

Elina Jokela

MASTER DATAN MERKITYS LIIKETOIMINTATIEDON HALLINNASSA

Johtamisen ja talouden tiedekunta
Kandidaatintyö
Tarkastaja: Jussi Myllärniemi
Joulukuu 2021

TIIVISTELMÄ

Elina Jokela: Master datan merkitys liiketoimintatiedon hallinnassa
The importance of master data in business intelligence
Kandidaatintyö
Tampereen yliopisto
Tietojohtaminen
Joulukuu 2021

Erään sanonnan mukaan organisaation tekemät päätökset ovat korkeintaan yhtä hyviä kuin data, johon päätökset perustetaan. Liiketoimintatiedon hallinnan perimmäisenä tavoitteena on luoda tukea päätöksentekoon tuottamalla oikeaa tietoa oikeaan aikaan sekä tuottaa datasta relevanttia informaatiota. Master data (suomeksi ydin-, avain- tai perustieto), kuten asiakas- tai tuotetiedot, kuvastaa sitä organisaation kohtalaisen muuttumatonta ja pitkäaikaista dataa, jota käytetään transaktio- eli tapahtumadatan rinnalla raportoinnissa osana liiketoimintatiedon hallintaa.

Maailmassa syntyvän datan määrä on kasvanut odotettua suuremmaksi viime vuosina ja datan määrälle odotetaan lähitulevaisuudessa yhä merkittävämpää kasvua. Näin merkittävä datan määrän kasvu on nostanut esille ajankohtaisen haasteen datan laadunhallinnassa. Kirjallisuuden mukaan prosessi datan laadun parantamiseksi tulee aina aloittaa master datan laadun parantamisesta. Tämän työn tavoitteena oli tutkia sitä, millainen merkitys master datalla on liiketoimintatiedon hallinnan kontekstissa. Tutkimus rajattiin käsittelemään liiketoimintatiedon hallintaa (BI) ja liiketoiminta-analytiikkaa (BA) yhdistävää termiä liiketoimintatiedon hallinta ja analytiikka (BI&A), sillä analytiikan rooli on muodostunut entistä tärkeämmäksi organisaatioissa ja käsitteitä on alettu yhdistämään myös muussa kirjallisuudessa.

Tutkimus toteutettiin kirjallisuustutkimuksena Finkin seitsenvaiheista prosessimallia hyödyntäen. Tutkimusaineisto kerättiin Andor-, Google Scholar- ja ProQuest-tietokannoista. Tutkimuksen keskeisimpiä käsitteitä: master dataa ja sen hallintaa, datan laatua ja sen hallintaa sekä liiketoimintatiedon hallintaa ja analytiikkaa, tarkasteltiin työn teoriaosassa. Teoriaosan jälkeen tutkimusaineiston perusteella pyrittiin tutkimaan, miten master data näkyy osana liiketoimintatiedon hallintaa ja analytiikkaa, sekä millaisia haasteita master datalla tässä kontekstissa on. Lisäksi tutkimusaineiston perusteella koottiin master datan optimaalisen hyödyntämisen tavat liiketoimintatiedon hallinnassa ja analytiikassa.

Tutkimus osoitti, että master data on osana koko liiketoimintatiedon hallinnan prosessia. Tutkimuksen mukaan master data toimii organisaation datan osalta liiketoiminnan perustana ja sen hallinnan avulla voidaan muun muassa tuottaa laadukkaampaa ja nopeampaa päätöksentekoa sekä parantaa organisaation tiedon laatua. Datun laatu, datan puute tai sen laadunhallinnan puute todettiin yleisimmiksi syiksi master datan hallinnan aloittamiseksi ylipäätään organisaatioissa. Lisäksi master datan omistajuuden määrittäminen oli tutkitusti yksi merkittävimmistä haasteista liiketoimintatiedon hallinnan ja analytiikan kontekstissa. Tutkimuksen perusteella jokaisen organisaation tulisi huomioida oman master datansa hallinnalliset toimenpiteet. Jotta master datasta voidaan hyötyä osana päätöksentekoa ja siten liiketoimintatiedon hallintaa ja analytiikkaa, on master datatarpeiden tuettava myös liiketoimintatavoitteita.

Avainsanat: Master data, master datan hallinta, datan laatu, liiketoimintatiedon hallinta, liiketoimintatiedon hallinta ja analytiikka

Tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck –ohjelmalla.

ALKUSANAT

Tämä kandidaatintyö on toteutettu Tampereen yliopiston tietojohdamisen koulutusohjelmaan syksyllä 2021. Tutkimuksen aiheen valinta perustui kiinnostukseeni liiketoimintatiedon hallintaa kohtaan. Lisäksi halusin päästä tutustumaan tarkemmin master dataan ja erityisesti sen hyödyntämiseen osana BI:tä, joten kandidaatintyön avulla pääsin syventämään osaamistani tällä alueella.

Haluan kiittää työni ohjaajaa ja tarkastajaa Jussi Myllärniemeä tuesta ja tsempeistä. Hänen kannustamana lähdin pohtimaan aihetta yhä syvällisemmin ja täten sain kehitettyä työtäni aina parempaan suuntaan. Haluan myös vielä kiittää ystäviäni ja perhettäni kannustuksesta ja työn oikolukemisesta.

Rovaniemellä, 18.12.2021

Elina Jokela

SISÄLLYSLUETTELO

1. JOHDANTO	1
1.1 Tutkimuksen tausta ja merkitys	1
1.2 Tutkimusongelma ja rajaukset.....	2
1.3 Työn rakenne.....	3
2. TUTKIMUSMENETELMÄ JA -AINEISTO.....	5
2.1 Tutkimusmenetelmä.....	5
2.2 Tutkimusaineisto.....	6
3. MASTER DATA	9
3.1 Master datan määritelmä	9
3.2 Master datan hallinta.....	10
3.3 Datan laatu ja sen hallinta.....	13
4. LIIKETOIMINTATIEDON HALLINTA.....	16
4.1 Liiketoimintatiedon hallinnan prosessi	17
4.2 Liiketoimintatiedon hallinta ja analytiikka.....	19
5. MASTER DATAN MERKITYS LIIKETOIMINTATIEDON HALLINNASSA JA ANALYTIIKASSA	21
5.1 Master data osana liiketoimintatiedon hallintaa ja analytiikkaa	21
5.2 Master dataan liittyvät haasteet.....	23
5.3 Master datan optimaalinen hyödyntäminen osana liiketoimintatiedon hallintaa ja analytiikkaa	25
6. PÄÄTELMÄT	27
6.1 Tutkimustulokset.....	27
6.2 Tutkimuksen arviointi	28
6.3 Jatkotutkimusmahdollisuudet	29
LÄHTEET.....	31

1. JOHDANTO

Johdantoluvussa lukija johdatetaan aiheeseen tutkimuksen taustoituksella ja lähtökoh-
tien sekä perusteluiden esittelyllä. Tämän lisäksi esitetään tutkimuksen rajaukset ja tut-
kimuskysymykset. Johdannon lopuksi kuvaillaan kandidaatintyön rakenne.

1.1 Tutkimuksen tausta ja merkitys

Maailmassa syntyvän datan määrä ei ole koskaan ollut yhtä suuri kuin vuonna 2020. Datan kokonaismäärä on kasvanut odotettua suuremmaksi viime vuosina, ja datan mää-
rälle odotetaan lähitulevaisuudessa vielä entistä merkittävämpää kasvua. (IDC, Statista
2021.) Näin merkittävä datan määrän kasvu on nostanut esille ajankohtaisen haasteen
datan laadunhallinnassa (Vilminko-Heikkinen & Pekkola 2017, s. 454). Erään sanonnan
mukaan organisaation tekemät päätökset ovatkin korkeintaan yhtä hyviä kuin data, johon
päätökset perustetaan (Dayton 2007; Haug *et al.* 2009). Liiketoimintadataan pohjautu-
vaa edistysellisempää päätöksentekoa ja siten myös tuloksekkaampaa liiketoimintaa
tavoitellaan liiketoimintatiedon hallinnan kontekstissa (Laihonen *et al.* 2013, s. 46). Datan
merkityksen kasvaessa myös liiketoimintatiedon hallinnan ja analytiikan käytänteet ja sii-
hen liittyvät työkalut ovat kasvattaneet suosiotaan organisaatioissa (Marjamäki 2017, s.
6). Liiketoimintatiedon hallinta voidaan kontekstista riippuen määritellä muun muassa
teoreettisena lähestymistapana, prosessina tai päätöksenteon työkaluna. Sen perimmäi-
senä tavoitteena on kuitenkin luoda tukea päätöksentekoon tuottamalla oikeaa tietoa oi-
keaan aikaan sekä tuottaa datasta relevanttia informaatiota. (Vitt *et al.* 2010; Laihonen
et al. 2013, s. 44.)

Laihosen *et al.* (2013, ss. 19–20) mukaan organisaatioiden data on valitettavan usein
heikkolaatuista – puutteet, virheet ja ristiriidat datassa heikentävät datan laatua. Väre
(2019) toteaa, että prosessi datan laadun parantamiseksi tulee aina aloittaa master da-
tan laadun parantamisesta. Master data (suomeksi ydin-, avain- tai perustieto) on liike-
toiminnalle kriittistä, luonteeltaan kohtalaisen muuttumatonta ja pitkäaikaista dataa, jota
hyödynnetään läpi organisaation. Organisaation master dataa ovat esimerkiksi asiakas-
, työntekijä-, toimittaja- ja tuotetiedot. (Loshin 2009, s. 6; Väre 2019, s. 16.) Heikkolaa-
tuinen master data tai sen epä johdonmukainen hyödyntäminen voivat vaikuttaa liiketoi-

mintaan monin tavoin. Virheelliset strategisen tai operatiivisen tason päätökset, epäonnistuneet toimitusprosessit tai päätöksenteon hitaus ovat esimerkkejä master dataan liittyvistä haasteista. (Silvola *et al.* 2011.) Sen sijaan onnistuneella master datan hallinnalla on monia hyötyjä: edistyneempi riskienhallinta ja organisaation kilpailukyky, parempi operationaalinen tehokkuus ja kustannustehokkuus, laadukkaampi ja nopeampi päätöksenteko, kehittyneempi tiedonlaatu sekä parempi asiakasymmärrys ja -tyytyväisyys ovat esimerkkejä onnistuneen master datan hallinnan hyödyistä (Loshin 2009, ss. 10–12).

Liiketoimintatiedon hallinnan ja analytiikan suosio on kasvussa ja toisaalta dataa syntyy odotettua enemmän. Tämä tuo esiin entistä voimakkaammin datan laadun merkityksen ja sitä myötä myös master datan merkityksen. Kuitenkin Harvard Business Review:n ja Profiseen toteuttaman tutkimuksen (2021) mukaan vain 28 % organisaatioista hyödyntää master datan hallintaa systemaattisesti osana liiketoimintaansa. Tämän johdannon myötä voidaan kuitenkin tehdä oletus, että master data ja sen hallinta on merkityksellistä liiketoimintatiedon hallinnan ja analytiikan kontekstissa. Haettaessa tutkimusaineistoa aiheeseen liittyvillä hakulausekkeilla (taulukot 1 ja 2), voidaan huomata, että master dataa tai sen hallintaa sekä liiketoimintatiedon hallintaa ja analytiikkaa käsitellään joissakin kirjallisuuslähteissä yhdessä (White 2007; Zoder 2011; Kekwaletswe & Lesole 2016; Bentley 2017; Vilminko-Heikkinen & Pekkola 2017; Vilminko-Heikkinen & Pekkola 2019; Väre 2019). Tutkimusaineiston perusteella ei kuitenkaan pystytä suoraan tai yksiselitteisesti vastaamaan tutkimusongelmaan. Tässä kandidaatintyössä pyritään siis kokoamaan tutkimusaineistosta vastaus tutkimusongelmaan ja täten täydentämään tutkimuskenttää aiheen osalta.

Tämän kandidaatintyön tavoitteena on selvittää master datan merkitys liiketoimintatiedon hallinnan kontekstissa. Työssä tutustutaan sekä master datan että liiketoimintatiedon hallinnan käsitteisiin ja tarkastellaan niiden välistä yhteyttä. Master datan tarkastelun myötä työssä tutkitaan myös datan laadunhallinnan periaatteita ja sen yhteyttä tutkimusongelmaan. Lisäksi työssä kootaan yhteen kirjallisuuden esittämät käytänteet ja toimitatavat, joita onnistuneessa master datan hallinnassa odotetaan hyödynnettävän.

1.2 Tutkimusongelma ja rajaukset

Tämän kandidaatintyön tavoitteena on selvittää master datan merkitys liiketoimintatiedon hallinnassa. Työn päätutkimuskysymys on seuraava:

- millainen merkitys master datalla on liiketoimintatiedon hallinnan kontekstissa?

Päätutkimuskysymyksen aihetta on rajattu master datan ja liiketoimintatiedon hallinnan käsitteiden näkökulmasta. Master datan yksi käytetyimmistä suomenkielisistä käännoiksi on ”ydintieto” (Laihonen *et al.* 2013, s. 20). Kontekstista riippuen master datasta voidaan suomessa käyttää myös käännoiksi ”perustieto” tai ”avaintieto”. Täten kandidaatintyössä on valittu käytettäväksi neutraalia ja kansainvälisesti tunnettua ilmaisua ”master data”. (Kuntasektorin kokonaisarkkitehtuuriryhmä 2013, s. 5.) Master datan käsite määritellään luvussa 3.

Liiketoimintatiedon hallinnan (business intelligence, BI) kontekstissa liiketoiminta-analytiikasta (business analytics, BA) on tullut organisaatioissa viime vuosina aikaisempaa tärkeämpää (Goodwyn 2019, s. 33). Liiketoimintatiedon hallinnan ja liiketoiminta-analytiikan käsitteitä on ryhdytty nykypäivänä yhdistämään niiden toisiaan täydentävän luonteensa vuoksi, joten tässäkin työssä liiketoimintatiedon hallinta on rajattu näitä konsepteja yhdistävään käsitteeseen ”liiketoimintatiedon hallinta ja analytiikka” (business intelligence and analytics, BI&A) (Corte-Real *et al.* 2014, s. 2). Liiketoimintatiedon hallinnan ja analytiikan käsite määritellään luvussa 4.

Päätutkimuskysymykseen pyritään löytämään vastaus alatutkimuskysymysten avulla. Työn alatutkimuskysymykset ovat seuraavat:

- Mitä haasteita master datan hyödyntämiseen liittyy?
- Miten master datan hallintaa voidaan hyödyntää nykypäivänä päätöksenteossa?

Alatutkimuskysymysten tueksi työssä määritellään master datan, master datan hallinnan, datan laadunhallinnan, liiketoimintatiedon hallinnan sekä liiketoimintatiedon hallinnan ja analytiikan käsitteet.

1.3 Työn rakenne

Johdannon jälkeen työn toisessa luvussa esitellään työssä käytetty tutkimusmenetelmä ja -aineisto. Työ toteutetaan kirjallisuuskatsauksena hyödyntäen Finkin (2005) seitsenvaiheista kirjallisuuskatsauksen prosessimallia. Luvuissa 3 ja 4 tarkastellaan valitun aiheen teoreettista näkökulmaa sisältäen master datan ja sen hallinnan, liiketoimintatiedon hallinnan, liiketoimintatiedon hallinnan ja analytiikan sekä datan laadun ja sen hallinnan käsitteiden määrittelyt. Lisäksi edellä mainituissa luvuissa esitetään prosessimallit master datan hallinnan kehittämiseksi sekä liiketoimintatiedon hallinnalle.

Luvussa 5 tuodaan yhteen master datan sekä liiketoimintatiedon hallinnan ja analytiikan käsitteet ja tarkastellaan master datan merkitystä tässä kontekstissa. Tässä luvussa esi-

tetään myös master datan hyödyt sekä haasteet liiketoimintatiedon hallinnan ja analytiikan kontekstissa sekä kuvataan master datan ja sen hallinnan sijoittumista liiketoimintatiedon hallinnan prosessimalliin. Luvun 5 lopussa kootaan yhteen kirjallisuudessa tunnistetut optimaaliset toimintatavat ja käytänteet master datan hallintaan. Työ päätetään tutkimustulosten yhteenvetoon, jonka jälkeen tutkimus arvioidaan ja jatkotutkimustarpeet esitetään.

2. TUTKIMUSMENETELMÄ JA -AINEISTO

Tässä luvussa esitellään kandidaatintyön tutkimusmenetelmä ja -aineisto. Tämän kandidaatintyön tutkimusmenetelmänä käytetään kirjallisuuskatsausta. Tutkimusaineisto koostuu tieteellisistä julkaisuista. Aineistosta on pyritty löytämään tutkimusongelmaan yleispäteviä ratkaisuja, ei niinkään organisaatiospesifejä toimintatapoja.

2.1 Tutkimusmenetelmä

Kirjallisuuskatsaus on täsmällinen, systemaattinen ja toistettavissa oleva menetelmä, joka perustuu valmiina olevan, korkealaatuisen tutkimusaineiston johtopäätöksien tunnistamiseen, arviointiin ja tiivistämiseen (Fink 2005, s. 3). Tässä kandidaatintyössä hyödynnetään seuraavaa Finkin seitsenvaiheista prosessimallia:

1. Tutkimuskysymyksen asettaminen
2. Tietokantojen ja muiden lähteiden valinta
3. Hakutermien valinta
4. Käytännön seulan asettaminen (esimerkiksi hakutulosten kieli tai ajanjakso)
5. Metodologisen seulan asettaminen (tieteellisen laadun arviointi)
6. Katsauksen tekeminen
7. Tuloksien syntetisointi.

Prosessi alkaa tutkimuskysymyksen asettamisesta (lukuissa 1.1 ja 1.2). Tässä työssä eri tietokantoja ja esimerkiksi WWW-sivustojen lähdemateriaalia on hyödynnetty tutkimuskysymyksen asettamisessa, mutta varsinainen työssä käytettyjen tietokantojen ja muiden lähteiden valinta (luvussa 2.2) on tehty tutkimuskysymyksen asettamisen jälkeen. Luvussa 2.2 esitetään myös Finkin prosessimallin vaiheessa kolme, neljä ja viisi kuvatut hakutermien valinta sekä käytännön ja metodologisen seulan asettamiset. Finkin (2005, ss. 3–5) mukaan kirjallisuuskatsauksessa tavoitteena on hyödyntää mahdollisimman laadukasta materiaalia. Kokonaiskatsauksen jälkeen kirjallisuuskatsaus päätetään tulosten syntetisointiin, jolloin löydösten selittämisen ja muun muassa tutkimuksen laadun arvioinnin päätteeksi toteutetaan laaduntarkkailua sekä tulosten laadullista ja määrällistä synteesiä (Fink 2005, ss. 3–5).

2.2 Tutkimusaineisto

Tutkimusaineisto kerätään Andor-, ProQuest- ja Google Scholar- tietokannoista. Tutkimusaineistoa haetaan master datan, master datan hallinnan, datan laadunhallinnan, liiketoimintatiedon hallinnan sekä liiketoimintatiedon hallinnan ja analytiikan käsitteille. Tutkimuksessa pyritään hyödyntämään mahdollisimman tuoretta tutkimusaineistoa etenkin master datan ja liiketoimintatiedon hallinnan välisestä yhteydestä, sillä tarkoituksena on saada ajankohtaista ja jopa ennustavaa informaatiota tutkittavasta aiheesta. Taulukossa 1 esitetään rajoittamattoman haun tulokset edellä mainituista tietokannoista.

Taulukko 1. Rajoittamattoman haun tulokset

Hakulauseke	Andor	ProQuest	Google Scholar
"master data"	55 296	34352	36 300
"master data management"	27 443	19 953	8 270
"business intelligence"	890 056	632 563	309 000
"business intelligence and analytics"	27 988	12 950	10 900
"master data" AND "business intelligence"	7 593	5 771	6 380
"master data" AND "business intelligence and analytics"	563	293	519
"data quality management"	7 234	4 059	13 000

Finkin prosessimallin vaiheen neljä mukaisesti asetetaan tulosten hakuun käytännön seula, jolla rajataan tutkimusaineistosta pois yli 10 vuotta vanha tutkimusaineisto eli valitaan aineisto aikaväliltä 2011–2021. Lisäksi otetaan tarkasteluun ainoastaan vertaisarvioidut aineistot. Taulukkoon 2 on koottu hakutulokset rajatulla haulla.

Taulukko 2. Rajoitetun haun tulokset

Hakulauseke	Andor	ProQuest	Google Scholar
"master data"	2 791	1 630	535
"master data management"	430	267	170
"business intelligence"	17 178	12 014	4 580
"business intelligence and analytics"	1 055	796	557
"master data" AND "business intelligence"	303	194	119
"master data" AND "business intelligence and analytics"	32	26	23
"data quality management"	1 501	957	411

Hyödyntämällä Finkin prosessimallin neljättä vaihetta ja rajaamalla hakutuloksia, voidaan huomata hakutulosten määrän pienentyneen selkeästi. Hakutulosten perusteella Andor tarjoaa eniten tutkimusaineistoa. Tutkimuksessa hyödynnettävät tietokannat tarjoavat kaikki toisistaan poikkeavia, tutkimukselle relevantteja tuloksia, joten tutkimuksessa on siten perusteltua hyödyntää kaikkia kolmea tietokantaa. Tutkimusaineistoa käydään läpi tietokantojen ehdottamien tulosten mukaisessa järjestyksessä, mistä kukaan hakua enintään 50 ensimmäisen tuloksen osalta. Otsikkotasolla tutkimukselle soveltuvista hakutuloksista tutkitaan tiivistelmä, sisällysluettelo sekä tarvittaessa johdanto ja yhteenveto. Relevantiksi osoittautunutta tutkimusaineistoa hyödynnetään työssä niin, että tutkimusongelmaan pyritään löytämään mahdollisimman oikeanmukaisia ratkaisuja.

Tutkimusaineisto koostuu artikkelijulkaisuista, e- ja fyysisistä kirjoista sekä konferenssi-julkaisuista. Taulukkoon 3 on koottu työn kannalta keskeisimmät lähteet.

Taulukko 3. Keskeisimmän tutkimusaineiston julkaisujen tekijät, julkaisuvuosi, otsikko ja lyhyt kuvaus

Tekijät	Julkaisuvuosi	Otsikko	Kuvaus
Loshin, D.	2009	Master Data Management	Käsittelee master dataa ja sen hallintaa sekä datan laadunhallintaa
Chen, H. <i>et al.</i>	2012	Business Intelligence and Analytics: From Big Data to Big Impact	Käsittelee liiketoimintatiedon hallintaa ja analytiikkaa (BI&A) sekä sen historiaa
Vilminko-Heikkinen, R. & Pekkola, S.	2013	Establishing an Organization's Master Data Management Function: A Step-wise Approach	Tarjoaa prosessimallin master datan hallinnan kehittämiseen
Kekwaletswe, R. & Lesole, T.	2016	A Framework for Improving Business Intelligence through Master Data Management	Esittelee käytänteitä onnistuneeseen master datan hallintaan BI&A-kontekstissa
Mahanti, R.	2019	Data Quality: Dimensions, Measurement, Strategy, Management, and Governance	Käsittelee datan laatua sen ulottuvuuksineen sekä datan laadunhallintaa
Vilminko-Heikkinen, R. & Pekkola, S.	2019	Changes in roles, responsibilities, and ownership in organizing master data management	Käsittelee master datan hallintaa ja erityisesti sen tiedonhallinnallista näkökulmaa
Väre, T.	2019	Master Data	Käsittelee master dataa ja sen hallintaa

Tutkimusaineisto on rajattu pääasiassa tuoreeseen, viimeisen 10 vuoden sisällä julkaistuihin lähdemateriaaliin. Kuitenkin esimerkiksi master datan ja sen hallinnan käsitteiden osalta on hyödynnetty Loshinin teosta yli kymmenen vuoden takaa, sillä kyseiset käsitteet on tunnustettu jo jopa 1980-luvulla ja pysyneet lähes muuttumattomana (Loshin 2009, s. 3).

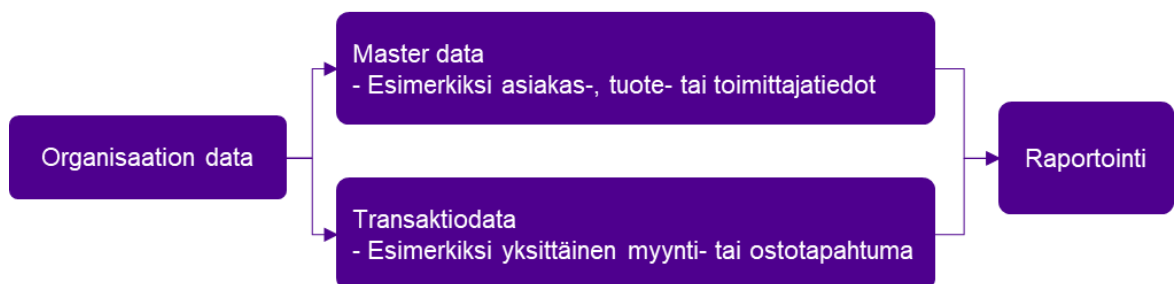
3. MASTER DATA

Datan hallinta käsitteellistyi 1960-luvulla ADAPSO:n (Association of Data Processing Service Organizations) tutkiessa datan laadun varmistamista ja siihen liittyvää ammatillista kouluttamista. Kyseisten koulutusohjelmien rinnalla kehittyi myös master datan käsite 1980-luvulla, ja pian sen jälkeen myös master datan hallinnan (MDM) käsite. (Foote 2019.) Nykyään master data käsitetään liiketoiminnalle kriittiseksi tiedoksi, jota hyödynnetään läpi organisaation (Loshin 2009, s. 6; Berson & Dubov 2011).

3.1 Master datan määritelmä

Master data (suomeksi ydin-, avain- tai perustieto) kuvastaa sitä organisaation dataa, joka toimii kriittisenä perustana liiketoiminnan prosesseille ja sovelluksille (Berson & Dubov 2011). Loshinin (2009, s. 6) ja Väre (2019, s. 16) mukaan master datalle ei ole olemassa täysin yksiselitteistä määritelmää, mutta se usein käsitetään kohtalaisen muuttumattomana ja pitkäaikaisena datana, jota hyödynnetään läpi organisaation. Bentley (2017, s. 140) mukaan master datan tavoitteena on toimia luotettavana ja yhtenä versiona totuudesta, johon organisaation päätöksenteossa voidaan tukeutua. Organisaation master dataa ovat esimerkiksi asiakas-, työntekijä-, toimittaja- ja tuotetiedot (Loshin 2009, s. 6; Väre 2019, s. 16.)

Master data ja sen hallinta on kriittistä organisaation päivittäiselle toiminnalle ja esimerkiksi liiketoimintatapahtumien luotettavalle raportoinnille (Väre 2019, s. 16). Organisaation data voidaan muun muassa Loshinin (2009, s. 8) ja Laihosen *et al.* (2013, s. 20) mukaan jakaa master dataan ja transaktiodataan. Täten organisaation datan kategorisointia aina raportointiin ja datan muuhun hyödyntämiseen voidaan havainnollistaa kuvan 1 mukaisesti.



Kuva 1. Organisaation datan kategorisointi ja datan hyödyntäminen (mukailtu Loshin 2009, ss. 6–8; Laihonen *et al.* 2013, s. 20; Väre 2019, ss. 14–15)

Liiketoimintatapahtumia, kuten ostotapahtumia, yhteydenottoja tai toimituksia, kuvaa käsite transaktiodata eli tapahtumadata. Transaktiodata sisältää master dataa, mutta transaktiodata ei itsessään ole master dataa. Master data sen sijaan toimii transaktioiden perustana. (Loshin 2009, s. 8; Väre 2019, ss. 14–15.) Haug ja Arlbjorn (2011, s. 289) huomauttavat, että master datalla on merkittävä vaikutus myös transaktiodatan luotettavuuteen ja siten myös raportoinnin luotettavuuteen.

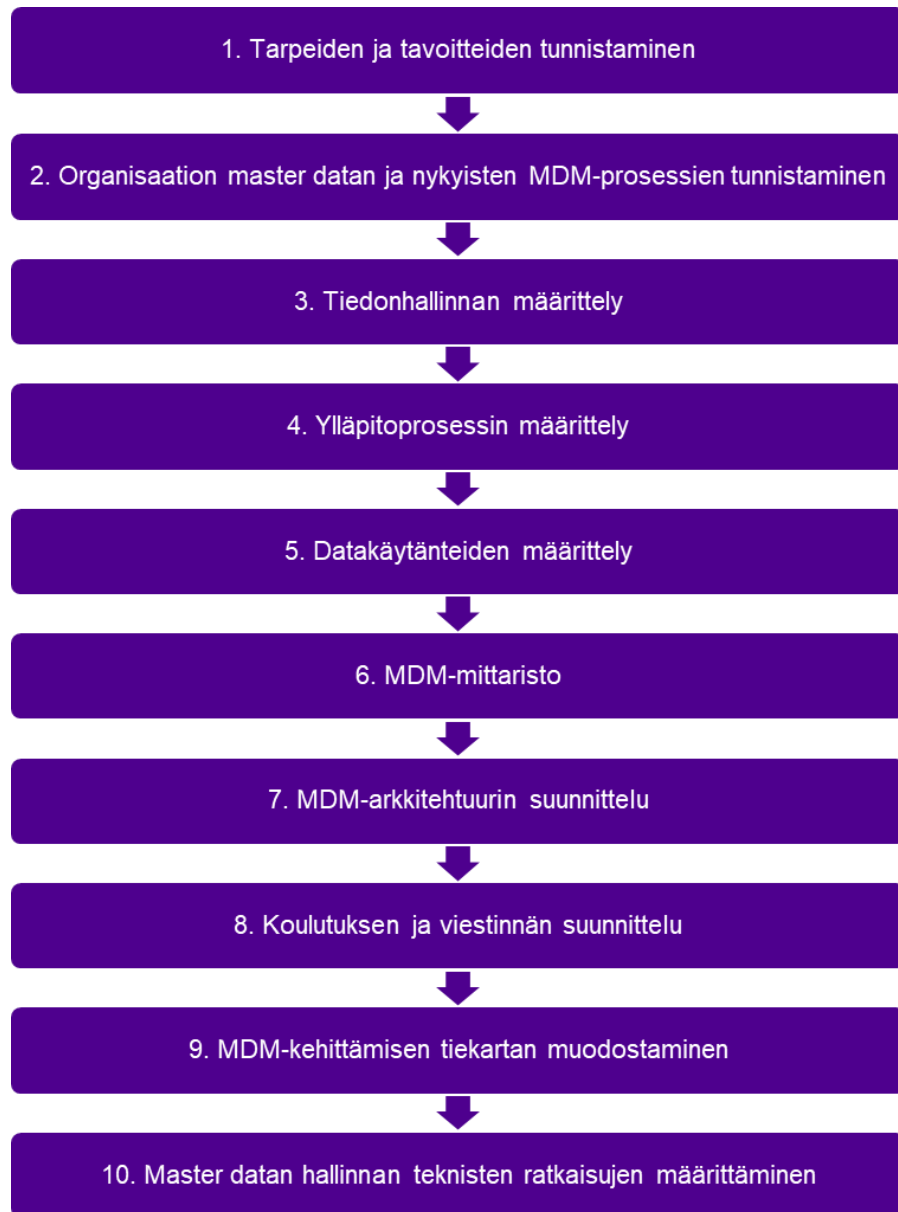
3.2 Master datan hallinta

Master datan hallinta (master data management, MDM) voidaan määritellä käytänteinä, toimintatapoina, sovelluksina ja teknologioina master datan harmonisointiin ja hallintaan läpi organisaation ja sen yksiköiden. MDM-käytänteet eivät ole yksiselitteisiä eri organisaatioiden kesken. Master datan hallinnalla voidaan kuitenkin yleisesti tavoitella master datan päällekkäisyyksien ja epäjohdonmukaisuuksien eliminointia, liiketoimintaprosessien tehokkuutta sekä liiketoimintamuutosten tukemista. (White 2007; Vilminko-Heikkinen & Pekkola 2017, s. 470.)

Master datan hallintaan liittyviä haasteita on kirjallisuudessa lähestytty muun muassa organisaation, datan laadun ja teknologioiden näkökulmasta. Organisaation valmius master datan hallintaan, tiedonhallinnan vastuut ja käytänteet sekä teknologiaintegroinnit toimivat esimerkkeinä organisaatiotason master datan hallinnan haasteista. Datan arviointi, integrointi ja laadunhallinnan varmistaminen taas ovat osa datan laadunhallinnan näkökulmaa. Sen sijaan teknologianäkökulma käsittää muun muassa datan mallinnuksen, standardisoinnin, käyttöoikeuksien hallinnan ja datan yhdistämisen. (Loshin 2009, ss. 15–18.) Täten onnistuneen master datan hallinnan tulee siis käsittää moniulotteinen kokonaisuus, joka yhdistää sekä organisaation käytänteet ja toimintatavat, datan laadunhallinnan sekä käytettävät teknologiat.

Master datan kokonaisuuden hallitsemiseksi on määritelty muun muassa kymmenvaiheinen kehittämis- ja käyttöönottoprosessi, joka perustuu vuosina 2010–2011 toteutettuun case-tutkimukseen (Vilminko-Heikkinen & Pekkola 2013). Myöhemmin julkaistussa tutkimuksessaan Vilminko-Heikkinen ja Pekkola (2019) sivuavat master datan hallinnan kehittämis- ja käyttöönottoprosessia erityisesti tiedonhallinnan näkökulmasta. Myös Väre (2019, s. 144) esittelee teoksessaan master datan hallintamallin sisällön, jota voidaan hyödyntää ylipäätään osana datanhallintaa. Väre (2019, s. 144) hallintamalli ei ole yhtä yksityiskohtainen verrattaessa Vilminko-Heikkisen ja Pekkolan (2013, ss. 4722–4725) prosessimalliin, mutta sitä voidaan käyttää tukena master datan hallinnan kehittämis- ja käyttöönottoprosessin esittämisessä. MDM-ratkaisut tulee määrittää organisaation

tiokohtaisesti, mutta kehittämisprosessia voidaan hyödyntää yleispätevästi apuna ratkaisujen kehittämiseen. Master datan hallinnan kehittämis- ja käyttöönottoprosessi on esitetty kuvassa 2.



Kuva 2. Master datan hallinnan kehittämis- ja käyttöönottoprosessi (mukailtu Vilminko-Heikkinen & Pekkola 2013, ss. 4722–4725; Vilminko-Heikkinen & Pekkola 2019, s. 79–83; Väre 2019, s. 144)

Master datan hallinnan kehittämisprosessi Vilminko-Heikkisen ja Pekkolan (2013, ss. 4722–4725; 2019, s. 79) mukaan alkaa master datatarpeiden ja tavoitteiden tunnistamisesta. Heidän mukaansa tavoitteet voivat kohdistua esimerkiksi työn tehostamiseen tai kehittyneempään raportointiin ja olennaista on ymmärtää master datan tai sen hallinnan tuomat mahdollisuudet tavoitteiden saavuttamisessa. Tuoreemmassa teoksessaan Vilminko-Heikkinen ja Pekkola (2017, s. 455) lisäävät, että tarpeiden tunnistaminen tulee

tehdä läpi organisaation, sillä master dataa hyödynnetään yleensä jokaisessa organisaation yksikössä.

MDM-kehittämisprosessin toisessa vaiheessa organisaation datan määrittelytoimenpiteeseen nimetyt tahot tutustuvat siihen, mitä master data on ja millaisissa tietoaineistoissa sitä organisaation sisällä esiintyy (Vilminko-Heikkinen & Pekkola 2013, s. 4723; Väre 2019, s. 144). MDM-prosessien ja ylipäätään organisaation käyttämän master datan tunnistaminen vaatii järjestelmien tuntemusta ja dataan liittyvien prosessien analysointia. Kolmannessa vaiheessa, tiedonhallinnan määrittelyssä, tavoitteena on tunnistaa ja määrittää roolit sekä vastuualueet master datan hallitsemiseen ja hallintamallin jalokauttamiseen. Master datan näkökulmasta hallinto käsittää muun muassa käytänteet, säännöt, datan ja käsitteen omistajuuden, vastuut ja roolit sekä roolien kuvaukset. (Vilminko-Heikkinen & Pekkola 2013, s. 4723; Vilminko-Heikkinen & Pekkola 2019, s. 79–80, Väre 2019, s. 144.) Tiedonhallinnan roolia pidetään merkittävänä tekijänä etenkin moniyksikköisessä johtamisessa. Onnistunut tiedonhallinta selkeyttää muun muassa sekä strategista että operatiivista toimintaa sekä edistää datan luotettavuutta ja läpinäkyvyyttä. (Girard 2020, ss. 3, 5.)

Ylläpitoprosessin määrittelyvaiheessa organisaation tulee nimetä vastuujat sekä tavat ja työkalut datan keräämiseen sekä laatia ohjeet datan käsittelyyn operatiivisessa työssä (Vilminko-Heikkinen & Pekkola 2013, s. 4724). Vilminko-Heikkisen ja Pekkolan samaisessa tutkimuksessa huomattiin, että organisaatioissa datan ylläpito on yleensä kallista ja tehottomat manuaaliset prosessit ovat lisäksi aiheuttaneet datan laadunhallinnallisia ongelmia, vaikka prosessit olisivat helposti korjattavissa. Viidennessä, datakäytänteiden määrittelyvaiheessa, data tulisi määrittellä sekä sisällöllisesti että mallintamalla master datan kokonaisuus. Vilminko-Heikkinen ja Pekkola (2013, s. 4724) jatkavat, että datan mallinnus mahdollistaa muutoksien tekemisen liiketoimintaympäristössä, sillä silloin tieto on ymmärrettävässä ja tulkittavassa esitysmuodossa.

Master datan hallinnan yksi tavoitteista on parantaa datan laatua, ja siksi master datan hallinnan mittaaminen ja siihen yhdistetyt mittarit ovat kasvattaneet suosiotaan. Master datan hallinnan ja siten datan laatua voidaan mitata muun muassa tarkastelemalla datan siivoamiseen, duplikaattien poistamiseen tai datan yhteensovittamiseen kulunutta aikaa. (Vilminko-Heikkinen & Pekkola 2013, ss. 4724–4725.) Datan laadun mittaaminen edellyttää myös organisaatioissa laadunhallintaa toteuttavien tahojen ja heidän vastuiden määrittämisen (Väre 2019, s. 144). Seitsemäs vaihe, MDM-arkkitehtuurin suunnitteluvaihe, keskittyy master dataan liittyvän arkkitehtuurin ymmärtämiseen ja suunnitteluun. MDM-arkkitehtuuri sisältää muun muassa siihen liittyvät sovellukset ja niissä liikkuvat

tietovirrat, järjestelmät, tiedonhallinnalliset käytänteet ja tietoturvallisuusnäkökulman. (Vilminko-Heikkinen & Pekkola 2013, ss. 4724–4725.)

Kommunikaatio eri sidosryhmien välillä on merkittävä tekijä onnistuneessa master datan hallinnassa (Vilminko-Heikkinen & Pekkola 2013, s. 4725; Väre 2019, s. 179). Vilminko-Heikkisen ja Pekkolan (2013, s. 4725) mukaan MDM-kehittämisen prosessin kahdeksas vaihe nähtiin koostuvan muun muassa seuraavista olennaisista asioista tai tehtävistä: tapaamisten ajoittaminen ja pitäminen, koulutuksien suunnittelu sekä yhteisten tavoitteiden ja master datan kriteereiden viestiminen. MDM-kehittämisen prosessin lopuksi Vilminko-Heikkisen ja Pekkolan (2013, s. 4725) mukaan tulee muodostaa MDM-kehittämisen tiekartta tai MDM-strategia. Heidän mukaansa tämä ottaa huomioon alussa määritetyt MDM-tavoitteet, mutta myös prosessin aikana kehittyneet täsmentävät tavoitteet. Täten master datan hallinnalle voidaan muodostaa sen kehittämisen pitkäaikainen tiekartta. Viimeisimpänä vaiheena MDM-kehittämisen prosessissa on master datan hallinnan teknisten ratkaisuiden määrittely, joka sisältää muun muassa teknisten ratkaisuiden ylläpitäjäroolien nimeämisen, sovellusten käyttöliittymän suunnittelun, datan käsittelyn mahdollisuudet sekä integraatiot toisiin sovelluksiin tai järjestelmiin. (Vilminko-Heikkinen & Pekkola 2013, s. 4725; Vilminko-Heikkinen & Pekkola 2019, s. 83.)

3.3 Datan laatu ja sen hallinta

Elokuussa 2021 julkaistussa podcastissa Nokian Renkaiden Information Management-yksikön vastuuhenkilö Joonas Julkunen toteaa, että master dataa ei ole olemassa ilman tiedon laadun käsitettä (Julkunen 2021). Myös Loshin (2009, s. 87) teoksessaan mainitsee, että master datan hallinta on nimenomaan saanut alkunsa ja suosionsa halusta tuottaa parempilaatua yritystietoa. Täten useat MDM-prosessit ovat datan laadunhallinnallisia toimenpiteitä (Loshin 2009, s. 87).

Datan laatu määritellään usein termillä ”fitness for use”, jolla viitataan datan sopivuuteen sen käyttäjävaatimukseen nähden (Strong & Wang 1996). Datan laadun määritelmä ei kuitenkaan ole yksiselitteinen, sillä ”fitness for use” voidaan käsittää eri organisaatioissa tai eri yhteyksissä eri tavoin. Odotukset datan osalta saattavat olla eriäviä organisaatioiden kesken ja jopa organisaatioiden sisällä. Lisäksi epäselvien tai hajanaisten odotusten myötä datan sopivuutta odotuksiin on haastavaa arvioida tai mitata. (Sebastian-Coleman 2013, s. 40.)

Datan laadun moniulotteisen määritelmän tueksi kirjallisuudessa on esitetty teoreettisia malleja datan laadun eri ulottuvuuksille. Teorioiden tavoitteena on ymmärtää ja hallita datan laatua sen eri ulottuvuuksien näkökulmasta, jolloin datan laadun käsite pilkkoutuu

pienempiin ja helpommin hallittaviin kokonaisuuksiin. Datan laadun ulottuvuuksia on esitetty ensimmäistä kertaa teoriatasolla Balloun ja Pazerin (1985) teoksessa, jossa ulottuvuuksiksi nimettiin tarkkuus, täydellisyys, oikea-aikaisuus ja johdonmukaisuus. Tuoreemmassa, vuonna 2019 julkaistussa teoksessa, Mahanti on koonnut Wangin *et al.* (1995) sekä Haugin ja Arlbjornin (2011) teorioiden pohjalta yhteenvedon datan laadun eri ulottuvuuksien esiintyvyydestä kirjallisuudessa. Teoksensa yhteenvetona Mahanti (2019, s. 77) tunnisti jopa 29 eri datan laadun ulottuvuutta. Myöhemmin myös Zhang ja Lee (2021, s. 16) toteuttivat kirjallisuustutkielman hyödyntäen 21 eri tutkimusta, minkä tuloksena tunnistettiin 24 eri datan laadun ulottuvuutta. Täten voidaan sanoa, että datan laadun käsite nähdään nykyään laajempina verrattuna datan laadun käsitteen historiaan. Mahantin (2019, s. 75) tekemän yhteenvedon perusteella kirjallisuudessa viitatuimpia datan laadun ulottuvuuksia ovat datan tarkkuus, luotettavuus, oikea-aikaisuus, relevanttius, eheys, ajantasaisuus ja johdonmukaisuus. Zhang ja Leen (2021, s. 16) toteuttaman tutkielman yhteenvetona tulokset viitatuimmista datan laadun ulottuvuuksista olivat melko yhtenevät. Zhang ja Leen (2021, s. 16) tekemässä yhteenvedossa ei noussut esiin relevanttius tai oikea-aikaisuus, kun taas Mahantin (2019, s. 75) yhteenvedossa ei huomioitu lainkaan datan saavutettavuutta tai datan määrää laatua ulottuvuutena.

Datan laadun tarkkuuden ulottuvuus viittaa siihen, kuinka oikeaa data on esitysmuodoltaan ja sisällöltään, eli kuinka hyvin data vastaa sitä kohdetta, mitä sen on tarkoitus mallintaa (Loshin 2009, s. 90; ISO 8000-2 2017; Mahanti 2019, s. 75). Datan tarkkuudenhallinnan haasteena on se, että vahvistaakseen datan oikeellisuuden, samalle datalle tulee olla oikeaksi vahvistettu datan vertailuarvo. Lisäksi tiedon muuttuva luonne luo haasteellisuutta datan tarkkuudenhallinnalle. (Loshin 2009, s. 90.) Mahantin (2019, s. 122) mukaan data on laadultaan luotettavaa, mikäli se on käyttötarkoitukseensa nähden relevanttia, tarkkaa, ainutlaatuista, johdonmukaista, eikä data ole puutteellista. Lisäksi luotettavuuden mittarina toimii sen jäljitettävyyys luotettavaan lähteeseen. Datan käyttötarkoituksesta riippuen datan ei tarvitse olla täysin virheetöntä ollakseen luotettavaa, vaan datan käyttäjä määrittelee rajat sille, kuinka paljon virheitä datassa saa olla. (Caballero *et al.* 2018, s. 11; Mahanti 2019, s. 122.)

Datan ajallista laatua ohjaavat oikea-aikaisuuden ja ajantasaisuuden ulottuvuudet. Oikea-aikaisuudella viitataan siihen, että data on laadukasta silloin, kun se on saatavilla tarvittuun ja odotettuun aikaan. Ajantasaisuus taas viittaa siihen, että data on ajan tasalla sen käyttöhetkellä. (Mahanti 2019, ss. 117–118.) Datan laadun relevanttiudella sen sijaan tarkoitetaan sitä, kuinka hyvin data vastaa sille osoitettua nykyistä tai potentiaalista käyttötarkoitusta. Datan laadun yhdellä yleisimmistä ulottuvuuksista, eheydellä, viitataan

siihen, onko datalle arvoa. Puuttuva data ei täten ole ehyttä ja esimerkiksi relaatiotietokannoissa epäehyt data esiintyy tyhjänä kenttänä tai arvona "null". Viimeisimpänä datan laadun johdonmukaisuudella pyritään siihen, että data on muodoltaan ja sisällöltään yhtenevää kaikissa dataseiteissä läpi organisaation. (Loshin 2009, s. 91; Geisler *et al.* 2016, s. 10; Mahanti 2019, ss. 79, 93, 114.)

Datan laadunhallinta on prosessien, käytäntöjen, standardien, ihmisten, teknologioiden ja datan hallintaa, jossa tavoitteena on parantaa organisaation datan laatua. Sen tavoitteena ei ole ainoastaan parantaa datan laatua, vaan myös ymmärtää korkealaatuisen datan vaikutukset tavoiteltuihin liiketoimintatuloksiin (Mahanti 2019, s. 317.) Mahantin (2019, s. 319) mukaan datan laadunhallintaa voidaan toteuttaa proaktiivisesti tai reaktiivisesti; proaktiiviset toimet keskittyvät datan laadun ongelmien eliminointiin ennen niiden ilmentymistä, kun taas reaktiivisilla toimilla ratkotaan jo syntyneitä ongelmia. Mahanti (2019, s. 319) toteaa proaktiivisen datan laadunhallinnan olevan reaktiivista lähestymistapaa parempi. Loshinin (2009, s. 102) mukaan datan laadunhallinta tarkoittaa kuitenkin master datan hallinnan osalta sekä ongelmien tai virheiden ennaltaehkäisyä että niiden tunnistamista ja korjaamista.

Datan laadunhallintaan on kirjallisuudessa esitetty erilaisia lähestymistapoja ja metodologioita. Zhang ja Lee (2021, s. 6) tutkimuksessaan tunnistivat yhteensä 32 eri datan laadunhallinnan metodologiaa, viitekehystä tai standardia. Zhangin ja Leen (2021, ss. 7–9) kokoaman yhteenvedon perusteella metodologeista ja viitekehyksistä voidaan tunnistaa samankaltaisia piirteitä: organisaation datan tunnistaminen, nykyisen datan laadun arviointi, liiketoiminnan ymmärtäminen, datan laadun mittaaminen, datan laadun kehittäminen ja datan laadun toteutumisen tarkkailu ovat toistuvia elementtejä eri viitekehysten välillä. Tutkimuksen perusteella Zhang ja Lee (2021, s. 17) huomasivat, että datan laadunhallinnan viitekehyksistä ja metodologeista suurin osa keskittyy strukturoidun tai semi-strukturoidun datan hallintaan ja harvat ottavat huomioon strukturoimatonta dataa. Joidenkin lähteiden mukaan jopa 80 % organisaatioiden datasta on strukturoimatonta dataa eli esimerkiksi tekstidokumentteja, sähköposteja, sosiaalisen median sisältöä tai kuvia (King 2019; Nelson 2020). Täten mahdolliset puutteet tai aukot datan laadunhallinnan viitekehyksissä tulisi ottaa huomioon strukturoimattoman datan osalta.

4. LIIKETOIMINTATIEDON HALLINTA

Liiketoimintatiedon hallinta (Business Intelligence, BI) voidaan kontekstista riippuen määritellä muun muassa teoreettisena lähestymistapana, prosessina tai päätöksenteon työkaluna. Sen perimmäisenä tavoitteena on kuitenkin luoda tukea päätöksentekoon tuottamalla oikeaa tietoa oikeaan aikaan sekä tuottaa datasta relevanttia informaatiota. (Vitt *et al.* 2010; Laihonen *et al.* 2013, s. 44.) Datan merkityksen kasvaessa liiketoiminnan kannalta, yhä useampi organisaatio on ottanut käyttöönsä liiketoimintatiedon hallinnan ja analytiikan käytänteitä (BI&A), ja täten BI&A ja siihen liittyvät työkalut ovat kasvattaneet suosiotaan (Gartner, Inc. 2017; Marjamäki 2017, s. 6).

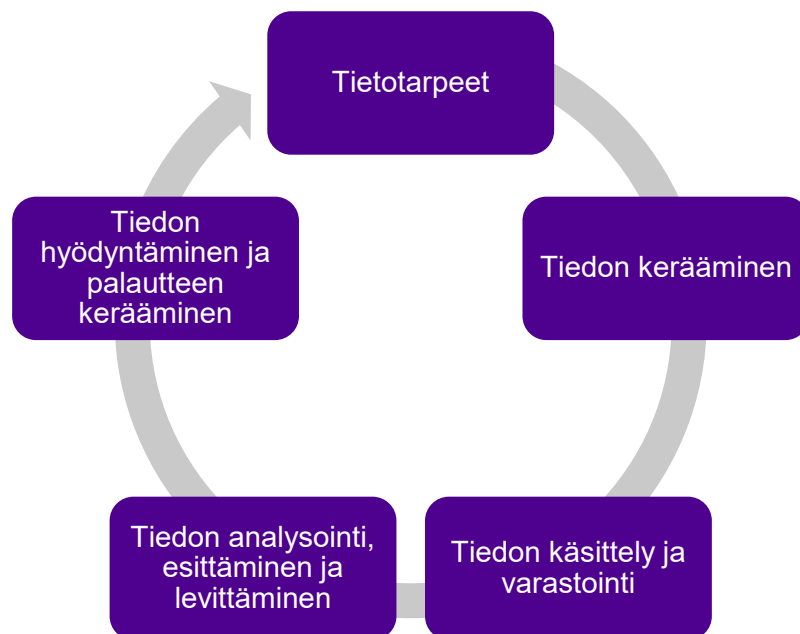
Liiketoimintatiedon hallinnan käsite ei ole säilynyt historian saatossa täysin muuttumattomana. Liiketoimintatiedon hallinnan ensimmäisissä määritelmissä BI nähtiin teknologioiden ja analytiikan työkalujen kokoelmana, jota käytetään tiedon, johtamisen ja sosiaalisten trendien hallinnassa tavoitellen tehokkaampaa organisaation päätöksentekoa. Tällöin BI-prosessi koostui datan keruusta, datan laadunarvioinnista, analysoinnista, datan varastoinnista sekä datan käytöstä. (Ghoshal & Kim 1986, s.127; Gilad & Gilad 1986, ss. 53–61.) Vuonna 2020 julkaistussa tutkimusartikkelissa (Kohtamäki *et al.* 2020, s. 31) liiketoimintatiedon hallinnasta voitiin tunnistaa neljä eri strategista näkökulmaa: BI tuotteenä tai BI prosessina palvelevat strategian toteutumista organisaation ulkoisesta näkökulmasta, kun taas BI järjestelmänä tai BI päätöksenteon työkaluna keskittyvät strategian toteutumiseen organisaation sisäisestä näkökulmasta. Muntean (2018) mukaan liiketoimintatiedon hallinta voidaan nykypäivänä nähdä sateenvarjoterminä teknologioille, järjestelmille ja strategioille, joita organisaatiot käyttävät monimuotoisen datan muokkauksessa relevantiksi informaatioksi ja tukena sekä operationaaliseen, taktiseen ja strategiseen päätöksentekoon.

Organisaatiot voivat saavuttaa monenlaisia hyötyjä toteuttaessaan liiketoimintatiedon hallintaa onnistuneesti. Sen onnistuneeseen toteuttamiseen liittyy kuitenkin myös useita haasteita. Munozin (2018, s. 5) mukaan onnistuneen liiketoimintatiedon hallinnan avulla organisaatiot voivat tehokkaammin tunnistaa liiketoiminnalle relevantin tiedon. Lisäksi BI tarjoaa parempaa ymmärrystä markkinatilanteesta, kuten tulevista trendeistä, joka voi auttaa strategian ja palveluiden selkeyttämisessä. Lisäksi asiakastuntemus, asiakasarvon ymmärtäminen, tiedon ajantasaisuuden ylläpito, liiketoimintakulujen vähentäminen ja liiketoimintatuottojen kasvattaminen sekä onnistunut strateginen liiketoiminnan suunnittelu ja siten kilpailukyvyyn ylläpito ovat esimerkkejä onnistuneeseen liiketoimintatiedon

hallintaan liittyvistä tuloksista tai tekijöistä. Liiketoimintatiedon hallintaan liittyvistä haasteista useampi on riippuvainen ihmisistä: puutteellinen osaaminen tai työntekijän ennakko-oletukset esimerkiksi tiedon keruuseen, oikeanlaiseen raportointiin tai sen analysointiin liittyen, heikko ymmärrys datan tai ylipäättään liiketoimintatiedon hallinnan relevanttiudesta tai datan hyödyntäminen oikeaan aikaan ja oikeassa tilanteessa ovat esimerkkejä ihmisiin liittyvistä liiketoimintatiedon hallinnan haasteista. Lisäksi teknologiaan tai BI-työkaluihin liittyvät puutteet sekä liiketoimintatiedon hallinnan kapea tai puutteellinen hyödyntäminen ovat tunnistettu haasteeksi liiketoimintatiedon hallintaan liittyen. (Munoz 2018, ss. 5–7.)

4.1 Liiketoimintatiedon hallinnan prosessi

Aiemmin luvussa 4 todettiin, että liiketoimintatiedon hallinnan tavoitteena on luoda datasta relevanttia informaatiota. Tätä prosessia varten kirjallisuudessa on esitetty useita prosessimalleja, jotka ovat hyvin samankaltaisia verrattuna yleisesti tiedonhallinnan prosessimalleihin (Hellsten & Myllärniemi 2019, s. 3). Täten kokoamalla yhteen Choon (2002) tiedonhallinnan prosessimallin sekä Helanderin *et al.* (2016) sekä Najdawin ja Patkurin (2021) liiketoimintatiedon hallinnan prosessimallit, voidaan liiketoimintatiedon hallinnan prosessimalli esittää kuvan 3 mukaisesti.



Kuva 3: Liiketoimintatiedon hallinnan prosessimalli (mukailtu Choo 2002; Helander *et al.* 2016; Najdawi & Patkuri 2021)

Liiketoimintatiedon hallinnan prosessi alkaa tietotarpeiden määrittämisestä. Tietotarpeella tarkoitetaan aukkoa nykyisen ja tarvittavan tiedon välillä. Tietotarpeiden määrittäminen on yksi prosessin tärkeimmistä vaiheista, sillä se ohjaa prosessin seuraavia vaiheita ja

toisaalta tietotarpeet on tunnistettava oikein, jotta liiketoimintatavoitteet voidaan täyttää. Tietotarpeiden määrittäminen ei ole yksinkertaista, sillä tietotarpeiden tunnistaminen muuttuvassa toimintaympäristössä ja organisaatiossa on hankalaa. Tietotarpeiden määrittämisessä tulee siis ottaa huomioon sekä liiketoiminnan historia, nykytilanne että myös tulevaisuuden trendit. (Pirttimäki 2007, s. 41; Laihonon *et al.* 2013, s. 25.)

Tietotarpeiden määrittämisen jälkeen tieto kerätään eri ulkoisista tai sisäisistä tietolähteistä, kuten operatiivisista tietojärjestelmistä, asiakkailta, mediasta tai työntekijöiltä. Tietoa voidaan hankkia joko säännöllisenä prosessina jatkuvasti tai ad hoc eli kertaluontoisesti jotakin tiettyä määriteltyä tarvetta varten. (Laihonon *et al.* 2013, ss. 25, 47; Najdawi & Patkuri 2021, s. 201.) BI-prosessin yhteydessä käytetään usein termiä ETL (Extract, Transform & Load), jolla tarkoitetaan datan poimimista tietolähteestä, datan muokkaamista ja lopuksi lataamista tietovarastoon. Tiedon keräämisvaiheessa data siis poimitaan tietolähteestä ETL-prosessin mukaisesti. Seuraavassa, tiedon käsittely- ja varastointivaiheessa eri lähteistä luettava data muokataan tietovaraston vaatimaan muotoon. (Hellsten & Myllärniemi 2019, s. 3; Najdawi & Patkuri 2021, s. 201.) Data yleensä tallennetaan tietokantoihin ja -varastoihin, joista tieto taas kerätään analysointia tai raportointia varten. Tietovarastot toimivat yhtenä olennaisimmista tietolähteistä liiketoimintatiedon hallinnan kontekstissa. (Laihonon *et al.* 2013, ss. 62, 66.)

Tietovarastoinnin jälkeen tietoa voidaan analysoida eli siitä tehdään merkityksellistä tarkasteltavan organisaation kontekstissa (Laihonon *et al.* 2013, s. 48). Analysoitaessa tieto yhdistetään jo organisaatiossa olevaan tietoon ja nykypäivänä tähän käytetään apuna esimerkiksi erilaisia BI-työkaluja (Hellsten & Myllärniemi 2019, s. 4). BI-työkalut mahdollistavat muun muassa datan käsittelyn, analysoinnin ja esittämisen visuaalisessa muodossa. BI-työkalujen suosio onkin ollut merkittävässä kasvussa ja suosion oletetaan kasvavan edelleen. (Munoz 2018, s. 79.) Tiedon analysoinnin tavoitteena on tuottaa tietotuote, josta on helposti ymmärrettävissä tiedon merkitys (Laihonon *et al.* 2013, s. 48). Jotta tietoa voidaan hyödyntää, on analysoidun ja esitettävässä muodossa olevan tiedon oltava saatavissa oikeaan aikaan ja oikeassa paikassa sitä tarvitsevilla tahoilla. Tiedon levittäminen voidaan toteuttaa monin eri tavoin ja eri kanavien välityksellä, eikä eri tavat tai kanavat määrittele itse jaetun tiedon arvoa. (Pirttimäki 2007, s. 75; Laihonon *et al.* 2013, ss. 48–49.)

Viimeisessä liiketoimintatiedon hallinnan prosessimallin vaiheessa tieto lopulta hyödynnetään ja prosessin onnistumista arvioidaan. Tällöin päätöksentekijä toimii saavuttamansa tiedon perusteella. Lopuksi BI-prosessin aikaisempia vaiheita ja niissä saavutettuja tuloksia voidaan arvioida ja mitata. On myös olennaista tarkastella, vastaavatko

alussa määritetyt tietotarpeet sitä tarvetta, joka BI-prosessin lopussa tietoa hyödynnetäessä realisoituu. (Pirttimäki 2007, ss. 75–76.) Iteratiivisessa prosessissa toiminta kehittyy aina palautteen keräämisen ja arvioinnin jälkeen, jolloin muutoksia voidaan tehdä jatkossa esimerkiksi tietotarpeiden määrittämisen osalta.

4.2 Liiketoimintatiedon hallinta ja analytiikka

Liiketoiminta-analytiikka (business analytics, BA) on liiketoimintatiedon hallintaan (BI) liittyvää analytiikkaa kuvaava käsite, joka on syntynyt 2000-luvulla (Chen *et al.* 2012, s. 1166). Goodwynin (2019) mukaan BA käsittelee sitä, miksi jotkut asiat ovat tapahtuneet. Sen sijaan BI ainoastaan kuvailee, mitä on tapahtunut (deskriptiivinen analytiikka). Liiketoiminta-analytiikan avulla voidaan siis selvittää, miksi jotkin asiat ovat tapahtuneet (diagnoosiivinen analytiikka). Lisäksi BA pyrkii ennustamaan, mitä tulee tapahtumaan (prediktiivinen analytiikka) ja ohjaamaan, miten olisi kannattavaa toimia ennustusten perusteella (preskriptiivinen analytiikka). Täten BA auttaa organisaatioita ymmärtämään liiketoiminnasta syntyneitä dataa. (Goodwyn 2019.) BI ja BA voidaan erottaa myös niiden tavoista hakea tietoa. Bentleyyn (2017, s. 3) mukaan BI hyödyntää liiketoimintadataa haikessaan muun muassa tietokantakyselyjä ja raportointia, kun taas BA hyödyntää tilastollisia ja määrällisiä työkaluja.

Liiketoimintatiedon hallinta ja liiketoiminta-analytiikka voidaan kuitenkin nähdä toisiaan täydentävinä käsitteinä ja siksi nykypäivänä konsepteista käytetään usein yhdistävää termiä liiketoimintatiedon hallinta ja analytiikka (business intelligence and analytics, BI&A) (Chen *et al.* 2012, s. 1166; Corte-Real *et al.* 2014, s. 2). BI&A-käsite voidaan esittää myös muun muassa muodossa analytiikka ja liiketoimintatiedon hallinta (Analytics and Business Intelligence, ABI) (Gartner IT-sanakirja 2021).

Liiketoimintatiedon hallinta ja analytiikka (BI&A) voidaan yleisesti määritellä liiketoimintatiedon analysoinnin tekniikoina, teknologioina, käytänteinä, sovelluksina ja metodologioina, joilla tähdätään parempaan liiketoimintaymmärrykseen ja oikea-aikaiseen päätöksentekoon (Chen *et al.* 2012). Gartnerin IT-sanakirjan (2021) mukaan ABI-käsite kattaa Chen *et al.* (2012) teorian lisäksi myös tiedon hallintaan liittyvän kokonaisinfrastruktuurin ja sen tavoitteena on optimaalinen päätöksenteko ja siten myös parempi liiketoiminnan suorituskyky.

Liiketoimintatiedon hallinnan ja analytiikan käsite ei ole pysynyt täysin muuttumattomana viimeisten vuosikymmenien aikana. Sen evoluutiosta voidaan erottaa ainakin kolme vaihetta: BI&A 1.0, BI&A 2.0 sekä BI&A 3.0. (Chen *et al.* 2012.) BI&A:n alkuvaiheissa 2000-

luvulla, BI&A 1.0-konsepti perustui suurimmaksi osaksi datan hallintaan ja tietovarastointiin. Tällöin liiketoiminnassa onnistuttiin hyödyntämään lähinnä strukturoitua dataa ja rakentamaan sen pohjalta kriittistä ja relevanttia tietoa liiketoimintapäätösten tueksi. 2010-luvulla BI&A 2.0-konseptin myötä liiketoimintatiedon hallinnan ja analytiikan avulla onnistuttiin hyödyntämään myös strukturoimatonta dataa. Internetin yleistyminen ja verkkosivustojen määrän kasvu tarjosi myös liiketoiminnalle uuden tietolähteen, jota pyrittiin hyödyntämään BI&A 2.0-aikana. Myöhemmin 2010-luvulla siirryttiin BI&A 3.0-konseptiin, jossa pääpaino kohdistui mobiili- ja sensorilaitteisiin ja niiden tuottamaan dataan. (Chen *et al.* 2012, ss. 1166–1168.) Tuoreemman, vuonna 2019 tehdyn tutkimuksen (Helander *et al.* 2019) perusteella liiketoimintatiedon hallinnan ja analytiikan tulevaisuus tulisi keskittymään aiempaa merkittävämmiin muun muassa ennustavan analytiikan ja sosiaalisen yhteiskehittämisen näkökulmiin, joista jälkimmäinen perustuu ulkoisen ja erityisesti avoimen datan käytön suosion kasvuun. Keskittyessä enemmän ennustavaan analytiikkaan, BI&A siirtyisi voimakkaammin perinteisestä liiketoimintatiedon hallinnasta analytiikan (BA) alueelle, mikä mahdollisesti toisi liiketoimintatiedon hallinnan ja analytiikan käsitteitä yhteen aiempaa merkittävämmiin.

5. MASTER DATAN MERKITYS LIKETOIMINTA-TIEDON HALLINNASSA JA ANALYTIIKASSA

Tässä luvussa käsitellään master datan merkitystä liiketoimintatiedon hallintaa ja liiketoiminta-analytiikkaa yhdistävän konseptin, liiketoimintatiedon hallinta ja analytiikka, kontekstissa. Luvussa 5.1 esitellään master datan hyödyt edellä mainitussa kontekstissa sekä kuvataan master datan ja sen hallinnan sijoittumista liiketoimintatiedon hallinnan prosessimalliin. Luvussa 5.2 tuodaan yhteen kirjallisuudessa havaitut haasteet master datan hallintaan liittyen. Lopuksi luku 5.3 kokoaa kirjallisuudessa tunnistetut optimaaliset toimintatavat ja käytänteet master datan hallintaan osana liiketoimintatiedon hallintaa ja analytiikkaa.

5.1 Master data osana liiketoimintatiedon hallintaa ja analytiikkaa

Master dataa voidaan hyödyntää sekä historiatiedon analysoinnissa että ennakoivassa analytiikassa. Master dataa voidaan siis hyödyntää myös osana liiketoimintatiedon hallintaa ja analytiikkaa (BI&A). (White 2007, s. 6.) Kuten aiemmin luvussa 3.1 todettiin, master dataa hyödynnetään läpi organisaation. Lisäksi tapahtumatieto, eli transaktiodata, sisältää master dataa ja tässä kontekstissa master data toimiikin perustana transaktioille. (Loshin 2009, ss. 6, 8; Väre 2019, ss. 14–16.)

Liiketoimintatietotarpeiden mukaisesti master dataa voidaan hakea eri sisäisistä ja ulkoisista tietolähteistä (Pansara 2021). Aiemmin esitetyssä kuvassa 3, eli liiketoimintatiedon hallinnan prosessimallissa, master data näyttäytyy siis tiedon keräämisvaiheessa, jossa tietotarpeeksi määritelty master data kerätään tietystä tietolähteestä. Täten master data on osana koko liiketoimintatiedon hallinnan prosessia, sillä kerätty tieto jatkaa aina prosessin seuraaviin vaiheisiin. Kirjallisuudessa on myös esitetty prosessimalleja master datan hallintajärjestelmän sijainnille kokonaisuuksiin, joissa tietolähteistä tietoa tuodaan joko tietovarastoihin tai aina ennakoivaan analytiikkaan saakka. Whiten (2007, ss. 7–8) prosessimallin mukaan master dataa voidaan hallita joko ennen tietovarastointia master datan hallintajärjestelmässä, tai vaihtoehtoisesti vasta tietovarastossa. Zoderin (2011, s. 9) mukaan master datan hallinta näkyy osana BI-prosessia tiedon käsittelyvaiheessa ja erityisesti ETL-prosessin ympärillä. Zoderin (2011, s. 9) prosessimallin mukaan tieto kerätään eri lähdejärjestelmistä, minkä jälkeen dataa käsitellään vaaditulla tavalla ETL-prosessin mukaisesti. Tämän jälkeen prosessimallissa esiintyy master datan hallinta, jonka jälkeen taas dataa käsitellään ETL-prosessissa. Edellä mainittujen vaiheiden jälkeen

data siirtyy tiedon analysointivaiheeseen, jossa prosessimallin mukaisesti dataa voidaan hyödyntää ennakoivassa analytiikassa. (Zoder 2011, s. 9.) Julkusen (2021) mukaan master datan hallintaan käytetyllä työvälineellä tai tietoteknisellä ratkaisulla ei kuitenkaan ole niin suurta merkitystä, kuin mitä itse prosessilla, jolla master dataa hallitaan.

Master datalla on monenlaisia vaikutuksia liiketoimintatiedon hallintaan ja analytiikkaan. Väreen (2019, s. 21) mukaan ilman master dataa ei ole liiketoimintaa – organisaation datan näkökulmasta master data toimii liiketoiminnan perustana. Myös Helsingissä julkaistussa Kuntasektorin MDM-viitearkkitehtuurissa (2013, s. 9) on master datan todettu kuuluvan organisaation kokonaisarkkitehtuurin jokaiselle osa-alueelle, jotka käsittävät organisaation toiminnan, tiedot, tietojärjestelmät sekä teknologiat.

Kuntasektorin tekemän MDM-viitearkkitehtuurin (2013, s. 13) perusteella heidän toiminnassansa master data ja sen hallinnointi näkyy muun muassa seuraavina hyötyinä liiketoimintatiedon hallinnassa ja analytiikassa: parempi raportoinnin laatu, useampien samansisältöisten rekisterien ylläpidon vähentäminen ja siten päällekkäisen työn vähentäminen sekä virheellisen tiedon aiheuttaman turhan työn vähentäminen. Jo aiemmin luvussa 3.3 todettiin, että datan laadun tarkkuus, eli tiedon oikeellisuus, on yksi viitatuimmista datan laadun ulottuvuuksista ja haasteeksi tämän kohdalla nähtiin datan vertailuarvon puutteellisuus ja tiedon muuttuva luonne. Kuntasektorin lisäksi myös Haug ja Arlbjorn (2011, s. 289) sekä Väre (2019, s. 16) teoksissaan totesivat, että master datalla on merkittävä vaikutus raportoinnin luotettavuuteen. Haugin ja Arlbjornin (2011, s. 289) mukaan tämä perustui siihen, että master data vaikuttaa myös transaktiodatan luotettavuuteen. Datan laadun tarkkuuden ulottuvuuden lisäksi datan laadun luotettavuus oli myös yksi viitatuimmista ulottuvuuksista. Mahantin (2019, s. 122) mukaan luotettava data näyttäytyy relevanttina, tarkkana, ainutlaatuisena, johdonmukaisena ja lähes puutteettomana. Whiten (2007) sekä Vilminko-Heikkisen ja Pekkolan (2017) teoksissa todetaan myös, että master datan hallinnalla pyritään epäjohdonmukaisuuksien eliminointiin. Datan laadun johdonmukaisuuden ulottuvuuden tarkoituksena onkin tavoitella tilannetta, jossa data on muodoltaan ja sisällöltään yhtenevää läpi organisaation (Loshin 2009, s. 91; Mahanti 2019, ss. 79, 93, 114).

Johdannossakin mainitun sanonnan mukaan organisaation tekemät päätökset ovat korkeintaan yhtä hyviä kuin data, johon päätökset perustetaan (Dayton 2007; Haug *et al.* 2009). Myös Loshinin (2009, ss. 10–12) mukaan onnistuneella master datalla ja sen hallinnalla voidaan tuottaa laadukkaampaa ja nopeampaa päätöksentekoa sekä parantaa tiedonlaatua. Edellä mainittujen kirjallisuuslähteiden perusteella voidaan siis vetää johtopäätös, että datan laadulla sen eri ulottuvuuksineen voidaan nähdä yhteys liiketoimintatiedon hallintaan ja analytiikkaan.

Loshinin (2009, ss. 10–12) ja Väre (2019, s. 49) mukaan onnistuneella master datan hallinnalla voidaan myös parantaa organisaation operationaalista tehokkuutta. Tällä tarkoitetaan sitä, että samaa master dataa ei syötetä tai päivitetä eri järjestelmiin useampaan kertaan. Tällöin tarvittava master data on valmiiksi saatavilla sillä hetkellä, kun sitä tarvitaan. (Väre 2019, s. 49.) Liiketoimintatiedon hallinnan prosessimallin näkökulmasta master dataa ei siis ole tarpeellista käsitellä ETL-prosessin mukaisesti joka käyttökerroilla, vaan master data on löydettävissä tietovarastosta käsiteltynä.

5.2 Master dataan liittyvät haasteet

Edellä luvussa 5.1 todettiin, että datan laadulla sen eri ulottuvuuksineen voidaan nähdä yhteys liiketoimintatiedon hallintaan ja analytiikkaan. Merkittävä datan määrän kasvu on nostanut esille ajankohtaisen haasteen datan laadunhallinnassa (Vilminko-Heikkinen & Pekkola 2017, s. 454). Väre (2019, s. 189) mukaan datan laatu, datan puute tai sen laadunhallinnan puute ovatkin yksi yleisimmistä syistä master datan hallinnan aloittamiseksi organisaatioissa. Toisaalta Kekwaletswe ja Lesole (2016, s. 4) tuovat teoksessaan ilmi, että yksi master datan hallintaan liittyvistä haasteista on toimittaa tarvittua ja laadultaan tarkkaa informaatiota oikeaan aikaan. Kuten aiemmin on todettu, tämä on yksi liiketoimintatiedon hallinnan päätavoitteista.

Kekwaletswe ja Lesole (2016, s. 4) teoksessaan toteavat, että useimmiten haasteet master datan osalta liiketoimintatiedon hallinnassa, ja erityisesti raportoinnissa, kohdistuvat muun muassa datan duplikaatteihin, datan heikkoon laatuun, datan epätarkkuuteen, epätäydellisyyteen tai sen epärelevanttiuteen. Datan duplikaatit on tunnistettu myös muussa kirjallisuudessa merkittäväksi master dataan liittyväksi haasteeksi (Loshin 2009, s. 251; Bentley 2017, s. 141; Vilminko-Heikkinen & Pekkola 2019, s. 79). Bentley (2017, s. 141) mukaan esimerkiksi yritysfuusioiden jälkeen master datan haasteet voivat näkyä juuri datan duplikaatteina, kun kahden organisaation dataa yhdistetään. Datan duplikaatit vievät ylimääräistä tilaa tietovarastoissa ja vaativat ylimääräistä ylläpitotyötä aiheuttaen lisäkustannuksia. Toisaalta eliminoimalla datan duplikaatit, voidaan parantaa datan laatua sekä kasvattaa tehokkuutta datan saatavuuteen (Iqbal 2020, ss. 2, 4.) Täten master datan laadunhallinta on osana liiketoimintatiedon hallinnan prosessia, etenkin tiedon käsittelyvaihetta. MDM-kehittämisen ja käyttöönottoprosessissa master datan laadunhallinta taas on osana muun muassa ylläpitoprosessin sekä datakäytänteiden määrittelyssä ja MDM-mittariston määrittelyssä.

Baghi *et al.* (2014, s. 3832) tunnistivat tutkimuksessaan, että organisaatiot kokevat haasteelliseksi määrittää juuri heille sopivan master datan hallintajärjestelmäarkkitehtuurin. Pansari (2021, s. 48) teoksessaan jatkaa, että haasteena nähdään myös master datan

ja sen hallintajärjestelmän yhteensopivuus muihin organisaation hyödyntämiin sovelluksiin tai järjestelmiin. Tutkimusartikkelissaan Baghi *et al.* (2014, s. 3831) ehdottavat yleispätevää mallia master datan hallintajärjestelmän valintaan, ottaen huomioon muun muassa sen kustannukset, riskit, organisaatiotyyppin sekä yhteensopivuuden muiden järjestelmien kanssa. Master datan ylläpitoprosessia ja integraatioita muihin järjestelmiin käsiteltiin aiemmin luvussa 3.2 mainitun MDM-kehittämisen vaiheissa 4 ja 10. Baghin *et al.* (2014, s. 3831) esittämää mallia järjestelmäarkkitehtuuriin liittyen voidaan mahdollisesti hyödyntää näissä MDM-kehittämisen vaiheissa, mikäli organisaatio kokee haasteita sopivan järjestelmäarkkitehtuurin suunnittelussa.

Vilminko-Heikkisen ja Pekkolan (2017) toteuttaman etnografisen tutkimuksen perusteella master datan hallintaan liittyy kahdeksan haastetta: terminologia ja datan käsitteet, master datan omistajuus ja sen yhteisymmärrys, tietohallinnon vastuut ja tehtävät, data-settien rakeisuus, lainsäädäntö sekä master datan hallinnan operationaalisen ja hallinnollisen omistajuuden roolien erottaminen. Yhteenvetona tutkimuksen perusteella havaituista haasteista voidaan todeta, että haasteet master datan hallinnassa kohdistuvat organisaation käytäntöihin, toimintatapoihin ja kulttuuriin, eikä niinkään tietotekniseen ulottuvuuteen. Myös Loshinin (2009, s. 16) mukaan master datan hallinnan onnistuminen perustuu organisaation toimintaan, eikä master datan hallintaa tulisi siis käsittää esimerkiksi työkaluna tai projektina. Samaa näkökulmaa puoltaa myös Julkunen (2021) podcast-haastattelussaan todetessaan, että master datan hallinnan prosessilla on suurempi merkitys verraten siihen käytettyihin työkaluihin tai teknisiin ratkaisuihin.

Aiemmin kuvassa 2 esitetyssä master datan hallinnan kehittämissä vaiheissa useat vaiheet koskettavat organisaation prosesseja, joita työstävät organisaatioon kuuluvat ihmiset. Master datan hallinnan kehittäminen voi aiheuttaa muutoksia näissä prosesseissa. Loshinin (2009, s. 16) ja Väreän (2019, s. 182) mukaan master datan hallinnan yhtenä haasteena voidaan nähdä organisaation heikko valmius muutokseen ja toisaalta muutosten sisäistäminen voi olla vaikeaa. Vilminko-Heikkisen ja Pekkolan (2017, s. 468; 2019, s. 80) taas nostivat esiin master datan omistajuuden määrittämisen ja roolin ymmärtämisen yhtenä merkittävimpänä master datan hallinnan haasteena. Master datan hallinnan kehittämissä vaiheissa master datan omistajuuden määrittäminen on osa tiedonhallinnan määrittelyä. Girardin (2020, ss. 3, 5) mukaan tiedonhallinnalliset määrittelyt, kuten master datan omistajuuden määrittäminen, vaikuttavat muun muassa operatiiviseen ja strategiseen toimintaan sekä datan luotettavuuteen ja läpinäkyvyyteen.

5.3 Master datan optimaalinen hyödyntäminen osana liiketoimintatiedon hallintaa ja analytiikkaa

Mikäli organisaatio ei ole vielä suorittanut toimenpiteitä master datansa hallintaan, tulisi siihen Julkusen (2021) mukaan kiinnittää huomiota mahdollisimman pian. Master datan hallinnan merkitys on Väre (2019, s. 233) mukaan helppo havaita aloittamalla datan laadun raportointi ja seuranta, sillä datan laadulliset haasteet paljastavat master datan hallinnan puuttumisen. Väre (2019, s. 192) kuitenkin huomauttaa, että vaikka laadunhallinnalliset toimenpiteet tulee kohdistaa koko organisaation dataan, on master data kuitenkin kriittisintä ja täten ensimmäiset toimenpiteet tulisi kohdistaa nimenomaan master dataan. Kun datan laadunhallinta ja master datan hallinta on otettu käytäntöön organisaatiossa, tulee hallintaa jatkaa sekä kehittää organisaatiossa jatkuvasti (Väre 2019, s. 192).

Kekwaletswe ja Lesole (2016, s. 3) esittävät master datan hallintaan kaksi lähestymistapaa: liiketoiminnallinen sekä tekninen lähestymistapa. Liiketoiminnallisen lähestymistavan mukaan, jotta master datasta voidaan hyötyä osana päätöksentekoa, on master datatarpeiden tuettava liiketoimintatavoitteita (Kekwaletswe & Lesole 2016, s. 11; Väre 2019, s. 39.) Täten master datan optimaalinen hyödyntäminen liiketoimintatiedon hallinnassa ja analytiikassa tulee aloittaa yhteensovittamalla master datan kehittämisprosessin sekä liiketoimintatiedon hallinnan prosessimallin ensimmäiset vaiheet. Kekwaletswe ja Lesole (2016, s. 3) huomauttavat, että master datan hallinnalliset toimenpiteet tulisi siis perustua liiketoiminnan tarpeisiin. Tekninen lähestymistapa taas keskittyy teknologioihin, teknisiin työkaluihin ja toimintatapoihin, joiden avulla voidaan tavoitella tehokasta master datan hallintaa ja master datan roolia yhtenä, luotettavana versiona totuudesta. Edellä mainittujen kahden, liiketoiminnallisen sekä teknisen, lähestymistavan yhdistäminen kuitenkin palvelee parhaalla mahdollisella tavalla master datan hallinnan onnistumista. Yhdistetty lähestymistapa tukee sekä tiedonhallinnallista, sosiaalista että datan laadun edistämisen näkökulmia, jotka kaikki nähdään välttämättöminä toimenpiteinä onnistuneen master datan hallinnan osalta. (Kekwaletswe & Lesole 2016, s. 3.)

Edellisessä luvussa 5.2 todettiin haasteiden kohdistuvan master datan hallinnan osalta muun muassa organisaation käytäntöihin, toimintatapoihin ja tiedonhallinnan määrittelyyn, eikä niinkään tietotekniseen puoleen. Kekwaletswe ja Lesole (2016, s. 9) tutkimuksensa pohjalta ehdottavat, että osana tiedonhallinnan määrittelyä organisaatioiden tulisi nimetä tietovastaavat (engl. data steward) vastaamaan datan eheydestä ja datan laadun mittaamisesta kuhunkin organisaation yksikköön. Uudemman, vuonna 2019 julkaistun tutkimusartikkelin (Vilminko-Heikkinen & Pekkola 2019, s. 80) mukaan tiedonhallinnalliset roolit MDM-kehittämiseen tulisi jakaa kolmeen: organisaatio-, tukitoiminto- sekä data-

alue (engl. data domain) -osiin, jotka käsittävät muun muassa roolit tietoturvan, datan laadunhallinnan, integraatioiden sekä ylläpidon osalta. Vilminko-Heikkinen ja Pekkola (2019, s. 80) jatkavat, että kyseiset roolit koettiin merkittävänä tekijänä MDM-kehittämisen prosessissa, mutta varsinaisen datan omistajuuden määrittäminen oli välttämätöntä.

Master datan laadunhallinnallisia optimaalisia käytänteitä ongelmatilanteissa käsitellään Väreen (2019, ss. 194, 197) teoksessa. Datan laadunhallinnallisten ongelmien kohdistuessa suoranaisesti ainoastaan dataan, toimenpiteinä käytetään lähinnä reaktiivisia tapoja, joilla virheet korjataan. Organisaatioissa tulisi kuitenkin pyrkiä proaktiivisiin toimiin, sillä jälkikäteen korjatut virheet ehtivät yleensä aiheuttaa ylimääräisiä kustannuksia. Täten master datan ja datan laadunhallinnan osalta virheet datassa tulisi korjata heti virheen syntyminen jälkeen, jotta ylimääräisiltä kustannuksilta voidaan välttyä. (Väre 2019, ss. 194, 197.) Myös Julkunen (2021) podcast-haastattelussaan toteaa, että master datan virheet tulisi korjata lähdejärjestelmään mahdollisimman pian, eikä vain siksi, että myöhemmin korjaaminen on kalliimpaa, vaan se on myös vaikeampaa. Tällöin data tulee siis korjata jo lähdejärjestelmään, josta data kerätään, eikä esimerkiksi raportointivaiheessa erikseen jokaiselle raportille.

Optimoidakseen MDM-kehittämisen jalkauttamisen organisaatiossa, Väre (2019, s. 179) on koonnut kolme periaatetta: jalkauttaminen osissa, kommunikoi ja tue rooleja tarpeeksi. Jotta MDM-kehittämisen käyttöönotossa voitaisiin siis onnistua, tulisi siihen liittyvät käytänteet ja toimintatavat jalkauttaa pienissä osissa tai esimerkiksi yksikkö kerrallaan, jotta MDM-prosessi ehditään sisäistämään ja sen eri osa-alueille ehditään tarjoamaan tukea. Tuen antaminen eri osapuolille mahdollistaa sen, että master dataa ylipäätään voidaan hallita onnistuneesti, sillä hallinnallisia toimenpiteitä ei voida suorittaa ilman sitä vaadittavaa osaamista. Viimeisimpänä muistutetaan aktiivisesta kommunikoinnista MDM-prosessin käyttöönoton yhteydessä. Sen tulisi käsittää muun muassa MDM-prosessin hyötyjen viestimisen, käyttöönottoprosessin etenemisen sekä prosessissa tapahtuneet muutokset. (Väre 2019, ss. 179–182.)

6. PÄÄTELMÄT

Tässä kirjallisuuskatsauksessa tutkittiin master dataa ja sen merkitystä suositaan kasvattaneeseen liiketoimintatiedon hallintaan ja analytiikkaan. Tässä luvussa käydään läpi tutkimuksen tulokset sekä arvioidaan tutkimuksen onnistumista. Luvun lopuksi pohditaan mahdollisia jatkotutkimuskohteita.

6.1 Tutkimustulokset

Tämän kirjallisuuskatsauksen tutkimusaineisto koostui niin artikkeleista, alan konferenssijulkaisuista sekä kirjoista, kuten myös ajankohtaisista lehdistötiedotteista ja podcast-haastattelusta. Tutkimusaineiston perusteella voitiin luoda kokonaiskuva master datan merkityksestä liiketoimintatiedon hallinnassa ja analytiikassa, sekä siihen liittyvistä haasteista ja optimaalisista käytänteistä master datan hallintaan osana liiketoimintatiedon hallinnan kontekstia. Jo johdannosta lähtien voitiin todeta, että master data, sen hallinta sekä lisäksi myös datan laatu ja sen hallinta ovat merkityksellisiä liiketoimintatiedon hallintaa harjoittaessa.

Tutkimuksen perusteella master data on osana koko liiketoimintatiedon hallinnan prosessia, joka on esitetty aiemmin kuvassa 3. Master data voidaan nähdä tietotarpeena, jota haetaan eri tietolähteistä, hyödyntäen sitä myöhemmin prosessimallin mukaisesti. Master dataa voidaan organisaatiosta riippuen hallita joko tietovarastoon tuotaessa ETL-prosessin mukaisesti, tietovarastossa tai jo ennen tietovarastointia master datan hallinnalle osoitetussa järjestelmässä. Tulosten perusteella kuitenkin master datan hallinnassa merkityksellisempää on se prosessi, jota master datan hallintaan käytetään, kuin siihen käytetyt työvälineet tai tietotekniset ratkaisut. Master datan hyödyiksi liiketoimintatiedon hallinnan osalta tunnistettiin muun muassa seuraavat: datan laadun ja raportoinnin luotettavuuden parantaminen, epäjohdonmukaisuuksien eliminointi, laadukkaampi ja nopeampi päätöksenteko, tiedon laadun parantaminen sekä operationaalisen tehokkuuden parantaminen.

Kirjallisuuskatsauksessa voitiin tutkimusaineistosta tunnistaa monia master dataan liittyviä haasteita liiketoimintatiedon hallinnan kontekstissa. Yhdeksi liiketoimintatiedon hallinnan kontekstin kannalta olennaisimmaksi haasteeksi voitiin todeta haasteet toimittaa tarvittua ja laadukasta informaatiota oikeaan aikaan. Master datan haasteiden yksi osa-alue koski datan laatua – haasteet esimerkiksi datan heikossa laadussa, duplikaateissa tai epärelevanttiudessa todettiin aiheuttavan muun muassa ylimääräistä työtä ja täten

myös ylimääräisiä kustannuksia. Tuloksissa korostuivat myös organisaation käytänteisiin, toimintatapoihin ja prosessiin liittyvät haasteet. Haasteet tässä kontekstissa koskivat muun muassa heikkoa muutosvalmiutta sekä master datan omistajuuden määrittämistä ja sen roolin ymmärtämistä.

Tämän työn toisen, ja samalla viimeisen, alatutkimuskysymyksen avulla pyrittiin selvittämään kirjallisuuden perusteella optimaaliset toimintatavat ja käytänteet master datan hallintaan osana liiketoimintatiedon hallintaa ja analytiikkaa. Tulosten perusteella master datan hallinnassa tulisi huomioida sekä liiketoiminnallinen että tekninen näkökulma. Liiketoiminnallisesta näkökulmasta master datatarpeiden tulee tukea liiketoimintatavoitteita, tiedonhallintaan osallistuvat tahot on nimettävä ja heidän roolinsa on määriteltävä, kommunikoinnin tulisi olla aktiivista prosessin aikana ja lisäksi jalkauttamiseen ja master datan hallintaan on tarjottava tukea sitä tarvitseville tahoille. Teknisestä näkökulmasta tärkeimpänä elementtinä tutkimuksesta ilmeni tarve reaktiivisille toimille. Tuloksissa nousi esille, että mikäli master datan hallintaan ei vielä ole kiinnitetty huomiota, tulisi se tehdä mahdollisimman pian.

6.2 Tutkimuksen arviointi

Tutkimuksen tavoitteena oli tunnistaa, millainen merkitys master datalla on liiketoimintatiedon hallinnan kontekstissa. Tavoitteena oli syventää aiheen ymmärrystä alatutkimuskysymysten avulla, jotka käsittelivät master dataan liittyviä haasteita ja sen optimaalista hyödyntämistä liiketoimintatiedon hallinnan ja analytiikan kontekstissa. Kirjallisuuskatsauksen avulla pystyttiin vastaamaan kaikkiin tutkimuskysymyksiin. Aiempaa kirjallisuutta suoranaisesti master datan tai sen hallinnan sekä liiketoimintatiedon hallinnan ja analytiikan yhteydestä oli saatavilla vain vähän. Kirjallisuudessa näiden käsitteiden välillä oli tehty ainoastaan yksittäisiä havaintoja, jotka tähän kandidaatintyöhön pyrittiin koamaan yhtenäiseksi kokonaisuudeksi. Läpi työn eri kirjallisuuslähteistä löydetty havainnot kuitenkin pääosin tukivat toisiaan, mikä parantaa tämän kandidaatintyön luotettavuutta.

Alkuperäisenä tavoitteena oli myös pyrkiä tuomaan master data osaksi liiketoimintatiedon hallinnan prosessimallia, joka esitettiin kuvassa 3. Tämän tavoitteena olisi ollut syventyä päätutkimuskysymykseen. Aiemman kirjallisuuden perusteella tätä ei kuitenkaan voitu toteuttaa, sillä luvussa 5.1 esitettyjen tulosten perusteella master datalle tai sen hallinnalle ei voitu määrittää yksiselitteistä sijaintia liiketoimintatiedon hallinnan prosessimalliin. Tämän työn teemana oli tuottaa yleispäteviä ratkaisuja, mutta organisaatio-spesifisti master datan tai sen hallinnan sijoittaminen liiketoimintatiedon hallinnan prosessimalliin olisi varmasti mahdollista.

Tässä työssä hyödynnettiin Finkin (2005, s. 3) seitsenvaiheista kirjallisuuskatsauksen prosessimallia. Työ eteni mallin mukaan vaiheittain aina tutkimuskysymyksen asettamisesta, tietokantojen ja muiden lähteiden valinnasta hakutermin valintaan, käytännön seulan asettamiseen sekä metodologisen seulan asettamiseen. Kuudennen vaiheen eli katsauksen tekemisen aikana työssä jouduttiin kuitenkin palaamaan takaisin ensimmäiseen vaiheeseen eli tutkimuskysymyksen asettamiseen. Tällöin yksi alatutkimuskysymyksistä, koskien master datan roolia liiketoimintatiedon hallinnan historiassa ja sen tulevaisuudessa, suljettiin pois tutkimuksesta. Aiemmassa kirjallisuudessa on käsitelty master datan sekä liiketoimintatiedon hallinnan historiaa erikseen, mutta ei varsinaisesti yhdessä. Kirjallisuuden perusteella ei voitu myöskään luoda tarpeeksi kattavaa ja laadukasta vastausta siihen, mikä näiden käsitteiden yhteys tulisi olemaan liiketoimintatiedon hallinnan tulevaisuudessa. Finkin malli esitetään prosessina, joka etenee systemaattisesti kohta kerrallaan tutkimuskysymyksen asettamisesta tuloksien syntetisointiin. Vaikka tässä työssä mallia sovellettiin projektin aikana, kokonaisuudessaan mallin soveltamisen voidaan uskoa tukevan Finkin prosessimallinkin tavoitetta korkealaatuisesta kirjallisuuskatsauksesta.

6.3 Jatkotutkimusmahdollisuudet

Luvussa 3.1, todettiin, ettei master datalle ole yksiselitteistä määritelmää. Master data kuitenkin määriteltiin muun muassa kohtalaisen muuttumattomana ja pitkäaikaisena datana, jota voi olla esimerkiksi asiakas- tai tuotetiedot. Myöhemmin luvussa 5.2 käsiteltiin master dataan liittyviä haasteita ja yhdeksi haasteeksi oli tunnistettu master datan käsite. Kun master data käsitetään kohtalaisen muuttumattomana datana, voidaanko esimerkiksi viikoittain päivitettäviä asiakastietoja kutsua master dataksi? Näin ollen jatkotutkimuksissa olisi perusteltua selkiyttää master datan käsitettä. Koska master datan käsite voi vaihdella organisaatioittain, olisi master datan käsitteen tutkiminen tarpeellista toteuttaa organisaatiokohtaisesti.

Vilminko-Heikkisen ja Pekkolan artikkelissa (2017, s. 471) sivutaan jo jatkotutkimusmahdollisuuksia sen suhteen, miten organisaation tyyppi vaikuttaa MDM-haasteisiin. Tässä kandidaatintyössä master dataa ja sen merkitystä liiketoimintatiedon hallinnan kontekstissa tarkasteltiin yleispätevästi, eikä niinkään organisaatiospesifisti. Koska master datan hallinnan käytänteet ja toimintatavat eriävät organisaatioiden kesken, olisi perusteltua tutkia tämän työn päättämiskysymystä, master dataan liittyviä haasteita sekä optimaalisia käytänteitä ja toimintatapoja tietyissä organisaatioissa tai organisaatiotyypeissä.

Kuten edellisessä luvussa, eli tutkimuksen arvioinnissa, todettiin, tässä kandidaatintyössä ei voitu tutkimusaineiston perusteella määrittää master datalle tai sen hallinnalle

yksiselitteistä sijaintia liiketoimintatiedon hallinnan prosessimallissa. Lisäksi aiemmin luvussa 5.2 todettiin haasteeksi ymmärtää master datan käsite ja käytänteet tai toimintatavat siihen liittyen. Master datan tai sen hallinnan sijoittaminen edellä mainittuun prosessimalliin voisi siis edistää master dataa hyödyntävien tahojen ymmärrystä siitä, millainen merkitys master datalla on liiketoimintatietoa käsitellessä eli mistä master data kerätään tai missä vaiheessa ja miten sitä käsitellään. Koska master datan hallinnan käytänteet eriävät organisaatioittain, olisi tämän aiheen tutkiminen perusteltua tehdä organisaatiospesifisti.

LÄHTEET

- Baghi, E., Schlosser, S., Ebner, V., Otto, B. & Oesterle, H. (2014). Toward a Decision Model for Master Data Application Architecture. 47th Hawaii International Conference on System Sciences, 3827–3836. Saatavissa (viitattu 14.12.2021): <https://doi.org/10.1109/HICSS.2014.475>.
- Ballou, D. & Pazer, H. (1985). Modeling Data and Process Quality in Multi-Input, Multi-Output Information Systems. *Management Science*, 31(2).
- Bentley, D. (2017). *Business Intelligence and Analytics*. Library Press, New York, United States.
- Berson, A., & Dubov, L. (2011). *Master data management and data governance*. Second edition. McGraw-Hill, New York, United States.
- Caballero, I., Perez-Castillo, R., Carretero, A.G., Rodriguez, M., Piattini, M., Mate, A., Kim, S. & Lee, D. (2018). DAQUA-MASS: An ISO 8000-61 based data quality management methodology for sensor data. *Sensors*, 18(9).
- Chen, H., Chiang, R. H. & Storey, V. (2012). Business Intelligence and Analytics: From Big Data to Big Impact. *MIS Quarterly*, 36(4), 1165–1188. Saatavissa (viitattu 24.10.2021): <https://doi.org/10.2307/41703503>.
- Choo, C.W. (2002). *Information management for the intelligent organization: the art of scanning the environment*. Information Today, Inc.
- Corte-Real, N., Oliveira, T. & Ruivo, P. (2014). Understanding the hidden value of business intelligence and analytics (BI&A). 20th Americas Conference on Information Systems (AMCIS 2014): Smart Sustainability: The Information Systems Opportunity. AIS/ICIS, Georgia, 20th Americas Conference on Information Systems, AMCIS 2014. Savannah, GA, United States.
- Dayton, M. (2007). Strategic MDM: the foundation of enterprise performance management. *Cutter IT Journal*, Massachusetts, United States, 20(9), 13-17.
- Fink, A. (2005). *Conducting Research Literature Reviews: From the Internet to the Paper*. Thousand Oaks: Sage Publications, Inc. California, United States.
- Foote, K. (2019). A Brief History of Master Data. Saatavissa (viitattu 3.11.2021): <https://www.dataversity.net/a-brief-history-of-master-data/#>.
- Gartner, Inc. (2017). Gartner Says Worldwide Business Intelligence and Analytics Market to Reach \$18.3 Billion in 2017. Saatavissa (viitattu 28.11.2021): <https://www.gartner.com/en/newsroom/press-releases/2017-02-17-gartner-says-worldwide-business-intelligence-and-analytics-market-to-reach-18-billion-in-2017>.
- Gartner IT Glossary. (2021). Analytics and Business Intelligence (ABI). Saatavissa (viitattu 30.11.2021): <https://www.gartner.com/en/information-technology/glossary/business-intelligence-bi>.
- Geisler, S., Quix, C., Weber, S. & Jarke, M. (2016). Ontology-based data quality management for data streams. *Journal of Data Information Quality*, 7(4), 1–34.

- Ghoshal, S. & Kim, S. (1986). Building effective intelligence systems for competitive advantage. *Long Range Planning*, 20(4).
- Gilad, B. & Gilad, T. (1986). SMR Forum: Business Intelligence - The Quiet Revolution. *Sloan Management Review*, 27(4).
- Girard, M. (2020). Helping Organizations Master Data Governance. Centre for International Governance Innovation.
- Goodwyn, G. (2019). Business Analytics. *The Armed Forces Comptroller*, 64(1), 33-34.
- Haug, A., Arlbjorn, J.S. & Pedersen, A. (2009). A classification model of ERP system data quality. *Industrial Management & Data Systems*, United Kingdom, 109(8), 1053-68. Saatavissa (viitattu 23.10.2021): <https://doi.org/10.1108/02635570910991292>.
- Haug, A. & Arlbjorn, J. (2011). Barriers to master data quality, 24(3), 288-303. *Journal of Enterprise Information Management*, United Kingdom.
- Harvard Business Review Analytic Services & Profisee. (2021). The Path to Trustworthy Data: Effective Strategies to Bolster Enterprise Data Management. Harvard Business School Publishing. Saatavissa (viitattu 15.12.2021): <https://profisee.com/harvard-survey-on-trustworthy-data/>.
- Helander, N., Hellsten, P. & Myllärniemi, J. (2016). Business Intelligence Process Model As A Learning Method. *TOJET Turk. Online J. Educ. Technol.*
- Helander, N., Ratia, M. & Myllärniemi, J. (2019). The potential beyond IC 4.0: the evolution of business intelligence towards advanced business analytics. *Measuring Business Excellence*, 23(4), 396–410. Saatavissa (viitattu 30.11.2021): <https://doi.org/10.1108/MBE-12-2018-0103>.
- Hellsten, P. & Myllärniemi, J. (2019). Business intelligence process model revisited. SCITEPRESS.
- IDC, Statista (2021). Volume of data/information created, captured, copied, and consumed worldwide from 2010 to 2025 (in zettabytes). Statista. Statista Inc. Saatavissa (viitattu 24.10.2021): <https://www.statista.com/statistics/871513/worldwide-data-created/>.
- ISO. (2017). ISO 8000-2:2017 Data Quality—Part 2: Vocabulary; Standard, International Organization for Standardization/TC 184/SC 4 Industrial Data. Geneva, Switzerland.
- Iqbal, M., Tahir, M., Naqvi, M. R. & Shahzad, S. K. (2020). Resolving Data De-Duplication issues on Cloud. 2020 International Conference on Engineering and Emerging Technologies (ICEET), 1–5. Saatavissa (viitattu 15.12.2021): <https://doi.org/10.1109/ICEET48479.2020.9048214>.
- Julkunen, J. (2021). Master Data ja tiedon laatu [Podcast-jakso]. Ajan ilmiöitä analytiikasta – Invenco podcast. Spotify. Saatavissa (viitattu 2.11.2021): <https://open.spotify.com/episode/6D2HiyA9KfgyHwEYiAawce?si=5b59e364b5a4483d>.
- Kekwaletswe, R. & Lesole, T. (2016). A Framework for Improving Business Intelligence through Master Data Management. *Journal of South African Business Research*, 1–12. Saatavissa (viitattu 1.12.2021): <https://doi.org/10.5171/2016.473749>.

- King, T. (2019). IDC: 80 Percent of Your Data Will Be Unstructured in Five Years. *Data Management Solutions Review*.
- Kohtamäki, M., Talaoui, Y. & Rajala, R. (2020). Seeking “Strategy” in Business Intelligence Literature: Theorizing BI as part of strategy research. *Technology Innovation Management Review*, 10(9). Saatavissa (viitattu 25.11.2021): <https://doi.org/10.22215/timreview/1387>.
- Kuntasektorin kokonaisarkkitehtuuryöryhmä (2013). Kuntasektorin MDM-viitearkkitehtuuri. Kuntaliitto, Helsinki.
- Laihonen, H., Hannula, M., Helander, N., Ilvonen, I., Jussila, J., Kukko, M., Kärkkäinen, H., Lönnqvist, A., Myllärniemi, J., Pekkola, S., Virtanen, P., Vuori, V., & Yliniemi, T. (2013). Tietojohtaminen. Tampereen teknillinen yliopisto - Tiedonhallinnan ja logistiikan laitos.
- Loshin, D. (2009). *Master Data Management*. Elsevier Science & Technology. Saatavissa (viitattu 23.10.2021): <https://doi.org/10.1016/B978-0-12-374225-4.X0001-X>.
- Mahanti, R. (2019). *Data Quality: Dimensions, Measurement, Strategy, Management, and Governance*. Quality Press.
- Marjamäki, P. (2017). *Evolution and Trends of Business Intelligence Systems: A Systematic Mapping Study*. Master’s Thesis. University of Oulu, Oulu.
- Munoz, M. (2018). *Global Business Intelligence*. Taylor and Francis. Saatavissa (viitattu 28.11.2021): <https://doi.org/10.4324/9781315471136>.
- Muntean. (2018). *Business Intelligence Issues for Sustainability Projects*. Sustainability (Basel, Switzerland). 10(2). Saatavissa (viitattu 27.11.2021): <https://doi.org/10.3390/su10020335>.
- Najdawi, A. & Patkuri, S. (2021). Modeling Business Intelligence Process: Toward Smart Data-Driven Strategies. 2021 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE).
- Nelson, P. (2020). Search and unstructured data analytics: 5 trends to watch in 2020. Saatavissa (viitattu 24.11.2021): <https://www.accenture.com/us-en/blogs/search-and-content-analytics-blog/search-unstructured-data-analytics-trends>.
- Pansara, R. (2021). Master Data Management Challenges. *International Journal of Computer Science and Mobile Computing*, 10(10), 47–49. Saatavissa (viitattu 2.12.2021): <https://doi.org/10.47760/ijcsmc.2021.v10i10.008>
- Pirttimäki, V. (2007). *Business intelligence as a managerial tool in large Finnish companies*. Tampere University of Technology, Tampere.
- Sebastian-Coleman. (2013). *Measuring data quality for ongoing improvement a data quality assessment framework 1st edition*. Elsevier.
- Silvola, R., Jääskeläinen, O., Kropsu-Vehkaperä, H., & Haapasalo, H. (2011). Managing one master data - challenges and preconditions. *Industrial Management & Data Systems*, United Kingdom, 111(1), 146–162. Saatavissa (viitattu 23.10.2021): <https://doi.org/10.1108/02635571111099776>.

Strong, D. & Wang, R. (1996). Beyond Accuracy: What Data Quality Means to Data Consumers. *Journal of Management Information Systems*. 12(4). Saatavissa (viitattu 20.11.2021): <http://www.jstor.org/stable/40398176>.

Vilminko-Heikkinen, R. & Pekkola, S. (2013). Establishing an Organization's Master Data Management Function: A Stepwise Approach. 46th Hawaii International Conference on System Sciences, 4719–4728. Saatavissa (viitattu 9.11.2021): <https://doi.org/10.1109/HICSS.2013.205>.

Vilminko-Heikkinen, R., & Pekkola, S. (2017). Master data management and its organizational implementation. *Journal of Enterprise Information Management*, United Kingdom, 30(3), 454–475. Saatavissa (viitattu 23.10.2021): <https://doi.org/10.1108/JEIM-07-2015-0070>.

Vilminko-Heikkinen, R. & Pekkola, S. (2019). Changes in roles, responsibilities, and ownership in organizing master data management. *International Journal of Information Management*, 47. Ss. 76–87. Saatavissa (viitattu 15.12.2021): <https://doi.org/10.1016/j.ijinfomgt.2018.12.017>.

Vitt, E., Michael. Luckevich, & Stacia. Misner. (2010). *Business Intelligence*. First edition. Sebastopol: Microsoft Press, United Kingdom.

Väre, T. (2019). *Master data 1. painos*. Alma Talent, Helsinki.

Wang, R., Veda, C. & Christopher P. (1995). A Framework for Analysis of Data Quality Research. *IEEE Transactions on Knowledge and Data Engineering*, 7(4).

White, C. (2007). *Using Master Data in Business Intelligence*. Version 1. BI Research, Redding, CT.

Zhang, J. & Lee, S. (2021). Data Quality Management in the Internet of Things. *Sensors* (Basel, Switzerland). 21(17). Saatavissa (viitattu 21.11.2021): <https://doi.org/10.3390/s21175834>.

Zoder, S. (2011). Analytical Master Data Management 2.0. *AMCIS 2011 Proceedings - All Submissions*, 40.