekä julkaisuaika ja -paikka. Internetissä olevien dokumenttien metadataan kuuluu merkitä, mikä organisaatio tai kuka henkilö on laittanut dokumentin internetiin ja minkä kokoinen dokumentti on. Asiasanat ovat metadatan osa, joka kuvailee dokumentin sisältöä eikä muodollisia ominaisuuksia ja tekee dokumentin löydettäväksi asiasanahauilla. Indeksoinnissa dokumentin sisällön olennaisimmat asiat kuvataan asiasanoilla eli termeillä, jotka jakautuvat erisnimiin ja muihin sanoihin. Monikolliset yleisnimet ovat kategorioita, joihin sisältyy erilaisia kategorian edustajia. (Hedman 2000, 31; Vakkari 2006, 22.)

Sisällönkuvailuun kuuluvat luokitukset ovat kirjastojen hyllypaikkamerkintöjä, jotka ilmaisevat mihin ryhmään teos kuuluu sisällöltään. Luokituskielten numeroja kirjainkoodeilla on mahdollista merkitä teosten aiheet hyvinkin tarkasti esimerkiksi desimaalien avulla.

Vakkari näkee luokittamisen enemmän indeksointia muistuttavana kuin Alaterä ja Halttunen. Hänen mukaansa luokituskaavan avulla valitut luokkatunnukset viittaavat käsitteisiin kuten asiasanat (Vakkari 2006, 22). Toisaalta indeksoija kuvaillee dokumentteja niistä etsimiensä keskeisten elementtien avulla toisin kuin luokittaja, jonka tehtävänä on sijoittaa dokumentti sille sopivaan luokkaan (Alaterä & Halttunen 2002, 80).

Asiasanoittaessa on tyypillistä antaa dokumentille useita kuvailutermejä. Luokituksiakin on periaatteessa mahdollista olla useita samalle dokumentille, mutta niitä ei perinteisesti ole annettu yhtä enempää jo kirjojen hyllyjärjestyksen selkeyden takia. Usean luokan antaminen samalle dokumentille sopii virtuaaliseen kirjastoon paremmin kuin perinteiseen. Asiasisällöt on mahdollista esittää joko aak-

kosellisena asiasanalistana tai aiheiden mukaan loogiseen järjestykseen kootuina, ikään kuin käsitekarttana. Indeksoinnissa eli asiasanoituksessa tuodaan esiin samoja näkökulmia kuin dokumentin luokituksessakin. (Suominen, Saarti & Tuomi 2009, 180–185.)

Indeksointi ja luokitus ovat dokumentaatiokieliä (englanniksi *controlled vocabulary*), jotka on kehitetty dokumenttien kuvailuun tiedonhaku ympäristöjä varten. Niiden tarkoitus on ratkaista luonnollisen kielen monimuotoisuudesta tiedonhauille aiheutuvia vaikeuksia, mutta dokumentaatiokielet pohjautuvat silti luonnolliseen kieleen, josta asiasanastojen sanat on valikoitu. Siksi tiedonhakijoiden on todennäköisesti helpompaa ymmärtää asiasanastoja kuin numeroilla merkityjä luokituksia. Tesauruksille eli termien suhteita ilmaiseville asiasanastoille on ominaista vahva sidos kieleen, jolla ne on laadittu. Siten tesaurukset ovat vähemmän soveltuvia kansainvälisiksi kuin luokituskielet, mutta monikielisiäkin tesauruksia on olemassa. Dokumentaatiokielten kontrolloiminen tarkoittaa, että jokin ryhmä tai organisaatio kehittää luokitus- ja indeksointikieliä ja tiedottaa muutoksista käyttäjille. (Alaterä & Halttunen 2002, 71, 75, 80.)

Sanaston määrittely (*vocabulary control*) on yksi tärkeimmistä tiedonhakua mahdollistavista asioista. Tiedonhaku järjestelmä pyrkii yhdistämään kävijän hakuruutuun kirjoittaman sanallisen kyselyn varastoitujen dokumenttien tietoihin. Jotta tietueet löytyisivät hakutermien avulla, niissä molemmissa tulee käyttää samoja termejä. Siitä johtuu standardoidun sanaston tarve. (Chowdhury 2010, 155.)

Dokumenttien aiheen analysoinnin tarkoituksena on luoda mielekäs tapa saada käyttöönsä informaatiota tiedonhakutyökalujen avulla ja säästää käyttäjän aikaa. Tietokantahaun tavoitteena on saada kattavasti tietoon kokoelman kaikki dokumentit kyseisestä aiheesta. (Taylor 2004, 241.) Aiheiden mukainen järjestys vaihtelee hakukomentojen mukaan; tekemällä tiedonhaun saa tulokseksi listan dokumentteja, joita hakukriteeri yhdistää.

3.2 Tiedonhaun tulosten arviointi

Tiedonhakua ja tallennusta evaluoidaan eli arvioidaan, jotta tiedon löytyvyyttä opittaisiin parantamaan. Tätä evaluointia tehdään perustutkimusta suorittavien laitosten lisäksi kirjastoissa ja tietopalveluissa. Indeksointia evaluoidessa tutkitaan muun muassa indeksoijien välisiä eroavaisuuksia ja indeksoinnin yksityiskohtaisuutta, spesifisyyttä, tyhjentyvyyttä, johdonmukaisuutta ja oikeellisuutta. (Alaterä & Halttunen 2002, 123–124.)

Spesifisyys tarkoittaa, miten hyvin dokumentissa esiintynyt käsite saa vastaavuuden kuvailussa. Jos dokumentti käsittelee esimerkiksi opiskelumotivaatiota, spesifi asiasana on 'opiskelumotivaatio', mutta vähemmän spesifissä indeksoinnissa voidaan asiasanaksi valita 'motivaatio'. Aiheeltaan yleisluonteisissa asiasanoissa riittävät vähemmän spesifit termit kuin aihekohtaisissa erikoissanastoissa. Indeksoinnin tyhjentyvyys on sitä suurempi, mitä enemmän asiasanoja valitaan kuvaamaan yhtä dokumenttia. Johdonmukaisuuden tulisi säilyä indeksoijan valitessa asiasanoja eri dokumenteille ja myös eri indeksoijien välillä. Eri henkilöt asiasanoittavat dokumentteja eri tavoin, vaikka olisivat kokeneita indeksoijia ja käyttäisivät samaa tesarusta. Dokumenttien keskeisten aiheiden esiin nostamisessa on onneksi enemmän johdonmukaisuutta kuin vähemmän olennaisissa asiasanoissa. Indeksoinnin oikeellisuuden mitta on se, etteivät tiedonhakijat saa hakuihinsa epärelevantteja tuloksia. (Alaterä & Halttunen 2002, 124–125.)

Tarkka sisällönkuvailu antaa tiedonhakijalle mahdollisuuksia rajata, mistä näkökulmasta hän haluaa etsimäänsä teemaa käsiteltävän aineistossa (Vakkari 2006, 22). Myös tiedonhakutaidoilla on merkitystä, sillä parhaat tulokset saa käyttämällä useita hyvin valittuja hakutermejä.

Hakutulosten relevanttiutta on jo pitkään arvioitu tutkimuksilla, joissa tehdään testihakuja ja lajitellaan niissä saadut hakuosumat relevantteihin, mahdollisesti relevantteihin ja epärelevantteihin. Hyötyjen saamiseksi tällaisista tutkimuksista eritellään epärelevanttien hakuosumien tyyppisiä ja syitä. Kristensenin ja Järvelinin käyttämässä luokittelussa virheelliset hakuosumat voivat johtua osumina olleista

teksteistä, hakustrategiasta tai testattavasta tesauruksesta. (Kristensen & Järvelin 1990, 6–7.) Vaikuttaa kärkevältä väittää epärelevanttien hakuosumien johtuvan itse teksteistä, sillä niiden sisältö ei ole tietokannan tekijöiden päätettävissä. Teksteissä ei välttämättä ole vikaa, vaan tietokannassa joka antaa hakutuloksiksi dokumentteja, jotka eivät sovi hakukyselyyn.

Kristensen ja Järvelin erittelevät teksteistä johtuvia virheitä. Joissakin tapauksissa monihaussa käytetyt hakusanat ovat molemmat tekstissä, mutta eivät liity toisiinsa. Haettujen termien suhde voi myös olla erilainen kuin käyttäjän tiedontarpeessa. Epärelevantit hakuosumat käsittelevät hakukomentossa olevaa asiaa vain marginaalisesti tai käyttäjän kannalta väärästä näkökulmasta. (Kristensen & Järvelin 1990, 7.) Asiasanastojen tekijät ja asiasanojen valitsijat voivat vaikuttaa tiedonhakujen onnistumiseen. Etsittyä aihetta marginaalisen vähän käsittelevät hakutulokset johtunevat pääosin liian yksityiskohtaisesta kuvailusta. Vaikeampaa on vaikuttaa siihen, että tietokanta poimii hakukomentoon sopivia sanoja dokumenttien otsikoista, joiden sanat eivät kaikissa tapauksissa ole relevantteja hakijalle vaikka usein ovatkin sitä.

Dokumenttien sisällön pelkistäminen mekaanisin keinoin tietokoneella voi epäonnistua muun muassa siitä syystä, että tietokone poimii tekstistä sanoja ymmärtämättä niiden merkitystä kyseisessä asiayhteydessä. Asiasanojen valintaan vaikuttaa, mihin tarkoituksiin käyttäjät etsivät tietoa. (Taylor 2004, 241.)

Sisällönkuvailun arvoa voi punnita sen tarkoituksenmukaisuudella, joka vaihtelee eri tilanteissa ja tiedontarpeissa. Sisällönkuvailussa ei ole vain yhtä standardoitua tapaa tehdä se oikein toisin kuin luetteloinnissa, mutta jokin sisällönkuvailu voi silti olla väärin tehty. Epäonnistumisen syitä ovat dokumentin ymmärtäminen väärin sekä käsitteiden ymmärtäminen väärin tai sekoittaminen toisiinsa. (Suominen, Saarti & Tuomi 2009, 294.)

Sisällönkuvailun epäonnistumista on, jos kuvailija antaa dokumentille sellaisia asiasanoja, joita dokumentissa ei käsitellä tarpeeksi paljoa tai ei osaa ajatella sellaista käsitettä, joka kuvailee hyvin dokumentin sisältöä ja vastaa tiedonhaki-

joiden hakukriteerejä. Melko helppoa on valita asiasanoiksi dokumentissa toistuvia sanoja, mutta tulkitessa tekstin teemaksi on mahdollista nostaa myös sellaisia sanoja, joita ei lue suoraan tekstissä ainakaan toistuvasti.

Asiasanoja valitessa pitää harkita, käsitelläänkö dokumentissa hakutermin ai-
hetta niin paljon, että aihe kannattaa ottaa asiasanaksi. On kuitenkin suhteellista, mikä on riittävästi eri tiedonhakijoille. Jotakuta voi kiinnostaa, missä yhteyksissä jostakin määrätystä aiheesta on kirjoitettu ja mitä siitä on nostettu esille. Moni tiedonhakija etsii mieluummin kiinnostuksenkohteestaan tietoa, jota hän ei ennestään tiedä. Asiasanaketjujen käyttö on yksi ratkaisu vaikeuteen ennakoida monihaun termien välisiä suhteita tietokantaa ja asiasanoitusta tehdessä, mutta Turun kansankirjaston digitoidun kokoelman teoksia kuvailevia asiasanoja ei ketjuteta. YSA:ssa on valittavissa kahden asiasanan ketjuja, esimerkiksi 'luonto – virkistyskäyttö' (Kansalliskirjasto 2015a. Finto – Suomalainen asiasanasto- ja ontologiapalvelu). Sanaketjuilla kuvailtujen dokumenttien on löydyttävä myös yksisanaisia termejä hakukomennossa käyttäen.

Tiedonhakutapa tietenkin vaikuttaa tiedonhaun onnistumiseen. Tyypillisiä hakustrategiassa tehtyjä virheitä ovat jonkin olennaisen termin jättäminen pois hakukomennosta sekä väärin tai liian laajojen termien käyttö. Tesaurus vaikuttaa hakuosumiin muun muassa siten, että huonolla tavalla esitetyt termien suhteet tesauruksessa aiheuttavat epärelevantteja hakuosumia. (Kristensen & Järvelin 1990, 7–8.) Tesauruksen tulee ohjata käyttämään virallisia hakutermejä ja ehdottaa laajempia, suppeampia ja rinnakkaisia hakutermejä. YSA:ssa nämä ominaisuudet toimivat. Tavallinen tiedonhakija ei kuitenkaan välttämättä käytä asiasanastoa eikä silloin saa sen tarjoamaa apua hakutermien valintaan.

3.3 Asiasanastojen historiaa

Ensimmäisiä asiasanahakemistoja käytettiin 1700-luvulla, mutta 1800-luvun jälkimmäisellä puoliskolla niissä alettiin käyttää entistä selkeämpää systematiikkaa

(Benito 1995, 78). Ensimmäinen pyrkimys perustaa yleistettävät säännöt aakkosellisten asiasanahakemistojen tekemiseen oli Charles Ammi Cutterin vuonna 1876 julkaistu *Rules for a dictionary catalogue* (Foskett 2007, 123).

Cutterin sanaston termeistä jotkut koostuivat monesta sanasta, jolloin jokaisen termien piti alkaa merkittävimmällä sanallaan. Cutterin mukaan termien tuli muistuttaa luonnollista kieltä. Asiasanaston aakkosjärjestyksestä huolimatta hän teki myös ristiviittauksia ilmaisemaan asiasanojen välisiä yhteyksiä, mutta ainoastaan laajemmista termeistä suppeampiin. Vaikka Cutterin asiasanaluettelon käyttö ei ole säilynyt sellaisenaan, siitä on tullut asiasanoituksen perusta Yhdysvalloissa. (Foskett 2007, 124–126.) Kansainvälisenä yhteistyönä muun muassa painotuotteille on kehitetty kansainvälisesti hyväksytyt luettelointisäännöt, jotka helpottavat aineiston löytämistä. Luettelointi ja sisällönkuvailu on kehitetty kirjoja varten, mutta niitä on laajennettu kattamaan myös muunlainen aineisto. Samaa tietämystä ja menetelmiä sovelletaan verkkoresurssienkin kuvailun kehittämisessä. (Vakkari 2006, 23.)

Englanninkielisissä maissa on yleistä käyttää yhdysvaltalaisista, laajaa *Library of Congress Subject Headings* -asiasanastoa (LCSH), joka on kontrolloitu asiasanasto luetteloita ja bibliografioita varten. Monet pienehköt yhdysvaltalaiskirjastot käyttävät *Sears List of Subject Headings* -asiasanastoa. (Olson 2002, 6; Chowdhury 2010, 159.)

3.4 Asiasanoittaminen valintoina ja vallankäyttönä

Kirjastojen asiasanoituksen pioneeri Cutter teki selkeän eron asiasanoiksi otettavien ja muiden termien välillä. Kun ilmiölle (*topic*) annettiin oma paikka järjestelmässä ja nimi, siitä tuli asiasanaksi sopiva asia (*subject*). Cutter pyrki ottamaan sanastoon yleisesti hyväksytyjä termejä, mutta ei ole selvää, miksi hän hyväksyi jotkin termit ja jätti toiset samankaltaiset asiasanaston ulkopuolelle. Asiasanoiksi sopivia olivat ilmiöt, joiden nimitykset tiesi vähintään enemmistö niistä ihmisistä,

joilla oli samantapainen koulutus kuin Cutterilla. (Foskett 2007, 123.) Tässä näkyy asiasanojen kytkeytyminen indeksoijien maailmankuvaan, tietotasoon ja kulttuuriin.

Asioiden nimeäminen kirjastoluetteloissa asiasanoilla ja luokituksilla voidaan nähdä problemaattisena ilmiönä, koska nimeäminen heijastaa arvoja eikä ole välttämättä neutraalia. Kirjastojen luettelointi ei vain heijasta ympäröivän yhteiskunnan vallitsevia arvoja, vaan näiden arvojen ilmaiseminen on valinta. Yleisiä oletuksia hyväksytään myös kyseenalaistamattomasti. Näin on ollut varsinkin 1900-luvun aikana, mutta vieläkin on parantamisen varaa ja asia on tärkeä tiedostaa. Nimeäminen on valtaa kontrolloida aiheiden esittämistä ja siten niiden löytyvyyttä. Asiasanastot ovat kirjastonhoitajien väline informaation nimeämiseen, joka voidaan nähdä informaation esittämisen lisäksi myös informaation rakentamisena. (Olson 2002, 2–6.)

Jokainen nimeää todellisuuden ilmiöitä perustuen kokemustensa pohjalta muodostuneisiin käsityksiin maailmasta. Asiasanoittaja valikoi, mitä hän nimeää ja mitä jättää nimeämättä. (Olson 2002, 4.) Tämä korostaa jokaisen asiasanoittavan ja luokittelevan informaatioalan ammattilaisen kuten kirjastotyöntekijän roolia ja henkilökohtaista näkemystä. Kirjastoihin saapuu keskitetyn luetteloinnin myötä valmiit ehdotukset teosten luokista, mutta niitä voidaan muuttaa kirjastoittain.

Asiasanastot ovat hyödyllisiä apuvälineitä indeksoinnissa, vaikka niissä on ominaisuuksia, joita voidaan kyseenalaistaa (Hedman 2000, 34–35). Yleisen suomalaisen asiasanaston käytön etuihin kuuluu, ettei jokainen kirjastonhoitaja voi valita asiasanoja omien mieltymystensä mukaisesti, koska mitkä tahansa sanat eivät kuulu viralliseen asiasanastoon. Siitä huolimatta kirjastonhoitajalla on liikkumavaraa siinä, mitä asiasanoja hän valitsee teokselle eli mitkä asiat hän näkee sen tärkeimpinä piirteinä.

1900-luvun alkupuolella feministinen kirjailija Virginia Woolf näki epäkohtana sen, millaista käsitystä naisista tiedonhakupjärjestelmä välitti. Etsiessään kirjastosta tietoa naisten köyhyydestä hän ei löytänyt asiallista tietoa siitä. 1970-luvullakin Yhdysvalloissa tehtiin huomioita asiasanoituksen yksipuolisuudesta, joka syrji muun

muassa naisia ja joitakin etnisiä ryhmiä. (Olson 2002, 1–2, 7.) Kuvailtava aineisto vaikuttaa tietysti paljon siihen, millaisia hakuosumia saadaan etsiessä aineiston joukosta tekstiä esimerkiksi naisista, vähemmistöistä ja ulkomaista. Naisia ei tule huomata sukupuolena ainoastaan silloin, kun heidän asemansa on heikko.

Turun kansankirjaston vanhan kokoelman asiasanoja valittaessa on harkittava, kannattaako antaa Martin Lutherin vaimon Katharina von Boran elämäkerralle (*Katariina von Bora. Martti Lutherin puolison lyhy elämä-kerta*) asiasanaksi 'naiset' vai onko se tarpeellista. Asiasanan 'naiset' käyttöä voisi perustella sillä, että heistä kertovaa aineistoa on vanhoissa kirjoissa vähemmän kuin miehistä kertovaa. Historiallista aineistoa lukiessa helposti unohtuu, että miehetkin ovat sukupuoli eivätkä ihmisen perustyyppi. Olisi silti liioittelevaa viedä sukupuolten aseman tiedostaminen niin pitkälle, että kirjoittaisi asiasanaksi 'miehet' kaikkiin niihin kirjoihin, joissa kaikki tai lähes kaikki nimeltä mainitut henkilöt ovat miehiä. Turun kansankirjaston kokoelman digitoiduista kirjoista suuri osa on sellaisia, mutta näissä teoksissa ei käsitellä erityisesti mieheyttä. Kriittinen tarkastelija voisi huomauttaa, että asiasanaa 'miehet' harvoin käyttävä indeksoija hyväksyy historiallisen aineiston miesvaltaisuuden itsestään selvyytensä.

3.5 Hyvän indeksoinnin ominaisuuksia

Hyvän sisällönkuvailutyön tunnuspiirteitä ovat sanastontuntemuksen ja muun yleissivistyksen hyödyntäminen, asiasanoitettavaan aineistoon perehtyminen ja johdonmukaisuus. Tekstin sisällön ymmärtämistä vaativat sisällönkuvailu on tulkinnallista, ja siinä on mahdollista suuntautua eri tavoin painottamalla eri asioita. Sisällön ymmärtämisen haastavuutta lisää, että dokumentit voivat edustaa monipuolisesti eri aloja, joista kaikista kuvailija ei tiedä kovin paljoa. Erityisesti aboutness-ilmion eli sisällönkuvailulle ominaisen pelkistämisen vuoksi kuvailijan on mietittävä, mistä kuvailtavassa dokumentissa on kyse. (Suominen, Saarti & Tuomi 2009, 293–295.) Pelkistäminen on tarpeellista, koska olisi käyttäjälle liikaa luvattu, jos dokumentti löytyisi asiasanalla, jota dokumentissa käsitellään erittäin

vähän. Toisaalta jo selkeyden vaatimus saa pitämään asiasanojen määrän kohtuullisena.

Sisällönkuvailussa tulee painottaa käyttäjä- ja hakusuuntautuneisuutta enemmän kuin dokumenttisuuntautuneisuutta. Tällöin otetaan huomioon, millaisilla hakutermeillä ja selauksella dokumenttia voisivat hakea ne tiedonetsijät, joille dokumentti on hyödyllinen. Hakusuuntautuneisuutta ei kuitenkaan voi pitää täysin dokumenttisuuntautuneiden vastakohtana niin, että pitäisi valita jompikumpi lähestymistavoista. Käyttäjäsuauntautunutkin sisällönkuvailu kuvailee dokumentin sisällöllisiä piirteitä. (Suominen, Saarti & Tuomi 2009, 296–297.)

Haku- ja käyttäjäsuauntautuneisuuden mahdollisuuksia vähentää, jos dokumenteilla on laaja ja heterogeeninen käyttäjäkunta tai sisällönkuvailijalla on tuskin lainkaan tietoa käyttäjäkunnalle tyypillisistä tarpeista. Heterogeeniselle käyttäjäkunnalle dokumentteja kuvaillaan dokumenttisuuntautuneemmin kuin jos on tiedossa tietynlainen, suppeahko dokumenttien käyttäjäkunta. (Suominen, Saarti & Tuomi 2009, 297.) Yleisten kirjastojen käyttäjäkunta on laaja ja moninainen, mutta voi olla, että Turun kansankirjaston kokoelman digitoituista aineistoista kiinnostuvat lähinnä historian harrastajat ja tutkijat. Turkulaiset osoitekalenterit voivat kiinnostaa sukututkijoitakin, mutta luultavasti enemmistö kirjaston digitoitujen kirjojen lukijoista on muita historiasta kiinnostuneita aikuisia. Digitoitujen kirjojen kielet rajaavat kohderyhmän suomen ja ruotsin kieliä osaaviin.

Tiedon löydettävyyteen vaikuttaa, osaako asiasanoittaja käyttää riittävän laajoja tai spesifejä asiasanoja. Suomalaisten asiantuntijoiden mukaan sekä spesifejä että yleistasoisia asiasanoja tarvitaan. Asiasanoitusta tekevässä organisaatiossa tehtävä sanastokontrolli säätelee myös kuvailun spesifyyttä. (Suominen, Saarti & Tuomi 2009, 295–296.) Yhdysvaltaisen *Sears` List of Subject Headings* -asiasanaston käyttöohjeissa on periaatteena, että käytetään kuvailtavista asioista niin spesifejä asiasanoja kuin listassa on. Esimerkiksi 'sillat' on tämän periaatteen mukaan riittävän spesifi asiasana, jota ei tule kuvailla ylimalkaisemmin 'infrastruktuuriksi'. (Chowdhury 2010, 162.)

Kuvailun tyhjentävyydellä tarkoitetaan sitä, missä määrin kuvailut huomioivat dokumentin sisältämiä asioita pelkistämisestä huolimatta. Tyhjentävyys kuvaa sitä, miten pitkälle kuvailu etenee dokumentin aiheen ytimestä ”tyhjentäen” sisältöä. Tyhjentävyyteen vaikuttaa, miten monta asiasanaa kullekin dokumentille annetaan ilman, että asiasanat ovat synonyymisiä. (Suominen, Saarti & Tuomi 2009, 296.)

Johdonmukaisuuden vaatimus näkyy indeksoijan työssä muun muassa siten, että hänen on käytettävä dokumentaatiokielen ilmaisuja yhtenäisellä tavalla ainakin kuvailemansa kokoelman sisäisesti. Tarvittaessa on katsottava tietokannasta esimerkkejä, miten siellä mahdollisesti on kuvattu samankaltaisia teoksia kuin uusi kuvailun kohde. (Suominen, Saarti & Tuomi 2009, 299.) Turun kansankirjaston kokoelman digitoidut kirjat ovat erillinen kokonaisuus, jolloin Turun kaupunginkirjaston antamaan toimeksiantoon kuuluu noudattaa sisäistä johdonmukaisuutta pieneen kokoelmaan kuuluvien sisällönkuvailujen kesken.

3.6 Asiasanojen valitseminen

Sisällönkuvailua on kuvattu kolmivaiheisena prosessina. Aluksi on tutustuttava dokumentin sisältöön, sitten analysoitava tätä sisältöä ja lopuksi käännettävä analysoinnin tulokset dokumentaatiokielelle. (Alaterä & Halttunen 2002, 81; Suominen, Saarti & Tuomi 2009, 293.) Asiasanoitustyöhön kuuluu dokumenttien sisällön käsitteellinen erittely (*conceptual analysis*) ja sen tulosten kääntämisen standardoiduiksi asiasanoiksi (Taylor 2004, 242–243; Chowdhury 2010, 155). Asiasanoittaminen täytyy kuitenkin aloittaa tutustumalla dokumentin sisältöön.

Kuvailtaviin kirjoihin on perehdyttävä, mutta kaikkia niitä ei tarvitse lukea läpi. Käsitteiden saamiseksi dokumenttien sisällöstä kannattaa käyttää erityisesti teosten johdantoja, johtopäätöslukuja, sisällysluetteloja, väliotsikoita, typografisesti harvennettuja tai lihavoituja kohtia sekä mahdollisia kuvia, diagrammeja, taulukoita ja niiden alla olevia tekstejä. Dokumentin nimi on voi auttaa näkemään teoksen pääasiaa ja näkökulmaa. (Alaterä & Halttunen 2002, 81; Rowley & Hartley 2008,

51; Suominen, Saarti & Tuomi 2009, 294–295.) Pelkkä dokumentin nimi ei kuitenkaan riitä sisällön määrittämiseen. Dokumentin sisällön määrittämisen apuvälineinä voi käyttää kyseistä dokumenttia käsitteleviä muita lähteitä, jos se on mahdollista. (Suominen, Saarti & Tuomi 2009, 294–295.) Ainakin klassikoita kuvaillessa tämä on tärkeätä ja onnistuu melko vaivattomasti.

Yksi niistä syistä, miksei dokumenttien otsikoissa esiintyviä termejä pidä automaattisesti poimia asiasanoiksi, on sanojen metaforisen käytön mahdollisuus. Joidenkin kirjojen nimet, etenkin yläotsikot, antavat mielikuvan jostakin muusta kuin itse aiheesta. (Hedman 2000, 32.)

Asiasanojen antamisessa käytetään induktiivista lähestymistapaa, jossa poimitaan yksityiskohtia todellisuudesta ja tehdään niiden perusteella päätelmiä kokonaisuudesta (Suominen, Saarti & Tuomi 2009, 189–190). Käsitteellisen analyysin voi tehdä ensin riippumattomana asiasanastosta ja sitten kääntää analyysin tulokset asiasanaston termeiksi (Taylor 2004, 242–243), vaikka käytännössä tottuneisuus asiasanaston kieleen vaikuttaa sisällönkuvailijan ajatteluun ja voi saada nopeasti hylkäämään asiasanoiksi sopimattomat kuvailut.

Yhdessä käsitteellisen analyysin menetelmässä termien valitsija tiivistää dokumentin kuvailun yhdeksi virkkeeksi, josta hän sitten poimii asiasanastoon kuuluvat hakutermit (Taylor 2004, 242–243). Voi kuitenkin olla työläämpää muotoilla sisältö virkkeeksi kuin sanoiksi, sillä toisinaan olisi vaikeaa tehdä sellainen virke, johon sisältyisivät kaikki yhdelle dokumentille annettavat asiasanat. Yksittäisiä asiasanoja voi lisätä listaan tekstiä lukiessa sitä mukaa kun niitä ilmenee. Sanaa laajemmissa kuvailuissa saattaa olla ilmauksia, jotka eivät sovi asiasanoiksi, sillä esimerkiksi 'nousu ja tuho' tai 'ryysyistä rikkauteen' voivat tiivistää dokumentin sisältöä, mutta eivät sovi tiedonhakuun.

3.7 Asiasanojen tyypit ja muoto

Asiasanastoihin kuuluvat yhdestä tai useammasta sanasta tai symboleista muodostuvat termit viittaavat käsitteisiin, jotka ovat abstrakteja mielikuvia jokaisen

omassa mielessä. Käsitteistä voidaan puhua yksiselitteisesti käyttämällä vakiintuneita termejä tai käsitteitä kuvaavia määritelmiä. (Suonuuti 2006, 11.) Yleisessä suomalaisessa asiasanastossa on yksisanaisten lisäksi useista sanoista koostuvia termejä, kuten 'Ruotsin vallan aika' (Kansalliskirjasto 2015a. Finto – Suomalainen asiasanasto- ja ontologiapalvelu).

Termit voivat olla yksittäisten sanojen lisäksi yhdyssanoja, sanaliittoja tai lyhenheitä. Sanastotyöstä kirjoittaneen Suonuutin mukaan sanastoissa termit ovat useimmiten yksikkömuotoisia substantiiveja, mutta sanastoissa käytetään myös monikkomuotoisia substantiiveja sekä verbejä, adjektiiveja ja adverbeja. (Suonuuti 2006, 32.) Tämä ei kuitenkaan koske kaikkia sanastoja, sillä Alaterän ja Halttusen mukaan verbejä ja adjektiiveja ei oteta asiasanastoihin (Alaterä & Halttunen 2002, 75). Yleisessä suomalaisessa asiasanastossa olevat termit ovat yleensä substantiiveja joko yksiköllisessä tai monikollisessa muodossa. Silti ainakin värit ovat YSA:n asiasanoja, vaikka ne kuuluvat adjektiiveihin.

Termien tulee olla virallisia sanoja, oikein kirjoitettuja, oikeassa muodossa ja mieluiten sellaisia, joita tiedonhakijat käyttävät. Tiedonhakijoillekin on mahdollista käyttää asiasanastoa tiedonhakukyselyjensä muotoilussa. (Chowdhury 2010, 155–156.) Termien pitää noudattaa kielen normeja ja sääntöjä, heijastaa kuvaamiensa käsitteiden piirteitä, sallia taivuttaminen ja johdokset sekä mieluiten olla lyhyitä (Suonuuti 2006, 32).

Asiasanat jakautuvat systemaattisiin ja vapaisiin. Systemaattisesti valitut, kontrolloidut asiasanat valitaan asiasanastoista, tesauruksista tai luokitusjärjestelmistä, kun taas vapaat asiasanat muotoillaan itse. (Hedman 2000, 34–35.) Tesauruksen tai asiasanaston johdannossa voidaan määritellä sallittuja vapaan indeksoinnin termejä eli sanaryhmiä, joihin kuuluvia sanoja saa käyttää vapaasti, vaikkei kaikkia näiden ryhmien sanoja löydy asiasanastosta (Alaterä & Halttunen 2002, 76–77). Yleisen suomalaisen asiasanaston käyttöohjeessa on lista vapaan indeksoinnin termien kategorioista, joihin kuuluvat muun muassa erisnimet, ammatit, kansallisuudet ja eläinlajit (Helsingin yliopiston kirjasto 1999. YSA:n käyttöohjeet).

Asiasanastoon hyväksyttävien asiasanojen määrittelyssä yhdistetään lähes samaa tarkoittavia sanoja eli niputetaan synonyymeja. Keskenään samalta näyttäviä sanoja eli homonyymeja erotellaan. (Chowdhury 2010, 156.) Suunnilleen samaa tarkoittavista sanoista valitaan jokin viralliseksi asiasanaksi ja noudatetaan tätä päätöstä järjestelmällisesti antaessa asiasanoja dokumenteille. Virallista termiä merkitykseltään muistuttavista termeistä on viittaukset siihen termiin, jota pitää käyttää. Homonymien tapauksessa on tarpeellista voida hakea monimerkityksisen sanan yhdellä merkityksellä. Jos sanaa käytetään useammassa kuin yhdessä merkityksessä, voidaan termin perään merkitä esimerkiksi kulmasulkeissa, minkä alan terminä ilmausta käytetään. Tämän käytännön mukaan homonyymi merkitään esimerkiksi 'vuoto <kasvitiede>'. (Suonuuti 2006, 28, 32–33; Suominen, Saarti & Tuomi 2009, 188–189, 195.)

Asiasanojen on oltava sellaisia, että tiedonhakijalla on mahdollisuus ennakoida, miten häntä kiinnostava asia on ilmaistu tiedonhakupöytäkirjassa. Selailuhaussa tämä tarkoittaa, että lukijan on oltava mahdollista hahmottaa ja ymmärtää dokumentaatiokieltä ja sen rakenteita. (Hedman 2000, 33; Suominen, Saarti & Tuomi 2009, 188.) Ennakoitavuutta lisäävät sanastokontrolli ja sen ylläpitämisen johdonmukaisuuden noudattaminen. Sanastokontrolli on käsitetyyppien, synonymian, homonymian ja spesifisyystason kontrollia. (Alaterä & Halttunen 2002, 75; Suominen, Saarti & Tuomi 2009, 188–189.)

On olemassa sääntöjä siitä, missä taivutusmuodoissa sanat ovat asiasanastossa tai tesauksessa. Yleinen lähtökohta on käyttää sanoja nomineina ja nominaalimuodoissa. Nominaalimuoto merkitsee verbien ja adjektiivien muuttamista substantiiveiksi, jolloin esimerkiksi sanoja 'juosta' ja 'köyhä' käytetään muodossa 'juokseminen' ja 'köyhyys'. Tekemiset, toiminnot, tapahtumiset ja ominaisuudet ilmaistaan nominaalimuodoissa. (Suominen, Saarti & Tuomi 2009, 193.) Laskettavia objekteja tarkoittavia sanoja käytetään asiasanastoissa yleensä monikko-muodossa. Sitä vastoin ainesanat ja abstraktit käsitteet ovat yksikkömuotoisia asiasanoja. (Hedman 2000, 33; Suominen, Saarti & Tuomi 2009, 193–194.)

Ruotsin kielessä yksi ratkaistava kysymys on, pitääkö asiasanoja käyttää määräisessä vai epämääräisessä muodossa. Onneksi Allärsin käyttö auttaa valitsemaan ruotsinkielisiä asiasanoja. Hedman neuvoo käyttämään abstrakteista käsitteistä yleensä määräistä muotoa, esimerkiksi 'naturen', mutta tarkoitettaessa täsmennettyä asiaa genetiivi-ilmauksissa käytetään epämääräistä muotoa, kuten 'Skånes natur'. Epämääräisen ja määräisen muodon käytöstä ei ole yleispätevää sääntöä esimerkiksi aatesuuntausten tapauksessa. (Hedman 2000, 33.)

3.8 Termien suhteet toisiinsa

Tesaurusten muodostamisesta ja käytöstä oppaan kirjoittaneiden Aitchisonin, Gilchristin ja Bawdenin mukaan tesaurus on kontrolloidun indeksointikielen sanasto, joka tekee termien väliset suhteet näkyviksi ja jota käytetään tiedonhaussa ja indeksoinnissa. Tesaurusten toissijainen tarkoitus on auttaa lisäämään ymmärrystä tiedon aihealueista. (Aitchison, Gilchrist ja Bawden 2004, 1.)

Sanastokontrolliin kuuluu viittausjärjestelmä, jossa viitataan käytettäviin sanoihin sellaisista samaa tarkoittavista sanoista, joita ei käytetä virallisina asiasanoina kyseisessä sanastossa. Niitä termejä, joista on viittaukset virallisiin asiasanoihin kutsutaan viittaustermeiksi. Asiasanoiksi suositeltavat termit tesauruksissa ovat niin sanottuja kuvailutermejä. (Suominen, Saarti & Tuomi 2009, 195.) Alaterä ja Halttunen käyttävät viittaustermeistä nimeä ohjaustermit ja muistuttavat, että asiasanastoihin kuuluvat sekä asiasanat että ohjaustermit. Ohjaus- tai viittaustermit ovat olennaisia tiedon löytyvyyden kannalta, sillä ne ovat käyttäjien mahdollisesti käyttämiä hakusanoja, joista tietokanta ohjaa virallisiin asiasanoihin. (Alaterä & Halttunen 2002, 75.)

Viittaukset ovat tarpeellisia muun muassa lyhenteiden ja lyhentämättömien ilmausten, lääkkeen kaupallisen tuotenimen ja tieteellisen nimityksen välillä ja puhtaasti suomenkielisten ilmaisujen ja samaa tarkoittavien lainasanojen välillä. Tarkkoja paikannimiä on mahdollista valita asiasanoiksi YSA:sta, jolloin termi alkaa sen kunnan nimellä, johon paikka kuuluu. Kunnan nimen jälkeen on ajatusviiva ja sen perässä paikan kuten tunturin tai järven nimi. (Suominen, Saarti &

Tuomi 2009, 196–197.) Yleiskäsitteiden, kuten 'tornit' lisäksi asiasanoina voidaan käyttää yksilökäsitteitä, jotka viittaavat ainoastaan yhteen tarkoitteeseen kuten termi 'Eiffel-torni' (Suonuuti 2006, 11). Turun kansankirjaston digitoidun kokoelman asiasanoina ei ollut tarvetta käyttää yksilötermejä ja tarkkoja paikannimiä.

Asiasanastoihin ja tesauruksiin kuuluvat aakkosellisesti järjestetyt termit sekä tarpeelliset ristiviittaukset merkitykseltään lähekkäisiin termeihin. Tesauukset sisältävät jokaisen termin yhteyteen asetettuina sitä laajempia, suppeampia ja muuten lähekkäisiä termejä. Nykyään sanoja asiasanasto ja tesaurus käytetään toistensa synonyymeina. (Alaterä & Halttunen 2002, 75–76; Chowdhury 2010, 157–159.) Tesauuksen tärkein erityispiirre asiasanastona on käsitteellisten suhteiden systemaattinen ilmaiseminen (Suominen, Saarti & Tuomi 2009, 190).

Asiasanastot voidaan jakaa yleisiin ja alakohtaisiin, aakkosellisiin ja hierarkkisiin (Alaterä & Halttunen 2002, 75). Sekä aakkosjärjestyksen että kirjastoluokituksen mukaisesti selattavissa olevia asiasanastoja on tarpeellista käyttää sekä sisällönkuvailussa ja luetteloinnissa että tiedonhaussa (Hedman 2000, 36). Tiedonhaun ei kuitenkaan tulisi koitua käyttäjille turhan monimutkaiseksi.

YSA:n luonne asiasanastona eikä tesauruksena näkyy siinä, ettei sitä voi selata ylä- ja alakäsitteiden mukaan, vaan samassa kirjaston luokituksessa olevat asiasanatkin esitetään aakkosjärjestyksessä. Toisaalta termien yhteyteen on lisätty niiden rinnakkaistermejä ja yläkäsitteitä eli sanasto esittää termien välisiä suhteita.

Termien välisten suhteiden esittäminen tesauruksissa auttaa tiedonhakijoita. Se helpottaa oikean asiasanan löytymistä ja muiden tarpeeseen sopivien hakutermin olemassaolon huomaamista sekä haun laajentamista tai tarkentamista. Termien jako kategorioihin aihepiireittäin ja tesauuksen käyttöön opastava johdanto ovat muita keinoja auttaa tiedonhakijoita. Kirjasto voi pyrkiä edistämään aineiston kuvailussa käyttämänsä tesauuksen hyödyntämistä tiedonhaussa sijoittamalla asiakaskäyttöliittymään linkin tesaurukseen. (Alaterä & Halttunen 2002, 76–77.)

Käsitteiden välillä on kolmenlaisia suhteita: hierarkkisia suhteita, koostumussuhteita ja funktiosuhteita. Hierarkkisessa suhteessa toisiinsa olevilla käsitteillä on

olennaisia yhdistäviä piirteitä, mutta alakäsitteellä on vähintään yksi erottava lisäpiirre. Alakäsitteen ala on pienempi kuin yläkäsitteen ala. Esimerkiksi 'havupuu' on 'puun' alakäsite, ja pienempi ala merkitsee, että havupuita on vähemmän kuin puita. (Suonuuti 2006, 13–14.)

Hierarkkiset suhteet ovat monitasoisia; yläkäsitteet kuuluvat johonkin vielä ylempiin käsitteisiin ja alakäsitteillä on omia alakäsitteitä ja niillä jälleen alakäsitteitä. Yläkäsite voidaan jakaa alakäsitteisiin useilla tavoilla. Saman kriteerin perusteella erotellut alakäsitteet ovat toistensa vieruskäsitteitä. Erottelu saman kriteerin perusteella tarkoittaa, että vain osa rinnakkaisista termeistä on toistensa vieruskäsitteitä. Esimerkiksi 'varjopuu' on 'valopuun' vieruskäsite, mutta 'ikivihreä puu' ei ole. (Suonuuti 2006, 13–14.)

Koostumussuhde merkitsee, että yläkäsite tarkoittaa kokonaisuutta ja alakäsite kokonaisuuden osaa. Hierarkkisten tavoin koostumussuhteisetkin käsitejärjestelmät voivat olla moniulotteisia ja monitasoisia. Moniulotteisuus tarkoittaa erilaisia mahdollisuuksia jakaa yksi käsite alakäsitteisiin. Funktiosuhteilla tarkoitetaan monenlaisia assosiaatioihin eli miellelyhtymiin perustuvia käsitteiden välisiä suhteita. Funktiosuhteita on muun muassa syiden ja seurausten, tuottajien ja tuotteiden, työkalujen ja käyttötarkoitusten sekä tapahtumien ja paikkojen välillä. Yhdessä käsitejärjestelmässä käytetään usein sekä hierarkkisia suhteita että koostumus- ja funktiosuhteita, koska käsitteet liittyvät toisiinsa useilla eri tavoilla. (Suonuuti 2006, 15–18.)

Alaterän ja Halttusen esittämä jaottelu käsitteiden välisistä suhteista on selkeä ja on käytössä Yleisessä suomalaisessa asiasanastossakin. Tesauruksessa esitettävät termien väliset suhteet voivat olla hierarkkisia tai assosiaatioon perustuvia, mutta kolmas tyyppi ovat ekvivalenssi- eli vastaavuussuhteet. Vastaavuussuhde on viittaus- tai ohjaustermin ja käytössä olevan asiasanan välillä. Hierarkkia osoittavat viittaukset laajempiin ja suppeampiin termeihin. Assosiaatiosuhteet ovat tyypillisiä rinnakkaistermien välillä, mikä on vähän eri asia kuin Suonuutin tarkoittama assosiativisuus. (Alaterä & Halttunen 2002, 76.) *Library of Congress Subject Headings* -sanastossa termien välisiä suhteita osoittavat viittaukset kus-

takin termistä niihin termeihin, jotka se korvaa, laajempiin ja rinnakkaisiin termeihin sekä muihin termeihin, joilla voi löytyä materiaalia aiheesta (Chowdhury 2010, 160).

Tesaurusta tehdessä sanaston keräämisen apuna käytetään lähteinä esimerkiksi valmiita asiasanastoja, tesauksia ja luokituskaavoja, hakuteoksia ja erityisalojen kirjallisuutta. Välineenä voidaan käyttää myös tehtyjä tiedonhakuja. Eri aloilla on erikoistuneita asiasanastoja ja tesauksia, jollaista tehdessä ei välttämättä ole olemassa aiempaa asiasanastoa samasta aiheesta. Erityisesti uusien termien puute on tyypillinen dokumentaatiokielten ongelma. (Suominen, Saarti & Tuomi 2009, 191.) Eri alojen erikoissanastoja voidaan käyttää Yleisen suomalaisen asiasanaston rinnalla (Alaterä & Halttunen 2002, 77).

4 DIGITOIDUN KOKOELMAN ASIASANOITUSTYÖ

4.1 Asiasanoitettava digitaalinen kokoelma

Yleisessä kirjastossa monet asiakkaat eivät pidä viitetietokannoista, vaan toivovat tiedonhakunsa tulokseksi kokotekstejä. Yhtenä ratkaisuna tähän on kirjastoaineistojen digitointi, joka tekijänoikeussyistä on mahdollista lähinnä vanhan painetun aineiston käyttöön saattamisen keinona. Digitoinnin ansiosta ei kaikkia vanhoja aineistoja tarvitse hakea asiakkaalle hakupyynnöksi seuraavaksi päiväksi. Heikkojen niteiden säilymistä edistää, ettei niitä tarvitse toistuvasti siirtää ja lukea.

Digitoitujen aineistojen on oltava riittävän helposti löydettävissä, jotta niistä on hyötyä asiakkaille. Sitä varten Kirjastot.fi ylläpitää Digi.kirjastot.fi -palvelua, joka kokoaa yleisten kirjastojen digitoitua aineistoa yhteen ja helpottaa sen löytämistä. Sivuston hyötyihin kuuluu, että se sisältää eri kirjastojen digitoimaa aineistoa, jolloin käyttäjän ei tarvitse rajoittua itseään lähimmän kirjaston aineistoon. Myös Turun kaupunginkirjasto tallentaa digitoimansa aineiston kyseiseen tietokantaan.

Digitoiduista teoksista on kirjattu Digi-tietokantaan nimeke, tekijän nimi, julkaiseva kirjasto, digitoidun kappaleen julkistamisvuosi, teoksen julkaisuvuosi, kieli sekä kirjaston kokoelma, johon teos kuuluu. Opinnäyteprojektissa asiasanoitetut Turun kansankirjaston kokoelmaan kuuluvat teokset ovat tietokannassa aakkosjärjestyksessä nimekkeen mukaan, jos hakukriteerinä on kokoelma. Useiden digitoitujen kirjojen tekijöiden nimiä ei tiedetä. Tietokirjoista on merkitty tietokantaan niiden aiheet ilmeisesti otsikkojen perusteella pääteltyinä, mutta kaunokirjallisilla teoksilla ei ole lainkaan asiasanoja ennen opinnäyteprojektia.

Enemmistö digitoidun kokoelman kirjoista on suomenkielisiä, vähemmistö ruotsinkielisiä. Suurin osa kirjoista on painettu fraktuurafontilla, mutta joukossa on myös uudemmallalla kirjoitustyyllillä painettuja teoksia. Digitoidut kirjat antavat autenttisen lukukokemuksen; lukija näkee kannen, aukeamat ja alkuperäisen tekstin. Asiasanoittamista vaikeuttaa, että monissa vanhoissa kirjoissa ei ole sisälly-

luetteloita. Toisaalta suurin osa näistä kirjoista ei ole pitkiä, vaan ne ovat pikemminkin vihkosia, joista lyhimmissä on vain noin kymmenen aukeamaa. Osa teoksista on kuitenkin paljon sitä laajempia.

Digitoituun kokoelmaan sisältyy kirjoja useista aiheista, sekä tieto- että kaunokirjallisuutta. Asiasanoitettavaan tietoaaineistoon kuuluu opaskirjoja terveydenhoidosta, lastenkasvatuksesta ja kotieläinten hoidosta. Aiheita ovat myös historia merkkihenkilöineen, tutkimusmatkat Afrikkaan ja Australiaan, kristillisyys sekä kesyt ja villit eläimet. Yksittäiset teokset kertovat ojituksesta, vaurastumisesta, avaruudesta ja kiinalaisista.

Turun kansankirjaston ajalta olevien kirjojen tekstile on tyypillistä selkeä arvottavuus; nykyisiin kirjoihin verrattuna huomiota herättää, miten selvästi historiateosten ja elämäkertojen kirjoittajat ylistävät joitakin henkilöitä ja mustamaalaavat toisia. Tämä ei kuitenkaan näy mitenkään asiasanoituksessa, jolla pyritään neutraaliuteen. Nykylukijoille kiinnostavaa 1800-luvun kirjoissa on tietosisällön lisäksi julkaisemisajan näkökulma ja kirjoittamistyyli.

Kristillisyys punoutuu usein sekä tietokirjojen että kaunokirjallisten teosten aiheisiin, minkä vuoksi kristinuskosta tulee helposti usein toistuva asiasana sisällönkuvailussa. Suuri osa kokoelman kaunokirjallisuudesta painottuu valistamiseen, mutta osa on viihteellisempää. Yleinen kirjallisuudenlaji kokoelmassa ovat valistavat kertomukset, jotka pyrkivät suostuttelemaan lukijoitaan raittiuteen, kristillisyyteen, ahkeruuteen ja säästäväisyyteen.

Suomessa alkoholiongelman yhteiskunnallisen käsittelyn peruselementit hahmottuivat 1840-luvulla. Alkoholinvastainen kirjoittelu lisääntyi erityisesti 1870-luvulla, jolloin Kohtuuden Ystävät -järjestö levitti raittiuskirjasia. (Sulkunen 1986, 19–20.) Raittiussivistyksessä työmiehiä leimattiin juopoiksi, kun sitä vastoin heidän vaimonsa kuvattiin uhreina, joiden kärsimyksen kuvailu korosti alkoholia käyttävien työmiesten moraalittomuutta. Alkoholismin vastustaminen muodostettiin hätähuudoksi perheiden, uskonnon ja isänmaan puolesta. (Sulkunen 1986, 33–34.) Turun kansankirjaston digitoituun kokoelmaan kuuluvissa valistuskirjassakin näkyvät selvästi nämä piirteet. Kirjaset on julkaistu 1850-luvulla, eli ne

ovat melko varhaista valistusmateriaalia. (*Kuwael(e)mia juopumuksesta ja raitteudesta...* 1854 & 1858; *Hywää tarkoittawia waroituksia wiinaa vastaan...* 1859.)

4.2 Työskentelyperiaatteet asiasanoittamisessa

Asiasanoittajan kannattaa käyttää hyödykseen väliotsikkoja, lihavoituja kohtia sekä joidenkin aineistojen luettelomaisia ja toistuvia piirteitä. Digitoidun kokoelman teoksista luettelomaisimpia ovat osoitehakemistot, kirjastoluettelot ja nimi-päiväkalenteri. Lihavoituja sanoja on ainakin historiateoksissa, joissa on lihavoituina erityisesti erisnimiä. Teosten lukujen aiheet sopivat asiasanoiksi, elleivät ne ole liian ilmiselviä. Esimerkiksi Ruotsi, Venäjä ja Turku ovat olennaisia Suomen historiassa, mutta se pitäisi ehkä olla pääteltävissä ilman asiasanojakin.

Yksi tämän opinnäytetyön tekijän pohtima kysymys asiasanoittamisessa on, kirjoitetaanko asiasanat tärkeys- vai aakkosjärjestykseen vai onko väliä sillä, missä järjestyksessä aiheet on käsitelty kirjassa. Turun kaupunginkirjaston edustaja suosittelee noudattamaan termien tärkeysjärjestystä, siis niiden olennaisuutta teoksen sisällön kuvaamisessa. Hänen mukaansa järjestyksen merkitystä vähentää, ettei sillä ole merkitystä asiasanahaun käytössä.

Toimeksiantajan mukaan ei tarvitse epäröidä käyttää sellaisiakaan hakusanoja, jotka voi päätellä nimekkeestä. Tätä perustelee ainakin tarve käyttää asiasanoja hakujen tekemisessä näkemättä etukäteen dokumenttien nimiä. Asiasanoja kirjoittaessa on tarpeellista kiinnittää huomiota oikeinkirjoitukseen. Projektin lopussa pitää tarkistaa, ettei asiasanoissa ole näppäilyvirheitä, sillä ne antaisivat huolimattoman vaikutelman ja estäisivät väärin kirjoitetuilla sanoilla kuvailtujen dokumenttien löytyvyyttä asiasanahauulla.

Joitakin termejä käytetään asiasanastossa sekä yksiköllisenä että monikollisena, jolloin tämän erovaisuuden avulla voidaan tehdä semanttisia eroja. Esimerkiksi 'elokuva' tarkoittaa elokuva-alaa ja 'elokuvat' elokuvateoksia. (Suominen, Saarti & Tuomi 2009, 193–194.) Historiaa käsitteleviin tietokirjoihin usein sopii monikomuotoinen asiasana 'sodat', koska Yleisessä suomalaisessa asiasanastossa

'sodat' viittaa nimettyihin sotiin ja 'sota' kuvaa yleisluonteisemmin sotaa ilmiönä (Kansalliskirjasto 2015a. Finto – Suomalainen asiasanasto- ja ontologiapalvelu).

Sisällönkuvailun suuntautumisia ovat ydinkuvailu, tarkempi sisällönmukainen kuvailu sekä käyttötarkoituksiin ja traditioihin kuten kirjallisuudenlajeihin suuntautuva kuvailu. Suurpiirteiset ydinkuvailut käsittävät vain harvoja termejä; ne kertovat teoksen aiheen. Tarkemmassa kuvailussa käytetään asiasanoina erisnimiäkin. (Suominen, Saarti & Tuomi 2009, 298.) Opinnäytetyöprojektissa tehty asiasanojen antaminen on tarkkaa sisällönmukaista kuvailua, mutta työn edellyttämä aiheen määrittely jokaiselle kirjalle on ydinkuvailua.

Sisällönkuvailu voi suuntautua dokumentin lisäksi sen ulkopuolelle dokumentin mahdollisiin käyttötarkoituksiin ja dokumentin diskursiivisten kontekstien ilmaistamiseen. Käsite intensionaalinen aboutness tarkoittaa dokumentin diskursiivisia konteksteja, esimerkiksi kysymyksiä, joiden käsittelyyn teksti osallistuu. Dokumentteja voidaan kuvailla esimerkiksi kirjallisuudenlajeilla tai aatesuuntauksilla, joita ne edustavat. (Suominen, Saarti & Tuomi 2009, 297–298.) Kansakoulun historiankirjalle (*Suomen historia kansakouluille. Lyhykäisesti kerrottu*) yhdeksi asiasanaksi sopii 'oppikirja', joka kuvaa kirjallisuustyyppiä ja käyttötarkoitusta. Käyttäjäsuaunutta sisällönkuvailua on esimerkiksi se, että käyttää asiasanaa 'autonomian aika', vaikkei se kuulunut aikalaiskieleen vielä kirjan julkaisuaikana 1800-luvulla. Kyseistä termiä ei esiinny kirjassa, mutta tiedonhakija voi tehdä haun sillä termillä.

Asiasanoja ei välttämättä tarvitse lisäillä tasaisesti koko dokumentin lukemisen ajan. Dokumentin alkua lukiessa kannattaa tarkkailla, mitä aiheita käsitellään niin paljon, että ne kannattaa merkitä asiasanoiksi. Lukemisen aikana voi merkitä muistiin joitakin mahdollisia termejä sulkumerkkeihin ja lukiessa tarkkailla, onko kirjassa näistä aiheista riittävästi asiaa, jotta ne sopivat asiasanoiksi. Kun yhden dokumentin lukeminen ei jakaudu pitkälle ajalle, saa paremmin kokonaiskuvan siitä, mitä se käsittelee. Dokumentin loppupuolta lukiessa voi huomata aiemmin esiin nousseiden teemojen toistuvan, jolloin asiasanojen listaa ei tarvitse kartuttaa yhtä paljon kuin siihen asti.

4.2.1 Kaunokirjallisuuden asiasanoittaminen

Kaunokirjallisuuden asiasanaston Verkko-Kaunokin hakukriteerit sopivat muistilistaksi siitä, mitkä asiat kaunokirjallisuuden sisällönkuvailussa ovat olennaisia. Näitä hakukriteerejä ovat genre, laji tai tyyli, teema, henkilöt, paikka ja aika. Paikan merkitsemistä vaikeuttaa, etteivät valtiot ja kansallisuudet eivät kuulu kaunokirjallisuuden asiasanaston Kaunokkiin. (Kirjastot.fi 2015b. Verkko-Kaunokki.) Toimeksiantajan eli Turun kaupunginkirjaston edustaja kertoi, että maita ja kansallisuuksia voi silti käyttää kaunokirjallisuuden sisällönkuvailussa.

Genretiedot antavat olennaista kokonaiskuvaa teoksesta, mutta genreä ei yleensä lue teoksessa. Kaunokirjallisuuden teemojen tulkitseminen etenee lukiessa, eivätkä kaikki tekstissä olevat tapahtumat ole yhtä olennaisia. Joku voi kuolla teoksessa, muttei kuolema silti osoittaudu teemaksi, jolloin sen ei välttämättä tarvitse olla kyseisen kirjan asiasanana. Teemojen tulkitseminen on yksilöllistä, jolloin eri sisällönkuvailijat voivat löytää samasta teoksesta toisistaan poikkeavia teemoja.

Asiasanojen tulee kuvata sisältöä, mutta kaunokirjallisuuden kuvailun erityispiirteisiin kuuluu, ettei ole hyvä jos loppu paljastuu asiasanoista. Useita tekstejä sisältävistä kokoomateoksista pitää lukea niiden kaikkia osia. Kokoomateokselle asiasanoja kirjoittaessa tekee mieli merkitä joka novellille erilliset asiasanat, elleivät ne ole kaikki lähes samanlaisia.

4.3 Ongelmia ja ratkaisuja

4.3.1 Sisällön pelkistäminen

Dokumenttien luokittelu yleistäviin kategorioihin edellyttää, että sisällönkuvailija näkee dokumenteissa käsitellyt asiat kategorian osina, jolloin hän tulkitsee esimerkiksi jotkut henkilöiden nimet urheilijoiksi ja useat eläinlajit lemmikkieläimiksi. Ongelmana on, että käyttäjä saattaa hakea spesifimmillä hakusanoilla kuin ne,

joiksi asiasanoittaja on kohteensa pelkistänyt. (Hedman 2000, 31–32.) On mahdollista esimerkiksi pelkistää kristinuskoon käännyttäminen, uskonpuhdistus, piispat ja hiippakunnat kaikki termin 'kristinusko' alle. Se pitää asiasanojen määrän kohtuullisena ja ehkäisee niitä hakuosumia, joissa on liian vähän tietoa asiasanan aiheesta. Kääntöpuoli on, ettei pelkällä 'kristinuskolla' asiasanoitettua kirjaa löydä asiasanahauulla 'uskonpuhdistus'. Oli harkittava, kannattaako historiateokselle (Hallstén, Alexander 1872. *Suomen historia*) antaa yhdeksi asiasanaksi 'kristinusko', vaikkei uskonto ole kirjan aihe ja varsinaisia kristillisiä kirjojakin on paljon. Silti kristinusko sopii kristinuskon asiasanoihin, koska Suomen historiankirjoitukseen kuuluvat kristinuskon tulo, piispat ja uskonpuhdistus.

Suomen historian yleisesityksissä on asiasanoituksen kannalta haastavaa, että niissä käsitellään lukuisia tapahtumia ja henkilöitä lyhyesti. Silloin on hankalaa valita, mitkä asiat teoksessa ovat tärkeimpiä. Asiasanoiksi sopivat kirjassa toistuvat aiheet, joita voivat olla esimerkiksi hallitsijat, sodat ja verotus. Voi käyttää myös aikakausien nimiä. Yksi vaihtoehto historiateoksen asiasanoittamistavaksi olisi keskittyä ilmiöiden sijasta listamaan tapahtumia, mutta tapahtumia on paljon eikä kaikilla niistä ei ole asiasanaksi sopivaa nimeä. Termien johdonmukaista käyttöä koskeva kysymys on, pitääkö esimerkiksi Kalmarin unioni ottaa kaikkien niiden kirjojen asiasanaksi, joissa unionia on käsitelty. Jos sisällönkuvailija valitsee joitakin tapahtumia asiasanoiksi kuvaillessaan aiheiltaan laajaa historiateosta, voidaan nähdä epäjohdonmukaisena, ettei kaikkia kirjoissa käsiteltyjä tapahtumia ole mainittu samalla tarkkuudella asiasanoissa.

Kotieläimistä ja hyötykasveista kertovassa kirjassa (*Kotieläimet ja kotitarvekasvut*) esitellyt eläin- ja kasvilajit ovat konkreettisia eliöitä, jotka oli helppoa poimia asiasanoiksi. Lajien luettelosta tuli epäyhtenäisen näköinen, koska YSA:ssa eläinlajit ovat yleensä yksikössä, mutta jotkut monikossa. Jos asiasanojen määrää pitäisi vähentää, ei ole yksinkertaista keksiä vaihtoehtoisia menetelmiä asiasanoittamiseen. Muina asiasanoina voi käyttää kirjoittajan näkökulmia kuvaavia sanoja kuten 'käyttö' ja 'alkuperä', mutta nämä ovat hyödyllisempiä teosten asiasanoja lukiessa kuin hakutermeinä. Erilaiset kotimaiset ja ulkomaiset viljalajit ovat koottavissa yläkäsitteeseen 'viljakasvit', joka on hyvä asiasana.

Toisessa luontokirjassa (*Luonnon oppi*) esitellään niin monia eläinlajeja, että luokituksellisten ryhmien luetteleminen on keino hillitä asiasanojen liiallista määrää, erityisesti koska kaikista lajeista ei kerrota paljoa. Eläinryhmiä ovat esimerkiksi matelijat, kalat ja pedot. Asiasanoittamisen ollessa kesken asiasanojen listaan kuului myös niitä eläinlajeja, joista kirjassa on eniten tekstiä. Lajit kuitenkin näyttivät irrallisilta ja oudoilta eläinryhmien joukossa ja sisältyivät ryhmiin.

Pelkistämiseen kuuluu, että asiasanojen on kuvattava suurta osaa dokumenttien sisällöstä pyrkimättä liialliseen yksityiskohtaisuuteen. Turkulaiset osoitekalenterit (*Adresskalender för Åbo stad...*) sisältävät esipuheen, jonka jälkeen sisältö on henkilöiden ja palveluiden nimiä ja osoitteita sekä mainoksia. Osoitekalenterien esipuheissa kerrotaan yllättävän tarkasti tilastotietoa Turusta ja sen nähtävyyksistä, tuomiokirkosta jopa hautakappeli kerrallaan. Asiasanoitustyön toimeksiantaja kuitenkin kehotti olemaan maltillinen osoitekalenterien asiasanoittamisessa, ettei sanojen runsaus aiheuttaisi pettymyksiä tiedonhakijoille, kun suuri osa sisällöstä koostuu osoitetiedoista ja mainoksista.

4.3.2 Sanojen merkitykset

Asiasanat voivat herättää useanlaisia mielikuvia. Raittiutta ja terveydenhoitoa edistävät suurelle yleisölle suunnatut kirjat ovat valistusmateriaalia. Asiasana 'valistus' voi tämän tekstilajin kuvaamisen lisäksi tuottaa miellelyhtymän valistuksen ajasta. Termin 'valistus' ohella YSA näyttää niin sanottuna assosiatiivisena terminä sanan 'kansanvalistus' (Kansalliskirjasto 2015a. Finto – Suomalainen asiasanasto- ja ontologiapalvelu). Tässä tilanteessa assosiatiivinen termi sopii tarkoitukseen paremmin kuin asiasanoittajan aluksi YSA:sta etsimä termi.

Kaunokirjallisuuden asiasanastossa Kaunokissa ei ole sanaa 'kansanvalistus', jota voi käyttää valistusmateriaalin asiasanoittamiseen YSA:lla (Kirjastot.fi 2015b. Verkko-Kaunokki). Termi 'opetukset' sopii valistavaan kaunokirjalliseen materiaaliin, mutta tällöin YSA:lla ja Kaunokilla asiasanoitetut toisiaan muistuttavat aineistot eivät löydy samalla hakusanalla.

Opinnäytteen tekijä on rajannut, mitä hän tarkoittaa asiasanalla 'historia'. Julkaisuaikansa kuvaavia teoksia ei tarvitse kuvailla sanalla 'historia', vaikka ne nykylukijoille sitä voivatkin olla. Termillä 'historia' kuvailtujen teosten aiheet ovat olleet julkaisuaikansakin lukijoille ainakin pääosin historiaa. Historia-termin käyttöä voi kaventaa myös käyttämällä sitä kuvaamaan vain esimerkiksi valtiollisen ja kirkkohistorian teoksia eikä mitä hyvänsä kertomuksia menneisyydestä. 'Historiaa' on mahdollista tarkentaa käyttämällä asiasanoina YSA:n termeiksi hyväksytyjä aikakausia sekä vuosisatoja ja vuosilukuja, koska YSA:n käyttöohjeiden mukaan numeraalisia ajanmääreitä saa käyttää vapaan indeksoinnin termeinä (Helsingin yliopiston kirjasto 1999. YSA:n käyttöohjeet).

Tositapahtumiin perustuva teksti kuten matkakertomus voi tuntua lukijalle seikkailulta. Tositapahtumien kuvailu seikkailuiksi olisi kuitenkin asenteellista, koska se esittäisi oikeat tapahtumat viihteellisinä. Dokumentista tulkittava kirjoittajan näkökulma aiheeseensa vaikuttaa asiasanojen valintaan. Esimerkiksi James Cookin matkoista kertovan kirjan (Hassel 1884. *Jaakko Cook`in matkat Tyynellä merellä*) kirjoittaja kuvailee luontoa, mutta vain katkelmallisesti ja silloinkin yleensä ihmisten ja elinkeinojen näkökulmasta. Tämän vuoksi on perusteltua jättää asiasana 'luonto' pois kyseisen teoksen asiasanoista ja käyttää mieluummin ihmisiin viittaavia sanoja. Nämä asiat ovat tulkinnallisia, sillä joku toinen lukija voi pitää 'luontoa' kuvaavana asiasanana kyseiselle teokselle.

Termien valintaa harkitessa joutuu arvioimaan, mitä Digi-tietokannan käyttäjä odottaa esimerkiksi termiltä 'merenkulku'. Laveasti ymmärrettynä tuo termi sopii muun muassa merimatkoja kuvaileviin kirjoihin, mutta suppeasti ymmärrettynä termi tarkoittaa merenkulkuelinkeinoon ja siinä käytettyjen menetelmien yksityiskohtaista käsittelyä. Asiasanastot eivät sisällä termien määritelmiä. Joissakin tapauksissa laajemmista, suppeammista ja rinnakkaisista termeistä voi päätellä, onko kyseinen termi kuvaavin valinta siihen asiayhteyteen, jossa sisällönkuvailija aikoo sitä käyttää.

Kun kaunokirjallisessa tekstissä (*Elehvantin kesyttäjä. Kertomus Intiasta*) on yhtenä aiheena eläimen lopettaminen, sille ei löydy Kaunokista kuvaavampaa ter-

miä kuin 'lopetus', vaikka se tuo mieleen kirjallisen teoksen taiteelliset lopetustavat. 'Lopetus' korvaa sanastossa sanan 'lopettaminen', eikä mikään Verkko-Kaunokin ehdottamista assosiativisista termeistä sovi tähän asiayhteyteen. (Kirjas-tot.fi 2015b. Verkko-Kaunokki.)

4.3.3 Väliotsikot ja vanhat sanat

Joidenkin vanhojen teosten väliotsikoita on vaikea hyödyntää asiasanoituksessa. Esimerkiksi Henry Stanley'n matkoista kertovan kirjan (Stanley, Henry 1887. *Kongo – uusi vapaavaltio mustien maanosassa. 2. osa*) väliotsikot ilmaisevat vain matkaa jostakin nimetystä paikasta toiseen, eivätkä nämä paikannimet ole todennäköisesti edes tuttuja suomalaiselle. Yksi mahdollinen tapa käsitellä vanhoissa teoksissa esiintyviä paikkojen nimiä on käyttää nykyisiä nimimuotoja vanhojen nimien tarkoittamista alueista. Se edellyttää yleensä ulkopuolisten lähteiden käyttöä. Asiasanoittajalle on hyödyllistä, että James Cookin matkoista kirjoittanut henkilö on kirjoittaessaan Uudesta-Hollannista merkinnyt alaviitteeseen sen tarkoittavan Australiaa (Hassel 1884, 27).

Hevosten hoidosta kertovan teoksen (Samuelsson 1861. *Jaloimman ja hyödyllisimmän luontokappalemme, hewoisen tautien tunteminen ja parantaminen*) väliotsikoiden perusteella pääsee selville sisällöstä. Hoito-opas on esimerkki sellaisesta tietoteoksesta, joka on kirjoitettu tiedonhaun välineeksi eikä luettavaksi kannesta kanteen. Otsikoista selviää, mihin tilanteisiin teksti tarjoaa apua. Hevosten sairauksien hoito-opaan sisältöä kuvaavia asiasanoja päättäessä vaikeutena on aiheen tuntemisen puute. Tietämättömyys lisää epävarmuutta siitä, millä käsitteillä niputtaa asioita. Toimeksiantaja huomautti, ettei hevosten ongelmista pidä käyttää sanaa 'mielenterveyshäiriöt'. Teoksessa käytetään termiä 'virma', joka ei kuulu asiasanastoon edes korvattuna terminä. Ratkaisuna on antaa asiasanaksi 'käyttäytymishäiriöt', koska tekstistä päätellen virma on sellainen.

Tiheästi toistuvista väliotsikoista huolimatta hevosten hoito-opaan rakenne ei ole niin selkeä kuin nykyisten tietokirjojen eivätkä väliotsikoiden sanat usein ole nykykielen mukaisia. Myöskään koirista kertovan kirjan (*Kertoelmia koirista,*

kuwapiirroksilla) väliotsikoissa olevista koiraroduista käytetyistä vanhoista nimi-tyksistä monet eivät ole enää nykyisin käytössä. Sama koskee lastentaudeista ja niiden hoidosta kertovaa kirjaa, jossa tautien nimet ovat erilaisia kuin nykyään eikä tauteja ole ryhmitelty (Roos 1857. *Kristillinen ja terveellinen lasten kasvataminen. Toinen osa...*). Siksi teos sai melko suurpiirteisen asiasanoituksen. Jo sisällysluettelosta näkee, että lääkekasveille ja ihotaudeille on omat päälukunsa, joten nämä sanat sopivat asiasanoiksi.

Runsaasti vanhoja termejä sisältävien kirjojen asiasanoittamisessa on riski valita asiasanoiksi vain niitä sanoja, joita asiasanoittaja ymmärtää ja jotka ovat helpoimmin tiivistettävissä asiasanoiksi. Riittävä tekstin lukeminen väliotsikkojen silmäilyn lisäksi auttaa havaitsemaan sen, jos väliotsikot eivät aina kuvaa sisältöä hyvin.

Eräässä teoksessa (*Kiinalaisista*) on selkeä väliotsikko kiinalaisista häistä, mutta kirjasta on revennyt muutama sivu, jolloin nykyaikaan säilyneessä dokumentissa ei kerrota häistä niin paljon kuin erillisen väliotsikon perusteella voi olettaa. Siksi väliotsikkojen silmäily saattaa johtaa harhaan valitessa asiasanoja joillekin vanhoille teoksille.

Lastenkasvatus ja -hoito-oppaan (Roos 1856. *Kristillinen ja terveellinen lasten kasvataminen. Edellinen osa...*) sisällöstä saa hyvän käsityksen jo sisällysluettelon perusteella, mutta asiasanoja valitessa on tarpeen pelkistää, sillä kaikkia sisällysluettelon otsikoita ei voi ottaa asiasanoiksi. ”Edustus” eli esipuhe ja ”lasten kasvattamisen ensimmäinen alku” eivät sovi asiasanoiksi. Pääluokun ”Nuoruuden kaunistus ja kaiken onnen perustus” sisältyvät alaluvut totuus, ahkeruus, säästäväisyys ja kohtuus ynnä muut sellaiset sisällysluettelossa ovat niputettavissa yläkäsitteeseen ’hyveet’ tai ’kasvatustavoitteet’. Seuraavista muutamasta otsikosta on tiivistettävissä asiasanaksi ’syntymä’. Lasten hoitamiseen opastavasta pääluvusta on helppo poimia asiasanoja, kuten ’syöminen’, ’uni’ ja ’liikkuminen’. ’Kristinusko’ kuuluu teoksen asiasanoihin, koska se on leimallista kasvatamiseen annetuille ohjeille.

Tähtitiedettä käsittelevälle ruotsinkieliselle oppikirjalle (Schubert 1858. *Naturkunnighet för skolor och familjer...*) on pitkä ja melko vaikeatajuinen, mutta helppo ja nopea keino valita asiasanat on päätellä ne otsikkojen perusteella. Asiasanoiksi ei kuitenkaan pidä ottaa kaikkia mahdollisia otsikkojen sanoja, koska esimerkiksi 'koko' ei kerro asiasanana mitään, vaikka se osana otsikkoa kertoisi sisällöstä. Kun otsikot ovat informatiivisia, yleistajuisia ja niissä olevat sanat lähellä nykykieltä, se helpottaa asiasanoittamista paljon. Jos otsikkoja ei ole, niistä ei saa riittävää käsitystä sisällöstä tai otsikoissa käytetään nykynäkökulmasta vanhentuneita sanoja, vie aikaa valita asiasanoja edes suomenkieliselle aineistolle.

Kokoelman haastavimpiin teoksiin asiasanoittamisen kannalta kuuluu Henrik Porthanin elämäkerta, jossa ei ole väliotsikoita (Rein 1864. *Henrik Gabriel Porthan*). Joku lukija on pyrkinyt erottamaan olennaista tekemällä alleviivauksia tähän ruotsinkieliseen kirjaan, joka ei kuulu kokoelman lyhimpiin teoksiin. Kirjan lukemisen lisäksi asiasanoittamisessa oli apuna Suomen merkkihenkilöitä esittelevän Kansallisbiografian artikkeli, joka tuo esiin Porthanin elämäntyötä historiantutkijana, filosofina, kirjastonhoitajana ja kansanrunouden tutkijana (Tarkiainen 2012).

YSA:n asiasanat ovat pääosin helppoja päätellä ja ymmärtää. Kuvailtavissa vanhoissa kirjoissa käytettyjen termien nykyisiä muotoja voi päätellä yleissivistyksen avulla. Vanhoissa kirjoissa käytetään esimerkiksi ilmaisua "Kalmarin yhteys", joka on nykykielellä Kalmarin unioni. Vapaan indeksoinnin termeiksi sallittujen käsite-ryhmien listassa ei ole historiallisia tapahtumia (Helsingin yliopiston kirjasto 1999. YSA:n käyttöohjeet). Niinpä kuvaillessa kirjaa *Tukholman veri-sauna eli kuningas Kristiani II:n hirmu-työt Ruotsissa w. 1520*, ei voi valita nimekkeen otsikon olennaisinta termiä 'Tukholman verilöylyä' asiasanaksi.

Dokumentin otsikkoa ei saa käyttää liian suoraan asiasanoittamiseen. Esimerkiksi kirjalla nimeltä "*Kristillinen päiväkirja. Ilmoittawa kristin uskokunnan historiasta tietoja pyhistä martyyreistä ja muista ihmisistä, joitten nimet ovat vuosittain almanakassa. Niin myöskin vähäinen tutkinto sielun huvituksesta*" ei ole mitään tekemistä sanojen 'huvitus' ja 'tutkinto' nykymerkitysten kanssa. On helppo antaa

asiasanat teokseen sisältyvälle nimipäiväkalenterille, jossa on tietoa nimien alkuperästä. Sen sijaan vanhaa hartaustekstiä voi olla hankala eritellä. Asiasanoiksi tulivat 'kristinusko', 'elämäntaito', 'synti', 'levottomuus' ja 'mielenrauha'. Termit 'elämäntaito' ja 'mielenrauha' eivät ole suoraan tekstistä, vaan kirjoittaja toistaa sanoja 'tyytyväisyys' ja 'tyytymättömyys'. Tekstin otsikot ja muu sisältö toistavat itseään, mikä auttaa varmistumaan keskeisistä teemoista ja pitämään asiasanojen määrän kohtuullisena.

4.3.4 Henkilöiden nimet asiasanoina

Henkilöiden nimiä voi valita vapaan indeksoinnin termeiksi, jolloin ne kirjoitetaan sukunimi ensin ja etunimi sen perään pilkulla erotettuna. On tarkistettava, mitä vakiintunutta muotoa käytetään erityisesti hallitsijoiden nimistä. Esimerkiksi Fredrik II on lisänimeltään Suuri, vaikkei lisänimeä käytetä digitoidussa vanhassa kirjassa. Tiedonlähteiksi nimien tarkistamisessa opinnäytetyön toimeksiantaja suosittelee Suomen kansallisbibliografiaa Fennicaa ja ulkomaisten henkilöiden osalta myös British Libraryn tietokantaa, joissa kummassakin käytetään asiasanoina myös erisnimiä.

Turun kansankirjaston digitoiduista kirjoista osan metatietojen joukkoon Digi-tietokantaan on merkitty aihe. Toimeksiantoon kuuluu tarkistaa, onko aiheiksi merkitty jotakin Yleiseen suomalaiseen asiasanastoon kuulumatonta. Kun *Ansgario*-nimisen teoksen aiheiksi on merkitty 'Ansgar, arkkipiispa', aiheen määrittely on sikäli hyvä, että se kertoo mistä teoksessa on kysymys. Kumpikaan aiheista ei kuitenkaan ole YSA:ssa, sillä sen sääntöihin kuuluu käyttää monikollista termiä 'arkkipiispat'. 'Ansgar, arkkipiispa' muistuttaa Fennican asiasanoja, joista joihinkin kuuluu henkilön nimen perään määrite kuten 'Ruotsin kuningas', mutta 'Ansgar' ei kuitenkaan ole Fennicaan kuuluva asiasana (Kansalliskirjasto 2015b. Suomen kansallisbibliografia Fennica). Koska teos kertoo yhdestä arkkipiispasta, on yksikkömuotoinen sana hyvä teoksen vieressä, mutta teos ei löydy jos joku hakee YSA:n mukaisilla asiasanahauilla 'arkkipiispat' tai 'piispat'.

Historiateoksissa on usein liikaa henkilöitä asiasanoiksi. Henkilön nimen ottamisessa asiasanojen joukkoon yhtenä kriteerinä voi pitää sitä, että asiasanana olevasta henkilöstä on oma väliotsikkonsa. Ehkä nimien käyttämisessä asiasanoina voi miettiä sitä, miten todennäköisesti joku hakee dokumentteja kyseisillä henkilönimillä. Tämän ajattelutavan ongelmana on, että se johtaa helposti itselle tutujen nimien pitämiseen tärkeämpinä kuin itselle vieraiden.

Teoksen *Yrjö Neumark* aiheena on saksalainen Georg Neumark. Kun tietää, että Yrjö on perinteinen suomalainen käänös nimille George ja Georg, asian voi tarkistaa British Libraryn tietokannasta, jossa on teoksia Georg Neumarkista (British Library 2015). Digitoidussa kokoelmassa on myös Martti Lutherin vaimon Katharina von Boran elämäkerta (*Katariina von Bora. Martti Lutherin puolison lyhy elämä-kerta*). Hänen nimestään British Libraryn tietokanta käyttää muotoa 'Luther, Katharina von Bora' (British Library 2015), jonka etuna on löytyvyys sekä puolison että omalla sukunimellä.

On selkeää käyttää ulkomaisessa muodossa ainakin niitä henkilöiden nimiä, jotka eivät ole perinteisesti vakiintuneet käytettäväksi suomennetussa muodossa. Kunninkaallisten nimet ovat tyypillisiä vakiintuneita käännöksiä. Uskonpuhdistaja Lutherin nimi saattaa olla Suomessa vakiintunut Martiksi, mutta toisaalta hänen oikea etunimensä on Martin. Termin valinnan ratkaisee, että Fennica-tietokannassa käytetään asiasanana suomennettua nimeä 'Luther, Martti', johon käyttäjä ohjataan korvatusta termistä 'Luther, Martin' (Kansalliskirjasto 2015b. Suomen kansallisbibliografia Fennica).

Martin Lutheria vähemmän tunnettujen henkilöiden Katharina Lutherin ja Georg Neumarkin nimet eivät kuulu Fennican asiasanoihin (Kansalliskirjasto 2015b. Suomen kansallisbibliografia Fennica). Vaikeimpia nimiä asiasanoina ovat ne ulkomaisten hallitsijoiden nimet, joita ei ole Fennican asiasanoissa ja joista ainakin asiasanoitettavissa teoksissa käytetään suomalaistettua nimeä. Opinnäytteen tekijä kirjoitti joitakin tällaisia nimiä asiasanoihin, mutta merkitsi kirjastolle tiedoksi, etteivät ne ole virallisia asiasanoja.

4.3.5 Kauno- ja tietokirjallisuuden rajanveto

Turun kansankirjaston asiasanoittamattomia vanhoja aineistoja lukiessa pitää päätellä, onko jokin teksti kaunokirjallisuutta. Sitä ei välttämättä tiedä etukäteen, koska ei ole apuna luokitusta eikä takakannen tekstiä. Pitää päättää myös se, antaako asiasanaksi jonkin genren.

Joitakin kertomuksia lukiessa pitää miettiä, ovatko ne kaunokirjallisuutta vai historiallisten tapahtumien kuvauksia kerrottuina viihdyttävän yksityiskohtaisesti. Kansanvalistukselle on ominaista tarinoiden muotoon puettu opettavaisuus. Tietokirjallisuuden ja kaunokirjallisuuden välimuotoa edustavat myös fiktiivisten henkilöiden dialogin muodossa annetut neuvot maatalouden menetelmistä. Teos *Mikko Nerosen rikastuminen, eli täydellinen osoitus kuinka apilaita viljellään* (Lundqvist 1848) alkaa novellimaisena, mutta monissa kohdissa sitä erottaa tietokirjallisuudesta ainoastaan vuoropuhelun käyttö kerrontatapana. Niinpä asiasanaksi tuli 'dialogi' kaunokirjallisuuden sijaan. Kaunokin käyttö asiasanastona osoittautui riittämättömäksi, koska siellä ei ole termejä 'lannoittaminen' ja 'rehu' (Kirjastot.fi 2015b. Verkko-Kaunokki). Yleiseen suomalaiseen asiasanastoon kuuluu tämän kirjan sisältöä kuvaavia maataloustermejä.

Turun kansankirjaston teos *Yrjö Neumark* vaikutti kaunokirjallisuudelta dialogisine kerrontatapoineen, mutta kirjoittaja paljasti lopussa, että virsikirjassa on kyseisen virsirunoilijan tekemiä virsiä. Siinä vaiheessa piti vaihtaa asiasanaksi 'kaunokirjallisuuden' sijaan 'elämäkerrat' ja tarkistaa, että asiasanat kuuluivat Kaunokin lisäksi Yleiseen suomalaiseen asiasanastoon.

4.4 Esimerkkejä kirjojen asiasanoittamisesta

4.4.1 Pitkä matkakertomus

Turun kansankirjaston kokoelmaan kuuluu tutkimusmatkailija Henry M. Stanley'n kirjoittama paksu teos *Kongo – uusi vapaavaltio mustien maanosassa. 2. osa* (Stanley 1887). Itsestään selvin asiasana teokselle on 'Afrikka', vaikkei sitä lue

nimekkeessä tai muutenkaan toistuvasti. Teoksessa ei myöskään toistu sana 'siirtomaat', mutta Stanley kertoo eurooppalaisista Kongossa ja heidän 'asemistaan'. Historiallisen kontekstin huomioon ottaen asiasanoittaja päätteli 'siirtomaiden' olevan tärkeä asiasana kyseiselle teokselle ja tulkitsi 'asemat' siirtokunniksi.

Henry Morton Stanley oli Kongon valtaamista valmistelevalle Belgian kuninkaan Leopold II:n palveluksessa. Kongosta tulikin Belgian siirtomaa. Leopold II peitteli aikeitaan erityisesti Iso-Britannialta siten, ettei hän antanut Stanleyille suoraan tehtäväksi vallata jotakin Afrikan osaa kuninkaan nimiin, vaan pyysi Stanleytä tekemään tutkimusmatkan, jolla perustettiin siirtomaan hyödyntämisessä tarvittavia tukikohtia. Leopold halusi naamioida siirtomaasuunnitelmiansa toteuttamisen näyttämään ulkopuolisille filantropialta ja tieteelliseltä tutkimukselta, ja sana "vapaavaltio" kuului tähän peittelyyn. (Hochschild 2004, 73, 81, 84.) Stanley oli ammattimaisena matkakirjailijana toimimisen edelläkävijä, jonka kirjat olivat myyntimenestyksiä (Hochschild 2004, 40–41).

Stanleyn kirjan erityispiirteenä on aiheiden listaaminen jokaisen luvun alkuun avainsanojen tyyliä. Nämä aiheet ovat kuin väliotsikkoja, vaikkei tekstissä ole muita väliotsikkoja kuin lukujen nimet. Osa tiivistävistä ilmaisuista on avuksi asiasanoittamisessa, mutta toiset eivät selvennä sisältöä lukematta sitä. Asiasanoiksi sopivat aiheet, jotka toistuvat useita kertoja otsikoissa ja sisällössä.

Kuvaliitteessä kuvana 1. on ensimmäinen sivu Kongosta kertovasta kirjasta. Koska nide on kaksiosaisen teoksen toinen osa, sen ensimmäinen luku on vanhalla numerointitavalla "kuudeskolmatta luku", jonka nimi on "Mustallejoelle". (Stanley 1887, 1; Kuva 1.) Tämä kuten muutkaan matkakertomusten lukujen nimet eivät ole asiasanoittamisen kannalta kovin informatiivisia.

Ensimmäisen luvun alussa on tiivistävinä ilmaisuina seuraavia: "Bolobon kauppa – Päämiesten rikkaus – Manguru, Bolobon Rotschild - Luonto Ylä-Kongolla – Ikkävä matkalla – Vertailua matkustuksiin muilla suurilla joilla – Ei yhtään ruoan vaihtelua – Eksymme tieltä – Hoi weijansit, sukulaiset Lukolelassa! – Kylmä vastaanotto – Tilapäiset markkinat..." (Stanley 1887, 1; Kuva 1). Näiden ilmaisujen perusteella voi päätellä tekstin käsittelevän muun muassa matkustamista,

luontoa ja kaupankäyntiä. Opinnäytteen tekijä olettaa, että weijansit ja nimet Bolobo ja Manguru eivät vapaan indeksoinnin termeinä kertoisi kovin olennaista käyttäjille, koska ne eivät ole asiasanoittajallekaan tuttuja. Fraaseille on ominaista, etteivät ne kerro paljoa sisällöstä ollessaan tiivistävinä ilmauksina erillään asiayhteydestään. Sellaiset asiat kuin eksyminen ja ikävä eivät todennäköisesti kuvaa kirjan sisältöä riittävän kattavasti tullakseen valituiksi asiasanoiksi.

Lukemisen aloittaminen vahvistaa vähitellen käsitystä aiheista. Kaupankäynnin toistuvuuden aiheena voi päätellä lukemisen lisäksi tiivistävistä ilmaisusta, joihin kuuluvat markkinat, rikkaus, ”Bolobon kauppa”, ”Irebu, etevimpien kauppiasten koto”, ”Suuret kauppatedot” ja ”Rikkauden kasvattamiskeinot”. (Stanley 1887, 1; Kuva 1). Lueskelu ja myöhempien lukujen tiivistävät ilmaisut vahvistavat kaupankäynnin keskeisyyden kirjassa. Tekstissä käsitellään toistuvasti muun muassa myytäviä luonnontuotteita. Yleisessä suomalaisessa asiasanastossa kaupankäynti ei kuitenkaan ole asiasana, vaan verkkosanasto tunnistaa sanan ja kehoittaa käyttämään korvaavaa termiä ’kauppa’ (Kansalliskirjasto 2015a. Finto – Suomalainen asiasanasto- ja ontologiapalvelu). Myös ’sota’ sopii asiasanaksi, koska ensimmäisen luvun alussa on ilmaisut ”Sota Irebussa – Minua pyydetään rauhaa välittämään – sodan syyt” (Stanley 1887, 1; Kuva 1). Myöhemminkin kirjassa on sota yhtenä aiheena, joka tulee esiin ajoittain.

Stanley oli ankara, afrikkalaisia muun muassa kahleilla kiduttava työnjohtaja, joka näki afrikkalaiset aseellisen hyväksikäyttönsä kohteina ja ruoantoimittajina. (Hochschild 2004, 78, 84–85). Näiden tietojen kanssa yhtenevää on, että kaupankäynti ja sota kuuluivat hänen kirjansa teemoihin.

Kongo-kirjan ensimmäisen luvun leipäteksti alkaa: ”Kuten Ibaka minulle kertoi, ovat beijansit vanhan tarinan mukaan Ubangista ---”. Siitä kappale etenee joen kuvailemiseen. (Stanley 1887, 1; Kuva 1.) Teksti siis alkaa vierailta nimillä ja kuvailulla, jotka eivät saa sisällönkuvaillijaa heti ehdottamaan mitään asiasanaa. Ensimmäisestä luvusta alkaen käsiteltyjä asioita kiteytyy koko teoksen silmäilemisen perusteella termeiksi ’heimot’, ’joet’ ja ’henkilöt’. ’Henkilöt’ ei tässä yhteydessä tarkoita kuuluisuuksia, vaan nimettyjen yksilöiden toiminnasta kertomista.

Aluksi asiasanaksi päätyi 'alkuperäiskansat', mutta termin 'heimot' käyttöä perustele, ettei se ota kantaa siihen, mitkä väestöryhmät määritellään alkuperäiskansoiksi.

Pitkien kirjojen asiasanoittamisessa on olennaista erottaa kokonaisuutta ja toistuvia teemoja kuvailevat termit muunlaisista. Joillakin aukeamilla voi olla olennaisena sisältönä aihe, jonka käsittely ei välttämättä jatku pitkään eikä toistu kirjassa. Kuitenkin asiasanoiksi sopivat termit, joiden aiheen käsittelyä tekstissä on pitkään, mutta vain yhdessä luvussa.

Stanley'n kirjoissa näkyy hänen matkoihinsa kuulunut jatkuva havainnointi, sillä hän on kirjoittanut kirjoihinsa taipaleiden pituuksia, lämpötiloja, järvien syvyyksiä sekä pituus- ja leveysasteita. Kantajat kuljettivat matkoilla ilmapuntareita, lämpö- ja askelmittareita. Stanley oli kuin maanmittari, joka kartoitti mannerta sen tulevien omistajien hyväksi. Hänen kirjoihinsa kuuluu luetteloita, kaavioita ja runsaasti yksityiskohtia kuten hintoja, koska matkakirjailija halusi viihdyttää lukijoitaan pikkuseikoilla maanosasta, josta lukijat eivät tienneet ennestään juuri mitään. Hänen kirjoitustyyliinsä on myös melodramaattisuutta hänen kertoessaan uhkarohkeasta löytöretkestä. (Hochschild 2004, 66, 41.)

Stanley käsittelee afrikkalaista ilmastoa seikkaperäisesti; hän muun muassa esittelee luokituksen siitä, millaista on kuuma ja erittäin kuuma, liittää kirjaansa taulukoita lämpötilasta, ilmapuntarin lukemista, tuulen suunnasta ja pilvisyydestä sekä kertoo kuumuuden aiheuttamista vaikeuksista, kuten kuumille maille tyypillisistä taudeista. Ilma on myös yhden luvun nimi kirjassa. (Stanley 1887, 234–259.)

Kongo-teoksessa on erilaisia osioita. Siinä kerrotaan henkilöiden toiminnasta ja kuvaillaan luontoa ja maisemia kiinnostavasti kuin kaunokirjallisuudessa. Loppupuolella on listoja, muun muassa taulukkoja ilmastosta ja sivu toisensa jälkeen lakitekstejä ja valtiollisia sopimuksia. (Stanley 1887.) Asiasanoiksi sopivat 'lait', 'julistukset', 'taulukot' ja 'kuvitukset', koska niiden kuulumisen kirjaan voi havaita jo silmäilyn perusteella. Tekstisisällön lisäksi kirjan visuaalinen ilme vaikuttaa siinä vaikuttaa asiasanojen valintaan.

Vaihtelevia lukuja sisältävä kirja sopii esimerkiksi siitä, ettei asiasanoitusta pidä tehdä lukemalla kirjaa vain jonkin verran alusta. Toisaalta yli 300-sivuista kirjaa ei tarvitse tehokkaassa työssä lukea kokonaan, vaan sieltä täältä lukeminen riittää. Opinnäytteen tekijä luki kirjan alkua, lukujen alussa olevia tiivistäviä ilmaisuja, lukujen alkua ja noin joka kymmenettä aukeamaa sekä selaili yleisilmettä nähdäkseen myös taulukot ja lakitekstit.

Varmuutta työhön antoi, että silmäilyssä tekstissä toistuivat samat aiheet, jotka oli jo aiemmin luetun perusteella valittu asiasanoiksi. Aiheeksi merkittiin 'Afrikka' ja muiksi asiasanoiksi 'siirtomaat', 'siirtokunnat', 'tutkimusmatkailijat', 'matkat', 'kauppa', 'luonto', 'joet', 'heimot', 'sodat', 'henkilöt', 'eurooppalaiset', 'ruoka', 'taudit', 'ilmasto', 'sopimukset', 'julistukset', 'lait', 'kuvitus' ja 'taulukot'.

'Tutkimusmatkailijat' on projektin lopussa lisätty asiasana, jonka valintaperusteena on kokoelman sisäinen yhdenmukaisuus termien käytössä. Kokoelmaan sisältyy toinenkin Stanleyestä kertova kirja, jonka aiheeksi tietokantaan on ennestään merkitty 'tutkimusmatkailijat' (Lagus 1889. *Henry M. Stanley'n nuoruus ja ensimmäiset tutkimusmatkat Afrikan sisämaissa*). Niinpä molempien kirjojen on löydettävä samalla olennaisella termillä.

4.4.2 Monta ihmiskohtaloa lyhyessä romaanissa

Esimerkkinä kansankirjaston asiakkaiden saatavilla olleesta kaunokirjallisuudesta on *Koivuluoto ja Pohjakoski. Hupainen kertomus muutamista hyvistä sydämmistä* (1861). Tässä yhteydessä hupaisuus ei tarkoita huumoria, vaan viihdekirjallisuutta, jossa on onnellinen loppu ja joka ei ole vaikeaselkoista korkeakirjallisuutta.

Kyseisessä pienuisromaanissa on muitakin nykyisestä kaunokirjallisuudesta poikkeavia ominaisuuksia. Teokseen ei ole merkitty kirjoittajan nimeä eikä nimi-merkkiä. Nimiösivulla lukee, että kirja on suomennettu, mutta kääntäjäkin on anonyymi. Suomentaja on muokannut tarinaa kotoisammaksi muuttamalla henkilöiden ja paikkojen nimet ovat suomalaisiksi.

Koivuluoto ja Pohjakoski jakautuu numeroituihin, nimettömiin lukuihin, joiden väliset siirtymät ovat usein yllättäviä henkilöiden ja teemojen vaihtumisen takia. Tällaista siirtymää kuvaa kuvaliitteessä oleva kuva 2. (*Koivuluoto ja Pohjakoski* 1861, 8–10.) Ensimmäisen luvun perusteella asiasanoiksi näyttivät sopivan 'lesket', 'orvot', 'palvelusväki' ja 'naiset'. Lopulliseen asiasanoitukseen ei kuitenkaan tullut 'leskiä' ja 'orpoja', koska asiasanoittaja ei arvioinut heitä riittävän keskeisiksi aiheiksi teoksessa.

Seuraava luku kertoo tehtaan konttorin henkilökunnasta, ja työnjohtaja on luvun keskeisin henkilö. Kuvailtaan muun muassa hänen keskusteluaan ruukin omistajan kanssa. Toisen luvun lukemisen jälkeen kirja tuntuu ikään kuin novellikokoelmalta, jota yhdistää yksi teema: rautaruukki. Jo ensimmäisessä luvussa esitelty palvelijattareksi menevä orpopyttö on ruukin sepän tytär. (*Koivuluoto ja Pohjakoski* 1861, 3–18.) Asiasanojen lista siis täydentyy 'ruukeilla' ja 'johtajilla'.

Kolmannen luvun aiheina ovat työnjohtajan koti, perhe, uskonnolliset perheenjäsenet ja ruukinsaarnaaja (*Koivuluoto ja Pohjakoski* 1861, 18–28). Asiasanoiksi tulevat 'perheet', 'papit' ja 'herännäisyys'. Neljännen luvun päähenkilö on työnjohtajan sukulainen, joka on saamaton mies vaikka yliopistossa opiskellut. Sitten tarina palaa alussa esiteltyyn palvelijattareen, joka tutustuu saamattomaan oppineeseen ja saa hänet opettamaan itseänsä. Sekä lukutaidottoman palvelijattaren että joutilaan maisterin elämä muuttuu, ja oppinut, hyvä palvelijatar pääsee isäntäväkensä suojatiksi. (*Koivuluoto ja Pohjakoski* 1861, 29–48.) Tiivistäen tätä voi kuvata sanoilla 'opiskelu' ja 'elämänmuutokset'. Suojattina olemiselle ei ole helppo keksiä asiasanaa, mutta sellainen voisi olla 'hyväntekeväisyys', kun kyseessä ei ole pienten lasten adoptointi. Harkinnan jälkeen 'hyväntekeväisyys' ei tullut valituksi asiasanojen joukkoon, koska se voisi antaa tietokannan käyttäjälle väärää käsitystä teoksen sisällöstä.

Romaanin loppupuolella esille astuu palvelijattaren isäntäväen poika, joka on ollut merimiehenä ja kadonnut haaksirikossa. Hän kertoo vaiheistaan Ranskassa ja avioituu opinhaluisen palvelijattaren kanssa. (*Koivuluoto ja Pohjakoski* 1861, 49–68.) Asiasanoihin kuuluvat edellisten lisäksi 'merimiehet' ja 'rakkaus'.

Pienoisromaani esittelee erilaisia henkilöitä, vaikka on kooltaan tiivis; vain 68-sivuinen. Kaunokirjallisuudelle ominaiseen tapaan siihen sisältyy pitkiä tunteiden ja näkymien kuvailuja, joita lukiessa asiasanojen lista ei täydenny. Asiasanoittamista helpottaa tekstin vaihtelevuus, sillä se ei tunnu tasaiselta massalta, josta ei erotu asioita. Päinvastaiset kirjat ovat hankalampia asiasanoitettavia.

5 LOPUKSI

5.1 Asiasanaprojektin lopputulos

Digitoidut Turun kansankirjaston kokoelman kirjat saivat asiasanoitusprojektissa tyypillisesti 8–13 asiasanaa kuvailemaan yksittäistä kirjaa. Vähimmillään kirjaa kuvailee viisi asiasanaa ja enimmillään noin kolmekymmentä, jos Suomen historian yleisesityksistä ottaa hallitsijoiden nimiä asiasanoiksi. Ilman hallitsijoiden nimilistoja asiasanojen määrä ei nouse enimmilläänkään yli kahdenkymmenen. Keskiarvona on laskettavissa, että tätä opinnäytettä tehdessä asiasanoitettu kirja sai keskimäärin noin 10,4 termin laajuisen kuvailun. Vähäisimmän määrän asiasanoja saaneet teokset ovat kaunokirjallisuutta ja lyhyitä maatalousoppaita. Paljon asiasanoja saivat esimerkiksi pitkät matkakirjat, lukuisia eläinlajeja esittelevä kirja sekä kolme pienoisoromaania sisältävä kaunokirjallinen kokoomateos.

Toimeksiantajan tilaaman työn tekeminen ja opinnäytetyön kirjoittaminen siitä on ollut projekti, jolla oli rajallinen laajuus ja joka päättyi tultuaan valmiiksi. Työn päätyttyä sitä arvioidaan. Projekti valmistui melko nopeasti; opinnäytetyön tekijä antoi kansankirjaston kokoelman digitoiduille kirjoille asiasanat runsaan kahden kuukauden aikana ja asiasanoitti toisinaan puolessa päivässä muutaman lyhyen kirjan. Raportointi vaati aikaa asiasanoittamisen jälkeenkin. Oli hyvä käytäntö järjestää asiasanoja mielekkääseen järjestykseen jo niitä valitessa. Toimeksiantaja salli indeksoijalle paljon omaa harkitsemisen varaa ja vastuuta lopputuloksesta antamalla vain niukasti kommentteja, joiden perusteella eräät asiasanat tulivat poistetuiksi.

Asiasanojen antamisen jälkeen indeksoijan oli helpompi kyseenalaistaa joidenkin kirjoittamiensa asiasanojen käytön perusteltavuutta kuin lisätä uusia termejä kuvailemaan kirjoja. Harkinnan jälkeen indeksoija poisti oma-aloitteisesti muutamia turhiksi arvioimiaan asiasanoja. Teoksen *Ei ole terveydeksi* (Berger 1886) asiasanoista karsittiin 'matkustaminen', koska vain pieni osa kirjasta sisältää matkustamista koskevia terveydenhoito-ohjeita. Lastentautien hoito-oppaan (Roos

1857. *Kristillinen ja terveellinen lasten kasvattaminen. Toinen osa...*) asiasanoista karsittiin 'aivot', koska niitä käsiteltiin kirjassa niin vähän. *Kiinalaisista-*teokselle aiemmin annettu asiasana 'virkamiehet' näytti kirjaa uudelleen lukiessa turhalta.

Kirjastojen kirjaluetteloissa uskonnollista kirjallisuutta luonnehtivat aluksi sanat 'uskonto' tai 'kristinusko'. Harkittuaan asiasanoittaja korvasi ne aiemmin käyttämiään sanoja tarkemmalla termillä 'kristillinen kirjallisuus', joka otettiin johdonmukaisuuden vuoksi kuvailemaan kaikkia kirjastoluetteloja. Kirjaluetteloiden oman jäsenyyksen mukaan niihin kuuluu uskonnollisia kirjoja, joista enemmistö on kristillisiä (esimerkiksi Folkbiblioteket i Åbo 1892. *Svenska böcker*; Turun kaupungin kansankirjasto 1897. *Suomalaisia kirjoja. II vihko*). Johdonmukaista on myös lisätä Stanleyistä Kongossa kertovan teoksen (Stanley 1887. *Kongo – uusi vapaa-valtio mustien maanosassa*) kuvailuun termi 'tutkimusmatkailijat', koska muutkin samanlaiset kirjat ovat saaneet sen asiasanan. Toinen kokoelman sisäistä johdonmukaisuutta lisäävä seikka on termin 'oppikirja' käyttäminen kaikista koulu-kirjoista eikä vain joistakin. Projektin lopussa tehtyjen korjausten perusteluina oli poistaa termejä, joiden aihetta teos käsittelee liian vähän sekä lisätä asiasanojen tarkkuutta ja niiden käytön johdonmukaisuutta.

Kokoelmien digitointi ja asiasanoittaminen on hyödyllistä. Digitoimalla vanhoja kokoelmiaan kirjastot tarjoavat asiakkaille mahdollisuuden lukea vanhoja kirjoja kotona riippumatta siitä, minkä paikallisen kirjaston asiakkaita he ovat. Digitointi myös edistää vanhojen teosten säilyvyyttä ja vähentää niiden varastosta hake-
misen vaatimaa työaikaa kirjastoissa.

Asiasanat ovat tiedon etsimisen apuväline, joka lisää tietokannan käyttäjien mahdollisuuksia hakea tietoa heitä kiinnostavista aiheista sekä hakutoimintoa käyttämällä että selailemalla. Selailijat saavat dokumentin asiasanoista nopeasti yleiskuvan teoksen sisällöstä ja voivat päätellä, kiinnostaako heitä lukea kirjaa. Asiasanastojen käyttö ehkäisee teosten katoamista tiedonhakijoiden ulottuvilta verrattuna siihen, jos jokainen indeksoija antaisi dokumenteille mielensä mukaisia ja joskus epävirallisiakin asiasanoja, joita muiden ihmisten on hankala päätellä.

5.2 Asiasanastojen käyttö vanhojen kirjojen kuvailuun

Yleinen suomalainen asiasanasto, Kaunokki, Allärs ja Bella säätelevät sanojen käyttöä asiasanoituksessa. Asiasanastojen avulla sisällönkuvailija voi varmistua, ettei hän käytä asiasanoina sellaisia teoksissa ilmeneviä sanoja, jotka eivät kuulu nykyiseen vakiintuneeseen kirjakieleen. Nykyajan ilmiöt ja nykykieli asiasanojen valinnan lähtökohtina voivat asettaa rajoituksia valitessa asiasanoja vanhoille kirjoille. Opinnäyteprojektissa tämä ei muodostunut huomattavaksi ongelmaksi, mutta vanhojen kirjojen asiasanoittamisessa olisi mahdollista hyödyntää menneisyyden aiheita kuvaavia erikoissanastoja. Opinnäytteen toimeksiantaja halusi kuitenkin käytettävän Yleistä suomalaista asiasanastoa ja Kaunokkia sekä niiden ruotsinkielisiä vastineita.

Nykyaikaisten sanojen käyttö voidaan usein nähdä käyttäjäsuuntautuneisuutena, koska asiasanoina on käytettävä tiedonhakijoiden ymmärtämiä sanoja. Sanaston ennakoitavuutta lisää, ettei tarvita erikoistermejä eikä harvoin käytettyä asiasanastoa. Toisaalta vanhaa aineistoa tietokannasta etsivät ihmiset saattavat yrittää käyttää vanhanaikaisiakin hakusanoja.

Menneisyyttä käsittelevien kirjojen huomioon ottamista osoittaa, että Kaunokkiin sisältyy esimerkiksi asiasana 'naimakaupat' (Kirjastot.fi 2015b. Verkko-Kaunokki). Esimerkki indeksoinnin kannalta haastavasta vanhasta sanasta on 'virma'. Tämä hevosten hoito-oppaan kirjoittajan käyttämä sana ei kuulu Yleiseen suomalaiseen asiasanastoon edes niin, että sanasta olisi viittaus suositeltavaan termiin (Kansalliskirjasto 2015a. Finto – suomalainen sanasto- ja ontologiapalvelu).

Vanhalla aineistolla on erityispiirteitä asiasanoitustyön kannalta. Vanhoja termejä ei voi ottaa sellaisinaan asiasanoiksi eikä aina ole helppoa tietää niiden vastinetta nykykielellä. Käytettäessä henkilöiden nimiä asiasanoina on joissakin tapauksissa tarpeen selvittää, onko nykyään vakiintunut asiasanaksi henkilön vieraskielinen vai suomennettu nimi. 1800-lukuisessa kirjakokoelmassa ei ole takakansi-

tekstejä eikä aina edes sisällysluetteloja ja väliotsikkoja tai jos on, ne eivät välttämättä jäsenä kirjaa samoin kuin uudemmissa teoksissa. Vanhan aineiston ominaisuuksiin kuuluu myös tieto- ja kaunokirjallisuuden liukuva raja.

5.3 Oppimista ja haasteita

Asiasanojen valinnassa on kyse teoksen sisällön tiivistämisestä termeihin. Tässä työssä on haasteita erityisesti silloin, jos teoksessa kerrotaan monista asioista vähän tai jos kaunokirjallisen teoksen sisällöstä suuri osa on kuvailua ja henkilöitä on monia. Asiasanojen antamiseen tarvitaan taitoa erottaa teksteistä niiden olennaisimmat asiat. Teoksissa käytettyjen sanojen poimimisen lisäksi on tarvittaessa valittava sisältöä kokoavia yläkäsitteitä ja teemoja, kirjallisuudenlajia kuvaavia termejä sekä vanhojen sanojen tilalla nykyisin käytettäviä ilmauksia. Vaikka näissä taidoissa on kehitettävää, opinnäytetyön tekijä havaitsi rutinoituvansa asiasanojen valintaan projektinsa edetessä.

Tavallisimpia pohdittavia kysymyksiä asiasanoja valitessa ovat, onko asiasanoja liian monta ja onko sisällönkuvailun spesifiyden tai pelkistämisen aste sopiva. Kyse on siis sisällönkuvailun tarkkuudesta ja tyhjentyvyydestä. Tavallinen kohdattu ongelma on sellainen yleisellä tasolla tehty asiasanoitus, jossa sisältöä kuvailevien aiheiden käsittely olisi helposti pääteltävissä jo nimekkeestä ja lyhyesti ilmaistusta aiheesta.

Hyvänä periaatteena voi pitää kokoelman asiasanojen sisäistä yhtenäisyyttä. Tavoitteena on, että samankaltaiset teokset löytyvät samoilla asiasanoilla eikä kuvailun yksityiskohtaisuus vaihtele suuresti kokoelman eri kirjojen välillä. Johdonmukaista termien käyttöä vaikeuttaa, etteivät kaikki termit kuulu sekä Yleiseen suomalaiseen asiasanastoon että Kaunokkiin, jolloin samaa termiä ei aina voi käyttää sekä tieto- että kaunokirjallisuuden kuvailuun. Termien käyttö on kuitenkin helpompaa pitää yhdenmukaisena kuin asiasanojen määrä.

Asiasanojen valintaprosessissa oli vaiheita, joissa asiasanoittaja luki tai silmäili jotakin teosta melko pitkään eikä kuitenkaan lisännyt uusia termejä kuvailuun. Se ei kuitenkaan ole välttämättä ongelma vaan kuuluu pelkistämiseen. Työskentelyä

voi arvioida tehokkuuden ja tarkkuuden näkökulmista. Tehokkuutta lisäisi, jos asiasanoittaja kävisi läpi lyhyitä ja itseään kiinnostavia tekstejä ja kaunokirjallisuutta silmäilytekniikalla, mutta pisimpien kirjojen sisällönkuvailussa tarkkuutta voisi lisätä lukemalla niiden tekstiä kattavammin.

Asiasanastojen käytöstä huolimatta mietittäväksi jää, mitä tiedonhakijat ymmärtävät termien sisällöksi ja voisiko jokin muu asiasana sittenkin kuvata kirjassa esiintyvää ilmiötä paremmin kuin indeksoijan mieleen ensimmäiseksi tuleva. Näissä valinnoissa auttavat asiasanastoihin kuuluvat rinnakkaiset termit ja assosiativiset käsitteet. Ehdotuksia jatkotutkimuksiksi ovat kahden eri indeksoijan samoille kirjoille antamien asiasanojen vertailu ja asiasanojen valinnan onnistuneisuuden arviointi sen perusteella, miten ne hyödyttävät Digi-tietokannan käyttäjiä. Voitaisiin tutkia tehtyjä hakukyselyjä ja tulosten perusteella pyrkiä kehittämään indeksointia vastaamaan yhä paremmin käyttäjien tarpeisiin.

LÄHTEET

Adresskalender för Åbo stad 1878. 1878. Turku: Åbo Boktryckeri-aktiebolag. DIGI - Yleisten kirjastojen digitoitu aineisto. Viitattu 17.4. 2015. Osoitteessa <http://digi.kirjastot.fi/items/show/8019>.

Adresskalender för Åbo stad 1880-1881. 1880. Turku: Åbo Boktryckeri-aktiebolag. DIGI - Yleisten kirjastojen digitoitu aineisto. Viitattu 17.4. 2015. Osoitteessa <http://digi.kirjastot.fi/items/show/8020>.

Adresskalender för Åbo stad 1884-1885. 1884. Turku: Åbo Boktryckeri-aktiebolag. DIGI - Yleisten kirjastojen digitoitu aineisto. Viitattu 17.4. 2015. Osoitteessa <http://digi.kirjastot.fi/items/show/8021>.

Adresskalender för Åbo stad 1886-1887. 1886. Turku: Åbo Boktryckeri-aktiebolag. DIGI - Yleisten kirjastojen digitoitu aineisto. Viitattu 17.4. 2015. Osoitteessa <http://digi.kirjastot.fi/items/show/8022>.

Adresskalender för Åbo stad 1888-1889. 1888. Turku: Åbo Boktryckeri-aktiebolag. DIGI - Yleisten kirjastojen digitoitu aineisto. Viitattu 17.4. 2015. Osoitteessa <http://digi.kirjastot.fi/items/show/8023>.

Ansgario. 1864. J.W. Lilljan ja kumpp. kustantama. DIGI - Yleisten kirjastojen digitoitu aineisto. Viitattu 17. 4. 2015. Osoitteessa <http://digi.kirjastot.fi/items/show/120283>.

Aitchison, Jean & Gilchrist, Alan & Bawden, David 2004: Thesaurus construction and use: a practical manual. 4. painos. London; New York : Europa Publications, 2004, cop. 2000.

Alaterä, Anu & Halttunen, Kai 2002. Tiedonhaun perusteet – osa lukutaitoa. Helsinki: BTJ Kirjastopalvelu Oy.

Benito, Miguel 1995. Bibliografisk kontroll. En introduktion till katalogisering, klassifikation och indexering. 2. painos. Borås: Tarancos bokförlag.

Berger, P 1886. Ei ole terveydeksi. Terveyden sääntöjä jokaiselle. Helsinki: K.E. Holm. DIGI - Yleisten kirjastojen digitoitu aineisto. Viitattu 20.4. 2015. Osoitteessa <http://digi.kirjastot.fi/items/show/7956>.

British Library. Kokoelmätietokanta. Viitattu 19.1.2015. <http://www.bl.uk/#>

Chowdhury, G.G. 2010. Introduction to Modern Information Retrieval. 3. painos. Lontoo: Facet Publishing.

Elehvantin kesyttäjä. Kertomus Intiasta.1893. Jyväskylä: Jyväskylän kirjapaino. DIGI - Yleisten kirjastojen digitoitu aineisto. Viitattu 13. 4. 2015. Osoitteessa <http://digi.kirjastot.fi/items/show/8115>.

Folkbiblioteket i Åbo 1892. Svenska böcker. Turku: Åbo Boktryckeri Aktiebolag. DIGI - Yleisten kirjastojen digitoitu aineisto. Viitattu 7. toukokuuta 2015. Osoitteessa <http://digi.kirjastot.fi/items/show/8025>.

Foskett, Antony Charles 2007. Subject approach to information. 5. painos. Lontoo: Facet Publishing.

Hallstén, Alexander Gustaf Julius 1872. Suomen historia. Turku: G.W.Wilénin ja kumpp. kirjapaino. DIGI - Yleisten kirjastojen digitoitu aineisto. Viitattu 13. 4. 2015. Osoitteessa <http://digi.kirjastot.fi/items/show/120309>.

Hassel, Antti Fredrik 1884. Jaakko Cook`in matkat Tynnellä merellä. Helsinki: Kansanvalistus-seura. Viitattu 11.3. 2015. DIGI - Yleisten kirjastojen digitoitu aineisto. Osoitteessa <http://digi.kirjastot.fi/items/show/7980>.

Hedman, Thomas 2000. Metadata. En bok om hur webbdokument kan katalogiseras och göras sökbara. Lund: Studentlitteratur.

Helsingin yliopiston kirjasto 1999. YSA:n käyttöohjeet. Viitattu 22.4.2015. <http://vesa.lib.helsinki.fi/ysa/ysakaytto.html>

Hochschild, Adam 2004. Kuningas Leopoldin haamu. Kertomus ahneudesta, terrorista ja sankaruudesta siirtomaa-ajan Afrikassa. Alkuteos ilmestynyt 1998.Helsinki:Tammi.

Hywää tarkoittavia varoituksia wiinaa vastaan, sekä Kertomus juopumuksen onnettomista seurauksista, muutamista totisista tapauksista esiin asetettuja. 1859. Turku: J.W. Lilljan kirjapaino. DIGI - Yleisten kirjastojen digitoitu aineisto. Viitattu 2.4. 2015. <http://digi.kirjastot.fi/items/show/8085>.

Kansalliskirjasto 2015a. Finto - Suomalainen asiasanasto- ja ontologiapalvelu. Viitattu 21.1.2015. Osoitteessa <http://finto.fi/ysa/fi/index>

Kansalliskirjasto 2015b. Suomen kansallisbibliografia Fennica. Viitattu 16.3.2015. Osoitteessa <https://fennica.linneanet.fi>

Kertoelmia koirista, kuwapiirroksilla. 1856. Vaasa: Gröhnin kirjapaino. DIGI - Yleisten kirjastojen digitoitu aineisto. Viitattu 17.4. 2015. Osoitteessa <http://digi.kirjastot.fi/items/show/7936>.

Kiinalaisista. 1859. Turku: J.W. Lilljan kirjapaino. DIGI - Yleisten kirjastojen digitoitu aineisto. Viitattu 3.4. 2015, <http://digi.kirjastot.fi/items/show/120285>.

Kirjastot.fi 2015a. Digi – yleisten kirjastojen digitoidut aineistot. Viitattu 21.1.2015.

Turun kansankirjaston kokoelman digitoitujen aineistojen osoite on: <http://digi.kirjastot.fi/items/browse?collection=25>

Kirjastot.fi 2015b. Verkko-Kaunokki. Viitattu 30.1.2015. <Http://kaunokki.kirjastot.fi/>

Koivuluoto ja Pohjakoski. Hupainen kertomus muutamista hyvistä sydämmistä. 1861. Turku: J.W. Lilljan kirjapaino. Viitattu 16.3. 2015. DIGI - Yleisten kirjastojen digitoitu aineisto. Osoitteessa <http://digi.kirjastot.fi/items/show/7977>.

Kotieläimet ja kotitarwekaswut, 1861. Turku: J.W.Lilljan kirjapaino. DIGI - Yleisten kirjastojen digitoitu aineisto, Viitattu 3. 4. 2015, <http://digi.kirjastot.fi/items/show/8086>.

Kristensen, Jaana & Järvelin, Kalervo 1990. Effectiveness of a searching thesaurus in free-text searching of a full-text database. Tampere: Tampereen yliopisto.

Kristillinen päiväkirja. Ilmoittawa kristin uskokunnan historiasta tietoja pyhistä martyyreistä ja muista ihmisistä, joitten nimet ovat vuosittain almanakassa. Niin myöskin vähäinen tutkinto sielun huvituksesta. 1857. Turku: Frenckellin kirjapaino. DIGI - Yleisten kirjastojen digitoitu aineisto, Viitattu 3.4. 2015, <http://digi.kirjastot.fi/items/show/7986>.

Kuwaelemia juopumuksesta ja raitteudesta. Ensimmäinen wihko: Juomari ja hänen tyttärensä. 1854. Helsinki: Frenckellin kirjapaino. DIGI - Yleisten kirjastojen digitoitu aineisto. Viitattu 2. 4. 2015, <http://digi.kirjastot.fi/items/show/8070>.

Kuwaelmia juopumuksesta ja raitteudesta. Toinen wihko: Ei, sanoi Marketta.1858. Helsinki: Frenckellin kirjapaino DIGI - Yleisten kirjastojen digitoitu aineisto. Viitattu 2.4. 2015, <http://digi.kirjastot.fi/items/show/8073>.

Lagus, F.H.B. 1889. Henry M. Stanleyn nuoruus ja ensimmäiset tutkimusmatkat Afrikan sisämaissa. Helsinki: Kansanvalistusseura. DIGI - Yleisten kirjastojen digitoitu aineisto. Viitattu 20.4. 2015. Osoitteessa <http://digi.kirjastot.fi/items/show/7978>.

Lundeqvist, Nils Wilhelm 1848. Mikko Nerosen rikastuminen, eli täydellinen osoitus kuinka apilaita viljellään”. Hämeenlinna: Nordensvan. DIGI - Yleisten kirjastojen digitoitu aineisto, viitattu 3.4. 2015, <http://digi.kirjastot.fi/items/show/7945>.

Luonnon oppi. 1865. Turku: J.W. Lillja ja kumppanit. DIGI - Yleisten kirjastojen digitoitu aineisto. Viitattu 3.4. 2015, <http://digi.kirjastot.fi/items/show/7979>.

Näätsaari, Inkeri 2012. Kaupungin valo. Turun pääkirjasto Siriuksen korttelissa. Turku: Enostone.

Olson, Hope A. 2002. The Power to Name : locating the limits of subject representation in libraries. Dordrecht : Kluwer Academic.

Rein, Gabriel 1864. "Henrik Gabriel Porthan." Helsinki: Finska Litteratursällskapetets tryckeri. DIGI - Yleisten kirjastojen digitoitu aineisto, Viitattu 2. 4. 2015, <http://digi.kirjastot.fi/items/show/7988>.

Roos, Samuel 1856. "Kristillinen ja terveellinen lasten kasvattaminen. Edellinen osa: Lasten terweinä pitäminen," Helsinki: Suomalaisen Kirjallisuuden Seura. DIGI - Yleisten kirjastojen digi-toitu aineisto, Viitattu 3. 4 2015, <http://digi.kirjastot.fi/items/show/8057>.

Roos, Samuel 1857. "Kristillinen ja terveellinen lasten kasvattaminen. Toinen osa: Lasten kiwut, niiden syyt ja holhous," Helsinki: Suomalaisen Kirjallisuuden Seura. DIGI - Yleisten kirjastojen digitoitu aineisto. Viitattu 3.4. 2015, <http://digi.kirjastot.fi/items/show/8058>.

Rowley, Jennifer & Hartley, Richard 2008. Organizing Knowledge. An Introduction to Managing Access to Information. 4. painos. Aldershot: Ashgate.

Samuelsson, Sven 1861. "Jaloimman ja hyödyllisimmän luontokappalemme, hewoisen tautien tunteminen ja parantaminen," Turku: J.W. Lilljan kirjapaino. DIGI - Yleisten kirjastojen digitoitu aineisto. Viitattu 3.4. 2015, <http://digi.kirjastot.fi/items/show/7987>.

Schubert, Gotthilf Heinrich von 1858. "Naturkunnighet för skolor och familjer. I, Stjernkunskap," DIGI - Yleisten kirjastojen digitoitu aineisto, Turku: J. W. Lillja. Viitattu 3. 4. 2015, <http://digi.kirjastot.fi/items/show/8013>.

Seppälä, Eila 1963. Turun kaupunginkirjasto 1863–1963. Turku: Turun kaupunki.

Stanley, Henry M. 1887. Kongo – uusi vapaavaltio mustien maanosassa. 2. osa. Porvoo: Werner Söderström. Viitattu 16.3. 2015. DIGI - Yleisten kirjastojen digitoitu aineisto. Osoitteessa <http://digi.kirjastot.fi/items/show/8017>.

Sulkunen, Irma 1986. Raittius kansalaisuskontona. Raittiusliike ja järjestäytyminen 1870-luvulta suurlakon jälkeisiin vuosiin. Helsinki: Suomen Historiallinen Seura.

Suomen historia kansakouluille. Lyhykäisesti kerrottu.1873. 3. painos. Turku: G.W.Wilénin ja kumpp. kirjapaino. DIGI - Yleisten kirjastojen digitoitu aineisto. Viitattu 13. 4. 2015. Osoitteessa <http://digi.kirjastot.fi/items/show/120310>.

Suominen, Saarti & Tuomi 2009. Bibliografinen valvonta. Johdatus luetteloinnin ja sisällönku-
vailun menetelmiin. Helsinki: BTJ Kustannus.

Suonuuti, Heidi 2006. Sanastotyön opas. 2. painos. Helsinki: Sanastokeskus TSK ry.

Tarkiainen, Kari 2012. Porthan, Henrik Gabriel (1739 - 1804) Suomalaisen Kirjallisuuden Seura: Biografiakeskus. Viitattu 29.3.2015. <http://www.kansallisbiografia.fi/kb/artikkeli/2599/>

Taylor, Arlene G. 2004. The Organization of Information. 2. painos. Westport, Conn.: Libraries Unlimited.

Tukholman veri-sauna eli kuningas Kristiani II hirmu-työt Ruotsissa w. 1520. 1862. Turku:J. W. Lilljan kirjapaino. DIGI - Yleisten kirjastojen digitoitu aineisto. Viitattu 17. 4. 2015, Osoitteessa <http://digi.kirjastot.fi/items/show/120307>.

Turun kaupungin kansankirjasto 1897. Suomalaisia kirjoja. II vihko. Turku: Åbo Tidningin kirja-
paino-osakeyhtiö. DIGI - Yleisten kirjastojen digitoitu aineisto. viitattu 7. toukokuuta 2015. Osoit-
teessa <http://digi.kirjastot.fi/items/show/7950>.

Turun kaupunki 2013. Turku.fi: Kirjaston historiaa Nummessa ja Runosmäessä. Viitattu 16.3.2015. Osoitteessa <http://www.turku.fi/public/default.aspx?contentid=453023&nodeid=4899>

Vakkari, Pertti 2006. Tiedonhankinnan tukeminen ja informaatiotutkimus. Teoksessa Ilkka Mäki-
nen (toim.) Tiedon tie. Johdatus informaatiotutkimukseen. 10. muuttamaton painos (1. painos
1999). Helsinki, BTJ Kirjastopalvelu Oy, 9–31.

Yli-Seppälä, Laura 2015. Turun kaupunginkirjaston vanhan kokoelman tutkimushankkeessa työkennelleen Suomen historian jatko-opiskelijan haastattelu. Elina Meriö haastatteli 23.2.2015.

Yrjö Neumark. 1859. Turku: J.W. Lilljan kirjapaino. DIGI - Yleisten kirjastojen digitoitu aineisto. Viitattu 3.4. 2015, <http://digi.kirjastot.fi/items/show/8171>.

Liite 1.

Lista digitoiduista ja asiasanoitetuista Turun kansankirjaston kokoelman kirjoista

Adresskalender för Åbo stad 1878. 1878

Adresskalender för Åbo stad. 1880-1881. 1880

Adresskalender för Åbo stad. 1884-1885 1884

Adresskalender för Åbo stad. 1886-1887 1886

Adresskalender för Åbo stad. 1888-1889 1888

Ansgario 1864

Benjamin Franklinin elämäkerta 1860

Berger, P: Ei ole terveydeksi. Terveyden sääntöjä jokaiselle 1886

Ei niin salattua, ettei ilmi tulisi. Tositapauksiin perustuva kertomus 1893

Elehvantin kesyttäjä. Kertomus Intiasta 1893

Esmarch, Johannes Friedrich August von: Ensimmäinen apu vahingoittuneelle

1884

Eurén, G.E: Kuinka nuoresta waimosta tuli kelpo ihminen 1856

Folkbiblioteket i Åbo. Svenska böcker 1892

Folkbiblioteket i Åbo. Svenska böcker. Första häftet 1874

Gossner, Johannes: Totinen kristillisyys. Lyhykäisissä tutkistelemuksissa jokaisena päivänä kuukaudessa, ja Rukouksia jokaisena päivänä viikossa (niille, jotka tahtowat lisääntyä Jesuksen Kristuksen tuntemisessa ja rakkaudessa) 1858

Hallstén, Alexander Gustaf Julius: Suomen historia 1872

Handbok för resande i Finland 1858

Hassel, Antti Fredrik: Jaakko Cook'in matkat Tyynellä merellä 1884

Hyvää tarkoittawia waroituksia wiinaa vastaan, sekä Kertomus juopumuksen onnettomista seurauksista, muutamista totisista tapauksista esiin asetettuja 1859

Isäntä ja palvelija. Siweys-opillinen tarina erästä rikkaasta, pahuuden tieltä jumalisuuteen hurskaalta palvelijaltansa saatetusta nuorukaisesta 1861

Kaksi kertomusta ihmiskunnan historiasta. Ihmisen julma kiittämättömyys ja pedon jalo kiitollisuus 1861

Katariina von Bora. Martti Lutherin puolison lyhy elämä-kerta 1860

Kertoelmia koirista, kuwapiirroksilla 1856

Kiinalaisista 1859

Koivuluoto ja Pohjakoski. Hupainen kertomus muutamista hyvistä sydämmistä 1861

Kolmenkymmenen wuotinen sota Saksassa 1859

Kosta' paha hywällä! Sydäntä liikuttawainen historia, kuinka eräs mies sai naapurinsa wähittäin kääntymään ilkeydestä hywyteen - nähtävästi tosi ja joka lualle sekä waroksi, että kehoitukseksi 1860

Kotieläimet ja kotitarwekaswut 1861

Kristillinen päiväkirja. Ilmoittawa kristin uskokunnan historiasta tietoja pyhistä martyyreistä ja muista ihmisistä, joitten nimet ovat vuosittain almanakassa. Niin myöskin vähäinen tutkinto sielun huvituksesta 1857

Kuwaelemia juopumuksesta ja raitteudesta. Ensimmäinen vihko: Juomari ja hänen tyttärensä 1854

Kuwaelmia juopumuksesta ja raitteudesta. Toinen vihko: Ei, sanoi Marketta 1858

Lagus, F.H.B: Henry M. Stanley'n nuoruus ja ensimmäiset tutkimusmatkat Afrikan sisämaissa 1889

Lammasten ruokkoomisesta, eli taitawa ja ymmärtäväinen Lammas-ukko 1853

Laulun woima. Murhe-tarina Itä-Suomesta 1859

Luonnon oppi 1865

Lundeqvist, Nils Wilhelm: Mikko Nerosen rikastuminen, eli täydellinen osoitus kuinka apilaita viljellään 1848

Lähteistä 1859

Myrtle, Harriet: Snögubben. Berättelser för små barn 1859

Oikea illan-wietto, eli Kortti-pelin muuttamisesta kirjain lukemiseksi 1857

Philp, Robert Kemp: Miten kopeekka kasvoi tuhanneksi ruplaksi. Kertomus Kei-
nolan Pekalle ja Kauppalan Paawolle rikastumista warten 1859

Rein, Gabriel: Henrik Gabriel Porthan 1864

Roos, Samuel: Kristillinen ja terveellinen lasten kasvattaminen. Edellinen osa:
Lasten terweinä pitäminen 1856

Roos, Samuel: Kristillinen ja terveellinen lasten kasvattaminen. Toinen osa: Las-
ten kiwut, niiden syyt ja holhous 1857

Samuelsson, Sven: Jaloimman ja hyödyllisimmän luontokappalemme, hewoisen
tautien tunteminen ja parantaminen 1861

Schalin, Zachris & Tucker, Charlotte & Ihle, Annette: Kyläkirjasto, lukemisia Suo-
men kansalaisille erinäisistä aineista. Kahdeksastoista vuosikerta

1890

Schubert, Gotthilf Heinrich von: Naturkunnighet för skolor och familjer. I, Stjernkunska
1858

Snell, Johan Peter Ludwig: Hywiä tapoja opettawaiset jutut yhteiselle kansalle ja
lapsille 1858

Stanley, Henry M: Kongo. Uusi vapaavaltio mustien maanosassa. 2. osa
1887

Suomen historia kansakouluille. Lyhykäisesti kerrottu 1873

Tukholman weri-sauna eli kuningas Kristiani II hirmu-työt Ruotsissa w. 1520
1862

Turun kaupungin kansankirjasto. Suomalainen kirjaluettelo (ei vuosilukua)

Turun kaupungin kansankirjasto. Suomalaisia kirjoja. II vihko 1897

Törmälä, Emanuel: Puujalka. Kertomus 1808 vuoden sodan ajoilta 1886

Wiattoman rukous ja Naiminen sillin tähden. Kaksi kertomusta 1856

Wirolainen 1862

Wolzogen, Ernst von: Kunniahousut 1893

Yhteinen ihmiskunnan historia. I. Frankkien kuningas, roomalainen keisari Kaarle Suuri ja Ruhtinas-sukunsa. 1861

Yhteinen ihmiskunnan historia. II. Paavi Gregorius VII ja Saksan keisari Henriikki IV 1862

Yhteinen ihmiskunnan historia. III. Kuningas Kaarle XII, selvästi kuvattu
1862

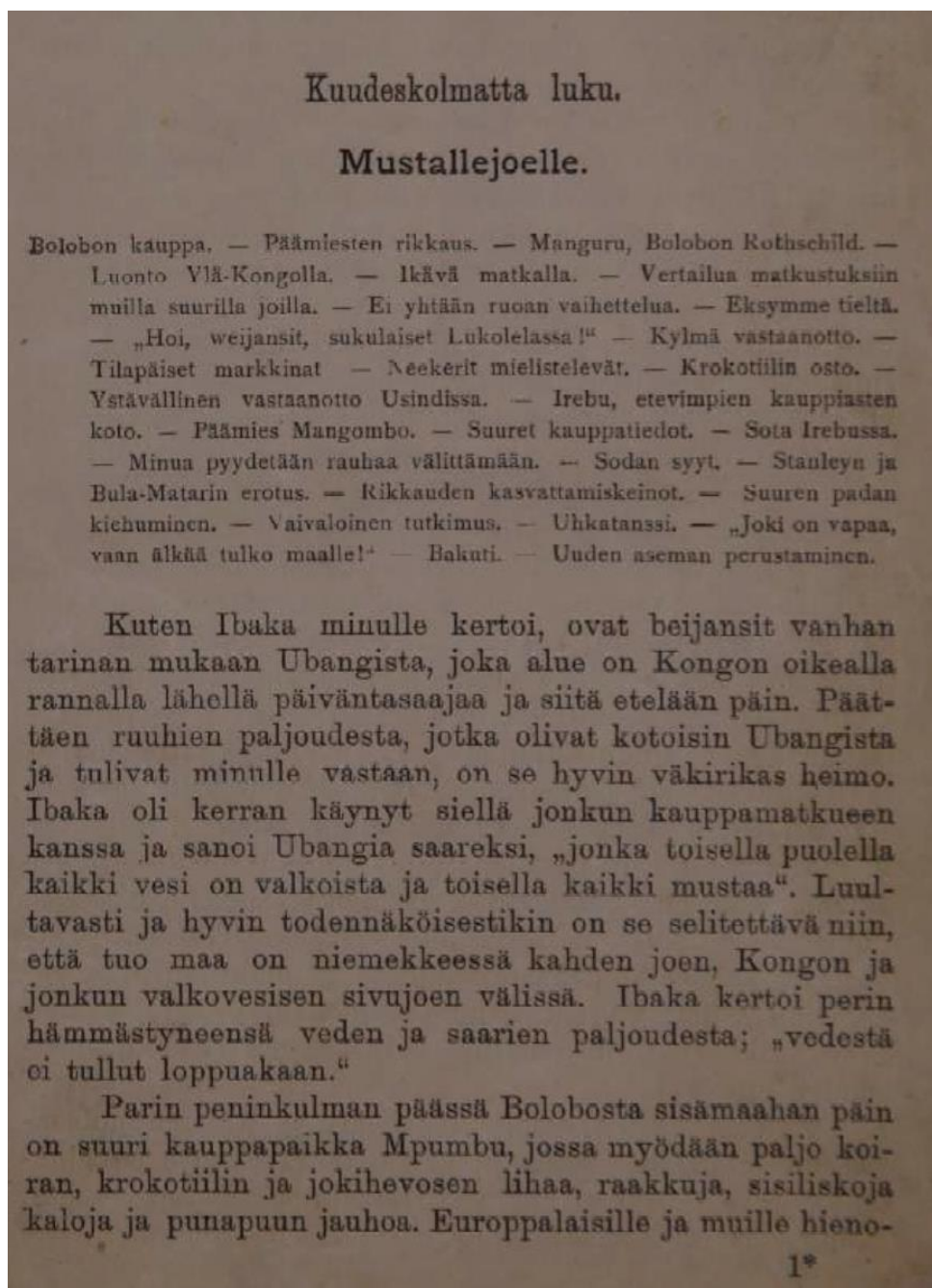
Yhteinen ihmiskunnan historia. IV. Wenäjän keisari Pietari Suuren wilpitön elämä-kerta. 1862

Yhteinen ihmiskunnan historia. V. Pröyessin kuningas Fredrikki II, kutsuttu ai-noaksi 1862

Yhteinen ihmiskunnan historia. VI. Wapaus-sota Pohjois-Amerikassa w. 1773-1783 ja yhdistettyin wapaawaltain nykyinen tila, ynnä lyhy wälikertomus Benjami Franklinista 1862

Yrjö Neumark 1859

KUVAT



Kuva 1. Ensimmäinen tekstisivu kirjasta *Kongo – Uusi vapaavaltio mustien maanosassa*. 2. osa (Stanley 1887, 1).

Arvoisella ei siis ollut tässä muuta tekemistä, mutta palattuansa konttoriinsa aukaisi hän sydämensä vaimolensa joristen: "Koska joka päivä istutaan herkullisesti syöden ja juoden pöydän ääressä, joka vaipuu herkkujen alla, niin mahdollista olla syytöin ja tietämätöin kuinka monta vilja-kappaa ihminen vuodessa tarvitsee ei nälkään nääntyäksensä." Kuitenkin sai se jäädä pätöin päätöksen mukaan, ja muori Rajunen, niinkuin toisetkin hänen tuttavistaan, näkivät niinmuodoin itsensä äkisti olevan ilman elatusta, josta he omain käsiensä avulla tähän asti köyhyydessä elivät. Lakkaamattomalla ahkeroitsemisella ja suuressa köyhyydessä eli Beata Rajunen niinmuodoin muutamat vuodet tyttärensä häviön-alaisessa tuvassa; mutta hän näki päivä päivältä sen ei taitavan niin aina olla.

Niin vastahakoisesti kuin hän siis tahtoi eroittaa itsensä ainoasta tyttärestänsä, ja niin kipiä kuin tämä aikomus oli hänelle, ymmärsi hän kuitenkin täytymisen tulevan Susannalle eritä kodostansa; osittain saadaksensa oppia muutamia vaimoväen askareita, joihin hänellä ei kotona ollut tilaa, osittain siitä että muori jo rupesi vanhentumaan ja ei voinut niin ahkerasti kuin ennen istua kangastuolissa ja kehruurukin vieressä, mintähden heidän palansa usein olivat kyllä pienet. Koska Rajusen muori ei koko maailmassa tuntenut ketään, kenen käsiin hän niin luottavaisesti taisi jättää Susannan onnen kuin Koivuluodon herrasväelle, niin vaelsi hän eräänä kauniina ja kirkkaana talvipäivänä sinne, ilmoit-

taen surunsa tyttärensä tulevaisesta ajasta ja palasi takaisin lohdutetulla sydämmellä ja iloisella toivolla, sillä se kunnioitettava rouva oli luvannut ottaa Susannan palvelukseensa.

2.

Nyt tahdomme mennä muutamat vuodet takaisin kertomuksessamme, kurkistaen sisään Pohjakosken ruukin konttoriin, missä me havaitsemme tirehtyöri Arvoisen istuvan viheriäisen pulpetin edessä ja lukevan kirjoja, jotka äsken posti-repussa olivat tulleet. Vastapäätä istui nuori, punaposkinen mies ahkerasti kirjoittaen, mutta katsoen salaa silloin tällöin tirehtyöriin puoleen saadaksensa hänen kasvoistansa havaita kirjain sisällepäiden — tämä oli kirjuri Pohjakoskella herra Salomo Flinkki. Äkisti kirkastuivat tirehtyöriin kasvat, ja tuskin näkyvä tytyväisyyden nauru läikkyy muutoin totisen suun ympäri.

— Hm, hm — sanoi äijä niinkuin itseksensä — se on ihana, se on hyvä ja lasi kirjan sen pulpetilla olevan marmori-painon alle. Sitten sanoi hän Flinkkille:

— Kuinka on laita Flinkki, ovatko kapineet valmiit?

— Kirjat, tarkoittanee tirehtyöri?

— Niin kyllä, mitä minä siis muuta tarkoittaisin.

— Ei juuri kokonansa — änkytti kirjuri hiukan hämille tullut.

— Mitä kuulen? Ei vielä kokonansa! Eikö tutkinnon aika ole sisällä, eikö herra sitä tiedä, hän?

Kuva 2. Henkilöiden ja aiheiden nopeaa vaihtelua teoksessa *Koivuluoto ja Pohjakoski* (J.W. Lilljan kirjapaino 1861, 8–9).