

Ydintietojen hallinnan kehittäminen

Tomi Rautiainen

Opinnäytetyö

Toukokuu 2020

Tekniikan ja liikenteen ala

Insinööri (ylempi AMK), digitaalinen toimitusketju

Tekijä(t) Rautiainen, Tomi	Julkaisun laji Opinnäytetyö, ylempi AMK	Päivämäärä toukokuu 2020
	Sivumäärä 99	Julkaisun kieli Suomi
		Verkkojulkaisulupa myönnetty: x
Työn nimi Ydintietojen hallinnan kehittäminen		
Tutkinto-ohjelma Insinööri (ylempi AMK), digitaalinen toimitusketju		
Työn ohjaaja(t) Franssila Tommi, Lehtola Pasi		
Toimeksiantaja(t) thyssenkrupp Aerospace Finland Oy, Inkiläinen Jari		
Tiivistelmä Digitalisaation aikakautena data on aiempaa isommassa roolissa, kun yrityksissä mietitään liiketoiminnan kehittämistä ja uusien digisovellusten käyttöönottoa. thyssenkrupp Aerospace Finland Oy:ssä aloitettiin syksyllä 2019 kilpailukyvyyn kehittämisprojekti, jonka tarkoituksena oli käytössä olevan toiminnanohjausjärjestelmän prosessien tehostaminen niiden digitalisaatioastetta nostamalla. Samalla tuli eteen tarve saattaa yrityksen käytössä oleva ydintieto ajan tasalle ja ottaa se paremmin hallintaan. Tehtävänä oli selvittää, mitä on kohdeyrityksen ydintieto ja kuinka sen hallintaa voisi kehittää siten, että ydintieto pysyisi laadullisesti riittävän hyvässä kunnossa, ottaen samalla huomioon toimintaympäristön sille asettamat vaatimukset. Työ toteutettiin kahdessa vaiheessa siten, että ensin selvitettiin haastattelututkimuksen avulla ydintiedon ja sen hallinnan nykytila, hyödyt, haitat, haasteet sekä tavoitteita kohdeyrityksessä. Toisessa vaiheessa käynnistettiin toimintatutkimus, jonka lopputuloksena ydintiedon kolmelle keskeisimmälle rekisterille (asiakas-, toimittaja- ja nimikerekisterit) luotiin datastandardit. Lisäksi haastattelututkimuksessa esille nostettujen dataongelmien mahdollisista syistä haettiin Six Sigma DMAIC-menetelmällä esiin ne juurisyyt, jotka poistamalla ydintiedon hallintaa parannettaisiin. Tämän tuloksena syntyi datan hallinnan operatiivinen malli, jossa määritellään ydintietoon ja sen hallintaan liittyvät roolit, vastuut ja velvollisuudet. Tulokset molemmista vaiheista olivat varsin hyödyllisiä ja niiden avulla kohdeyrityksen ydintieto on paremmin hallinnassa ja koska ydintiedon hallinta sisällytetään päivittäisiin prosesseihin, on ydintiedon laadun säilymiselle hyvät edellytykset.		
Avainsanat (asiasanat) Ydintieto, ydintiedon laatu, ydintiedon hallinta, datan hallintamalli		
Muut tiedot (Salassa pidettävät liitteet) Liite 5 on salassa pidettä, ja se on poistettu julkisesta työstä. Salassapidon perusteena on viranomaisten toiminnan julkisuudesta annetun lain (621/1999) 24 §:n kohta 17: yrityksen liike- tai ammattisalaisuus. Salassapitoaika on viisi (5) vuotta. Salassapito päättyy 31.5.2025.		

Author(s) Rautiainen, Tomi	Type of publication Master's thesis	Date May 2020 Language of publication: finnish
	Number of pages 99	Permission for web publication: x
Title of publication Development of (Master) Data Governance Possible subtitle		
Degree programme Engineer (M.Sc. Eng.), Digital Supply Chain		
Supervisor(s) Franssila Tommi, Lehtola Pasi		
Assigned by thyssenkrupp Aerospace Finland Oy, Inkiläinen Jari		
Abstract <p>In the era when digitalization is having a bigger role than ever before, and when companies develop their own businesses and digital solutions, data is rightly so also being recognised as an important asset which needs to be governed well.</p> <p>In autumn 2019 thyssenkrupp Aerospace Finland started a competitiveness development project which aims to increase productivity by increasing digitalization level on their processes. This highlighted the need to also update their current master data and data governance policies in such a way that master data quality could be maintained in high enough level with current available resources.</p> <p>The study was done in two phases. At first by interviewing data stakeholders and getting to know the current state, challenges and hoped future state of master data and data governance in the company.</p> <p>At second stage information from the interviews was used to start a development project, which lead to the creation of Data Standards for the three main domains (customer, supplier and material). In addition, for the possible causes of data issues gathered from the interviews, a Six Sigma DMAIC-method was used to find those root causes that by getting rid of, would improve master data quality. As a result, an Operating Model for Data Governance was created with clear set of roles, obligations and responsibilities to help maintaining good quality master data.</p> <p>Results for both phases were very useful and with them the organization is able to manage their master data and because data governance is incorporated to standard operating procedures there is good chance that master data will be well maintained also in the future.</p>		
Keywords/tags (subjects) Master Data, Master Data Quality, Master Data Management, Data Governance Model		
Miscellaneous (Confidential information)		

1 Sisällys

1	Johdanto	4
1.1	Tavoitteet ja rajaus	5
1.2	Tutkimuskohde	6
2	Käytetyt tutkimusmenetelmät	8
2.1	Tutkimustyyppi	8
2.2	Aineistonkeruu	8
3	Tietoperusta	10
3.1	Ydintieto (Master Data)	10
3.2	Ydintiedon hallinta (Master Data Management)	16
3.3	Ydintiedon laatu (Master Data Quality)	25
3.4	Datan hallinta (Data Governance)	30
4	Työn toteutus	49
5	Tulokset	51
5.1	Ydintiedon nykytila	51
5.2	Ydintiedon hallinnan haasteet	55
5.3	Ydintiedon hallinta jatkossa	57
5.4	Ydintiedon hallinnan kypsyyss	59
6	Johtopäätökset.....	61
6.1	Liiketoiminnan vaatimukset ydintiedolle	62
6.2	Datan hallintamalli	73
6.3	Datan laadun seuranta	79
7	Pohdinta.....	82
7.1	Tutkimusprosessi.....	82
7.2	Tulokset	84
7.3	Tutkimuksen luotettavuus.....	87
7.4	Suositteluja jatkotoimenpiteitä.....	88

Lähteet	90
----------------------	-----------

Liitteet	92
-----------------------	-----------

Liite 1. Datan hallinnan kypsyydesti	92
Liite 2. Nimikerekisterin Datastandardi	96
Liite 3. Asiakasrekisterin Datastandardi	99
Liite 4. Toimittajarekisterin Datastandardi	100
Liite 5. Datan hallinnan vastuut ja roolit .. Virhe. Kirjanmerkkiä ei ole määritetty.	

Kuviot

Kuvio 1. thyssenkrupp Aerospace toimipaikat 2019	7
Kuvio 2. Ydintieto (master data) ja muu data	11
Kuvio 3. Ydintiedon appelsiinimalli	14
Kuvio 4. Ydintiedon hallinnan järjestelmäkartta	23
Kuvio 5. Kolme erityyppistä MDM-arkkitehtuuria	24
Kuvio 6. 10 datan hallinnan aluetta	31
Kuvio 7. Ydintiedon hallinnan 80/20 -malli	32
Kuvio 8. Datan hallinnan periaatteet	34
Kuvio 9. Datan hallinnan operatiivisen mallin roolit ja vastuut	41
Kuvio 10. Tyypillinen datan hallintamallin rakenne	46
Kuvio 11. Tutkimusprosessi	49
Kuvio 12. Järjestelmäkartta	63
Kuvio 13. Aktiiviset vs lakkautetut nimikkeet	68
Kuvio 14. Ydintiedon täyttöaste 11/2019	71
Kuvio 15. Ishikawa - kalanruotokaavio	74
Kuvio 16. Datan hallinnan operatiivinen malli	77
Kuvio 17. Esimerkkejä "Data Stop" nappien sijainnista excel -työkirjassa	81
Kuvio 18. Aktiivisten tietueiden ydintietojen täyttöasteen kehitys	81
Kuvio 19. Ydintietojen määritysten mukaisuus	82

Taulukot

Taulukko 1. Tietovarantotaulukko.....	9
Taulukko 2. RACI -taulukko	45
Taulukko 3. Tutkimuksen yhteydessä haastatellut henkilöt	50
Taulukko 4. Ydintiedon määritelmä	52
Taulukko 5. Ydintiedon nykytila	53
Taulukko 6. Hyvälaatuisen ydintiedon edut	55
Taulukko 7. Ydintiedon ja sen hallinnan haasteet	56
Taulukko 8. Ydintiedon hallinta tulevaisuudessa	58
Taulukko 9. Osiot 7 ja 8 liiketoimintatiedon hallinnan kypsyyskyselystä.....	59
Taulukko 10. Esimerkkejä nimikerekisterin ydintiedoista.....	66
Taulukko 11. Asiakasrekisterin ydintietoja.....	69
Taulukko 12. Aktiiviset ja ei-aktiiviset tietueet	71
Taulukko 13. VOB - mahdolliset syyt - juurisyy	75

1 Johdanto

Opinnäytetyön aiheena oli kohdeyrityksen ydintiedon hallinnan kehittäminen. Työn lähtökohtana oli ensinnäkin selvittää, mitä ydintieto kohdeyrityksessä tarkoittaa ja kuinka sen nähdään vaikuttavan yrityksen toimintaan. Tämän lisäksi työn tarkoituksena oli selvittää ydintietoon ja sen hallintaan liittyvät haasteet, sekä millä tavalla ydintiedon hallintaa voitaisiin yrityksessä kehittää.

Kohdeyrityksenä oli thyssenkrupp Aerospace Finland Oy, joka on osa kansainvälistä thyssenkrupp AG -konsernia ja sen Aerospace liiketoiminta-alueetta. Kohdeyrityksessä oli jo muutamia vuosia tutkittu mahdollisuutta ottaa käyttöön konsernin yhteinen toiminnanohjausjärjestelmä SAP (*ByD tai S4/HANA*), mutta teknisistä ja taloudellisista syistä johtuen kumpaakaan järjestelmää ei lopulta otettu käyttöön. Toiminnanohjausjärjestelmän vaihtosuunnitelmista luovuttiinkin vuoden 2019 alussa.

Tämä johti siihen, että kyetäkseen paremmin vastaamaan nykyiseen enenevässä määrin digitalisoituvaan toimintaympäristöön kohdeyrityksen oli panostettava nykyisen, jo vuodesta 2008 asti käytössä olleen, toiminnanohjausjärjestelmäversion kehittämiseen. Syksyllä 2019 aloitettiin kehitysprojekti, jonka tarkoituksena oli parantaa kohdeyrityksen kilpailukykyä kehittämällä toiminnan tehokkuutta mm. lisäämällä digitalisaation osuutta toiminnanohjausjärjestelmän ja liiketoimintaprosessien eri toimintoihin.

Kuitenkin, jotta edellä mainittu olemassa olevan toiminnanohjausjärjestelmän digitalisointi ja toiminnallisuuksien kehittäminen onnistuisi, tulisi yrityksen datan ja erityisesti sen ydintietojen olla kunnossa.

Kohdeyrityksen ydintietoa ja sen hallinnan kehittämistä mietittäessä tuli huomioida myös thyssenkrupp Aerospace:ssa samanaikaisesti käynnistetty projekti (alfred.B-

EYE), jonka avulla pyritään yhtenäistämään kaikkien Aerospace liiketoimintayksikön paikallisyksiköiden toiminnanohjausjärjestelmien data ja siten mahdollistamaan lähes reaaliaikainen datamassa käytettäväksi (big) datan analysointiin ja sitä kautta tehostamaan päätösten tekoa.

Yksi Alfred.B-EYE projektin osaprojekti liittyi myös ydintietoon ja sen hallintaan siten, että esim. nimikkeille tultaisiin jatkossa muodostamaan ns. Global Material Master ID, jonka avulla pyritään yhdistämään eri toiminnanohjausjärjestelmissä olevat saman sisältöiset tietueet, jolloin datamassaa pystytään analysoimaan paremmin. Lisäksi tietueiden vanhentumiselle, sekä tuplatietuiden havainnoimiseen ja poistamiseen tulisi yhdenmukaiset määritelmät, jolloin tiedetään datamassassa olevan ainoastaan aktiivista ja oikeasti käytössä olevaa dataa.

Opinnäytetyön aiheeksi valikoitunut ydintietojen hallinnan kehittäminen oli opinnäytetyön tekijälle aiempiin työtehtäviin nähden varsin luonnollinen jatkumo, sillä hän oli ollut aktiivisesti mukana SAP-selvitystyössä ja oli myös ko. selvitystyön aikana ollut tuomassa esiin ydintiedon ja sen laadun merkitystä järjestelmämuutoksen sujuvan jalkauttamisen tärkeänä osatekijänä.

Työn avulla opinnäytetyön tekijälle muodostui vahva tietoperusta kohdeyrityksen ydintiedosta nykyisessä liiketoimintamallissa ja toiminnanohjausjärjestelmässä, sekä sen hallinnasta. Työtä varten kerättyä ja tuotettua tietoa voidaan hyödyntää kohdeyrityksessä myös jatkossa, kun datan hallintaa kehitetään.

1.1 Tavoitteet ja rajaus

Opinnäytetyön tavoitteena oli auttaa kohdeyritystä määrittelemään liiketoiminnalle kriittinen ydintieto, selvittää sen nykytila ja haasteet, sekä miettiä keinoja ydintiedon hallinnan kehittämiseksi, jotta yritys sen avulla saisi parhaan mahdollisen hyödyn irti toiminnanohjausjärjestelmään ja prosesseihin tehtävistä muutoksista. Lisäksi opinnäytetyön tekijän työlleen asettama epävirallinen tavoite oli *datatietoisuuden lisääminen kohdeyrityksessä*.

Opinnäytetyössä selvitetään mitä on kohdeyrityksen ydintieto, määritellään se, sekä selvitetään ydintiedon hallinnan haasteet ja luodaan ydintiedolle kohdeyrityksen organisaatorakenteeseen ja -kulttuuriin parhaiten sopiva ydintiedon hallinnan malli (Data Governance Operating Model).

Opinnäytetyössä ei käsitellä nykyisen ydintiedon siivoamiseen ja harmonisointiin liittyviä asioita, mutta ne voidaan silti joissain asiayhteyksissä mainita. Lisäksi opinnäytetyössä ydintietojen määrittely ja hallinnan kehittäminen rajataan koskemaan vain osaa kohdeyrityksen ydintietorekistereistä (nimike, asiakas, toimittaja). Tämä siitä syystä, että kaiken ydintiedon ja sen hallinnan sisällyttäminen yhteen opinnäytetöihin ei ole tarkoituksenmukaista.

Opinnäytetyö pyrkii kuitenkin kuvaamaan ydintiedon ja sen hallinnan määrittelyprosessit, sillä tarkkuudella, että niitä voidaan hyödyntää jatkossa muidenkin ydintietorekistereiden määrittelemisessä ja hallinnoimisessa.

Opinnäytteen avulla pyritään vastaamaan seuraaviin kysymyksiin:

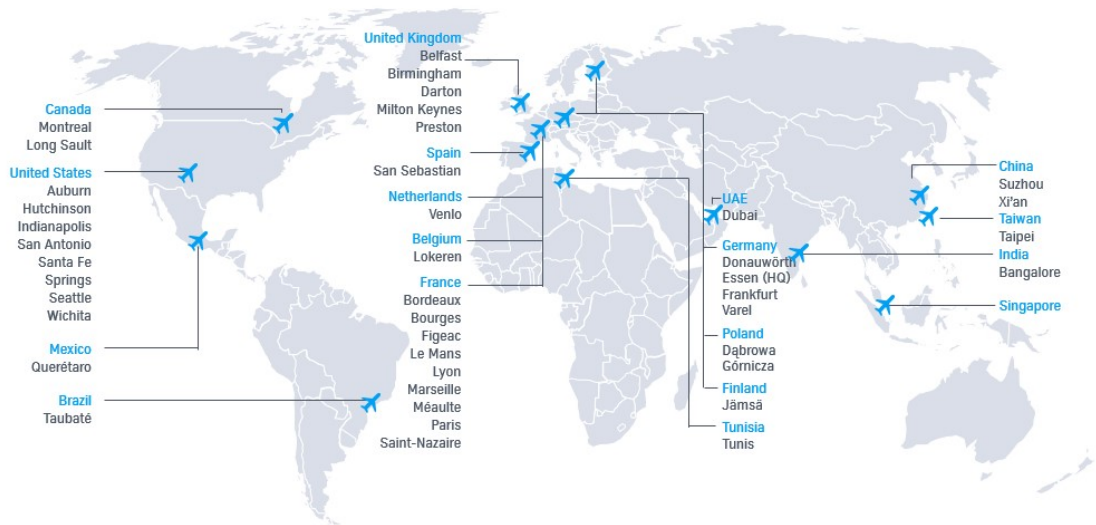
- Mitä on ydintieto kohdeyrityksessä?
- Mitä haasteita ydintietoon ja sen hallintaan liittyy?
- Minkälainen organisaatorajat ylittävä ydintiedon hallintamalli olisi kohdeyritykselle toimivien?

1.2 Tutkimuskohde

Toimeksiantaja oli thyssenkrupp Aerospace Finland Oy (tkA Finland Oy), joka on osa thyssenkrupp AG -konsernin Materials-liiketoiminta-alueeseen kuuluvaa Aerospace liiketoimintayksikköä, joka toimii maailmanlaajuisesti 20 eri maassa (Amerikoissa, Tyynenmeren alueella, Euroopassa ja Pohjois-Afrikassa) ja yli 40 toimipaikassa (kuvio 1). Materials-liiketoiminnassa tilikaudella 2018/19 työskenteli 20 340 henkilöä ja liikevaihtoa oli 13 880 milj. €.

Close to our customers

We operate globally but stock and process locally



Kuvio 1. thyssenkrupp Aerospace toimipaikat 2019

Suomen toimipiste sijaitsee Keski-Suomessa, Jämsänkoskella, jossa työskenteli vuoden 2020 alussa 34 henkilöä.

tkA Finland Oy on lentokone- ja konepajateollisuuden materiaalitoimittaja, jonka päätoimialue on Skandinavia. Pääosa yrityksen liikevaihdosta tulee metallimateriaalien (mm. alumiini, titaani, ruostumaton teräs) myynnistä kotimaisille konepajoille, sekä puolustusvälineteollisuudelle Skandinavian alueella.

Tilikaudella 2018/19 Suomen yksikön liikevaihto (23,5 milj. €) jakautui näiden kahden päätoimialan kesken suhteessa 45/55, puolustusvälineteollisuuden osuuden ollessa hienoisesti isompi.

tkA Finland Oy noudattaa toiminnassaan ISO900- ja EN9100-laatu järjestelmien vaatimuksia.

2 Käytetyt tutkimusmenetelmät

2.1 Tutkimustyyppi

Opinnäytetyön tarkoituksena on tutkia kuinka ydintiedon hallintaa voisi kohdeyrityksessä kehittää ja minkälainen organisaatorajat ylittävä ydintiedon hallintamalli olisi kohdeyritykselle paras ja realistisin, jotta sen avulla varmistetaan ydintiedon laadun ja hallinnan säilyvyys ydintiedon määrittelyn jälkeen tehtävän ydintietorekisterin siivous- ja harmonisointiprojektin jälkeenkin.

Ydintieto on kohdeyrityksessä aiemmin tutkimaton ilmiö, joten ensin oli selvitettävä mistä ydintiedossa ja ydintiedon hallinnassa on kyse, nimenomaan kohdeyrityksen ja sen työntekijöiden kannalta katsottuna. Ja koska ydintietoa ei ollut yrityksessä aiemmin määritelty, eikä sen laadullisia tekijöitä siten ollut käytettävissä kvantitatiivisen tutkimuksen pohjaksi, niin opinnäytetyön tutkimustyyppiksi valikoitui kvalitatiivinen tutkimus (Kananen, 2010).

Kvalitatiiviseen tutkimukseen yhdistyy kohdeyrityksen tapauksessa vahvasti myös toimintatutkimus, sillä opinnäytetyön tarkoituksena on määritellä ydintieto sekä dokumentoida sen ominaisuudet datastandardeihin, sekä ydintiedon ja sen hallinnan nykytilan ja haasteiden selvittämisen jälkeen luoda datan hallinnan operatiivinen malli, sekä sisällyttää ydintiedon hallinta olemassa oleviin prosesseihin.

2.2 Aineistonkeruu

Tutkimusaineiston keruu ja lähtötilanteen analysointi tapahtui haastattelemalla kaikkien niiden ydintiedon sidosryhmien edustajia, jotka pääasiallisesti luovat, käyttävät ja ylläpitävät yrityksen ydintietoja osana jokapäiväistä työtään. Haastateltavaksi valikoitui riittävä määrä henkilöitä, jotta opinnäytetyön tekijälle syntyi selkeä kuva siitä, kuinka ilmiö, ydintieto, kohdeyrityksessä tiedetään ja tunnistetaan sekä kuinka hyvin on selvillä sen hallintaan liittyvät haasteet ja mahdollisuudet. Tämän lisäksi haastateltavien määrään vaikutti se, että saaduista vastauksista ei enää noussut esiin uusia

ilmiöön vaikuttavia tekijöitä, vaan vastauksissa esiin tulleet asiat alkoivat toistamaan itseään (Kananen, 2014).

Haastattelut suoritettiin teemahaastatteluina. Haastateltaville ilmoitettiin etukäteen haastattelun aihepiiri (sekä joitain ohjaavia kysymyksiä), mutta itse haastattelutilante toteutettiin siten, että haastattelija saattoi poiketa aihepiiristä haastateltavalta saamiensa vastausten perusteella ja siten viedä keskustelua eri suuntiin, palaten myöhemmin takaisin alkuperäiseen teemahaastattelun aiheeseen. Tällä opinnäytetyön tekijä pyrki varmistamaan sen, että haastateltavat saivat tuotua mahdollisimman laajasti esiin kaikki ydintietoon ja sen hallintaan liittyvät seikat.

Taulukko 1. Tietovarantotaulukko

Datatyyppe	Määrä	Alkuperäinen datan lähde	Alkuperäinen kohdeyleisö
Haastattelut	7	Haastatellut henkilöt	Opinnäytetyön tekijä
Kysely tiedonhallinnan kypsyystä (Data Maturity)	3/7	Kyselyyn vastanneet henkilöt	Opinnäytetyön tekijä
Six Sigma charter	1	Opinnäytetyön tekijä	Datan hallinnan projektiryhmä
Palaverimuistiot	20	Opinnäytetyön tekijä	Opinnäytetyön tekijä
BI Strategy@tkAerospace & baMX-esitykset (.ppt)	3	Konserni HQ	Konsernin ja paikallisyhtiöiden johto
Tutkimuspäiväkirja (muistiinpanot opinnäytetyön tekijän havainnoista ja ajatuksista)	1	Opinnäytetyön tekijä	Opinnäytetyön tekijä

Haastattelutilanteessa vastauksista kirjattiin ylös vain olennainen, eli haastatteluiden kirjaaminen ja litterointi tapahtuu propositiotasolla (Kananen 2010). Yksilöhaastatteluiden ja niiden litteroinnin jälkeen haastatelluille pidettiin työpaja, jonka yhteydessä he vielä näkivät kuinka heidän omat näkemyksensä ydintiedosta ja sen hallinnasta peilautuvat muiden näkemyksiin. Tämän työpajan jälkeen tutkimus muuntui toimintatutkimukseksi, joka toteutettiin Six Sigma DMAIC-prosessia soveltaen.

Haastatteluiden lisäksi haastatelluille tehtiin kysely kohdeyrityksen liiketoimintatiedon hallinnan kypsyystasosta. Liiketoimintatietolähtöistä kypsyysmallia käytetään yleisen liiketoimintatiedon hallinnan sekä yksittäisten osastojen ja liiketoimintayksiköiden kypsyystasoa arvioinnissa.

3 Tietoperusta

3.1 Ydintieto (Master Data)

Ydintieto on tietoa, joka kuvaa liiketoimintatapahtumien tekijöitä, ”kuka”, ”mitä” ja ”missä”. Ydintietoja ovat esimerkiksi asiakas-, työntekijä- ja toimittajatiedot (kuka), osa- ja tuotetiedot (mitä), tuotteen tai varaston sijaintitieto (missä) (Cerwin, 2019; Väre, 2019).

Ydintiedosta on liikkeellä monia erilaisia määritelmiä, mutta yleisin määritelmä kuvaa ydintiedon liiketoiminnan ytimessä olevaksi luonteeltaan kohtalaisen pysyväksi tiedoksi, jota käytetään laajalti eri sovelluksissa läpi koko organisaation, yhdessä niihin liittyvien metadata-, attribuutti-, -määrittely-, liittymä- ja luokittelutietojen kanssa (Loshin, 2009; Väre, 2019). Ydintieto on sitä organisaatiolle merkittävää kriittistä dataa, jota käytetään toiminnanohjausjärjestelmissä, mitataan ja raportoidaan raportointijärjestelmissä ja analysoidaan analysointijärjestelmissä (Loshin, 2009) ja jota ilman organisaatio ei voi toimia (Väre, 2019). Cerwin (2018) lisää, että ollakseen tehokasta ja hyödyksi liiketoiminnalle, ydintiedon tulee olla tarkoitukseen sopivaa, aivan kuten lääkkeen sitä tarvitsevalle ihmiselle.

Ydintiedolla on tapana sijaita useammassa kuin yhdessä toimialueessa organisaation sisällä, jolloin esimerkiksi sama asiakas voi esiintyä niin myynti- kuin laskutusjärjestelmässä. Ydintietoja käytetään tapahtumien, eli transaktioiden sisällä, mutta verrattuna transaktiodataan, niiden taustalla oleva data on luonteeltaan pysyvää, eikä siis muutu usein (Loshin, 2009). Jonker ym. (2011) kehottaa kuitenkin huomioimaan, että ydintiedolla on voimakas vaikutus transaktiodatankin eheyteen.

Väreen (2019) mukaan ydintieto voidaan jakaa myös erilaisiin tarkentaviin luokkiin (kuvio 2). Rakenteellinen ydintieto sisältää usein organisaation taloudelliset rakenteet ja hierarkiat. Joidenkin määritelmien mukaan on myös väliaikaista tai aikaan sidottua ydintietoa, joka voi sisältää esimerkiksi hintatiedot. Näitä hän kutsuu ajalliseksi ydintiedoksi, joka on transaktiodatan ja ydintiedon välimuoto.



Kuvio 2. Ydintieto (master data) ja muu data (Väre, 2019)

3.1.1 Ydintiedon komponentit

Ydintiedosta (master data) puhutaan usein yhtenä käsitteenä, mutta todellisuudessa se sisältää runsaasti erilaisia (ala)käsitteitä.

Domain, kertoo yhden liiketoiminnallisesti loogisen kokonaisuuden. Se on rekisteri yksilöllisesti tunnistetuista entiteeteistä, joka muodostaa selkeärajaisen kokonaisuuden, jonka joku voi omistaa ja jota voi kehittää itsenäisesti (Väre, 2019). Asianmukaisen hallinnan ja valvonnan avulla voidaan varmistaa domainin tietojen yhteneväisyys

ja johdonmukaisuus, jolloin käytössä on dataa, johon kaikki sovellukset voivat luottaa johdonmukaisen ja korkealaatuisen tiedon saamiseksi. (Loshin, 2009)

Entiteetti, todellisen elämän digitaalinen versio, yhtä todellisen elämän asiaa luokitteleva käsite, joka voidaan selvästi yksilöidä ja tunnistaa (Väre, 2019).

Attribuutti, entiteettiä kuvaava tai selittävä asia. Jokainen attribuutti kertoo entiteetistä yhden asian. Näiden attribuuttien eli ominaisuuksien avulla entiteetti voidaan tunnistaa ja yksilöidä (Väre, 2019).

Väreen (2019) mukaan attribuutti on ydintietoa, silloin kun se on:

- **Uniikki**, nämä attribuutit kertovat, että kyseinen tietue on ainutlaatuinen, yksilöllisesti tunnistettavissa. Se tarkoittaa, että näiden attribuuttien sisältö muodostaa ainutlaatuisen yhdistelmän. Niiden muodostama datakokonaisuus ei saa esiintyä kahta tai useampaa kertaa tismalleen samanlaisena. Tämä on kaiken ydintiedon minimisisältö.
- **Raportoitu**, Nämä attribuutit ovat oleellisia raportoinnissa tai liiketoimintanalyseissä. Niiden avulla tiedetään, mitä organisaatiossa tapahtuu, ne ohjaavat toimintaa tai kertovat tuloksesta. Raportointijärjestelmissä näitä kutsutaan esimerkiksi dimensioiksi, ulottuvuuksiksi tai raportointihierarkioiksi. Ne eivät kerro, mitä raportoidaan, vaan miten raportoidaan.
- **Usein käytetty**, Nämä attribuutit toistuvat usein eri liiketoimintaprosesseissa. Niitä hyödyntää eri organisaation osat, eri tiimit, eri järjestelmät, toinen toisensa jälkeen. jotkut saattavat muokata niitä, toiset vain käyttävät dataa. Koska tätä dataa käytetään usein, on kustannustehokasta, että se jaetaan läpi koko organisaation, kaikkien dataa käyttävien tahojen kesken.
- **Tärkeä**, nämä attribuutit ovat oleellisia organisaation päivittäiselle toiminnalle. Tärkeät toiminnot, kuten vaikka myynti tai laskutus, eivät onnistu ilman niitä. Jos data puuttuu, jonkin täytyy kenties keskeyttää normaali prosessi ja lähteä etsimään dataa jostain muualta. Myöskään prosessien automatisointi ei onnistu, mikäli tämä data puuttuu (Väre, 2019)

Relaatio, kahden entiteetin välinen suhde. relaatio kertoo, millä kahdella entiteetillä on jonkinlainen suhde. Relaatiolla voi myös kuvata, minkälainen riippuvuus näiden käsitteiden välillä on (Väre, 2019).

Metadata, eli kuvailu- tai määrittelydata. Metadata on dataa kuvailevaa ja määrittelevää dataa, tai kuten usein sanotaan ”tietoa tiedosta”. Metadata liittyy ydintiedon kaikkiin osiin. Sanallisten määritelmien lisäksi erilaiset muodolliset määritelmät ovat oleellisia ydintiedolle (Väre, 2019).

David Loshinin (2009) mukaan yleinen määritelmä kuvata metadataa ainoastaan ”tiedoksi tiedosta” on kuitenkin turhan laimea ilmaisutapa, eikä se tuo tarvittavaa syvyyttä ymmärrykseemme metatiedon tärkeydestä. Metatieto yhdistettynä liiketoiminnan ydintietoon tekee enemmän kuin vain kuvaa tiedon kokoa ja tyyppiä. Eri tietueiden integrointi voidaan tehdä vain, kun on selvää, että dataelementeillä on sama merkitys ja että niiden data-alueet ovat yhdenmukaisia, tietueet edustavat samanlaisia tai samoja reaali maailman asioita, eli metatietoja (Loshin, 2009).

Monet datan laatuongelmat juontavat juurensa metatiedoista. Ja vaikka metatiedot helpottavat ihmisiä löytämään ja yhdistämään tarvitsemansa tiedon, sekä tekemään niiden avulla asianmukaisia johtopäätöksiä, on metatiedon virheillä suuri vaikutus organisaatioiden toimintaan (Redman, 2013).

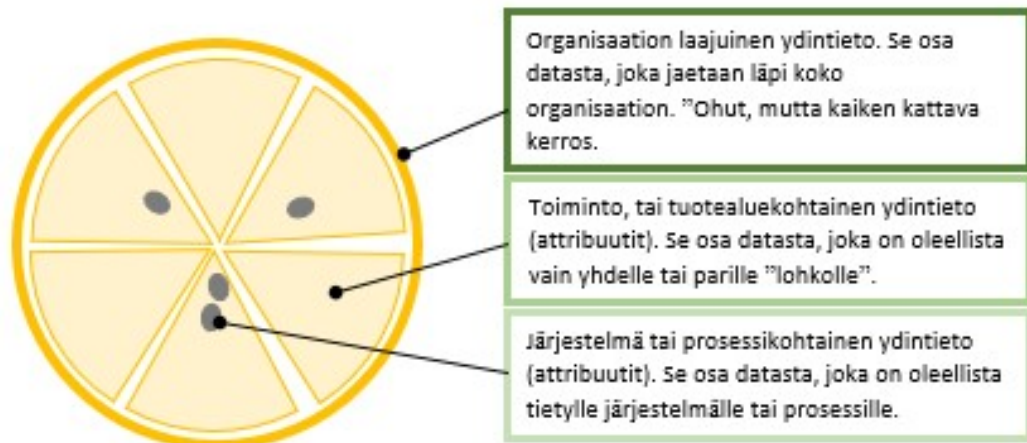
Referenssidata, eli luokittelu- ja viitetieto. Referenssidata ryhmittelee tai luokittelee muuta dataa. Se on lista ennalta määriteltäviä tietoja, joita käytetään muun tiedon yhtenäisyyden varmistamiseksi. Referenssidata voi olla vain yhteen domainiin kuuluva listaus tai se voi olla domain riippumaton yleislistaus (Väre, 2019)

3.1.2 Ydintiedon tasot

On olemassa ydintietoja, joita käytetään vain yhdessä ainoassa tapahtumassa (transaktiossa), jaetaan eri sovelluksissa tai jaetaan koko organisaatiossa, mutta joka tapauksessa on ydintietoa. Cerwinin (2019) mukaan liian rajoitettu näkemys siitä, mikä on ydintietoa, voi johtaa ydintietojen siiloutumiseen ja rajautumiseen käytettä-

väksi ainoastaan sovelluskohtaisena ydintietona, rajattuna pois koko organisaation käytössä olevasta yhtenäisestä ydintietojoukosta. Tämä voi johtaa tiedon laatuongelmiin koko organisaatiossa.

Kirjassaan *Master Data* Taru Väre hahmottaa ydintiedon kerroksellisuutta kuvaavasti appelsiinimallin (kuvio 3) avulla seuraavasti:



Kuvio 3. Ydintiedon appelsiinimalli (Väre, 2019)

Kuorikerros: tämä on se ohut ydintiedon kerros, joka on jaettu läpi koko organisaation ja jonka avulla kyseinen data voidaan yksilöllisesti tunnistaa. Mikäli tätä kerrosta ei ole ydintiedon hallinnassa, ei ydintietoa todellisuudessa hallita.

Lohkokerros: Kun appelsiini kuoritaan, sen lohkot on helppo irrottaa toisistaan. Samoin lohkokerroksen paikallisempi ydintieto on erillään toisista lohkoista. Vain kuorikerros kiinnittää eri lohkojen tiedot samaan tietueeseen. Nämä lohkot edustavat eri liiketoimintayksiköiden omaa ydintietoa. Nämä erilaiset lohkot myös esittävät omat vaatimuksensa ydintiedolle eli määrittelevät, mitä attribuutteja juuri ne tarvitsevat kyseisestä datasta.

Siementaso: Siemenkaltaista prosessi- tai järjestelmäkohtaista ydintietoa on organisaatiosta riippuen hiukan tai ei lainkaan. Tiukimmat ydintiedon määrittelijät eivät laske tätä kerrosta lainkaan ydintiedoksi. On kuitenkin tavallista, että organisaatiossa on joku tärkeä ydinprosessi tai globaalisti käytetty järjestelmä, joka esittää omia vaa-

timuksiaan datalle. Vaikka se ei välttämättä ole ”virallista” ydintietoa, se on kuitenkin tärkeää dataa, jotta päivittäinen toiminta onnistuu. Ottamalla tämäkin data mukaan ydintiedon hallintaan voidaan paremmin varmistaa datan laatu ja prosessien toimivuus.

Nämä kaikki liiketoimintaprosessi- tai järjestelmäkohtaiset tiedot on tärkeitä, kun tehdään datastandardia ja tiedon tulee löytyä sieltä. (Väre, 2019)

3.1.3 Datastandardi

Datastandardilla tarkoitetaan ydintietoa tarkemmin määrittelevää dokumentaatiota, jossa kuvataan ydintieto siten, että liiketoiminnan eri osa-alueet voivat puhua samaa kieltä. Dokumentaatiossa listataan kaikki ydintiedoksi luokiteltavat attribuutit, niiden määritelmät, sisältö- ja laatuvaatimukset sekä liiketoimintasäännöt. (Väre, 2019)

Datastandardin sisältö muodostuu yksityiskohdiltaan organisaation tarpeen mukaan, mutta siinä tulisi olla mukana vähintään seuraavat tiedot (Väre, 2019):

- Kaikki ydintiedoksi määritellyt entiteetit ja niiden attribuutit
- Jokaiselle entiteetille ja attribuutille yhteisesti sovittu nimi sekä määritelmä
- Mahdolliset synonyymit eri käsitteille (esim. miten kyseistä dataa kutsutaan jossain toisessa järjestelmässä).
- Tiedon tyyppi, eli onko data vapaasti kirjoitettavaa tekstiä, numero vai ennalta määrättyltä listalta valittava arvo.
- Tiedon muoto, mikäli se on tärkeä tieto (sähköposti @ jne.).
- Pakollisuustieto. Onko tieto aina pakollinen, vai pakollinen vain tietyn tyyppiselle datalle?
- Liiketoimintaprosessit, joissa kyseistä dataa syntyy, päivitetään, käytetään tai poistetaan. (CRUD)

3.2 Ydintiedon hallinta (Master Data Management)

Historiallisesti tarkasteltuna liiketoimintasovellukset suunniteltiin täyttämään tarkoin määritellyt liiketoiminta-alueen operatiiviset tarpeet (Loshin, 2009; Panian, 2010). Tällöin myös sovellusarkkitehtuuri kehittyi tukemaan vain tiettyä liiketoiminta- aluetta ja sen tarvetta. Sen seurauksena yritysten toimintaympäristö saattaa muodostua useasta eri sovelluksesta, jotka linkittyvät moninkertaisiin, usein toisistaan irrallisiin datavarastoihin, joiden on kuitenkin tarkoitus edustaa samaa tai samankaltaista käsitettä. (Loshin, 2009; Smith, H., McKeen, J. 2008)

Siirtymisen sovelluksiin, jotka on tarkoitettu vastaamaan koko yrityksen tarpeita, on perustuttava johdonmukaiseen kuvaan useista eri lähteistä kerätystä datasta. Ja jotta tätä dataa voidaan käyttää sekä operatiivisiin että analyttisiin prosesseihin, organisaation on kyettävä selkeästi määrittelemään liiketoimintakonseptit sekä tunnistamaan eri tavat, joilla datajoukot edustavat näitä liiketoimintakonsepteja, integroidaan nämä tiedot yhtenäiseksi näkymäksi ja tuomaan tämä näkymä koko organisaation käytettäväksi. (Loshin, 2009)

3.2.1 Ydintiedon hallinnan määritelmä

Ydintiedon hallinta (Master Data Management, MDM) on syntynyt tarpeesta saada eri järjestelmiin sirpaloituneet, moneen kertaan talletetut ja usein eri tasoilla olevat ydintiedot parempaan hallintaan (Hovi, 2018; Smith ym. 2008). Ydintiedon hallinta on joukko operatiivisia prosesseja, joiden avulla ihmiset ja teknologia ylläpitävät luotettavaa, hallittua ja tarkoituksenmukaista ydintietoa, jota voidaan hyödyntää transaktioprosessien lähteenä (Cerwin, 2019; Jonker, R.A., Kooistra, F.T., Cepariu, D., van Etten, J., Swartjes, S. 2011).

Ydintiedon hallinta on kriittistä organisaation päivittäisen toiminnan sujumiselle sekä organisaation liiketoiminnan tapahtumien luotettavalle raportoinnille (Väre, 2019).

1980-luvun tietokonevallankumouksesta syntynyt ”tiedonhallinnan” käsite on joutunut kehittymään vuosien varrella teknisemmistä tietokonetietojen tallennukseen

liittyvistä rakenteista käsittämään kaikki tiedon hallintaan liittyvät aiheet. DAMA Internationalin mukaan tiedon hallinta (data management) on sellaisten arkkitehtuurien, käytäntöjen ja menettelytapojen kehittämistä ja toteuttamista, joiden avulla hallitaan asianmukaisesti yrityksen datan koko elinkaaren tarpeita. (Mosley, M. 2010; Väre, 2019; Cerwin, 2019)

Ydintietojen hallinta (MDM) käsittää kokoelman tiedon hallinnan parhaita käytäntöjä, jotka auttavat yrityksen sidosryhmiä sisällyttämään liiketoimintasovelluksensa, datan hallintamenetelmät ja -työkalut pyrkimykseen parantaa ja ylläpitää virheetöntä, tarkkaa, ajantasaista ja täydellistä ydintietoa eri sovelluksissa. (Loshin, 2009; Väre, 2019; Kekwaletswe, R. & Lesole, T. 2016)

Ydintiedon hallintaa (MDM) voikin ajatella eräänlaisena ”*sateenvarjo*”-terminä niille käytänteille, jotka liittyvät tiedon ja datavarantojen kehittämiseen ja valvontaan. Datan hallintaa (data governance) tulee pitää elintärkeänä osana ydintiedon hallintaa (master data management).

Ydintiedon hallinta on tarkoitettu tukemaan organisaation liiketoiminnan tarpeita tarjoamalla pääsyn yksilölliseen ydintietoon koko toiminnallisen infrastruktuurin alueella. Ydintiedon hallinta hallinnoi menetelmiä, työkaluja, dataa ja palveluita, jotta niiden avulla voidaan mm. tunnistaa liiketoiminnan menestykselle oleelliset ydintiedot, joita käytetään eri sovelluksissa, ja jotka voisivat hyötyä tiedon keskittämisestä, sekä käynnistää asianmukaiset datan hallintakäytännöt ja -menetelmät yritys- ja toimialatasoilla korkealaatuisen ydintietoaineiston varmistamiseksi. (Loshin, 2009)

Väreen (2019) mukaan ydintiedon hallinnan yleiset tavoitteet pitää sovittaa omassa organisaatiossa liiketoiminnan tavoitteisiin ja tarpeisiin. Näihin tavoitteisiin pyritään hänen mukaansa kuuden eri osa-alueen kautta, jotka yhdessä muodostavat koko organisaation kattavan ydintiedon hallinnan. Nämä osa-alueet ovat:

1. Ydintiedon hallinnan taustavaikuttajat: arkkitehtuuri ja teknologiat. Ydintiedon hallinta sovitetaan aina olemassa olevaan organisaatioon. Miten ydintietoa voi hallita, minkälaisia rajoituksia on huomioitava tai mitä mahdollisuuksia on? Näihin vaikut-

taa se, minkälaista teknologiaa löytyy, minkälainen organisaatio on toimintatyyliltään ja mitä tavoitteita liiketoiminnalla on (Väre, 2019; Smith ym. 2008). Ja vaikka ydintiedon hallinta (MDM) ei ole IT-hanke, on teknologialla kuitenkin merkittävä rooli onnistuneen ydintiedon hallinnan toteuttamisessa. (Smith ym. 2008)

2. Ydintiedon hallinnan raamit: datan hallintamalli. Eryyisen tärkeä ydintiedon hallinnalle on oikeanlaisen hallintamallin luominen. Hallintamalli asettaa raamit, joiden puitteissa dataa hallitaan. Nämä raamit määrittelevät, ketkä ydintiedon hallintaan osallistuvat ja minkälaisia vastuita ja velvollisuuksia heillä siihen on. Datan hallintamalli tekee ydintiedon hallinnan näkyväksi koko organisaatiossa. (Väre, 2019; Smith ym. 2008)

3. Ydintiedon hallinnan kehityskaari: visio, tavoitteet ja strategia. Jotta ydintiedon hallinnassa on suunta eteenpäin, sille tarvitaan selkeät tavoitteet. Näiden tavoitteiden pitää liittyä organisaation liiketoimintatavoitteisiin.

4. Ydintiedon taulukot ja rakenteet: tiedon määrittely ja standardointi. Ydintiedon hallinta lähtee siitä, että tiedetään, mistä datassa on kyse ja minkälaista sen pitää olla. Ydintieto täytyy siis määritellä. Ydintiedon rakenteet kuvataan tietomallien avulla, jotka kertovat, miten ydintiedot liittyvät toisiinsa ja liiketoiminnan tapahtumatietoon. Ydintiedolle määritellään täsmälliset kuvaukset, säännöt, rajoitukset ja muut liiketoimintatarpeet.

5. Ydintiedon hallinnan rattaat: tiedon hallinnan prosessit ja tiedon elinkaari. Liiketoimintaprosessit ovat ne, joissa ydintieto syntyy, elää ja kuolee. Siksi ydintiedon hallinnassa täytyy varmistaa, että nämä prosessit tuottavat hyvänlaatuista ydintietoa. Ydintiedon hallinnan prosessien tavoitteena on, että dataa käsitellään mahdollisimman vähän manuaalisesti ja sen saatavuus eri järjestelmiin ja liiketoimintoihin varmistetaan mahdollisimman tehokkaasti.

On tärkeää muistaa, että ydintieto luodaan ja ylläpidetään liiketoiminnan prosesseissa. Kun halutaan varmistaa ydintiedon laatu, tarkastelu kohdistetaan liiketoimintaprosesseihin, joissa data luodaan tai muokataan. Näistä prosesseista poimitaan kaikki

ne toiminnot, jotka liittyvät dataan ja muodostetaan niistä yksi kokonaisuus. Tästä kokonaisuudesta voidaan nähdä, kuinka data syntyy tai muokkautuu prosessin edessä (Väre, 2019). Tutkimuksessaan Kekwaletswe ym. (2016) tulivat myös siihen lopputulokseen, että voidakseen toimia kunnolla liiketoiminnan tulee määritellä menetelmät ja teknologiat, joiden avulla ydintiedon hallinta säilyy linjassa liiketoiminnan tavoitteiden ja sääntöjen kanssa.

Mikäli prosessista tehdään laadunvarmistuksen nimissä liian monimutkainen ja hidas, prosessia toteuttavat henkilöt tulevat kehittämään omat epäviralliset toimintatavat. Siksi prosessin kehittämisessä kannattaa noudattaa Lean-ajattelun periaatteita (Väre, 2019).

Leanissa on keskeistä tunnistaa ja eliminoida hukka nopeasti ja tehokkaasti, pienentää kustannuksia sekä parantaa laatua. Hukalla tarkoitetaan ylimääräisiä, tuottamattomia toimintoja, jotka hidastavat prosessia tai tuottavat tarpeettomia kustannuksia. Hukka on seurausta prosesseissa tapahtuvista vioista ja virheistä, jotka vaihtelu aiheuttaa. Jos poistetaan vain hukkaa, hukka tulee aina uudestaan, koska hukkan syyt ei ole poistettu, vain "oire". Arvoa tuottamattomiksi toiminnoiksi tai turhiksi asioiksi lasketaan:

- Kuljetukset (**T**ransportation)
- Varastointi (**I**nventory)
- Siirtyminen (**M**otion)
- Odottaminen (**W**aiting)
- Ylituotanto (**O**ver Production)
- Yliprosessointi (**O**ver Processing)
- Virheet (**D**efects)

6. Ydintiedon laadun mittaaminen ja parantaminen. Yksi datan hallinnan tärkeimmistä tavoitteista on tuottaa hyvää datan laatua. Datan laatua voidaan tarkastella monelta eri näkökulmalta, joita tyyppisesti kutsutaan ulottuvuuksiksi.

3.2.2 Ydintietojen hallinnan hyödyt

Koska organisaatioiden toiminnalla on lähtökohtaisesti joku tarkoitus, niin ydintiedon hallinta (MDM) tulisi myös sitoa tuohon tarkoitukseen. Toisin sanoen ydintiedon hallinnan liiketoiminnalliset perusteet olisi liitettävä asioihin, jotka hyötyvät saatavilla olevasta, yhteneväisestä ydintiedosta. On myös huomioitava, että ydintietojen integrointi ja hallinta eivät ole tiedon hallinnan perimmäinen tavoite; pikemminkin ne ovat keino, jolla mm. asiakassuhteiden hallintajärjestelmien (CRM), tuotetietojen hallintajärjestelmien (PIM) ja yrityksen toiminnanohjausjärjestelmien (ERP) lupaamat muut strategiset ja operatiiviset tavoitteet saavutetaan onnistuneesti. (Loshin, 2009; Väre, 2019). Lisäksi, jotta yritys pystyisi parantamaan liiketoiminnan analysointia ja raportointia (BI), on tärkeää, että ydintiedon hallinnan eri osa-alueet on ymmärretty ja analysoitu riittävän hyvin (Kekwaletswe ym. 2016).

Liiketoiminnan raportoinnin ja analysoinnin (BI) avulla on tarkoitus saada parempi käsitys mm. yrityksen suorituskyvystä sekä asiakkaiden ja tuotteiden kannattavuudesta. Jonker ym. (2011) muistuttaakin artikkelissaan, että näiden raporttien avulla tehdään yrityksissä tärkeitä päätöksiä ja että huonolaatuinen data vaikuttaa niiden lopputulemaan välittömästi johtaen helposti vääriin johtopäätöksiin.

Experian Information Solutionin teettämässä tutkimuksessa (McCarthy, 2019), jossa haastateltiin yli 1000 henkilöä eri organisaatioista, todetaan, että 69% vastanneista on samaa mieltä siitä, että virheellinen data heikentää olennaisesti heidän kykyään tarjota hyvä asiakaskokemus.

Ydintiedon hallinnan hyödyt löytyvät organisaation toiminnasta ja Väre (2019) liittää näihin hyötyihin kolme näkökulmaa:

1. Datasta uutta liiketoimintaa. Uusi liiketoiminta tarkoittaa jonkin sellaisen tuotteen tai palvelun tarjoamista, mitä ei aiemmin ole tehty. Se voi tarkoittaa uusia, innovatiivisia tapoja palvella nykyisiä asiakkaita tai aivan uuden asiakasryhmän tavoittamista (Väre, 2019).

2. Datalla parempaa liiketoimintaa. Parempi liiketoiminta tarkoittaa nykyisen liiketoiminnan tehostamista, siinä paremmin onnistumista tai olemassa olevien asiakkaiden sitouttamista. Hyvin hoidetulla asiakasdatalla ja sen kattavalla hyödyntämisellä voidaan parantaa asiakastuntemusta ja -osaamista, sekä sitä kautta myös asiakaspalvelua ja -tyytyväisyyttä. (Loshin 2009; Väre, 2019; Smith ym. 2008)
3. Datalla tehokkuutta liiketoimintaa. Tehokkuus tarkoittaa nopeampaa operatiivista toimintaa, asioiden käsittelyn helppoutta ja nopeutta sekä työajan säästämistä oikeisiin asioihin. Ydintiedon osalta se tarkoittaa sitä, että samaa dataa ei syötetä tai päivitetä useampaan kertaan eri järjestelmiin. Se tarkoittaa, että tarvittava data on saatavilla silloin, kun sitä tarvitaan. (Väre, 2019; Loshin, 2009; Smith ym. 2008).

Dataversityn vuonna 2019 liike-elämän eri osa-alueiden edustajille tekemässä kyselytutkimuksessa vastaajilta kysyttiin, mitkä ovat pääasialliset motiivit ydintiedon hallinnan toteuttamiseen organisaatioissa. Kyselyyn vastanneista 80% mainitsi motiiviksi paremman näkemyksen saamisen raportoinnin ja analytiikan kautta, 68% vastaajista motivoi kustannussäästöt ja parantunut toiminnan tehokkuus ja lisäksi 59% asiakas-tyytyväisyyden parantaminen (Burbank, D., Knight, M. 2019).

Ydintiedon hallinta (MDM) auttaa myös riskien hallinnassa ja erilaisiin lakivaatimuksiin mukautumisessa (Loshin, 2009; Jonker ym., 2011). Toukokuussa 2018 voimaan astunut uusi EU:n tietosuoja-asetus (*General Data Protection Regulation, GDPR*) edellyttää organisaatioita tekemään asioita, jotka ovat osa hyvää datan hallintaa. Se edellyttää yrityksiä muun muassa lisäämään tiedon läpinäkyvyyttä, saatavuutta ja virheiden oikaisua.

GDPR, on EU:n uusi tietosuoja-asetus, joka korvaa ja yhtenäistää eri maiden nykyiset säännöt. Se määrittelee, miten henkilötietoja saa käsitellä EU:ssa. Virallisesti asetus tunnetaan nimellä Asetus (EU) 2016/679.

Artikla 16

“Rekisteröidyllä on oikeus vaatia, että rekisterinpitäjä oikaisee ilman aiheutonta viivytystä rekisteröityä koskevat epätarkat ja virheelliset henkilötiedot.”

Artikla 17

“Rekisteröidyllä on oikeus saada rekisterinpitäjä poistamaan rekisteröityä koskevat henkilötiedot ilman aiheetonta viivytystä, ja rekisterinpitäjällä on velvollisuus poistaa henkilötiedot ilman aiheetonta viivytystä”.

Artikla 20

“Rekisteröidyllä on oikeus saada häntä koskevat henkilötiedot, jotka hän on toimittanut rekisterinpitäjälle, jäsennellyssä, yleisesti käytetyssä ja koneellisesti luettavassa muodossa, ja oikeus siirtää kyseiset tiedot toiselle rekisterinpitäjälle sen rekisterinpitäjän estämättä, jolle henkilötiedot on toimitettu”. (<https://findwise.com/en/gdpr-fi#1>)

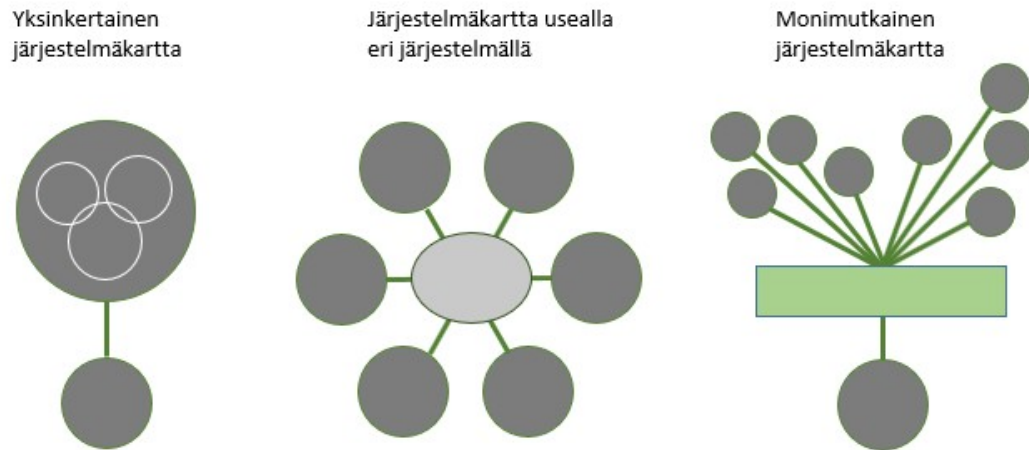
3.2.3 Ydintiedon hallinnan arkkitehtuuri

Jotta ydintiedon hallinnan hyödyt saadaan käyttöön, tieto pitää olla helposti hyödynnettävissä ja jaettavissa läpi organisaation. Onnistuminen tässä vaatii huolellisesti valittua mallia järjestelmien keskinäisestä arkkitehtuurista. Näitä malleja on monia erilaisia, joista jokainen sopii erilaiseen organisaatioon ja tarpeeseen. (Väre, 2019)

Arkkitehtuurimallit sisältävät päivittäisessä toiminnassa käytettävät tietojärjestelmät, joita kutsutaan operatiivisiksi järjestelmiksi. Operatiivisten järjestelmien lisäksi malleissa käsitellään MDM-järjestelmää. Tällä tarkoitetaan tietojärjestelmää, joka toimii ydintiedon pääasiallisena varastona ja jossa ydintieto on luotettavinta. Esimerkkinä mainittakoon yrityksen käyttämä toiminnanohjausjärjestelmä (ERP) tai asiakashallintajärjestelmä (CRM). (Väre, 2019; Maedche, 2010)

Vaikka ydintiedon hallintaa (MDM) ei tule pitää teknologiaprojektina, on selvää, että sitä ei voida tehdä ilman työkalujen ja tekniikan hyödyntämistä. Ydintietojen hallinnan tekniset näkökohdat perustuvat työkaluihin, jotka mm. tukevat tiedon tuottamista, tiedon analysointia, metatietojen hallintaa, tiedon mallintamista, integrointia, standardisointia, tietueiden yhdistämistä, sovittamista sekä puhdistamista (Loshin, 2009).

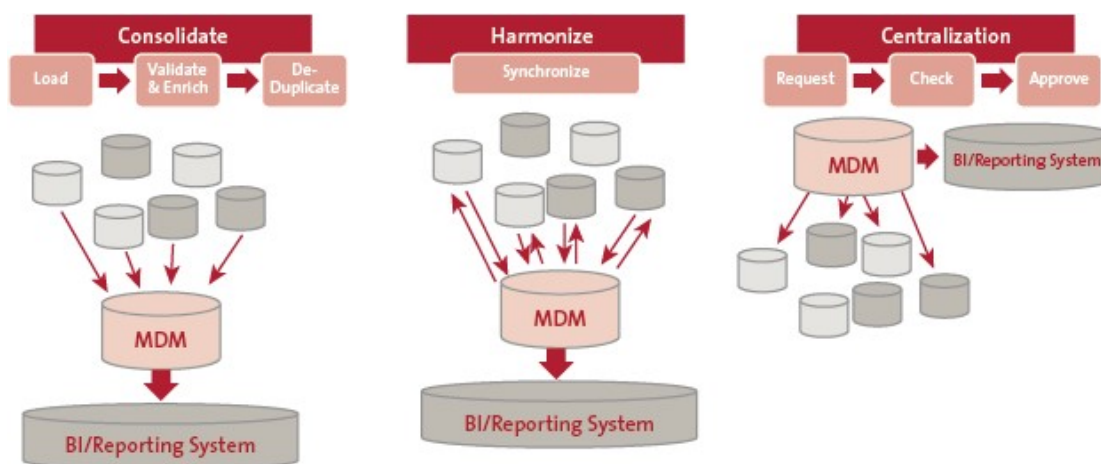
Ydintiedon hallinnan näkökannalta valittavaan arkkitehtuurimalliin (kuvio 4) vaikuttavat eniten organisaation käyttämien järjestelmien määrä ja monipuolisuus (Väre, 2019).



Kuvio 4. Ydintiedon hallinnan järjestelmäkartta (Väre, 2019)

1. Yksinkertainen järjestelmäkartta. Organisaatiossa on vain vähän eri järjestelmiä käytössä.
2. Järjestelmäkartta usealla eri järjestelmällä. Eri domainien data luodaan eri järjestelmiin ja ylläpidetään jossain toisissa järjestelmissä.
3. Monimutkainen järjestelmäkartta. Organisaatiossa on käytössään runsaasti erilaisia järjestelmiä. Samaa ydintietoa luodaan ja ylläpidetään eri järjestelmissä riippuen liiketoiminoista tai maasta. Ydintietoa on hyödyntämässä kymmenittäin erilaisia pienempiä ja suurempia järjestelmiä.

Jokaisessa ryhmässä on omanlaisensa vaihtoehdot järjestelmäarkkitehtuurille ydintiedon hallinnan näkökannalta ja jokaisessa järjestelmäryhmässä voidaan ydintiedon hallintaa tarkastella kolmella eri vaihtoehtoisella tavalla (kuvio 5).



Kuvio 5. Kolme erityyppistä MDM-arkkitehtuuria (van der Staaij, A.J., de Goeij, A.S.M., van Unen, C.J.W.A., Swartjes, S., 2012)

1. Keskitetty hallinta (*Centralization*). Ydintiedon luominen, ylläpito ja poistaminen tehdään keskitetysti yhdestä järjestelmästä käsin. Ydintieto siirretään tästä järjestelmästä muihin järjestelmiin integraatiolla. Muissa järjestelmissä ydintietoon voidaan lisätä muuta dataa, mutta perustietoa ei voida muuttaa tai poistaa. Yksinkertainen malli tiedon laadun hallinnan kannalta.

Yksinkertaisen järjestelmäkartan omaavan organisaation ei tarvitse useinkaan suunnitella keskitettyä ydintiedon hallintaa, vaan se saattaa olla jo valmiiksi keskitetty.

2. Hajautettu hallinta (*Consolidate*). Ydintiedon luominen, ylläpito ja poistaminen tehdään täysin hajautetusti eri operatiivisissa järjestelmissä. MDM-järjestelmä toimii vain tiedon välittäjänä eri järjestelmien kesken.
3. Yhdistelmähallinta (*Harmonize*). Ydintietoa luodaan, ylläpidetään ja poistetaan niin operatiivisissa kuin MDM-järjestelmässäkin. Teknisesti monimutkaisin malli, mutta mahdollistaa hyvän tiedon laadun hallinnan. (Väre, 2019)

Yksinkertaisella järjestelmäkartalla erillinen MDM-järjestelmä tuo vain turhaa monimutkaisuutta, mikäli olemassa olevat järjestelmät voivat hallita dataa riittävällä tasolla. MDM-järjestelmän perusominaisuus on kerätä dataa useammasta lähteestä ja toimittaa yhtenäistettyä dataa edelleen. Jos datalle ei ole useaa eri lähdettä eikä useampaa kohdetta, MDM-järjestelmä olisi kuin monikaistainen monitasoristeys keskellä kaksikaistaista suoraa tietä. (Väre, 2019)

Ennen minkäänlaisen ydintiedon hallintatyökalun hankkimista on syytä olla hyvin selvillä oman organisaation tarpeista ydintiedon hallinnan suhteen sekä nykyisistä käytössä olevista työkaluista ja niiden mahdollisuuksista ja rajoitteista näihin tarpeisiin nähden. MDM-työkalut eivät yksistään ole keino ratkaista ydintiedon hallinnan haasteita, vaan enemmänkin keino, jolla organisaatiot voivat päästä ydintiedon hallinnan tavoitteisiinsa (van der Staaij ym., 2012).

3.3 Ydintiedon laatu (Master Data Quality)

Laadullisesta näkökulmasta vain kahdella hetkellä on merkitystä tiedon elinkaaren aikana: toinen on hetki, kun se luodaan ja toinen on hetki, kun sitä käytetään (Redman, 2013).

Ydintietojen hallinnan (MDM) menestys riippuu korkealaatuisesta tiedosta ja tiedon laatua voidaan puolestaan parantaa ydintietojen hallintaohjelmalla (Loshin, 2009). Mikäli organisaatiossa ei keskitytä ydintietojen laadun hallintaan, ja tiedon laadun seuranta ja parantamista tehdään ainoastaan satunnaisesti osana yksittäisiä hankkeita, tiedon laatu palautuu aiemmalle tasolle hyvinkin pian tehtyjen toimenpiteiden jälkeen. Tiedon laatu on kriittinen ydintietojen hallinnan menestyksen kannalta (Loshin, 2009; Väre, 2019; Smith ym. 2008), ja jos laatuvirheiden juurisyihin ei puututa, mikään ei muutu. Jotta tiedon laadussa voi saada aikaan pysyvän muutoksen parempaan, tarvitaan organisaatiossa pysyvä toimintamalli tiedon laadun parantamiseksi (Väre, 2019; Jonker ym., 2011).

Smith ym. (2008) lisää, että ilman hyvää tiedon laatua ei voida parantaa luottamusta dataan ja ilman luottamusta käyttäjien sitoutuminen datasidonnaisiin kehityshankkeisiin jää huonoksi ja siten mahdolliset hyvän datan tuomat hyödyt jäävät saavuttamatta.

3.3.1 Ydintiedon laadun mittaaminen

Ydintietojen laatuodotukset on jaoteltu tiedon laatuominaisuuksiin niiden määrittelyn, mittaamisen ja validoinnin yksinkertaistamiseksi (Loshin, 2009). Mikäli tiedon

laadun halutaan parantavan organisaation toimintaa, vaaditaan syvällistä perehtymistä tiedon kontekstiin ja käyttötapauksiin, eli miten tiedon laatu puutteet todella vaikuttavat organisaatiossa. Väre (2019) kehottaa pohtimaan, mitkä ovat niitä liiketoiminnallisia haasteita, joita tiedon laadun puutteet aiheuttavat.

Pohjimmiltaan datan laadun vaatimustenmukaisuuden mittaukset määrittelevät keskeiset suoritusindikaattorit (KPI) yritykselle itselleen, joten olisi hyvä, jos laadun mittaukseen ja raportointiin valikoituisi ne ongelmat, jotka ovat kaikkein kipeimpiä ja siksi tärkeimpiä ratkaista (Loshin, 2009; Väre, 2019). Mittarit kannattaa suunnitella myös edistyviksi ja kehittyviksi. Eli jos päätät aloittaa yhdellä mittarilla, siinä ei tarvitse pysyä loppuikä (Väre, 2019). Hän lisää, että hyvä mittari ei kerro vain arvoa, vaan se kertoo, miksi luku on merkityksellinen.

Artikkelissaan van der Staaij, A.J., de Goeij, A.S.M., van Unen, C.J.W.A., Swartjes, S. (2012) mainitsevat tyypillisiksi ydintiedon laadun mittareiksi mm. tuplatietuiden määrän, ydintiedon attribuuttien täyttöasteen ja näiden attribuuttien vaatimustenmukaisuus määritelyihin datastandardeihin nähden. Ydintiedon hallinnan laadullisia mittareita voisivat heidän mukaansa olla mm. kuinka nopeasti datan laatuvirheet korjataan sen jälkeen, kun ne havaitaan tai kuinka moni ydintiedoista ei ole enää aktiivisesti käytössä.

Datan laadun hallinta on toimintamalli, jolla organisaatiossa huolehditaan datan laadusta, sen ohjaamisesta, seurannasta ja laadun parantamisen prosesseista. Siihen liittyy erittäin tärkeänä osana tietoisuuden kasvattaminen datan laadun merkityksestä organisaatiolle. Datan hyvä laatu on se tavoite, johon datan hallinnasta vastaavat henkilöt pyrkivät kaikella toiminnallaan.

Datan laadun hallinnan on tarkoitus olla jatkuva ja kehittyvä prosessi datan hyvän laadun varmistamiseksi. Sen lähtökohta on asiakkaiden tarpeet ja organisaation tavoitteet. Parhaimmillaan datan laadun hallinnan huomio keskittyy siihen, miten hyvä laatu rakennetaan osaksi päivittäisiä toimintatapoja. Silloin datan laatu otetaan huomioon jo siinä, kun dataa luodaan ja aina kun sitä muokataan.

Datan laadusta, samoin kuin ydintiedon hallinnasta on puhuttu jo yli kahdenkymmenen vuoden ajan. Ja vaikka datan laadun merkityksen ymmärrys on parantunut viime vuosina, datan laatu ei ole. (Väre, 2019)

Experian Information Solutionin raportissa vuodelta 2018 kysyttiin datan laadun hallinnan kypsyystilasta organisaatioissa. Kyselyn tuloksista oli huomattavissa selkeä muutos aiempiin vuoden 2016 ja 2017 tuloksiin verrattuna. Vuonna 2018 kyselyyn vastanneista 29% näki oman organisaation kypsyystilan olevan alhaisin mahdollinen, kun samoin ajattelevien osuus oli aiempina vuosina ollut 17%. (Experian: The 2018 Global data management Benchmark report, 2018)

Tämä ja muut vastaavat raportit osoittavat, että samalla kun ymmärrys datan laadun merkityksestä on kasvanut, on havaittu hyvin haastavaksi saada se toteutumaan.

Yli puolet tutkimuksessa haastatelluista organisaatioista (52%) ilmoittaa, että niiden laadukkaan datan ylläpitämisen strategiat perustuvat ensisijaisesti tarpeeseen parantaa toiminnan tehokkuutta ja saada aikaiseksi kustannussäästöjä.

Jos datan laadun katsotaan olevan heikko, ihmiset yleensä reagoivat siihen ohittamalla datan virheet tai korjaamalla virheet itse. Työntekijät, rutiininomaisesti, kiertävät tai korjaavat valtaosan datavirheistä päivittäisessä työssään. Redman (2013) viittaa artikkelissaan tutkimuksiin, jotka osoittavat, että tietotekniikan työntekijät tuhlaavat jopa 50% ajastaan etsiessään dataa, tunnistamalla ja korjaamalla virheitä ja etsiessään varmennuslähteitä tiedoille, joihin he eivät luota. Lisäksi hän huomauttaa, että ne virheet, jotka pääsevät ”läpi” aiheuttavat vahinkoa asiakkailta organisaation ulkopuolella ja siten datavirheiden mainehaitat voivat olla varsin merkittäviä.

Redman (2013) jatkaa, että kun data ei ole luotettavaa, johtajat menettävät nopeasti uskonsa siihen ja palaavat takaisin luottamaan pelkästään intuitioonsa tehdessään päätöksiä, ohjatessaan yrityksiä ja toteuttaessaan strategiaansa. 50 vuotta ilmaisun ”roskaa sisään, roskaa ulos” jälkeen, kamppailemme edelleenkin datan laatuongelmien kanssa. Redman uskoo kuitenkin, että ongelman korjaaminen ei ole niin vaikeaa

kuin moni luulee, eikä ratkaisu ole myöskään teknologiaan turvautuminen, vaan tiedonkulun parantaminen datan luojien ja käyttäjien välillä.

Experian Information Solutionin raportin (2018) mukaan ihmisten tekemät virheet ovat suurin tekijä datan laadun näkökannalta. Väre (2019) lisää, että myös motivaation puute sekä virheelliset, ristiriitaiset tai puutteelliset ohjeet aiheuttavat virheitä. Hyvän ja selkeän ohjeistuksen laatiminen vaatii aikaa ja vaivaa, minkä lisäksi ohjeet täytyy myös jalkauttaa, eli niistä pitää kommunikoida runsaasti. Hän mainitsee myös, että huonosti suunniteltu käyttöliittymä voi aiheuttaa tilanteita, joissa ihminen ei ymmärrä, mitä dataa hänen pitäisi mihinkin kenttään syöttää. (Väre, 2019)

Experian Information Solutionin raportissa (2018) todetaan, että näiden ongelmien ratkaisemiseksi organisaatioiden on otettava käyttöön asianmukaiset datan hallinnan käytännöt sääntöjen noudattamisen varmistamiseksi sekä investoitava tarvittavaan tietotekniikkaan, joka auttaa tunnistamaan ja ennalta ehkäisemään virheitä, joita syntyy dataa luotaessa ja ylläpidettäessä.

3.3.2 Ydintiedon laadun ulottuvuudet

Tiedon laatu on todella moniulotteinen asia. Jotkut näistä ulottuvuuksista on helppoja tulkita, toiset vaikeampia. Eivätkä kaikki ole aina yhtä merkityksellisiä laadun kannalta (Väre, 2019). Tiedon laatua voidaan siis tarkastella monista eri lähtökohdista ja eri lähteissä korostetaan eri ulottuvuuksien tärkeyttä, mutta useimmin esiin tuleviksi tiedon laadun ulottuvuuksiksi mainitaan ainutlaatuisuus, oikeellisuus, yhteneväisyys, kattavuus ja ajanmukaisuus (Loshin, 2009; Väre, 2019). Väre, DAMA DMBOK: a muokailen listaa kahdeksan merkityksellisintä ulottuvuutta seuraavasti:

Sisällön kattavuus (Completeness) kertoo, onko dataa olemassa. Puuttuva data on suuri ongelma. Jos sitä ei ole, tietokanta on tyhjä ja liiketoimintaprosesseilla ei ole tarvitsemaansa dataa. Prosessit keskeytyvät tai hidastuvat puuttuvan tiedon etsimiseksi. Analyysit ja raportit eivät näytä mitään tuloksia (tai vähintäänkin virheellisiä tuloksia).

Tiedon oikeellisuus (Accuracy) on kaikkein tärkein tiedon laadun ulottuvuus. Se kertoo, onko olemassa oleva data oikein, eli vastaako se tosielämää. Tiedon oikeellisuus on tärkein, mutta myös yksi vaikeimmin todennettavia tiedon ulottuvuuksia.

Ainutlaatuisuus (Uniqueness) on ydintiedon hallinnassa erittäin oleellinen laadun ulottuvuus. Tieto on ainutlaatuista silloin, kun se esiintyy yhdessä tietokannassa tai järjestelmässä vain kerran. Mikäli tieto ei ole ainutlaatuista, ongelmat tulevat tyypillisesti esiin raportoinnin tai analyysien epäselvyytenä. Saman ydintietotietueen toistuvat ilmentymät eli tuplatietueet aiheuttavat piileviä ongelmia liiketoimintaprosesseissa. Tiedon ainutlaatuisuus on kohtalaisen helppo varmistaa. Mitä parempi on sisällön kattavuus ja sen oikeellisuus, sitä helpompi ainutlaatuisuuden tarkistaminen on.

Kelpoisuus/vaatimustenmukaisuus (validity) viittaa siihen, että tieto täyttää sille asetetut muodolliset ja sisällölliset vaatimukset. Muodollisten vaatimusten täyttyminen edistää tiedon teknistä tulkittavuutta sekä siirrettävyyttä eri järjestelmien välillä. Sisällöllisten vaatimusten täyttyminen tarkoittaa, että tieto täyttää sille annetut liiketoiminnalliset vaatimukset.

Yhteneväisyys (Consistency) tarkoittaa tiedon sisällön samankaltaisuutta kaikkialla missä tietoa on. Puutteet tiedon yhteneväisyydessä aiheuttavat usein hämmennyksen organisaation sisällä, mutta erityisesti sen ulkopuolella. Yhteneväisyyden puute voi aiheuttaa asiakaspalveluongelmia ja maineenmenetystä.

Eheys (Integrity). Tiedon eheydellä ydintiedon yhteydessä tarkoitetaan, ovatko eri järjestelmien väliset viittaukset oikein, eli voidaanko yksi ydintiedon tietue yhdistää eri järjestelmien välillä luotettavasti. Tietokannan sisällä eheys tarkoittaa tietokannan sisäisiä viittauksia.

Järkevyys (Reasonability) ilmaisee, onko tieto malliltaan järkevää tai oletusten mukaista. Mikäli tieto ei ole järkevää, siitä aiheutuu helposti samankaltaisia ongelmia kuin puutteista tiedon vaatimuksenmukaisuudessa (esim. tulevaisuuteen merkitty kuolinpäivä tai menneisyyteen merkitty toimituspäivä).

Ajanmukaisuus (*Timeliness*) kertoo, onko tieto oikea-aikaista, eli se ei ole liian vanhaa tai että se on saatavilla oikeaan aikaan. Ydintiedossa harva tieto on ajanmukaisuuden näkökulmasta merkityksellistä, mikäli huolehditaan tiedon oikeellisuudesta (Väre, 2019).

3.4 Datan hallinta (Data Governance)

Hallintatarve syntyy, kun joukko ihmisiä kokoontuu yhteen saavuttaakseen yhteisen tavoitteen. Hallintaa (Governance) käsittelevä kirjallisuus tunnistaa sille useita eri määritelmiä, mutta kuitenkin suurin osa pohjautuu kolmeen pääkohtaan: auktoriteetti, päätöksenteko ja vastuuvollisuus. Hallinta määrittelee, kenellä on valta, kuka tekee päätökset ja kuinka kaikki sidosryhmät/toimijat saavat äänensä kuuluviin ja kuinka vastuuvollisuus jakautuu. (Institute of Governance, 2019)

Cerwinin (2017) mukaan on olemassa kolme (ihmis)ryhmää, jotka muodostavat tiedon hallinnan yrityksissä. On niitä, jotka luovat sääntöjä, niitä, jotka toteuttavat sääntöjä ja niitä, jotka päättävät säännöistä. Datan hallinta muistuttaakin monelta osin (länsimaisten) valtioiden hallintoa, jossa on erilliset instanssit lainsäädännölle, lakien toimeenpanolle ja niiden seurannalle (oikeuslaitos). Tiedon hallinnan ”lakiosasto” kehittää ja välittää säännöt siitä, ketkä, missä ja milloin tietoja voi käyttää. Tiedon hallinnan johtoryhmä määrittelee, miten säännöt toteutetaan, suorittaa ne joko manuaalisten tai automatisoitujen prosessien muodossa ja valvoo sääntöjen noudattamista. Ja kun rikkomuksia havaitaan, oikeuslaitos tuomitsee sääntöjen rikkomisista ja määrää seuraamuksia.

Tiedon hallinnan (Data Governance) lähtökohtana on halu arvioida ja hallita monenlaisia yritystietokannassa piileviä riskejä ja vähentää valvonnan puutteesta aiheutuvia seurauksia (Loshin 2009). Ydintiedon hallinta (MDM) edellyttää datan hallintaa (DG). Frank Cerwin toteaa haastattelussaan, että ydintietoa ei voi hallita ilman datan hallintaa. Ja että juuri datan hallintaan luodut säännöt takaavat ydintiedon laadun. (Cerwin, 2019; Jonker ym., 2011)

Kirjallisuudessa on olemassa useita eri käsityksiä siitä, mitä tarkoitetaan termillä “datan hallinta” (data governance). Yleisesti ottaen tiedon hallinnan odotetaan kuitenkin varmistavan, että data vastaa liiketoiminnan sille asettamia vaatimuksia tiedon hallinnan, omistajuuden, sääntöjen noudattamisen, yksityisyyden, turvallisuuden, tietoriskien, metatietojen hallinnan ja ydintiedon hallinnan (MDM) osalta. (Weber, K., Otto, B., Österle, H. 2009; Loshin 2009; Mosley, M. 2010; Panian, 2010)

Jokainen tiedon hallinnan näkökohta (kuvio 6) liittyy tiedon hallinnan (MDM) menettelytapojen määrittelyyn, jotka heijastavat liiketoiminnan tarpeita ja odotuksia, samoin kuin prosessit näiden menetelmien noudattamisen seuraamiseksi (Loshin 2009; Mosley, M. 2010).



Kuvio 6. 10 datan hallinnan aluetta (Data Management Association)

Yksi yleisimmin tehtävä virhe määriteltäessä datan hallintaa (DG) on rinnastaa se ydintiedon hallintaan (MDM). Siinä missä datan hallinta on ydintiedon hallinnan menettelytapojen ja niihin liittyvien roolien ja velvollisuuksien tekemistä, niin ydintiedon hallinta on näiden menettelytapojen toteuttamista. (Dyché, J., Nevala, K., 2017)

Cerwinin (2019) mukaan datan hallinta (DG) ei tarkoita datan hallintaa, koska tiedolla itsessään ei ole kognitiivista tai päätöksentekokykyä sääntöjen noudattamiseksi. Au-

tomaation luominen tai sääntöjen suorittaminen manuaalisesti vaatii toimiakseen ihmisiä, joten datan hallinta tarkoittaa oikeammin ihmisten käyttäytymisen hallintaa.

Ydintiedon omistaa liiketoiminta, joten vastuu ydintiedosta ja sen kehittämisestä kuuluu liiketoiminnalle, joko sopivalle roolille tai kollektiivisesti ryhmälle rooleja. Toki lopullisessa vastuussa on aina organisaation ylin johto. Ydintieto on olemassa liiketoimintaa varten ja koska liiketoiminta on se, joka luo, ylläpitää ja käyttää sitä, sen täytyy myös ottaa vastuu sen kunnossapidosta. Ja jotta organisaation ydintieto on varmasti hyväkuntoista ja palvelee liiketoimintaa parhaalla mahdollisella tavalla, vaaditaan liiketoiminnan johtajuutta tiedon kehittämisessä. Ydintiedon hallinta on 80 prosenttisesti (kuvio 7) liiketoiminnan toimintatapojen kehittämistä ja ihmisten johtamista. (Väre, 2019)

Ydintiedon hallinnan 80/20 -malli



Kuvio 7. Ydintiedon hallinnan 80/20 -malli (Väre, 2019)

Nokkala, T., Salmela, H., Toivonen, J. (2019) ovat tutkimuksessaan nostaneet kirjallisuudesta esiin viisi tiedon hallinnan pääperiaatetta: tiedon periaatteet (kuinka data yhdistetään liiketoimintaan, mitkä ovat mahdollisuudet, oikeudet ja vastuut), tiedon laatu (esim. tarkkuus, ajantasaisuus, täydellisyys ja uskottavuus; niihin liittyvät vastuut), metatiedot (tiedot tiedosta), tiedon käyttöoikeudet (kuka voi luoda, nähdä, muokata tai poistaa tietoja) ja tiedon elinkaari (tietojen määrittely, tuottaminen, poistaminen ja arkistointi). Edelleen kuudentena tärkeänä periaatteena he nostavat

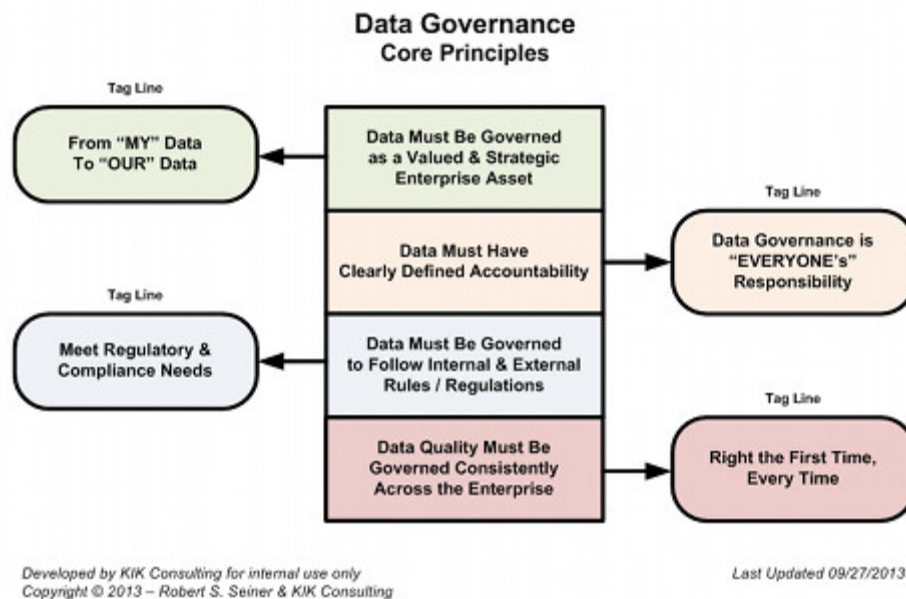
kirjallisuudesta esiin tiedon hallinnoinnin (*data stewardship*), joka on tiedon hallinnan vastuiden suorittamista.

Kuten aiemmin ydintiedon hallintaa (MDM) käsiteltäessä mainittiin, niin nykyisenä säädösten kulta-aikana, syntyy jatkuvasti uusia säädöksiä, esim. GDPR. Tämä on johtanut siihen, että enää ei riitä, että dataa vain hallinnoidaan tehokkaasti, nyt on tärkeää, että yritys hallitsee sitä, kuka voi päästä käsiksi ja käyttää tietoaineistoja.

Riskien hallinta on suurin yksittäinen tekijä, joka ajaa tarpeeseen hallinnoida dataa aiempaa paremmin. Sekä liiketoimintariskit että säädösten noudattamisen riskit ohjaavat hallintaa (Loshin, 2009; Panian, 2010). Panian lisää, että datan hallinta on ensisijaisen tärkeää, jotta yritykset pystyvät vastaamaan alati kasvaviin vaatimuksiin.

Useimmat organisaatiot pitävät datan hallintaa tavanomaisen työponnistuksen lisäksi tehtävänä työnä, joka uhkaa organisaation nykyistä työkuilttuuria. Robert s. Seinerin (2014) mukaan sitä ei kuitenkaan tulisi ajatella niin.

Seinerin (2014) mukaan datan hallinta on tiedon hallinnan valtuuksien toteuttamista ja täytäntöönpanoa, jonka avulla datan käyttäjien vastuut tunnustetaan ja tunnustetaan, virallistetaan ja otetaan käyttöön heidän nykyisten vastuidensa mukaisesti sen sijaan, että heille määrättäisiin lisää työtä. Seiner (2014) kuvaa lisäksi datan hallinnan neljä pääperiaatetta seuraavasti (kuvio 8).



Kuvio 8. Datan hallinnan periaatteet (Robert S. Seiner, 2014)

- Periaate 1: Data on tunnistettava arvokkaana ja strategisena yrityksen omaisuutena. Täsmällinen ja ajantasainen data on perusta tehokkaalle päätöksenteolle ja asiakaspalvelulle. *From "MY" Data to "OUR" Data.*
- Periaate 2: Datalla on yrityksen sisällä arvoa useammallekin kuin ainoastaan yhdelle sitä käyttävälle sovellukselle. Datalle on oltava selkeästi määritelty vastuuvollisuus (accountability). *Data Governance is "EVERYONE's" Responsibility*.
- Periaate 3: Dataa on hallittava sisäisten ja ulkoisten sääntöjen/määräysten mukaisesti. Datan hallintaorganisaatio on vastuussa yksittäisten käyttäjien perehdyttämisestä ja vastuuttamisesta. *Meet Regulatory & Compliance Needs*
- Periaate 4: Datastandardit on määriteltävä ja dokumentoitava hyvin. Yrityksen kriittisen datan vastaavuutta datastandardeihin on jatkuvasti seurattava (mitattava). Lisäksi ymmärrettävät datastandardit tulee saattaa jokaisen käyttäjän ulottuville. *Right the First Time, Every Time.*

Datan hallintaa (DG) sovelletaan olemassa oleviin toimintatapoihin, käytäntöihin ja menetelmiin sen sijaan, että otettaisiin käyttöön uusia prosesseja tai menetelmiä. Tiedonhallinnan tarkoituksena on olla läpinäkyvä soveltamalla datan hallintaa olemassa oleviin prosesseihin, sen sijaan, että se ohjaisi organisaation ajattelemaan,

että kaikki hallittavat prosessit johtuivat tiedon hallinnan toiminnasta. Hyvin toteutettuna datan hallinta keskittyy saamaan oikeat henkilöt mukaan nykyisiin prosesseihin oikeaan aikaan ja oikeasta syystä. (Seiner, s. 13)

3.4.1 Tiedon hallinnan (DG) haasteet

Suunniteltaessa datan hallintaa tulee se sovittaa kulloisenkin organisaation tarpeisiin, kulttuuriin ja organisaatorakenteeseen ja päätöksentekoprosesseihin (Dyché ym., 2017). Jokainen organisaatio on erilainen ja jokainen organisaatio on eri kypsyydellä, kun mietitään tiedon hallinnan haasteita. Siitä johtuen se miten haasteisiin vastataan, riippuu täysin siitä, millä kypsyydellä kulloinenkin yritys on ja mikä on yrityksen tavoitetaso tiedon hallinnolle. Seiner (2019) kuvaa yleisiksi tiedonhallinnan haasteiksi seuraavat asiat (ei kuitenkaan näihin rajoittuen): tietojohdamisen puute, tiedonhallinnan liiketoiminnallisen arvon ymmärtäminen ja tiedon omistajuuden puute.

Tietojohdamisen puute on haaste, jonka monet organisaatiot kohtaavat, koska ne ovat vasta hiljalleen heräämässä ajatukseen, että ne tarvitsevat ihmisiä, jotka ottavat vastuun datasta, ja tiedon hallinnan rakenteet, joiden avulla ihmiset pystyvät hallinnoimaan dataa (Panian, 2010).

Tiedonhallinnan liiketoiminnallisen arvon ymmärtäminen puolestaan auttaisi yrityksiä saamaan paremmin hyödyt irti erilaisista digitalisaatioon liittyvistä hankkeista, koska ilman hyvälaatuista ja hallintoitua dataa ei mikään tekninen ratkaisu toimi täydellä teholla, ja mikään tekninen ratkaisu ei itsessään paranna datan laatua ja arvoa. Data ei kykene hallinnoimaan itse itseään. (Seiner, 2019)

3.4.2 Organisaation kyvykkyys (Data Governance Maturity model)

Siirtyminen organisaatioon, joka luottaa yksinomaan ydintietoihin, ei tapahdu yhdessä yössä, pikemminkin se kehittyy useiden siirtymävaiheiden kautta. Ydintiedon hallinnan saavuttamisen ensiaskel on sen tosiasian tunnustaminen, että prosessiin kuuluu muutakin kuin ohjelmistopakettien ostaminen tai ulkopuolisten ratkaisutoimittajien palkkaaminen. (Loshin, 2009)

Suunnitellessaan tiedon hallinnan (DG) kehitystä, monet yritykset käyttävät kypsyyssmallia hallitakseen muutosta määrittämällä ensin mikä taso on sopiva yritykselle ja sen käyttämälle teknologialle, sekä miten ja milloin edetä tasolta toiselle (Seiner, 2014; Jonker ym., 2011).

Mitä pienempi organisaation kypsyys ja ymmärrys tiedon hallinnosta on, sitä suurempi muutos tarvitaan ja sitä kauemmin kehittyminen vaatii aikaa (Väre, 2019).

Capability Maturity Model (CMM) on menetelmä, jota käytetään kehittämään ja tarkentamaan organisaatioiden kehitysprosesseja. Malli kuvaa viiden tason evoluutiopolkua entistä organisoituneemmista ja systemaattisesti kypsemmistä prosesseista. CMM menetelmän on kehittänyt Software Engineering Institute (SEI) tutkimus- ja kehityskeskus, jota tukee Yhdysvaltain Puolustusministeriö (DoD).

CMM on samanlainen kuin ISO 9001, joka on yksi kansainvälisen standardisointijärjestön (ISO) määrittelemistä ISO 9000 -standardisarjoista. ISO 9000 -standardit määrittelevät tehokkaan laatu järjestelmän valmistus- ja palveluteollisuudelle. ISO 9001 käsittelee erityisesti ohjelmistojen kehittämistä ja ylläpitoa. Tärkein ero näiden kahden järjestelmän välillä on niiden tarkoituksissa. ISO 9001 määrittelee minimaalisen hyväksyttävän laatutason ohjelmistoprosesseille, kun taas CMM luo puitteet jatkuvalle prosessien parantamiselle ja on selkeämpi kuin ISO-standardi määriteltäessä välineet, joita käytetään tähän tarkoitukseen.

(<https://searchsoftwarequality.techtarget.com/definition/Capability-Maturity-Model>)

Kypsyyssmalleissa toiminnan kehittäminen jaotellaan tasoihin, joita kiivetään ylöspäin. Kariikoidusti esitettynä epäkypsässä organisaatiossa prosessit ovat kertaluonteisia, ad hoc -tyyppisiä ja improvisoituja. Kypsissä organisaatioissa prosessit ovat dokumentoituja, toistettavia ja johdettuja. Organisaatioiden kehittyminen tasolta toiselle ei tapahdu hetkessä, vaan se vaatii organisaatiolta kypsyyttä pitkäjänteiseen kehitystyöhön.

Alkutaso (taso 1). Tällä tasolla tapahtuvat datan hallinnan prosessit ovat tyypillisesti dokumentoimattomia ja dynaamisen muutoksen tilassa. Tällaisia prosesseja ohjataan

yleensä käyttäjien tai tapahtumien tapauskohtaisella, hallitsemattomalla ja reaktiivisella tavalla. Tämä tarjoaa kaottisen tai epävakaan ympäristön prosesseille. Tason 1 organisaatioilta puuttuu tiukat säännöt tai menettelytavat tietojen hallinnasta. Hulluudessa ei ole nähtävissä näkyvää menetelmää. Raportit kehitetään lennossa liiketoimintayksiköiden pyynnöstä, eikä datan omistajuutta ole määritelty. (Mosley, M. 2010)

Toistettava taso (taso 2). Jotkut prosessit ovat toistettavissa, mahdollisesti yhtenäisin tuloksin. Prosessien kurinalaisuus ei todennäköisesti ole tiukka. Siirtyäkseen tasolta 1 tasolle 2 organisaation on alettava noudattaa tiedonhallinnan parhaita käytäntöjä. Vaikka tason 2 organisaatiot seuraavat jonkinlaista datan hallintaohjelmaa, niiden on yleisesti ottaen vielä vakiinnutettava se. Tason 2 organisaatiot alkavat perustamaan tiedonhallinnan käytäntöjä, jotka keskittyvät tietyn tyyppisiin tietoihin, joita käytetään liiketoimintayksiköiden raportointiin.

Määritelty taso (taso 3). Tämä taso sisältää määriteltyjä ja dokumentoituja standardiprosesseja, jotka on perustettu ja joita voidaan parantaa jonkin verran ajan myötä. Vakioprosessit ovat käytössä ja niitä käytetään prosessien suorituskyvyn yhdenmukaisuuden varmistamiseen koko organisaatiossa. Tason 3 organisaatiot valvovat ja testaavat varmistaakseen, että tiedon laatuvaatimukset määritetään ja täytetään.

Hallittu taso (taso 4). Organisaatio seuraa ja ohjaa omia prosessejaan tiedonkeruun ja analyysien avulla. Johto voi tehokkaasti hallita vakioprosesseja, esim. ohjelmistojen kehittämiseen.

Optimoitu taso (taso 5), Tämän tason prosessit keskittyvät prosessin suorituskyvyn jatkuvaan parantamiseen asteittaisten ja innovatiivisten teknologisten muutosten ja parannusten avulla. (Seiner, 2014)

3.4.3 Datan hallintamalli

Historiallisesti tarkasteltuna yksi suurimmista tiedon hallinnan ongelmista on seurannan puuttuminen. Organisaatioilla voi olla hyvinkin määritelty hallintatapa, mutta ne eivät ehkä ole luoneet taustalla olevaa organisaatorakennetta hallintatavan mahdollistamiseksi käytännössä (Loshin, 2009). Datan hallintamalli on koko ydintiedon hallinnan peruskehikko ja se on elintärkeä tiedon hallinnan onnistumiselle (Väre, 2019).

Datan hallintamalli ei ole projekti tai hanke, eikä myöskään päälle rakennettuja prosesseja, vaan se on mielletävä pysyväksi ja systemaattiseksi toimintatavaksi kaikille tiedon parissa työskenteleville, ja joka on keskittynyt dataan organisaation voimavarana, tukien kaikkien sidosryhmien tarpeita yrityksessä (Loshin, 2009; Väre, 2019, Dyché ym., 2017).

Datan hallintamalli kattaa kaikki ne osa-alueet, joissa tehdään päätöksiä dataan liittyen. Datan hallintamalli kertoo, mitkä ovat jokaisen dataa käsittelevän henkilön vastuut ja velvollisuudet. Se kertoo, kenen puoleen voi kääntyä, kun vastaan tulee tiedon laatuongelma. Siinä myös määritellään vastuut siitä, miten asiat viedään käytäntöön ja miten niiden toteutumista seurataan. (Väre, 2019)

Dataa koskevia päätöksiä tehdään, oli organisaatiossa virallista hallintamallia tai ei. Hallintamallin hyödyt tulevat usein välillisesti tiedon laadun paranemisen tai prosessien tehostumisen kautta. Mutta mikäli organisaation johdossa on tunnustettu, että datalla on strategista merkitystä tai jopa rahallista arvoa, hallintamalli on kriittinen työkalu. Datan hallintamallin jalkauttaminen hyvin tehtynä on paras keino kasvattaa ymmärrystä tiedon merkityksellisyydestä läpi koko organisaation. (Väre, 2019)

Datarpeiden muuttuessa ja tietomäärien jatkuvasti kasvaessa, täytyy myös tiedon hallinnan pystyä sopeutumaan muuttuviin tilanteisiin ja sitä varten yrityksillä tulee olla järjestelmällinen ja pysyvä toimintatapa tiedon hallinnan menettelytapojen ja datasääntöjen kehittämiseen (Dyché ym., 2017).

Soveltaen DAMA DMBOK:n viitekehystä Taru Väre (2019) kertoo kestävän datan hallintamallin perustuvan kuuteen pääkulmakiveen.

Johtaminen ja strategia. Johdon sitoutuminen on elintärkeää datan hallintamallille. Myös ylin johto tekee dataan liittyviä päätöksiä ja heidät pitää saada mukaan hallintamalliin.

Liiketoimintavetoisuus. Ydintiedon omistajuus kuuluu liiketoiminnalle. Siksi myös datan hallintamallin vetovastuu tulee olla erityisesti liiketoiminnalla. Liiketoiminnan tehtävä on ohjata dataan liittyviä teknologiapäätöksiä samoin kuin se ohjaa liiketoimintaan vaikuttavia päätöksiä.

Vastuun jakaminen. Ydintietoa käytetään kaikkialla organisaatiossa. Datan hallintamallin täytyy jakaa myös vastuut tiedon hallinnasta sopivasti, niin että kaikki, joihin data vaikuttaa, voivat vaikuttaa dataan.

Monitasoisuus. Datan hallintamalli määritellään sopimaan organisaation eri tasoihin ja vastaamaan ydintiedon eri tasoista. Jos organisaatiossa on hyvin matala hierarkia, hallintamallissa ei pidä olla useampaa hierarkkista tasoa.

Viitekehys. Hallintamalli ei aina voi olla täsmälleen samankaltainen eri domainien kohdalla. Viitekehukseksi muodostettu hallintamalli antaa mahdollisuuden sille, että hallintamallin vastuut ja roolit voidaan sovittaa kuhunkin tarpeeseen sopivaksi.

Periaatteellisuus. Datan hallintamallin sovittaminen helpottuu, kun sille laaditaan toimintaa ohjaavat periaatteet. Hyvin laaditut periaatteet vähentävät sääntöjen tarvetta, lisäävät joustavuutta, mutta antavat kuitenkin lujan perustuksen yhtenäiselle toiminnalle. Esimerkkejä hyvistä periaatteista: ”ei minun dataani, vaan meidän dataamme”. Data on yhteistä ja siksi kaikkien tarpeita pitää kuunnella tiedon määrittelyssä. ”Oikein ensimmäisellä kerralla, aina”. Paras hyöty saadaan, kun data luodaan heti ensimmäisellä kerralla oikein. (Väre, 2019)

3.4.4 Roolit ja vastuut

Roolit ja vastuut ovat onnistuneen datan hallintaohjelman selkäranka. Jos tiedon hallinnan halutaan toimivan tehokkaasti ja datan laatuun vaikuttavasti, tarvitaan organisaation kulttuurille sopivien roolien määrittely ja käyttöönotto. Vastuullisuuden virallistaminen siitä, miten ihmiset määrittelevät, tuottavat ja käyttävät dataa, vaatii rooliperusteiset käytännöt (Seiner, 2019). Panian (2010) kuvaa omassa tutkimuksessaan organisaation itsessään olevan avainasemassa tiedon hallinnan onnistumisessa. Saavuttaakseen tiedon hallinnalle asettamansa tavoitteet, organisaatioiden täytyy määritellä datalle omistajuus, datastandardit ja sovittujen menettelytapojen noudattamista tulee valvoa.

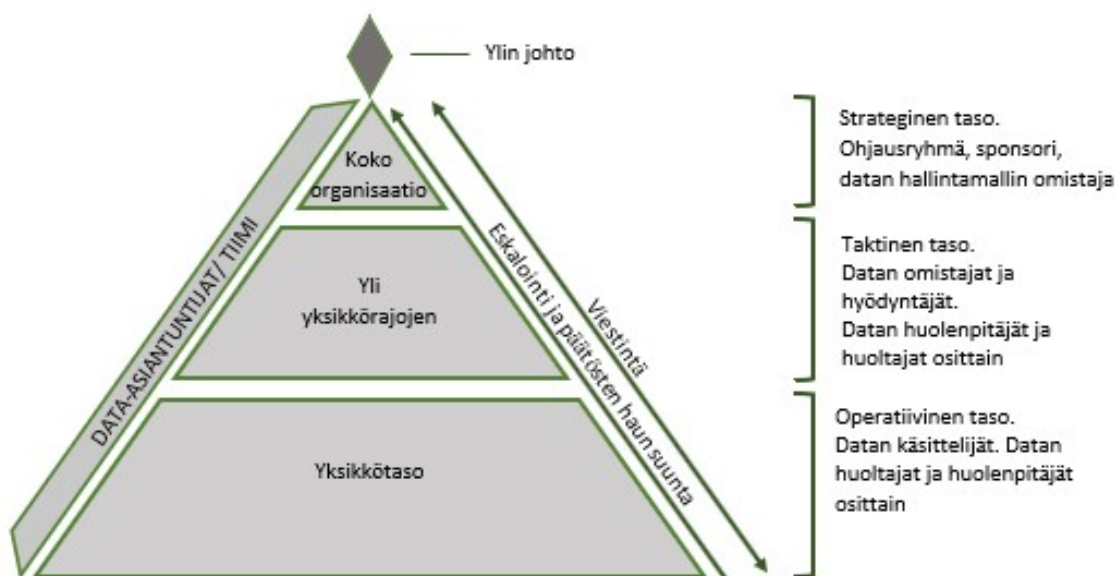
Seinerin mukaan paras tapa visualisoida hallintamalli tai roolien ja vastuiden viitekehys on pyramidikaavio (kuvio 9). Tila pyramidin jokaisen kerroksen sisällä edustaa tiedon päätöksentekotasoa. Päätökset olisi tehtävä operatiivisella tasolla, jos päätökset vaikuttavat vain kyseiseen tasoon.

Kun päätökset nousevat liiketoiminta-alueiden yli, nämä päätökset tehdään pyramidin taktisissa tai strategisissa kerroksissa, joissa yksilöillä ja osastoilla on valtuudet tehdä yritystä koskevia päätöksiä tietyistä aihealueista tai tietorekistereistä (domain). Monien organisaatioiden mielestä oikeiden, toimialueita ylittävien, roolien identifiointi on yksi vaikeimmista asioista mietittäessä tiedon hallinnan rooleja (Seiner, 2019). Tällä tasolla datasiilot rikotaan ja data jaetaan läpi koko organisaation.

Datan hallinta on koko organisaation laajuinen asia, joka vaatii tietotekniikan ja liiketoiminnan välisten esteiden poistamista ja korvaamista organisaation liiketoiminta-alueiden ja teknisten alueiden tarkoin määritellyillä rooleilla ja vastuulla. Tällöin kysymykset, kuka tekee, mitä ja milloin, ovat tärkeämpiä kuin missä. (Seiner, 2014)

Robert S. Seiner toteaa myös, että käytännöllisissä datan hallintaratkaisuuksissa yhdelläkään liiketoiminta-alueella ei voi olla ylivaltaa tai valtuuksia muun organisaation suhteen (Seiner, 2014).

Pyramidin oikean kyljen nuolet kuvaavat kaksisuuntaista kommunikaatiota ja ongelmien eskaloitumista. Huomattavaa ongelmien eskaloitumisessa on, että vain hyvin harvoin ja vaikeissa tapauksissa ongelmien ratkaisuja haetaan yritysten ylimmältä johdolta. (Seiner, 2019)



Kuvio 9. Datan hallinnan operatiivisen mallin roolit ja vastuut (Robert S. Seiner, 2019; Väre, 2019)

3.4.5 Datan hallintamallin yleinen rakenne

Operatiivisella tasolla ovat ne roolit, jotka käytännössä käsittelevät ydintietoa osana normaalia työtään (Seiner, 2014; Väre, 2019) He ovat henkilöitä, jotka perustavat uusia asiakkaita, tuotteita, toimittajia tai mitä tahansa muuta ydintietoa. Datan hallintamallissa operatiivisilla rooleilla on jonkin verran vastuuta, mutta ei välttämättä määräysvaltaa datasta ja sen käytöstä. Vastuu heidän kohdallaan tarkoittaa lähinnä yleisimmin vastuuta noudattaa tiedon omistajien ohjeistuksia ja toimintamalleja tiedon käsittelyssä. Tiedon käsittelijöillä on vastuu tiedon laadusta siltä osin kuin he itse sitä luovat ja ylläpitävät (Väre, 2019; Seiner, 2014). Lisäksi operatiivisten roolien vastuulla on kertoa omat tarpeensa tiedon omistajalle, joka huolehtii, että data tulee jatkossa vastaamaan näihin tarpeisiin (Väre, 2019).

Uutta dataa luovien ihmisten olisi oltava muodollisesti vastuussa siitä, että he tarkastavat etukäteen, mitä tietoja on jo olemassa, ennen kuin he tuottavat esimerkiksi tuotteesta (nimikkeestä) toisen version (Seiner, 2018).

Rooleja: Datan käyttäjä, Operatiivinen tietovalvoja, Operational Steward/Data Steward

Taktisen tason tekeminen on avainroolissa kaikkien tiedon hallinnan operatiivisten tehtävien ohjaamisessa. Taktisella tasolla huolehditaan tiedon määrittelystä, tiedon hallinnan prosessien laatisemisesta ja linkittämisestä olemassa oleviin liiketoimintaprosesseihin, arkkitehtuuri- ja teknologiavalinnoista sekä tiedon laadun seurannasta. Taktisella tasolla ovat ne henkilöt, jotka kuuluvat tietyn toimialan tai liiketoimintayksikön piiriin ja joilla on koko organisaation laajuinen näkemys toimialueestaan, *domainista*. He vastaavat kyseisen domainin määrittelystä, kehittämisestä ja erityisesti sen laadusta kokonaisuudessaan. (Seiner, 2014; Väre, 2019; Jonker ym., 2011)

Datan hallinta taktisella tasolla toimialarajoja rikkoen edellyttää, että tietyssä asemassa olevalla henkilöllä on vastuu kyseisestä toimialojen rajoja ylittävästä domainista operatiivisella tasolla. Siksi pitäisikin olla selvää, että yksi henkilö ei voi hallita kaikkea toimialarajoja ylittävää dataa ja onkin tärkeä erottaa data pienemmiksi, liiketoimintatiedon osa-alueiksi, eli domaineiksi. (Seiner, 2014).

Ydintiedon kehittämisessä tarvetta ja priorisointia ohjaavat liiketoiminnan ja asiakkaan tarpeet. Taktisen tason roolien tärkein tehtävä on tunnistaa nämä tarpeet ja viedä niiden mukaisesti ydintiedon hallinnan osa-alueita eteenpäin, huomioiden kuitenkin, että koko organisaation yhteinen etu menee heidän omien liiketoimintalueidensa edun edelle. (Väre, 2019, Seiner 2014). Tämä voi johtaa jonkin asteiseen kipuiluun, koska monet organisaatiot ovat tottuneet toimimaan silloissa ja siirtymisen toimialarajat rikkovaan liiketoimintaan voikin tuoda mukanaan poliittisia taisteluja, mielipide-eroja ja paljon työtä (Seiner, 2014).

Rooleja: Taktinen tietoalueenvalvoja, Data Domain Steward, Data owner, Subject Matter Expert (SME)

Apurooleja: Datan huolenpitäjä, Data Steward

Strategisella tasolla tehdään päätökset, joita ei voida tehdä operatiivisella tai taktisella tasolla. Strategisella tasolla on vastuu koko organisaation laajuisesta datastrategiasta sekä koko ydintiedon hallinnan ja sen merkityksen tunnettuuden edistämisestä (Väre, 2019). Strategiset päätökset edellyttävät, että päätöksiä tekevillä ihmisillä on asianmukaiset tiedot auttamassa heitä tekemään oikeita päätöksiä (Seiner, 2014).

Tiedon hallinnassa on oltava tietyt auktoriteetit. Jonkun on oltava vastuussa vaikeidenkin päätösten tekemisestä yrityksen tiedon suhteen. *Ohjausryhmä* on ylin toimielin dataa koskeville päätöksille ja sen tulisi olla muodollinen ja sisältää kaikki tarvittavat liiketoiminta-alueet yrityksestä. Asiat, jotka koskevat kaikkia domaineja vaativat ohjausryhmän hyväksynnän. (Väre, 2019; Seiner, 2014)

Rooleja: Sponsorit, Datan hallinnan ohjausryhmä/Data Governance Council

Tiedon hallinnan johtotasolla (executive level) ei ole muuta erityistä toimintoa tiedonhallinnossa sen lisäksi, kuin että se tukee, sponsoroi ja ymmärtää sitä (Seiner, 2014).

Tiedon hallinnan tiimi (**Data Governance Team, DGT**) jäsenet ovat organisaatiossa jo työskenteleviä henkilöitä, joiden työajasta on osa tarkoitettu tiedon hallinnan määrittämiseen, kehittämiseen ja käyttöönottoon yrityksessä.

Tyypillisesti datan hallintatiimin vastuu sisältää mm:

1. Tiedon hallinnan kehittämisen ja toteuttamisen valvonnan
2. Organisaation tiedonhallinnan roolien, vastuiden ja viestinnän tarkastelun ja dokumentoinnin, sekä tiedon hallinnan läpimenon varmistaminen organisaatiossa.
3. Tiedon hallinnan mittareiden määrittäminen, suosittaminen ja hyväksynnän saaminen datan hallintaneuvostolta.
4. Varmistaa, että datastandardimääritelmät, -menetelmät ja -mallit ovat paikallaan tietojen riskien, laadun ja käytettävyyden hallinnan ylläpitämiseksi ja parantamiseksi.

5. Dataongelmien hallinta, mukaan lukien puuttuvat tai virheelliset dataraportit ja dataan pääsyongelmat.
6. Tarkastaa satunnaisesti dataroolien sääntöjen noudattaminen.

On syytä huomata, että datan hallintatiimin jäsenet eivät ole taktisia tai operatiivisia tietojen hallitsijoita, eivätkä he ole vastuussa olemassa olevien, epävirallisten, datan hallintaa koskevien aloitteiden muuttamisesta. Datan hallintatiimi on vastuussa datan hallintaohjelman toiminnan määrittelystä, toimittamisesta ja ylläpidosta.

Sillä ei ole väliä, millä nimellä organisaatiossa näitä, tai muitakaan hallintamallin rooleja kutsutaan. Mikä on tärkeintä, kun luodaan datan hallintamallia, on se, että vastuut ja velvollisuudet määritellään ja kommunikoidaan selkeästi läpi koko organisaation. (Seiner, 2014)

Ydintiedon hallinnassa roolien ja vastuiden täytyy istua juuri kyseiseen organisaatioon ja sen kulttuuriin. Onnistunutta datan hallintamallia ei voi vain liimata olemassa olevan organisaation päälle, vaan se täytyy huolellisesti sovittaa istumaan siihen. Ja jotta datan hallintamalli sopii organisaatioon, sen muodostamisessa on tärkeä huomioida yrityksen tehtävä ja tavoitteet, työyhteisön normit ja uskomukset (Väre, 2019; Weber ym., 2009). Tavoite mallin luomisessa pitäisi kuitenkin olla aina se, että päätökset tehdään mahdollisimman lähellä niitä, joihin ne vaikuttavat. On myös muistettava, että mikäli asioiden päättämiseksi vaaditaan liian hidasta tai monimutkaista toimintatapaa, organisaatioon syntyy epäviralliset tavat päättää asioista. (Seiner, 2014)

Vastuiden jakaminen konkreettisesti eri roolien kesken voi joskus olla haastavaa. Tärkeää on kuitenkin varmistaa, ettei mikään ydintiedon hallinnan osa-alueista unohdu vastuista, mutta vastuussa ei myöskään saa olla useita eri henkilöitä. Selkeä vastuunjako helpottaa viestintää, työn suorittamista ja työn valvontaa huomattavasti (Väre, 2019). On järkevää, että organisaatiot erittelevät henkilöt, joilla on vastuu tiedon hallinnosta kullakin toimialueella. Kaikissa organisaatioissa tämä asema kuitenkin ei tarkoita, että kyseinen henkilö olisi kaikkietävä ylivalentias kyseisestä aihealueesta.

Tätä henkilöä pidetään kuitenkin organisaatiossa riittävän suuressa arvossa, jotta voidaan luottaa siihen, että hän hallinnoi dataa asianmukaisesti. (Seiner, 2014)

Yksi hyväksi havaittu tapa on ryhmitellä roolit ja tarvittavat tekemiset taulukon 2 esimerkin mukaiseen RACI -taulukkoon.

Taulukko 2. RACI -taulukko

Tehtävät	Rooli #1	Rooli #2	Rooli #3
Tehtävä #1	A	R	C
Tehtävä #2	A	R	C
Tehtävä #3	I	A	R

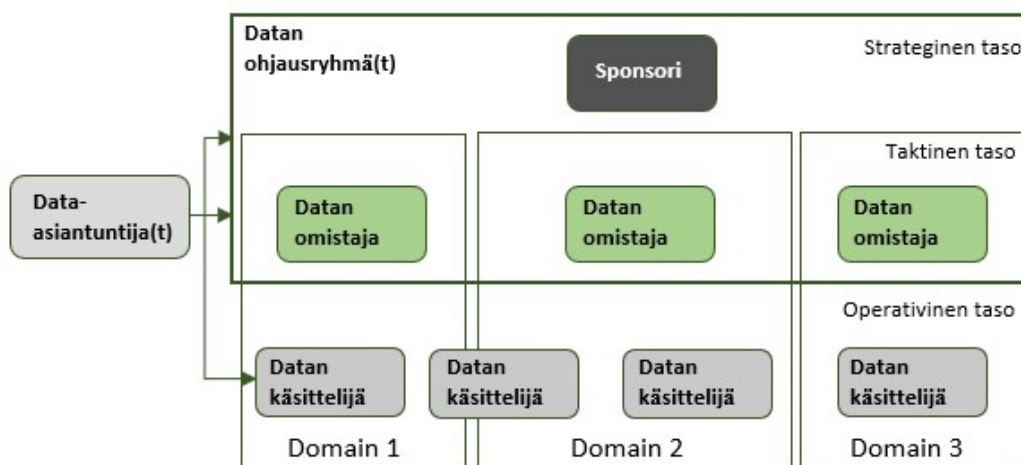
RACI -mallissa on neljä vastuun tasoa. A (*Accountable*), vastuussa oleva. Vastuussa olevalla on velvollisuus huolehtia siitä, että sovitut asiat tehdään kuten on sovittu ja sovitussa aikataulussa. R (*Responsible*), vastuullinen henkilö on velvollinen hoitamaan annettu työ ohjeistuksen mukaisesti. C (*Consulted*), neuvonantaja ei ole vastuussa kyseisen tehtävän aikaansaamisesta, mutta hänen asiantuntijuuttaan kuitenkin tarvitaan tehtävän tekemiseen. I (*Informed*), tiedotettava on henkilö, jolle pitää kertoa päätöksen tai tekemisen lopputulokset, jotta hän voi saadun tiedon avulla esim. suorittaa oman työnsä.

Cheong, L. ja Chang, V. (2007) tulivat omassa tutkimuksessaan tiedon hallinnan tarpeellisuudesta siihen lopputulokseen, että ydintiedon hallinta ei ole tehokasta ilman muodollista datan hallintamallia. Syynä tähän he mainitsivat selkeiden roolien ja vastuiden puutteen datan käyttäjien keskuudessa ympäristössä, jossa datan hallintamalli ei ole käytössä.

3.4.6 Datan hallintamallin tyypilliset tuotokset

- *Tiedon hallinnan toimintaperiaatteet*. Dokumentaatio kertoo tiedon hallinnan periaatteet (periaatteellisuus).

- *Operatiivinen malli*. Tuotos kattaa datan hallintamallin roolit ja vastuut. Operatiivinen malli suunnitellaan organisaatiolle sopivaksi (kuvio 10).
- *Data sanasto*. Tuotos käsittää kaikki oleellimmat liiketoiminnan käsitteet ja niiden määritelmät.
- *Tiedon (Data) periytyminen kaavio*. Kaavion tarkoitus on antaa helposti ymmärrettävä kuva siitä, miten data tulee organisaatioon ja miten se siirtyy organisaation sisällä prosessista toiseen.
- *Tiedon laatusäännöt*. Tärkeä tuotos, jolla kuvataan, millä säännöillä voidaan varmistaa tiedon hyvästä laadusta. Syntyy osana datastandardin muodostamista.
- *Tiedon laadun raportit*. Säännöistä voidaan laatia tiedon laadun mittarit, joita seurataan raporteilla. Raporttien merkitys on tehdä tiedon laatu näkyväksi sekä osoittaa laatuongelmien ratkaisumallin toimiminen käytännössä.
- *Tiedon laatuongelmien ratkaisumalli*. Tuotoksessa on kyse prosessista ja toimintamallista, miten raportoinnissa ja toiminnassa löydettävät tiedon laadun virheet ratkaistaan ja korjataan. (Keep it simple)
- *Riskienhallinnan suunnitelma*. Edistyneemmissä datan hallintamallissa voidaan osana tuotoksia tunnistaa dataan liittyvät riskit ja miten ne voidaan minimoida.



Kuvio 10. Tyypillinen datan hallintamallin rakenne (Väre, 2019)

Väre korostaa kirjassaan *Master Data*, että edellä luetelluista tuotoksista osa kehittyy pikkuhiljaa ajan myötä, mutta on muutama, jotka olisivat oleellisia heti hallintamallin käynnistyessä (Väre, 2019).

- Tiedon hallinnan toimintaperiaatteet ohjaamaan kaikkea toimintaa
- Operatiivinen malli ohjaamaan henkilöiden etsimistä rooleihin ja heidän perehdyttämistään vastuisiin.

Mahdollisimman pian olisi hyvä saada myös nämä tehtyä:

- Datanasto antamaan yhteisen kielen
- Tiedon laatusäännöt ja laadun raportit kertomaan, mihin tiedon hallinnassa pitäisi kiinnittää huomiota.
- Tiedon laatuongelmien ratkaisumalli

3.4.7 Datan hallinnan sisällyttäminen liiketoimintaprosesseihin

Tiedon hallinnan parhaat käytännöt muodostavat perustan ja ohjenuoran tiedon hallinnan toteuttamiselle. Datan hallintaa onnistuneesti toteuttavat organisaatiot aloittavat sen määrittelemällä pienen joukon parhaita käytäntöjä. Kun he ovat määritelleet parhaat käytännöt, he suorittavat kuiluanalyysin (gap-risk analysis) tunnistakseen erot (kuilu/gap) sen välillä, mitä he määrittelevät tiedonhallinnan parhaiksi käytännöiksi ja nykyisiin käytäntöihin, sekä näihin eroihin liittyviin nykyisiin ja potentiaalisiin riskeihin. (Seiner, 2014; Jonker ym., 2011)

3.4.8 Datan hallintamallin jalkautus

Datan hallintamallin uusien käytäntöjen, sääntöjen ja vastuiden hyödyt saavutetaan vain sillä, että kaikki noudattavat niitä. Tämä vaatii toimintatapojen jalkauttamisen käytäntöön. Hallintamallin jalkauttaminen oikeasti toimivaksi käytännöksi on ydintiedon hallinnan vaikeimpia asioita – onhan kyse organisaation toiminnan ja sen toimintakulttuurin muuttamisesta jollain tavalla.

Ensimmäinen askel kohti datan hallintamallia on määritellä sen rakenne, roolit, vastuut ja velvollisuudet. Tämä rakenne kertoo läpinäkyvästi kaikille datan käyttäjille, kuka vastaa mistäkin datan osa-alueesta (Cheong ym., 2007). Ja sovitettuna kulloisenkin yrityksen toimintakulttuuriin ja tavoitteisiin, se voidaan pitää juuri sopivan

yksinkertaisena. Datan hallintamallin tarkoitushan ei ole luoda lisää byrokratiaa, vaan tuoda datan vastuut ja velvollisuudet kaikille näkyväksi.

Taru Väre (2019) kertoo, että datan hallintamallin jalkauttamisessa on kolme kriittistä periaatetta, joista ei voi tinkiä.

Periaate 1: Jalkauta hallintamalli osissa. Hallintamallista suunnitellaan ennen sen jalkauttamista kaikki siihen käytettävät roolit ja niiden vastuut, eli hallintamallin operatiivinen malli. Operatiivisen mallin tarkoitus on kattaa lopulta kaikki eri domainit, joihin hallintamalli aiotaan ulottaa. Tämä ei kuitenkaan tarkoita, että operatiivinen malli olisi heti ensimmäisestä versiosta alkaen kiveen hakattu tuotos. Sen pitää olla elävä suunnitelma ja sallia joustavuutta roolien ja vastuiden sovittamisessa eri organisaation osille (Väre, 2019).

Yksi tärkeä syy, miksi hallintamalli pitää jalkauttaa osissa on se, että datan hallintamallin sisäistäminen vaatii kypsymistä ydintiedon hallintaan (MDM) yleisesti. Tämä vie aikaa. Hallintamallin jalkauttaminen kannattaa pilkkoa domaineittain ja rooleittain (Väre, 2019).

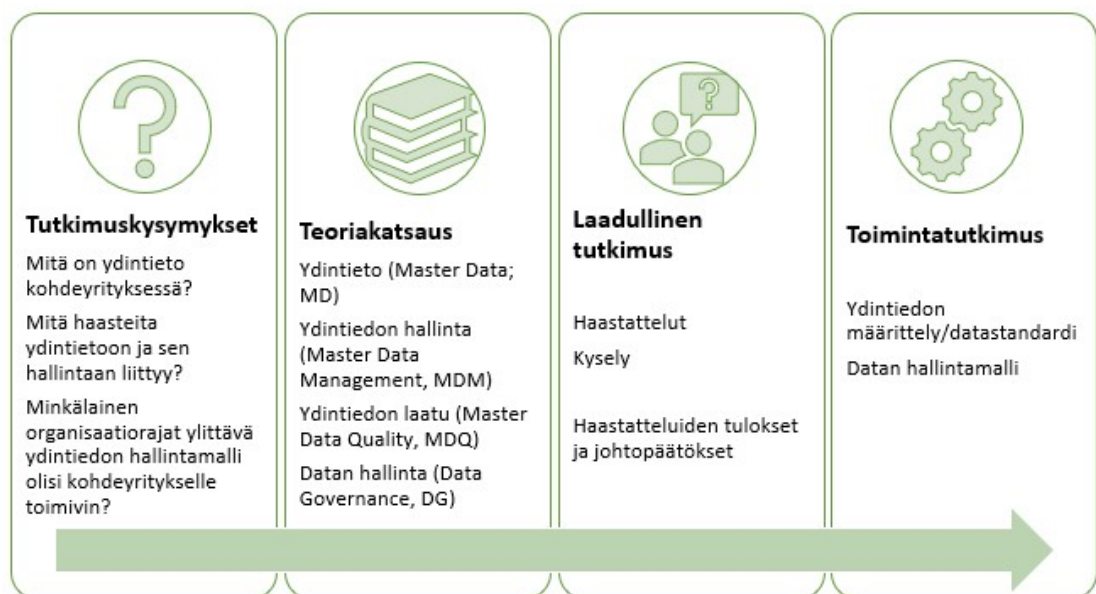
Periaate 2: Kommunikoij, kommunikoij, kommunikoij. Viestintää ei ole koskaan liikaa, vaan useimmiten kaikista ponnisteluista huolimatta sitä on liian vähän. Viestikää datan hallintamallin merkityksestä ja hyödyistä. Kommunikoikaa jalkauttamisen etenemisestä. Julkaiskaa tärkeimmät roolit nimityksinä siinä missä muutkin henkilöstön roolimutokset. Juhlikaa onnistumisia näkyvästi (Väre, 2019).

Periaate 3: Tue uusia rooleja niin kauan kuin he sitä tarvitsevat. Eri ihmiset sisäistävät uudet asiat eri tahdissa. Kun uudet roolit ja niiden sisältämät vastuut viedään ihmisille organisaatiossa, jotkut tarvitsevat tukea roolinsa oppimiseksi kauemmin kuin toiset (Väre, 2019).

4 Työn toteutus

Pystyäkseen vastaamaan asetettuihin tutkimuskysymyksiin opinnäytetyön tekijä tutustui ensin mahdollisimman laajasti ja monipuolisesti aiheesta olemassa olevaan kirjallisuuteen, artikkeleihin ja tutkimuksiin.

Kirjallisuuskatsauksessa opinnäytetyön tekijä pyrki käyttämään mahdollisimman tuoretta materiaalia, mutta myös aiheesta kirjoitetut vanhemmat ns. perusteoksetkin otettiin tarkasteluun. Oli huomattava, että aiheesta ei ole kirjoitettu montakaan suomenkielistä teosta, joten valtaosa opinnäytetyön teoriaosassa käytetyistä viittauksista on englanninkielisistä lähteistä. Mainitsemisen arvoinen suomenkielinen ja tuore teos aiheesta on Taru Väreen vuonna 2019 kirjoittama ”Master Data”, jossa on hyvin selkeästi kuvattuna ydintietoon ja sen hallintaan liittyvät peruskäsitteet.



Kuvio 11. Tutkimusprosessi

Seuraavassa vaiheessa opinnäytetyön tekijän oli selvitettävä mitä on ydintieto, sen haasteet ja ydintiedon hallinnan nykytila toimeksiantajayrityksessä. Tätä selvitettiin haastattelemalla kohdeyrityksen ydintiedon eri sidosryhmistä niitä henkilöitä, jotka ovat päivittäin tekemisessä ydintiedon kanssa ja jotka ovat siis ydintiedon hallinnan kannalta avainasemassa.

Haastatteluiden avulla opinnäytetyön tekijä pyrki selvittämään ydintiedon ymmärrettävyyden käsitteenä, eli mikä on haastateltujen mielestä yrityksen ydintietoa ja mitä haasteita ydintiedon nykytila asettaa toiminnalle. Lisäksi haastatelluilta kysyttiin kuinka heidän mielestään ydintiedon hallintaa tulisi kehittää, jotta nykyisiin ydintiedon ja sen hallinnan haasteisiin pystyttäisiin vastaamaan.

Haastatteluita tehtiin niin monelle henkilölle, että opinnäytetyön tekijällä oli se käsitys, että tutkittavaan ilmiöön ei tulisi oleellista uutta tietoa haastatteluja jatkamalla. Kohdeyrityksen tapauksessa tämä tarkoitti seitsemän henkilön (taulukko 3) haastatteleminen (20% yrityksen henkilöstöstä).

Opinnäytetyön tekijä kirjasi haastattelut propositiotasolla ylös haastatteluiden yhteydessä, koska keskusteluiden nauhoittaminen ei tuntunut luontevalta, eikä haastattelutilanteen sujuvuus tai tehtyjen muistiinpanojen luotettavuus vaarantunut tästä johtuen.

Taulukko 3. Tutkimuksen yhteydessä haastatellut henkilöt

Haastateltava	Ajankohta	Haastattelun kesto
Tuotantoinsinööri	25.7.2019	60 min
Tuotantopäällikkö	8.8.2019	60 min
Myyntipäällikkö	14.8.2019	45 min
Laatupäällikkö	14.8.2019	60 min
Hankintapäällikkö	15.8.2019	60 min
Toimitusjohtaja	21.8.2019	60 min
vt. Talouspäällikkö/Taloushallinnon asiantuntija	22.8.2019	60 min

Haastatelluille teetettiin myös kysely kohdeyrityksen liiketoimintatiedon hallinnan kypsyydestä (Capability Maturity Model), jolla opinnäytetyön tekijä halusi selvittää, mikä on yrityksen ydintiedon ja sen hallinnan taso arvioitaessa sitä asteikolla 1-5.

Lopuksi opinnäytetyön tekijä teki haastatteluista yhteenvedon, jota käytettiin toimintatutkimuksen alussa muodostamaan Six Sigma DMAIC-prosessin määrittelyvaiheen Voice of Business (VOB). VOB:n avulla pyrittiin toimintatutkimuksen edetessä löytämään mahdolliset syyt ydintiedon nykytilan ongelmille, sekä mahdollisia syitä edelleen tutkimalla, löytämään ongelmien juurisyy(t).

5 Tulokset






Keskeiset tutkimustulokset ja löydökset esitetään tässä kappaleessa. Tulosten avulla pyritään selvittämään mikä haastateltujen mielestä on ydintietoa kohdeyrityksessä, mitä haasteita ydintietoon ja sen hallintaan liittyy, sekä mikä heidän mielestään olisi paras tapa ydintiedon hallintaan jatkossa.

5.1 Ydintiedon nykytila

Tutkimus suoritettiin teemahaastatteluna niille ydintietoa käyttävien sidosryhmien edustajille, joiden katsottiin olevan eniten tekemisissä ydintiedon ja sen hallinnan kanssa. Tehtyjen haastatteluiden tuloksista kävi hyvin ilmi se, että kaikki haastatellut olivat vastauksissaan samoilla linjoilla siitä, mitä ydintieto (master data) yleisellä tasolla tarkoittaa. Eli ydintiedon koetaan olevan *”liiketoiminnalle tärkeää (lähes) muuttumatonta, päivittäin tarvittavaa tietoa, jonka avulla prosessit voivat toimia halutulla tavalla”* (taulukko 4).

Lisäksi ydintiedon koettiin olevan tietoa, jota käytetään liiketoiminnan eri toimintojen analysoinnissa ja raportoinnissa.

Taulukko 4. Ydintiedon määritelmä (tkA Finland Oy)








Mitä on ydintieto		%	Kommentteja
Ydintieto on liiketoiminnalle tärkeää (lähes) muuttumatonta, päivittäin tarvittavaa tietoa, jonka avulla prosessit voivat toimia halutulla tavalla.		100 %	esim. osto- ja myyntilausten nimikkeet, asiakas- ja toimittajatiedot
Ydintieto on perustietoa/tapahtumia, joita raportoidaan/analysoidaan.		29 %	
Dataa, jota tarvitaan pääprosessien suorittamiseen.		14 %	lähtökohta toiminnalle
Toisten toimintaan vaikuttavaa tietoa		14 %	
Päivitettyä ja ajantasaista tietoa		14 %	

Kun haastatelluilta kysyttiin mikä heidän mielestään on kohdeyrityksen ydintiedon ja sen hallinnan nykytila (taulukko 5), niin iso osa (43 %) vastaajista tunnisti ydintiedon laadussa olevan puutteita, eikä siihen siten voi nykyisellään luottaa. Myöskään esimerkiksi katelaskelmia ei voida tehdä luotettavasti nykyisen datan avulla. Tämän lisäksi monet toiminnanohjausjärjestelmästä otettavat raportit eivät sellaisenaan ole luotettavia, vaan niiden sisältöä joudutaan erikseen varmentamaan. Tähän kuluu paljon työaikaa ja datan luotettavuuden parantaminen nähtiinkin tärkeänä asiana ydintietojen hallintaa kehitettäessä.

Syinä nykyiseen datan epäluotettavuuteen nähtiin yhtenäisten, kaikkien sidosryhmien sisäistämien, yhteiseen näkemykseen pohjautuvien, sääntöjen puuttuminen ydintiedon luomisessa ja ylläpidossa, sekä olemassa olevien yleisten ydintiedon ominaisuuksien (mukaan lukien metatiedot) dokumentoimattomuus, eli datastandardin puuttuminen.

Lisäksi osa vastaajista koki, että ihmisillä ei oikein ole ymmärrystä siitä, kuinka mikäänkin data vaikuttaa toiminnanohjausjärjestelmän eri toimintoihin oman välittömän lähipiirin ulkopuolella. Osa vastaajista oli sitä mieltä, että tämä on johtanut siihen, että jokainen työntekijä tai osasto muodostaa omat määrityksensä ydintiedosta ja toteuttaa niitä sitten käytäntöön, sen sijaan että noudatettaisiin olemassa olevia ohjeistuksia. Tämä puolestaan tekee ydintiedosta sisällöllisesti epä johdonmukaista, sekavaa ja epäluotettavaa.

Taulukko 5. Ydintiedon nykytila (tkA Finland Oy)

Ydintiedon ja sen hallinnan nykytila		%	Kommentteja
Tietoa ei luoda/ylläpidetä samojen sääntöjen mukaan.		43 %	Puuttuu yhteinen näkemys datalle asetetuista vaatimuksista (määrittely ja säännöt). Yksilöt määrittelevät ydintiedon, kaikilla oma käsitys ydintiedosta.
Ydintiedon määrittelyt puuttuvat		43 %	
Tiedetään kyllä mitä kenenkään pitäisi tehdä, mutta kaikkia tarvittavia tietoja ei siltikään lisätä esim. nimikkeille, koska saattaa hetkellisesti hidastaa omaa tekemistä.		43 %	Välinpitämättömyyttä? Ei ymmärretä oman tekemisen vaikutusta muihin? Ei ymmärretä mitä dataa muut toiminnot tarvitsevat.
Ydintiedon hallinta yleisesti on parantunut aikaisemmasta, mutta edelleenkin puuttuu yhteinen näkemys tiedon laadun merkityksestä liiketoimintaan.		43 %	Datan avulla voidaan pyörittää liiketoimintaa, vaikkakin väliin hieman vaikeasti ja ylimääräisiä kustannuksia aiheuttaen.
Nykyinen data estää luotettavan raportoinnin.		43 %	KPI:t eivät ole luotettavia
Datan oikeellisuutta ei kunnioiteta.		29 %	Ei ole koulutettu oikeita tapoja.
Ydintiedon laatuongelmia ei tiedosteta ongelmiksi.		14 %	Datavirheiden korjaaminen osa jokapäiväistä tekemistä > Prosessi ei pääse muuten maaliin.
Nykyinen ydintiedon hallinta ei vastaa tulevaisuuden tarpeisiin/haasteisiin.		14 %	Nykyisellä linjalla ei ole mitään perusteita jatkaa.
Asiakkaiden, toimittajien ja nimikkeiden jne. kaikki tieto ei ole Oscarissa, vaan yksittäisten ihmisten takana (exceleissä tai päässä).		14 %	Ohjeistusten tekeminen on hankalaa, kun ohjeisiin ei voi laittaa yksiselitteisesti mistä oikean tiedon voi hakea (tieto kun ei aina ole Oscarissa).
Nykyiset käytännöt eivät tue hyvälaatuisen datan syntymistä/säilymistä.		14 %	Esim. nollahintainen myynti (hyvityksien yhteydessä). Katelaskenta (ei mahdollista ilman perustietoja, kuten työkustannuksia).
Nykyisellään data ei mahdollista yrityksen kasvua.		14 %	Mahdollinen lisävolyymi ja kasvu toisi mukanaan myös lisää virheellistä dataa ja sen korjaamiseen jouduttaisiin käyttämään entistä enemmän työaika, ja ehkä jopa palkkaamaan lisää henkilöitä.
Ydintiedon laadun seuranta ja hallinta puuttuu.		14 %	

Moni haastateltu nosti esille myös ehkä hieman huolestuttavasti sen seikan, että ydintiedon, tai ylipäätään minkään toiminnanohjausjärjestelmässä olevan datan oikeellisuutta ei tunnuta kunnioitettavan ja että ihmiset ovat välinpitämättömiä datan laadun suhteen.

Vaikka on yleisesti tiedossa ydintiedon ja sen hallinnan puutteet, niin kenelläkään ei ole ollut halua korjata tilannetta. Tämän arveltiin ainakin osin johtuvan siitä, että

nykyisen toimintaympäristön datalle asettamat vaatimukset ovat olleet ja ovat ehkä edelleenkin hieman epäselviä.

Haastatteluista kävi ilmi, paitsi että paljon on vielä työtä tehtävänä, jotta ydintieto ja sen hallinta saadaan paremmalle tasolle, niin myös se, että monen vastaajan mielestä ydintiedon laatu on kuitenkin viime vuosina parantunut aikaisempaan verrattuna. Mutta edelleenkin ydintietoa luovilta ja ylläpitäviltä henkilöiltä puuttuu yhteinen, koko organisaation tarpeet kattava näkemys ydintiedon ja sen laadun merkityksestä liiketoimintaan ja sen eri prosesseihin.

Ydintiedossa on nykyisellään paljon virheitä, joten niille on ikään kuin sokeuduttu. Tämä ilmenee siten, että päivittäin eteen tulevia pieniä, yksittäisiä datavirheitä ei huomioida virheiksi, vaan ne koetaan osaksi normaalia prosessia. Tällaisissa tapauksissa käy helposti niin, että datavirhettä ei korjata sinne mihin se pitäisi korjata, eli siihen tietueeseen ja datakenttään, jossa virheellinen ydintieto sijaitsee, vaan se korjataan ainoastaan transaktiodataan esim. myynti- tai ostotilaukseen. Tämä aiheuttaa saman virheen toistumisen uudestaan, kun viallista dataa seuraavan kerran käytetään.

Monet (43%) haastatelluista nostivat esille hyvälaatuisen ydintiedon vaikutuksen toimintojen nopeuttamisessa ja tehostamisessa (taulukko 6). Kun ydintieto olisi riittävän laadukasta, pystyisivät prosessit toimimaan ilman turhia katkoksia. Samoin liiketoiminnan ja sen prosessien kehittäminen onnistuisi aiempaa paremmin, kun olemassa olevaa dataa pystyttäisiin luotettavasti mittaamaan ja analysoimaan sekä analysoinnin perusteella tekemään faktaan (dataan) perustuvia ratkaisuja toiminnan kehittämiseksi.

Eräs ehkä hieman yllättävä vastaus kysyttäessä hyvän ydintiedon ja sen hallinnan hyödyistä oli, että hyvälaatuinen data parantaisi työmotivaatiota. Motivaation paraneminen sinällään olisi kyllä luonnollista, koska nykyisellään data tai oikeammin sen virheet ja puutteellisuus aiheuttavat päivittäin tilanteita, jossa työ pitää keskeyttää oikean tiedon etsimisen ja varmistamisen ajaksi. Tämä jos mikä aiheuttaa turhautumista ja työmotivaation alenemista.

Taulukko 6. Hyvälaatuisen ydintiedon edut (tkA Finland Oy)

Mitä hyötyä hyvällä ydintiedolla saavutetaan		%	Kommentteja
Parempi data nopeuttaisi/tehostaisi toimintaa.		43 %	Esim. raportoinnissa ei hukattaisi aikaa tietojen varmistamiseen, kun data olisi luotettavasti oikein.
Parempi data mahdollistaisi liiketoiminnan kehittämisen.		29 %	Mittarit > analysointi > kehitys
Lisää työmotivaatiota, kun olisi vähemmän käsin korjattavia asioita.		14 %	
Mahdollistaisi ”kerralla valmiiksi” toimintatavan.		14 %	Turhat tiedustelut vähenisivät.

5.2 Ydintiedon hallinnan haasteet

Kysyttäessä ydintiedon hallinnan haasteista, haastatellut kokivat, että jos ydintiedon määrittelyssä ei keskitytä olennaiseen, liiketoiminnan kannalta kriittiseen ydintietoon, vaan haalitaan ”ydintiedon” alle liikaa ei kriittistä dataa hallittavaksi, niin se teettää turhaa lisätyötä ja aiheuttaa byrokratian lisääntymistä.

Osa haastatelluista toikin esiin huolensa mahdollisesti lisääntyvästä byrokratiasta, jos ydintietojen ylläpito tehdään liian kankeaksi. Tämä voisi johtaa pahimmassa tapauksessa siihen, että ydintietojen ylläpito jäisi huonolle hoidolle ja näin ollen ydintieto olisi edelleen epäluotettavaa ja huonolaatuista.

Se, kuinka ydintietoihin valikoituisi optimaalisen minimaalinen joukko attribuutteja, onkin sitten eri asia. Monien haastateltujen vastauksista pystyi selvästi aistimaan, että varsinkin nimikerekisterin ydintietojen määrittely koetaan haasteelliseksi, koska eri osastoilla on niin valtavan erilainen kokemuspohjainen näkemys siitä, mitä dataa voidaan lukea sellaiseksi liiketoiminnalle kriittiseksi ydintiedoksi, jonka ylläpitämiseen kaikki voisivat sitoutua.

Osa haastateltavista koki haasteeksi myös sen, kuinka toimintojen siiloutumista pystyisi vähentämään tai estämään ja kuinka aiempi eri osastojen tekemä osaoptimointi olisi mahdollista korvata kokonaisvaltaisella ydintiedon kehittämisellä ja optimoinnilla.

Lisäksi valtaosa (71%) haastatelluista tunnisti isoimmaksi haasteeksi ydintiedon hallinnalle sen, kuinka uuden toimintamallin toiminnot pystytään ylläpitämään (taulukko 7). Kuinka saada ihmiset muuttamaan totuttuja toimintatapoja ja sitoutumaan ydintietojen kokonaisvaltaiseen hallintaan ja kehittämiseen ottamalla se mukaan jokapäiväiseen tekemiseen?

Taulukko 7. Ydintiedon ja sen hallinnan haasteet (tkA Finland Oy)

Ydintiedon ja sen hallinnan haasteita		%	Kommentteja
Toimintojen ylläpitäminen (datan hallintamallissa)		71 %	Sovittujen pelisääntöjen noudattaminen. Muutokseen sitoutuminen
Roolitus ja käyttöoikeuksien selkeä määrittely		43 %	Kuka tekee ja mitä? esim. raaka-ainanimikkeiden perustaminen
Periaatteen: "Kerralla kuntoon" ymmärtäminen		43 %	
Mahdollinen käyttöoikeuksien rajaaminen saattaisi parantaa datan laatua, mutta prosessi saattaisi hidastua.		29 %	Kuinka varmistetaan riittävä vasteaika, jos tekemistä rajoitetaan?
Datan hallinnasta ei tule tehdä liian byrokraattista (hidasta ja kankeaa).		29 %	Ei pakotettuja muutoksia, vaan annetaan kaikkien vaikuttaa määrittelyihin omalta osaltaan, muut prosessit huomioiden
Liiallinen tieto saattaa aiheuttaa turhaa työtä (jos ei osata määritellä mikä on oleellista ja liiketoiminnalle kriittistä tietoa).		29 %	Voi aiheuttaa johtamisen keskittymistä väärin asioihin. Työläs ylläpitää = ei ylläpidetä
Ydintiedon hallinnan kannalta on hankalaa, ettei tiedetä mitä tietoja järjestelmään pitää täyttää.		14 %	Pakkokentät (kriittiselle) ydintiedolle. Uuden tiedon luomiselle pitäisi olla tarkastus/hyväksyntäprosessi.
Kuinka tekeminen pidetään riittävällä tasolla?		14 %	Pakkokenttien käyttö liiketoiminnalle kriittiselle datalle. Riittävän (laatu)tason määrittely?
Johtaminen		14 %	Uskalluksen puute. Vastuun pakoilu
Tulevien datatarpeiden tunnistaminen		14 %	
Soveltuvien mittareiden rakentaminen ydintiedolle		14 %	
Kyky ymmärtää oikein tekemisen hyödyt		14 %	
Datan hallinnan tärkeyden ymmärtäminen ja sisäistäminen		14 %	

Edelleen valtaosan mielestä sovittujen pelisääntöjen noudattaminen on haaste, johon ydintietojen hallinta tulee törmäämään. Olisi ensiarvoisen tärkeää datan hallinnan ja ydintiedon laadun kannalta, että yrityksen kaikki työntekijät pystyisivät sitoutumaan sovittuihin sääntöihin. Täytyy nimittäin muistaa, että kohdeyrityksessä on toimittu yrityksen perustamisesta lähtien ilman ydintietojen hallintaa, joten mahdollinen datan hallinnan toimintamallin jalkauttaminen tulee olemaan erittäin suuri or-

ganisatorinen haaste, jonka voittamisessa yritysjohdon riittävä tuki ja esimerkin näyttäminen uusia toimintamalleja käyttöönotettaessa on avainasemassa.

Tässä yhteydessä ydintietojen hallinnan haasteeksi mainittiin myös tulevien (ydintieto) tarpeiden tunnistaminen ja esiin tuominen, eli kuinka saada määriteltyä ydintiedot ja niiden hallinta sellaiseksi, että se pystyy reagoimaan liiketoiminnan ja yrityksen toimintaympäristön muuttuviin vaatimuksiin ja tarpeisiin siten, ettei liiallinen määrittely tee ydintietojen ylläpidosta ja hallinnasta liian hidasta tai byrokraattista.



Haasteeksi koettiin myös se, kuinka datan hallinta ja ydintiedon määrittelyt saadaan jalkautettua käytännön tekemiseen. Ihmisille pitäisi tuoda hyvin selkeästi esille oikein tekemisen hyödyt ja että ydintiedon hallinta ei ole minkään yksittäisen osaston osaa-optimointia, vaan osa liiketoiminnan perusedellytysten kehittämistä ja tulevaisuuden haasteisiin varautumista.

5.3 Ydintiedon hallinta jatkossa

Haastateltujen mielestä ydintietojen hallintaan tulee löytää oikea roolitus ja vastuut, koska ainoastaan jos vastuut datan suhteen on selkeästi ja läpinäkyvästi kerrottu, on ydintietojen hallinnalla onnistumisen mahdollisuus. Tämä koettiin myös haasteelliseksi toteuttaa siten, että kaikki datan hallinnan sidosryhmät voisivat niihin sitoutua.

Osa haastatelluista oli sitä mieltä, että aiempaa tarkempi ja tiukempi roolitus on tarpeen, minkä lisäksi osa haastatelluista oli myös sitä mieltä, että jatkossa ydintietojen hallinnalle tulisi määritellä paljon nykyistä tiukemmat säännöt. Ja nykyinen malli, jossa ”kaikki” saavat esim. luoda uusia nimikkeitä ei toimi jatkossa. Toisaalta iso osa haastatelluista (57%) oli myös sitä mieltä, että lisäkoulutuksen avulla nykyinenkin toimintatapa saattaisi olla ihan hyvä vaihtoehto (taulukko 8). Tätä näkemystä perusteltiin mm. liiallisen byrokratian ehkäisemisellä.

Taulukko 8. Ydintiedon hallinta tulevaisuudessa (tkA Finland Oy)

Ydintieto ja sen hallinta jatkossa		%	Kommentteja
Ihmiset tulee vastuuttaa datan hallinnan rooleihin.		57 %	On määriteltävä, kuka tekee, mitä tekee ja miten tekee.
Pitää lisätä koulutusta, jotta voidaan varmistaa datan laadun riittävä taso.		57 %	On ymmärrettävä, miten omat toimet vaikuttavat muiden tekemiseen. Datan ylläpitäjillä pitää olla ymmärrys tehtyjen ylläpitotoimien vaikutuksesta kaikkiin toimintoihin.
Tulisi tehdä listaus (määrittelyt) tiedoista, jotka ovat ydintietoa.		29 %	Kaikilla olisi yhtenäinen näkemys asiasta
Ainoastaan liiketoiminnan tarvitsema tieto pakollisena		29 %	Pakkokenttien käyttö ylläpidettävälle tiedolle
Mittarointi > virheet esiin		29 %	Datavirheillekin tulisi laittaa hintalappu.
Kaikkien tulisi noudattaa sovittuja pelisääntöjä.		14 %	
Datan hallinnassa laatu ennen tehokkuutta		14 %	
ERP:n pitäisi tukea paremmin ydintiedon ylläpitoa.		14 %	Asiakkaan erityisvaatimukset, jotka poikkeavat normista → linkitys asiakastiedoista.
Tiedottamista dataan liittyvissä asioissa lisättävä		14 %	

Ydintiedon ja muunkin toiminnanohjausjärjestelmässä olevan datan ylläpitäjille tulisi koulutuksen avulla tarjota siten aiempaa parempi ymmärrys koko organisaation ydintietotarpeista ja ydintietoihin tehtyjen toimenpiteiden vaikutuksesta omaan ja muiden toimintaan. Kaikkien datan käyttäjien ja ylläpitäjien tulee ymmärtää, että toiminnanohjausjärjestelmän data ei ole olemassa ainoastaan heidän, vaan koko organisaation ja myös koko konsernin tarpeita varten.

Ydintiedon ylläpitäminen nykyisessä toiminnanohjausjärjestelmässä koettiin hieman haasteelliseksi, koska nykyisellään järjestelmä ei tue kovin hyvin esim. pakkokenttien käyttöä. Ja jatkossa tulisikin, yhdessä järjestelmätoimittajan kanssa, miettiä niitä keinoja, joilla toiminnanohjausjärjestelmän toiminnallisuuksia kehittämällä ihmisten tekemien virheiden määrää pystyisi pienentämään.

Vastauksista kävi ilmi myös, että jatkossa ydintiedolle tulisi luoda sellaisia toiminnan kannalta merkityksellisiä ydintiedon laadun mittareita, joiden avulla kaikille olisi selkeästi näkyvissä mikä on yrityksen liiketoimintakriittisen ydintiedon tila, jotta siten voitaisiin tarvittaessa tehdä korjausliikkeitä ko. datan laadun parantamiseksi. Mitta-

reiden tulisi kertoa myös se, miksi jokin ydintiedon laadullinen ominaisuus on tärkeä, sekä mitä kyseisen ydintiedon laadullinen parantaminen merkitsee yritykselle ja sen eri toiminnoille. Tästä esimerkkinä mainittiin mm. nimikkeiden CN-koodi (tullinimike). Jos se on väärin, saattaa siitä aiheutua yritykselle ylimääräisiä tullauskustannuksia maahantuonnin yhteydessä. Tai jos asiakkaan verkkolaskutunnus on väärin, niin asiakkaille ei lähde verkkolaskut ajallaan ja näin yrityksen kassavirta heikkenee.

Mittareiden luotettava käyttö edellyttää kuitenkin, että ydintiedot on tarkoin määritelty ja että määrittelyt on kirjattu yhteiseen kaikkien saatavilla olevaan datasanastoon, josta kaikki voivat tarvittaessa käydä varmistamassa miten ja missä muodossa ydintieto pitää järjestelmään syöttää ja päivittää.

5.4 Ydintiedon hallinnan kypsyys

Haastatelluille teetettyyn kyselyyn kohdeyrityksen liiketoimintatiedon hallinnan kypseydestä (Capability Maturity Model) vastasi 43%, eli kolme seitsemästä haastatellusta. Vaikka vastausprosentti jäi melko alhaiseksi, olivat kyselyyn vastanneiden tulokset keskenään hyvin samankaltaiset ja tukevat myös hyvin sitä yleistä käsitystä, joka opinnäytetyön tekijällä on kohdeyrityksen ydintietojen hallinnan tilanteesta. Kohdeyritys on monessa kohdassa vasta alkutaipaleella, kun puhutaan yritystiedon hallinnasta. Kyselyn kaikkien kahdeksan osion tulosten keskiarvo (1,6) kuvaa hyvin sitä toiminnan tasoa, jossa tällä hetkellä ollaan. Kyselytulokset on nähtävissä kokonaisuudessaan liitteessä 1.

Tähän opinnäytetyön tulososioon on otettu mukaan kyselystä osiot 7 ja 8 (taulukko 9), jotka käsittelevät juuri ydintietojen hallintaa (Master Data Management, MDM) ja datan hallintaa (Data Governance, DG).

Taulukko 9. Osiot 7 ja 8 liiketoimintatiedon hallinnan kypsyyskyselystä

Data Governance Disciplines	Score (1 to 5)	Importance to business (1 to 5)
-----------------------------	-------------------	---------------------------------------

<p>7. Master Data Management/Ydintietojen hallinta:</p> <p>Organisaatiomme tiedostaa, että ydintietojen hallinta (MDM) on yksi tehokkaimmista ja tärkeimmistä datan aloista, joista nykyään puhutaan. Olemme tunnistanee ihmiset hallitsemaan MDM-aloitettamme ja olemme alkaneet myös tunnistamaan tekniikoita, jotka auttavat meitä hallitsemaan ja jakamaan ydin- ja referenssitietojamme. Kun täydennämme MDM-ympäristöämme, päätöksentekoa hallitaan kurinalaisesti isäntä- ja resurssijärjestelmän, metatietokomponentin sekä ydintietojen viestinnän ja saatavuuden suhteen. Olemme hyvässä asemassa ydintietoaloitteen saattamiseksi päätökseen budjetissa ja aikataulussa.</p>	2	5
<p>8. Data Governance & Data Stewardship/Datan hallinta:</p> <p>Meillä on datan hallintaohjelma, joka määrittelee selkeästi roolit ja vastuut operatiivisella, taktisella, strategisella ja tukitasolla. Ohjelmamme keskittyy hyödyntämään tietovastaavien nykyistä tietämystä datasta. Johtajamme, tietovastaavat, liiketoiminta- ja teknologiahenkilöt ovat omaksuneet valintamme lähestymistapaan ja siinä käsitellään datan hallintaa ennakkoivasti ja reagoivasti. Tietojenhallintaohjelmamme on ensisijainen tekijä menestyksessämme kaikilla tässä testissä luetelluilla aloilla.</p>	1,3	4

Ydintiedon hallinnan (MDM) koettiin kyselyn tulosten perusteella olevan hieman paremmalla tasolla kuin datan hallinnan (DG), eli ydintiedon hallinnan kypsyyksi oli *toistettavalla tasolla*, jossa jotkut prosessit ovat toistettavissa, mutta prosessien kurinalaisuus ei todennäköisesti ole tiukka. Tason 2 organisaatiot ovat alkamassa perustamaan tiedon hallinnan käytäntöjä, jotka keskittyvät tietoihin, joita käytetään liiketoimintayksiköiden raportointiin (Mosley, M. 2010).

Datan hallinnan (DG) taso kohdeyrityksessä on kyselyyn vastanneiden mielestä tasolla 1, jota kuvataan termillä tiedonhallinnan kypsyyden *alkutaso*, tällä tasolla tapahtuvat datan hallinnan prosessit ovat tyypillisesti dokumentoimattomia ja dynaamisen muutoksen tilassa. Tällaisia prosesseja ohjataan yleensä käyttäjien tai tapahtumien tapauskohtaisella, hallitsemattomalla ja reaktiivisella tavalla. Tason 1 organisaatioilta puuttuu tiukat säännöt tai menettelytavat tietojen hallinnasta. Raportit kehitetään lennossa liiketoimintayksiköiden pyynnöstä, eikä datan omistajuutta ole määritelty (Mosley, M. 2010).

6 Johtopäätökset

Nykytila-analyysistä on tehtävissä useita johtopäätöksiä sen suhteen, mitkä seikat ovat olleet vaikuttamassa kohdeyrityksen ydintiedon ja sen hallinnan kehitykseen. Analyysistä kävi hyvin selville, että yksi suurimmista syistä, miksi ydintiedon nykytila on monien mielestä huonolla tasolla, on liiketoimintakriittisen ydintiedon määrittämättömyys sekä datan omistajuuden puuttuminen. Tämä johtuu ainakin osin siitä, että yritys on kasvanut ydintiedon hallinnan näkökulmasta hallitsemattomasti, eikä ydintietojen määrittely ja hallinta ole kuulunut kenenkään vastuualueelle. Tämä puolestaan johtuu monien haastateltujen mielestä siitä, että dataa ei aiemmin ole koettu kovin tärkeäksi osaksi jokapäiväistä tekemistä ja eteen tulleet virheet datassa on korjattu pikaisesti, tai jopa jätetty korjaamatta, ilman sen suurempaa miettimistä tai virheiden syihin puuttumista.

Yritys on, kuten eräs haastateltavakin mainitsi ”pystynyt toimimaan puutteellisellakin datalla, joskaan ei aina kovin tehokkaasti”. Mikä sitten on ollut oikeasti esteenä sille, ettei datan hallintaan tai sen laatuun ole kukaan aikaisemmin puuttunut. Ehkä juuri se, ettei datan hieman puutteellinen laatu ole ollut esteenä kohdeyrityksen normaalille liiketoiminnalle ja että vasta viime vuosina, kun digitalisaation mukanaan tuomat vaatimukset ydintiedolle ovat kasvaneet, on myös ydintiedossa olevat ongelmat tiedostettu paremmin.

Nykytila-analyysin pohjalta kohdeyrityksessä käynnistettiin Six Sigma DMAIC-kehitysprojekti, joka koostuu viidestä osasta. Ensimmäisenä ongelma tulee määrittellä (*Define*). Määrittelyyn liittyy myös kehitystavoitteiden asetanta, projektin tavoitteiden rajaaminen sekä resurssien määrittely. Toisessa vaiheessa ongelman laajuus ja vakavuus tulee mitata (*Measure*). Ongelman mittaamisen jälkeen analysoidaan (*Analyze*) ongelman juurisyitä ja mietitään ongelmaan sen poistamiseksi parhaiten soveltuvia ratkaisuja (*Improve*). Viimeisenä vaiheena prosessissa on varmistaa (*Control*), että tehdyt toimenpiteet ovat olleet oikeita ja ongelman ratkaisun kannalta merkityksellisiä, sekä varmistaa, että tehdyt toimenpiteet vakiintuvat käytännön tekemiseen.

Tämän kehitysprojektin kohdalla DMAIC-prosessia jouduttiin hieman soveltamaan, koska sen kohteena oli yrityksessä aiemmin tuntematon tekijä, eli ydintieto ja sen laatu. Siitä syystä, ennen ongelman syvällisempää analysointia, oli ydintieto määriteltävä, jotta saataisiin oikeasti selvitettyä ja mitattua, kuinka kaukana ydintiedon laatu on nykyisen toimintaympäristön sille asettamista vaatimuksista. Opinnäytetyön tekijälle tulikin toimintatutkimuksen ensivaiheessa tehtäväksi, yhdessä ydintiedon sidosryhmien kanssa, määritellä liiketoiminnan vaatimukset ydintiedolle, dokumentoida ne, sekä mitata ydintiedon laatutaso.

Näin ollen DMAIC-prosessi muuttui muotoon DIMAIC. Projektin määrittelyn jälkeen tuli siis ensin kehittää (Improve) toimintaa, eli määritellä ydintieto ja sen jälkeen jatkaa normaalin DMAIC-prosessin mukaisesti mittaamalla ydintiedon taso.

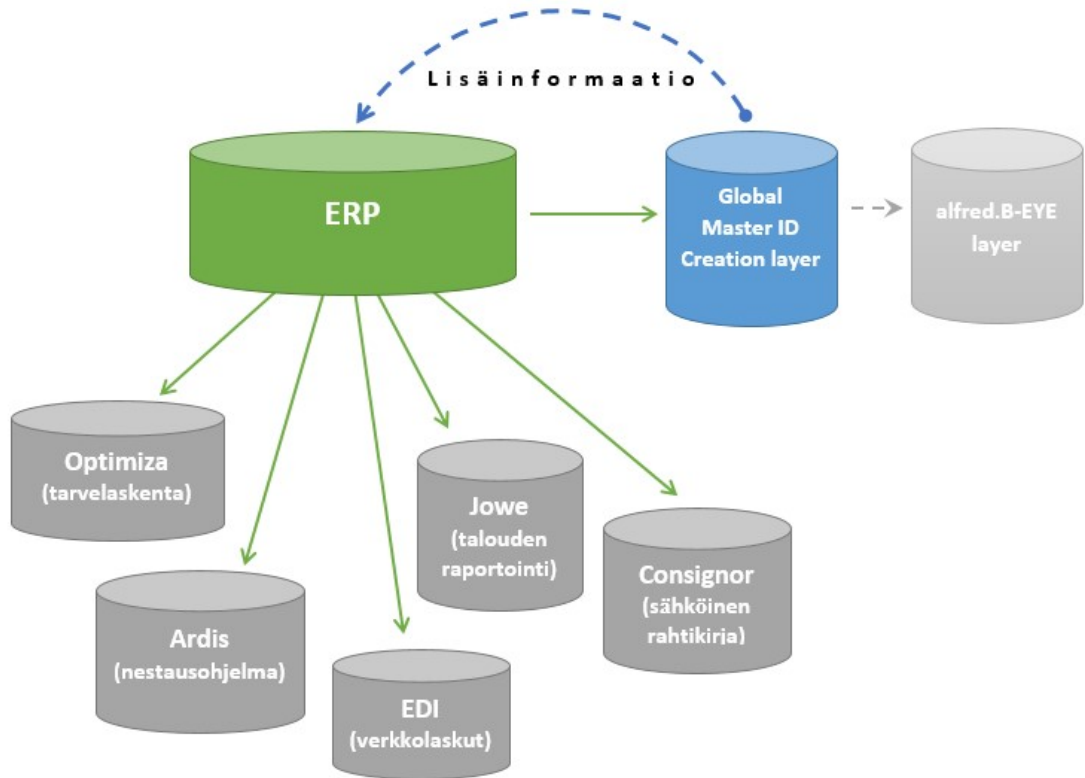
Six Sigma -prosessin ensimmäisessä, eli määrittelyvaiheessa sovittiin, että ydintietojen määrittely ja ydintiedon hallinnan kehittäminen rajoitetaan (alkuvaiheessa) kolmeen keskeiseen datarekisteriin (domainiin), eli nimike-, asiakas- ja toimittajarekistereihin. Tämä siitä syystä, että kaikkien datarekisterien ydintietojen samanaikainen määrittely ja hallintaan ottaminen oli liian iso urakka tehtäväksi yhdellä kertaa. Myöskään opinnäytetyötä ajatellen ei olisi ajallisesti ollut mahdollista ottaa enempää rekistereitä työn alle. Tarkoituksena on kuitenkin hyödyntää näiden kolmen datarekisterin määrittelyprosessin oppeja jatkossa myös muiden esim. tuotannon datarekisterin ydintietojen määrittelyyn ja hallintaan.

Datan hallintamallin sisäistäminen uutena asiana vaatii kypsymistä ydintiedon hallintaan yleisesti ja tämä vaatii aikaa (Väre, 2019) .

6.1 Liiketoiminnan vaatimukset ydintiedolle

Kohdeyrityksen järjestelmäkartta on melko yksinkertainen (3.2.3 Ydintiedon hallinnan arkkitehtuuri, 25), eli käytössä on vain yksi keskusjärjestelmä, tässä tapauksessa toiminnanohjausjärjestelmä (ERP), jossa hallitaan keskitetysti (Centralized) ydintiedon luomista, ylläpitoa ja poistamista. Ydintietoa siirretään muihin järjestelmiin eritasoisilla integraatioilla ja sisään luku -ajoilla. Näissä muissa järjestelmissä ydintie-

toon voidaan lisätä dataa, sitä voidaan muokata ko. järjestelmän sisällä, mutta sitä ei voida syöttää takaisin tai poistaa toiminnanohjausjärjestelmästä (järjestelmäkartta kuvattuna kuviossa 12).



Kuvio 12. Järjestelmäkartta (tkA Finland Oy)

Yksinkertaisesta järjestelmäarkkitehtuurista johtuen ydintietojen määrittelyprosessi oli melko suoraviivaista, eikä määrittelyitä tehtäessä tarvinnut huomioida toiminnanohjausjärjestelmään sen ulkopuolelta tulevaa ydintietoa. Tähän ehkä pienenä poikkeuksena ja mikä ydintietojen määrittelyvaiheessa tuli huomioida oli thys-senkrupp Aerospace -liiketoiminnassa samanaikaisesti käynnissä ollut kehitysprojekti (alfred. B-EYE), jonka tarkoituksena on parantaa läpinäkyvyyttä ja data analytiikan edellytyksiä liiketoiminnan kehittämisen tarpeisiin. Halutun läpinäkyvyyden saavuttamiseksi kaikkien konserniyritysten käytössä olevien eri toiminnanohjausjärjestelmien välille luodaan mm. nimikkeille yhteinen Global Material Master ID, joka tuo myös kohdeyrityksen ydintietojen määrittelyyn omia vaatimuksia mm. pakollisten kentätietojen (attribuuttien) osalta. Lisäksi kohdeyrityksenkin kaavailemassa tieto-

kannan vanhentuneiden sekä tuplatietueiden arkistoinnissa/poistamisessa tuli ottaa huomioon konsernin näille asettamat vaatimukset.

6.1.1 Datastandardien määrittely

Liiketoiminnan ydintiedolle asettamia vaatimuksia lähdettiin selvittämään datarekisteri (domain) kerrallaan kunkin sidosryhmän kanssa käydyissä työpajoissa siten, että kustakin organisatorisesta liiketoiminta-alueesta oli paikalla riittävästi ihmisiä tuomassa oman näkemyksensä siihen, mikä on juuri heidän osa-alueelleen kriittistä ydintietoa kyseiselle rekisterille, ja jota ilman päivittäinen toiminta olisi vaikeaa tai jopa mahdotonta.

Osana ydintietojen määrittelyä kohdeyrityksessä muodostettiin toiminnanohjausjärjestelmän rekistereistä (nimike-, asiakas- ja toimittajarekisterit) datastandardit, joihin kirjattiin mm. eri rekistereissä aktiivisessa käytössä olevat ominaisuudet (attribuutit), niiden metatiedot, pakollisuustieto sekä mahdollisten liittyvien sovellusten kenttävastineet.

Datastandardin sisältö muodostuu yksityiskohdiltaan organisaation tarpeen mukaan, mutta siinä tulisi olla mukana vähintään seuraavat tiedot (Väre, 2019):

- Kaikki ydintiedoksi määritellyt entiteetit ja niiden attribuutit
- Jokaiselle entiteetille ja attribuutille yhteisesti sovittu nimi sekä määritelmä
- Mahdolliset synonyymit eri käsitteille (esim. miten kyseistä dataa kutsutaan jossain toisessa järjestelmässä)
- Tiedon tyyppi, eli onko data vapaasti kirjoitettavaa tekstiä, numero vai ennalta määrättyltä listalta valittava arvo.
- Tiedon muoto, mikäli se on tärkeä tieto (sähköposti @ jne.).
- Pakollisuustieto. Onko tieto aina pakollinen, vai pakollinen vain tietyn tyyppiselle datalle?
- Liiketoimintaprosessit, joissa kyseistä dataa syntyy, päivitetään, käytetään tai poistetaan. (CRUD)

6.1.2 Nimikerekisteri

Nimikerekisteri on rekisteri, jota käyttää poikkeuksetta kaikki kohdeyrityksen liiketoiminta-alueet ja siksi sen ydintietojen hallinta on erityisen tärkeässä asemassa koko yrityksen toimintaa silmällä pitäen. Nimikerekisterin nimikkeistö koostuu useista erityyppisistä nimikkeistä, mutta pääasiassa kuitenkin raaka-ainenimikkeistä ja valmistetavista nimikkeistä. Näille molemmalle nimiketyypille on olemassa omat sisällölliset vaatimukset, joita niille asettaa niin yrityksen liiketoimintaympäristö, toiminnanohjausjärjestelmä kuin myös käytettävät ulkoiset sovellukset.

Nimikerekisteri oli lopputyön tekijän toiminnanohjausjärjestelmästä tekemien havaintojen sekä haastateltujen henkilöiden antamien lausuntojen perusteella ydintiedon hallinnan näkökulmasta katsottuna kaikkein huonoimmassa kunnossa nyt tarkastelun alle olleista datarekistereistä.

Tähän on johtanut mm se, että kohdeyrityksen toimintatavat ydintiedon ja datan hallinnan suhteen ovat peräisin ajalta, jolloin digitalisaatio ei ollut vielä muuttanut toimintaympäristöä ja jolloin lähes kaikki työntekijät saivat toimia vapaasti toiminnanohjausjärjestelmässä, uusia tietueita luoden ja päivittäen niitä ilman kunnollista ohjeistusta. Lisäksi toiminnanohjausjärjestelmän perehdytys on jäänyt jokaisen omalle vastuulle, jolloin mm. erilaisten kenttäliitännäisyyksien vaikutusten ymmärtäminen eri toimintojen välillä ei todennäköisesti ole onnistunut ja siten myös ymmärrys oman tekemisen vaikutuksesta muiden tekemiseen on vajavaista.

Lisäksi käytössä oleva toiminnanohjausjärjestelmä sallii käyttäjille valtavasti vapauksia nimikkeiden luomisessa. Käyttäjä voi esim. jättää täyttämättä pakolliseksikin määriteltyjä kenttätietoja, tai kirjoittaa kyseisiin kenttiin vapaata tekstiä. Tällöin on ollut käytännössä mahdotonta pelkästään ohjeistuksien avulla varmistaa syötettävien tietojen oikeellisuus ja samankaltaisuus.

Nimikkeitä on usein perustettu kopioimalla vanhasta lähes samanlaisesta nimikkeestä, mutta kuitenkin kopioitaessa ei ole aina huomattu tai muistettu tarkistaa uudel-

le nimikkeelle periytyviä tietoja. Näin ollen uudelle nimikkeelle on saattanut kopioitua turhaa ja virheellistä tietoa, joka saattaa haitata muiden käyttäjien toimintaa sekä aiheuttaa esimerkiksi laatukustannuksia myöhemmässä vaiheessa.

Usein nimikkeitä perustettaessa tai ylläpidettäessä on käynyt myös niin, että se, joka ei tarvitse esim. hankintatoimen nimikkeelle määrittelemiä tarvelaskentaan vaikuttavia ominaisuuksia ts. parametrejä (toimitusaika, minimi tilausmäärä jne.), ei niitä myöskään nimikkeille laita, jolloin nimikkeen tiedot jäävät siltä osin vajavaisiksi. Osin tämä johtuu siitä, että aina ei tiedetä mitä kenttiä tai miten ko. kenttiä pitäisi täyttää, mutta isolta osin tämä johtuu myös siitä, ettei ko. parametritieto kosketa nimikettä ylläpitävää tai perustavaa henkilöä tai toimintoa.

Nimiketietojen laadun parantamisen kannalta on ehdottoman tärkeää luetteloida kaikkien tarvitsemat ydintiedon ominaisuudet yhteiseen kaikkien saatavilla olevaan dokumenttiin, josta ne voidaan käydä tarvittaessa tarkistamassa ja jonka avulla henkilöstöä voidaan kouluttaa oikeiden tietojen ylläpitämiseen toiminnanohjausjärjestelmän eri rekistereissä. Rekistereissä ja tietueissa oleva muu, ei pakollinen ydintieto kirjattiin mahdollisesti pakolliseksi ydintiedoksi, jos se oli sellaista ydintietoa, jota toiset toiminnot tarvitsevat pystyäkseen toimimaan. Nämä tiedot edustavat eri liiketoimintayksiköiden omaa ydintietoa (Väre, 2019).

Alla olevaan taulukkoon (taulukko 10) on koottu osa nimikerekisterin ydintiedoiksi määritellyistä attribuuteista. Kaikki ydintiedoksi määritellyt tiedot löytyvät liitteestä 2.

Taulukko 10. Esimerkkejä nimikerekisterin ydintiedoista

Kenttä	Tyyppi	Määritelmä	Esimerkki
Tunnus	VARCHAR2(20 CHAR)	Nimiketunnus	PL508302000A
Nimi1	VARCHAR2(80 CHAR)	Nimikkeen kuvausteksti	PLATE 20 mm 5083 H111
Ryhmä	VARCHAR2(16 CHAR)	Materiaaliryhmä	AL, TE
Alaryhmä	VARCHAR2(8 CHAR)	Materiaalin seostieto	5083, 7075
Nimikelaji	VARCHAR2(1 CHAR)	Materiaalin käyttöluokittelu	R, V, P
Tilastoryhmä	VARCHAR2(10 CHAR)	Tilastoryhmä (raportointi)	PL, FB, RB
Standardi	VARCHAR2(30 CHAR)	Materiaalstandardi jonka mukaisena nimikettä hankitaan ja myydään	AMS4085

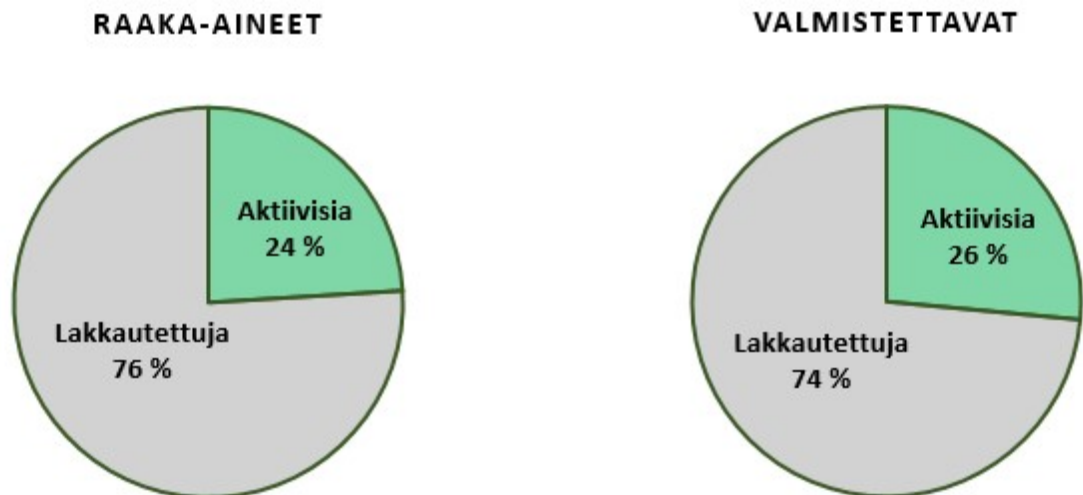
Tässä yhteydessä on huomattava, että esim. nimiketunnuksen kohdalla laatukriteeriksi nykyisessä järjestelmäympäristössä riittää, että se on olemassa. Mikäli kohdeyrityksessä haluttaisiin, että kaikki nimiketunnukset olisivat yhtenäisen kaavan mukaisesti muodostettuja, tarkoittaisi se sitä, että nykyinen nimikkeistö pitäisi uusia ja siirtää esim. juoksevaan numerointiin nimiketunnuksissa. Tämä puolestaan ei tällä hetkellä ole toteutettavissa siitä saatavaan hyötyyn nähden kohtuullisin kustannuksin, koska se tarkoittaisi mm. kaikkien varastossa olevien materiaalien uudelleen tarroittamista (materiaaliin kiinnitettävän erätarran päivitys). Tällä hetkellä materiaali-eriä, joihin nimikemuutos pitäisi päivittää on varastossa n. 6000 kpl.

Käytössä olevat nimiketunnukset on perustettu vuosien saatossa ilman kunnon ohjeistusta ja valvontaa. Lisäksi iso osa (72%) nimiketunnuksista on periytynyt kohdeyrityksen pääasiakkaan käyttämistä tunnuksista ja siten niiden harmonisointi vastamaan yrityksen omaa ohjeistusta tilanteessa, jossa toiminnanohjausjärjestelmää ei olla vaihtamassa, ei ole järkevää.

Nimikerekisterin määrittelyn jälkeen nimikkeistö siivottiin tehtyjen määritysten mukaisesti, kuitenkin siten, ettei siivouksen ja harmonisoinnin laajuus aiheuttaisi liiallisia kustannuksia saatavaan hyötyyn nähden. Tämä tarkoitti käytännössä sitä, että siivoustoimenpiteet kohdistuivat ensisijaisesti niihin aktiivisiin raaka-ainenimikkeisiin, joille oli avoimia osto- ja myyntitilauksia, tuotannon työmääräimiä tai saldoa varastokirjanpidossa. Näiden aktiivisten nimikkeiden tiedettiin myös jäävän tuotantotietokantaan (aktiivinen käytössä oleva tietokanta) myös suunnitellun datan arkistoinnin/poistamisen jälkeenkin.

Raaka-ainenimikkeiden lisäksi tietokannassa on hyvin paljon valmistettavia nimikkeitä, joista kuitenkin valtaosa perustetaan pitkälle automatisoitua nimikkeen perustamisprosessia hyödyntäen, jolloin niiden nimikerakenne on identtinen ja vastaa hyvin niille asetettuja laadullisia vaatimuksia. Näidenkin nimikkeiden arkistoinnissa tullaan käyttämään samoja periaatteita kuin raaka-ainenimikkeiden kohdalla, eli aktiiviseen tietokantaan ei jätetä vanhentuneita tietueita.

Yksi suunnitellun datan arkistoinnin tarkoitus on nopeuttaa tietokannan hakutoimintoja siirtämällä ei-aktiivista datamassaa erilliseen arkistotietokantaan. Tarkasteltaessa aktiivisten nimikkeiden osuutta kokonaisuudesta sekä raaka-aine-, että valmistettävien nimikkeiden osalta, voidaan jälleen todeta 80/20 -säännön paikkansapitävyys (kuvio 13).



Kuvio 13. Aktiiviset vs lakkautetut nimikkeet

Poistamalla n. 75% nimikkeistöstä nopeutettaisiin nykyisen järjestelmän toimintaa, niin päivittäisten prosessien suorittamisessa, kuin myös harvemmin otettavien liiketoiminnan raporttien osalta. Lisäksi datan hallinta helpottuu, kun tietokannasta poistetaan kaikki ylimääräinen datamassa, jonka päivittämiseen ei kenelläkään ole tarvetta/mielenkiintoa.

6.1.3 Asiakas- ja toimittajarekisterit

Asiakas- ja toimittajarekisterit ovat olleet pääosin myynti- ja markkinointiorganisaation (asiakasrekisteri) sekä hankintaorganisaation (toimittajarekisteri) hallinnoinnin alaisina. Mutta aivan samoin kuin nimikerekisterinkin kohdalla, niin myöskään asiakas- ja toimittajarekistereiden ylläpidolle ei ollut olemassa selkeitä sääntöjä ja ohjeistuksia. Ennen ydintietojen määrittelyiden tekemistä olikin toiminnanohjausjärjestelmää havainnoimalla todettavissa, että molempien rekisterien ylläpitokäytännöt olivat olleet varsin kirjava, mikä viittasi vahvasti siihen, että näiden rekistereiden datan ja datan hallinnan kehittäminen ei ole ollut kenenkään vastuulla.

Taulukkoon 11 on koottu esimerkinomaisesti joitain asiakasrekisterin ydintietoja. Kokonaisuudessaan asiakasrekisterin datastandardi on kuvattuna liitteessä 3 ja toimittajarekisterin datastandardi liitteessä 4.

Taulukko 11. Asiakasrekisterin ydintietoja

Kenttä	Tyyppi	Määritelmä	Esimerkki
Tunnus	VARCHAR2(12 CHAR)	<i>Asiakastunnus</i>	104764
Ryhmä	VARCHAR2(8 CHAR)		1
ALV-tunnus	VARCHAR2(20 CHAR)	<i>ALV tunnus</i>	FI12345601
Y-tunnus	VARCHAR2(11 CHAR)	<i>Y tunnus</i>	123456-01
Nimi	VARCHAR2(40 CHAR)	<i>Laskutusasiakkaan nimi</i>	Asiakas A
Postiosoite	VARCHAR2(40 CHAR)	<i>Laskutusasiakkaan postiosoite</i>	Jalostamontie 1
Postinumero	VARCHAR2(10 CHAR)	<i>Laskutusasiakkaan postinumero</i>	42300
Kaupunki	VARCHAR2(40 CHAR)	<i>Laskutusasiakkaan postitoimipaikka</i>	JÄMSÄNKOSKI
Maa	VARCHAR2(3 CHAR)	<i>Laskutusasiakkaan maakoodi</i>	FI
Maa (nimi)	VARCHAR2(16 CHAR)	<i>Laskutusasiakkaan maa</i>	FINLAND
Yhteyshenkilö	VARCHAR2(30 CHAR)	<i>Laskutusasiakkaan yhteyshenkilö</i>	Antti Asiakas
Puhelin	VARCHAR2(30 CHAR)	<i>Ensisijainen puhelinnumero</i>	+358 40 123 4567
Sähköposti	VARCHAR2(60 CHAR)	<i>sähköpostiosoite</i>	antti.asiakas@hotmail.com
Maksuehto	VARCHAR2(3 CHAR)		1
Toimitusehto	VARCHAR2(3 CHAR)		2
Toimitustapa	VARCHAR2(3 CHAR)		2
Valuutta	VARCHAR2(3 CHAR)		
Kielikoodi	VARCHAR2(3 CHAR)		EN
Myyntitili	VARCHAR2(8 CHAR)		
Myyntisaamistili	VARCHAR2(8 CHAR)		

Tarkasteltaessa asiakas- ja toimittajarekistereitä vuoden 2019 lopulla oli havaittavissa, että esimerkiksi asiakkaiden nimi- ja yhteystiedot olivat osin vanhentuneita tai niiden vaatimustenmukaisuus ei ollut kovin hyvällä tasolla.

Kelpoisuuden ja vaatimustenmukaisuuden mittarilla tarkasteltaessa, voidaan huomata, että kyseisten rekisterien tieto ei täytä sille asetettuja muodollisia ja sisällöllisiä vaatimuksia. Muodollisten vaatimusten täytyminen edistää tiedon teknistä tulkittavuutta sekä siirrettävyyttä eri järjestelmien välillä. Sisällöllisten vaatimusten täytyminen tarkoittaa, että tieto täyttää sille annetut liiketoiminnalliset vaatimukset (Väre, 2019).

Lisäksi on huomioitava, että koska monien asiakkaiden ja toimittajien yhteyshenkilöt, puhelinnumerot ja sähköpostiosoitteet olivat väärin tai vanhentuneita, niin se nyky-

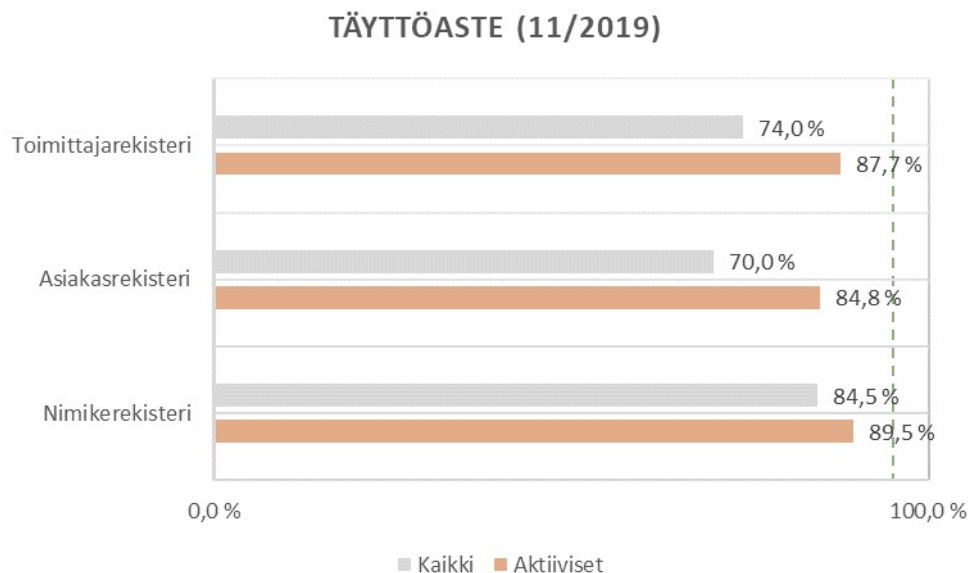
senä GDPR-asetuksen aikana on erittäin vahingollista ja saattaisi aiheuttaa yritykselle lisäkustannuksia ja mainehaittaa jatkossa.

6.1.4 Ydintiedon laatu

Yllä mainittujen asiakas-, toimittaja- ja nimikerekisterien määrittelyiden jälkeen toiminnanohjausjärjestelmän aktiivisten tietueiden täyttöaste mitattiin (DMAIC – *Measure*). Mittauksen avulla saatiin lisätietoa siitä, kuinka isosta ongelmasta on kyse. Ja vaikkakaan täyttöaste itsessään ei kerro ydintiedon muista laatutekijöistä paljoakaan, antaa se kuitenkin jonkinlaisen yleisarvion ydintiedon laadusta ja samalla se kertoo mihin ensimmäisenä kannattaa huomio keskittää, kun datan laatua lähdetään parantamaan.

Teoriaosaan viitaten, van der Staaij, A.J., de Goeij, A.S.M., van Unen, C.J.W.A., Swartjes, S. (2012) mainitsevatkin tyypillisiksi ydintiedon laadun mittareiksi mm. tuplätietuiden määrän, ydintiedon attribuuttien (ominaisuuksien) täyttöasteen ja näiden attribuuttien vaatimustenmukaisuuden määriteltyihin datastandardeihin nähden.

Marraskuussa 2019 tehty ydintietojen täyttöasteen mittaus osoitti, että yleisesti ottaen täyttöaste oli ihan kohtuullisella tasolla ollen nimikerekisterin osalta 84,5/89,5%, asiakasrekisterin osalta 70/84,8% ja toimittajarekisterin osalta 70/87,7% (kuvio 14).



Kuvio 14. Ydintiedon täyttöaste 11/2019 (tkA Finland Oy)

Ydintietojen täyttöastetta tarkasteltiin kahdella eri kriteerillä. Ensinnäkin kaikkien tietokannassa olevien tietueiden osalta ja toiseksi ainoastaan aktiivisten tietueiden osalta. Taulukossa 12 on esitetty miten tietueet jakautuvat määrällisesti näihin kahteen kategoriaan. Oheisesta taulukosta näkyy hyvin myös, että asiakas-, toimittaja- ja nimikerekisterien tietueissa toteutuu tuttu 80/20 -periaate, eli karkeasti 20% koko tietuemassasta on aktiivisia.

Taulukko 12. Aktiiviset ja ei-aktiiviset tietueet (tkA Finland Oy)

Rekisteri	Aktiivisia	%	Ei-aktiivisia	%	Yht.	%
Nimikerekisteri	12359	26 %	35348	74 %	47707	100 %
Asiakasrekisteri	587	13 %	4108	87 %	4695	100 %
Toimittajareksiteri	420	27 %	1140	73 %	1560	100 %

Nimikerekisterin osalta aktiiviset nimikkeet erotellaan nimikerekisterissä olevan Lak-kautettu-määrittelyn avulla. Nimike on (alfred.B-EYE/GMM) määritelmän mukaisesti merkittävä lakkautetuksi, mikäli sillä ei ole, eikä ole ollut osto- ja myyntitilauksia tai muita avoimia dokumentteja toiminnanohjausjärjestelmässä viimeisen 12 kk:n aikana, sillä ei ole saldoa varastokirjanpidossa, eikä se sisälly hankinnan tarvelaskennassa käytettäviin asiakasennusteisiin. Näiden lisäksi nimike pitää olla perustettu yli 12 kk ennen tarkasteluhetkeä.

Asiakas- ja toimittajarekisterin aktiiviset tietueet puolestaan erotellaan laskutustietojen perusteella siten, että ne asiakkaat ja toimittajat, joille tai joilta on ollut laskutusta 1.1.2018 jälkeen ovat aktiivisia tietueita. Tässä yhteydessä on syytä huomauttaa, että edellä mainittu päivämäärä koskee ainoastaan tämän opinnäytetyön aikana tehtyjä laadullisia mittauksia. Jatkossa tarkastelujaksona on kiinteän päivämäärään sijasta parempi käyttää esim. edeltävää 24 kk:n ajanjaksoa.

Kuitenkin, kuten edelläkin jo mainittiin, niin täyttöaste itsessään ei vielä kerro tietueen laadullisista ominaisuuksista läheskään kaikkea. Sen takia määriteltyjen datarekistereiden tietueille tehtiinkin vielä tarkempi muihin laadullisiin ominaisuuksiin keskittyvä satunnaistarkastus, jossa 2% nimikkeistä, sekä asiakas- ja toimittajatiedosta tarkastettiin sen suhteen, kuinka hyvin niiden tiedot vastasivat ydintietojen määrittelyssä ydintiedolle annettuja tarkempia määrittelyitä. Tämän satunnaistarkastuksen tuloksena oli, että nimikerekisteristä ainoastaan 2%, asiakasrekisteristä 0% ja toimittajarekisteristä 17,5% vastasi täysin annettuja määrityksiä. Saatu tulos puhuu selvää kieltään ydintietojen määrittelyn ja datan paremman hallinnan puolesta.

Väre (2019) korostaa, että tiedon laatua voidaan tarkastella monista eri lähtökohdista ja eri lähteissä korostetaan hieman eri ulottuvuuksien tärkeyttä, mutta useimmin esiin tuleviksi tiedon laadun ulottuvuuksiksi mainitaan ainutlaatuisuus, oikeellisuus, yhteneväisyys, kattavuus ja ajanmukaisuus.

Ainutlaatuisuus (Uniqueness) on ydintiedon hallinnassa erittäin oleellinen laadun ulottuvuus. Tieto on ainutlaatuista silloin, kun se esiintyy yhdessä tietokannassa tai järjestelmässä vain kerran. Mikäli tieto ei ole ainutlaatuista, ongelmat tulevat tyypillisesti esiin raportoinnin tai analyysien epäselvyytenä. Saman ydintieto tietueen toistuvat ilmentymät, eli tuplatietueet, aiheuttavat piileviä ongelmia liiketoimintaprosesseissa. Tiedon ainutlaatuisuus on kohtalaisen helppo varmistaa. Mitä parempi on sisällön kattavuus ja sen oikeellisuus, sitä helpompi ainutlaatuisuuden tarkistaminen on.

Kohdeyrityksessä tuplatietueita on syntynyt varsinkin nimikerekisteriin, nykyisten osin puutteellisten toimintamallien takia, koska ne eivät ole mahdollistaneet esimer-

kiksi yrityksen pääasiakkaan käyttämien nimiketunnusten jättämistä pois yrityksen omasta toiminnanohjausjärjestelmästä. Tällöin jo olemassa oleva nimike on saatettu joutua perustamaan samansisältöisenä uudelleen. Tuplanimikkeet ovatkin aiheuttaneet jatkuvasti ongelmia niin myynnissä, hankinnan tarvelaskennassa kuin myös tuotannon materiaalisuunnittelussa.

Tuplatietuiden poistamiseksi on suunnitteilla toimenpiteitä, kuten esim. asiakashinnaston luominen pääasiakkaalle, jonka avulla mahdollistetaan heidän automatisoidun myyntitilausten sisään luku -toiminnon muokkaaminen siten, että asiakkaan omaa raaka-ainanimikettä ei välttämättä enää tarvita kohdeyrityksen toiminnanohjausjärjestelmässä. Kohdeyrityksen omien toimenpiteiden lisäksi myös käynnissä oleva thyssenkrupp:n alfred.B-EYE -projekti tulee asettamaan omat vaatimuksensa tuplatietueiden poistamiselle. Maaliskuussa 2020 tuplatietueita oli raaka-ainanimikkeiden aktiivisista tietueista n. 2%.

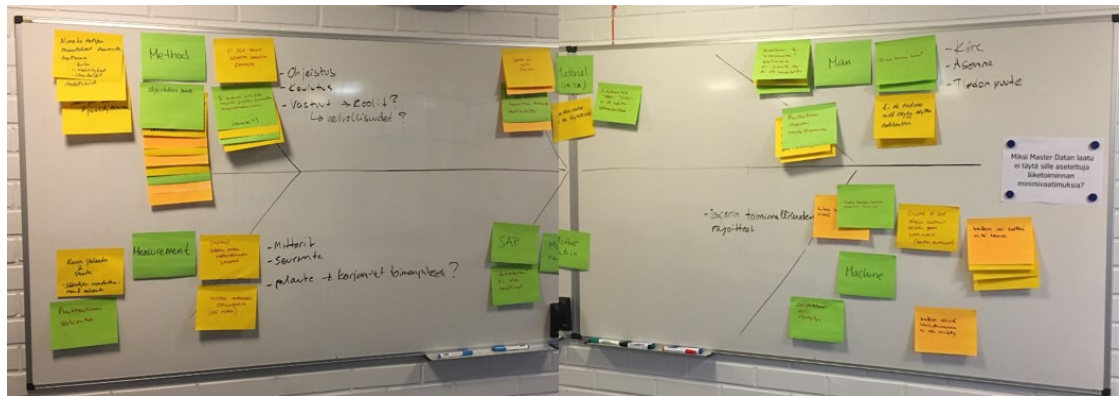
Tehtyjen tarkempien laadullisten mittausten luotettavuutta voidaan kyseenalaistaa, koska mittaus, tai lähinnä tehtyjen määritysten ja sääntöjen vertailu ydintietoihin, tehtiin manuaalisesti, jolloin tietojen oikeellisuuden arviointi ei välttämättä ole ollut tasalaatuista tai täysin virheetöntä. Siitä huolimatta voidaan todeta, että pelkästään täyttöastetta mittaamalla ei pystytä arvioimaan ydintiedon hyvyttä, vaan datan laadun mittaamiseen on otettava käyttöön myös muita ydintiedon laatuominaisuuksia seuraavia mittareita.

6.2 Datan hallintamalli

Ydintietojen määrittelyiden ja ydintiedoille tehdyn täyttöasteen ja muiden laatuominaisuuksien mittaamisen jälkeen voitiin todeta, että vaikka ydintiedon laatu on täyttöasteen osalta kohtalaisen hyvä, niin ydintietojen muille laatuominaisuuksille tehty satunnaistarkastus osoitti, että niiden sisällössä on vielä todella paljon parannettavaa. Vaikka nykyinen tietokanta siivotaan ja harmonisoidaan vastaamaan paremmin nyt tehtyjä määrittelyksiä, tulee ydintiedon kunnossa pitäminen olemaan jatkossa aivan yhtä huonolla tasolla kuin tähänkin saakka, jos ydintiedon huonolle laadulle ei löydetä juurisyitä sekä keinoja niiden poistamiseen.

Edellä mainituista syistä jo aiemmin käynnistetyn Six Sigma DMAIC-prosessin seuraavassa vaiheessa, haastatteluista kerättyjen vastausten (VOB) pohjalta pidettiin työpaja, jossa havaittuja mahdollisia syitä analysoimalla pyrittiin etsimään niitä juurisyitä, jotka poistamalla voitaisiin parantaa ydintiedon laatua (kuvio 15).

Voice of Business, eli liiketoiminnan ääni (VOB) muodostui kysymykseksi: *”Miksi ydintiedon (Master Data) laatu ei täytä sille asetettuja liiketoiminnan minimivaatimuksia?”*



Kuvio 15. Ishikawa - kalanruotokaavio

Työpajan, sekä tehtyjen havaintojen ja johtopäätösten perusteella suurin yksittäinen syy sille, miksi ydintiedot eivät täytä niille asetettuja vaatimuksia oli, että kohdeyrityksen ydintieto ei ole kenenkään vastuulla. Kukaan ei ole vastuussa eri datarekistereiden ydintietojen määrittelyistä, ylläpidosta, seurannasta tai ylipäätään niiden järjestelmällisestä kehittämisestä (taulukko 13).

Tämä on johtanut siihen, että ihmiset tekevät ja ylläpitävät kukin omanlaistaan epävirallista määritelmää liiketoiminnalle tärkeästä datasta. Tämä on johtanut ydintiedon huonoon laatuun, ja sitä kautta järjestelmädatasta on tullut epäluotettavaa. Lisäksi ydintietoon ja ylipäätään toiminnanohjausjärjestelmän dataan tehtäviä muutoksia ei ole viestitty muille osastoille, josko ei aina omallekaan osastolle.

Taulukko 13. VOB - mahdolliset syyt - juurisyy

VOB	Mahdollinen syy	Juurisyy
”Miksi ydintiedon (Master Data) laatu ei täytä sille asetettuja liiketoiminnan minimi-vaatimuksia?”	Onko kenttien asettelu oikein? (sekava etc.)	Kukaan ei vastaa datan laadusta/ omista dataa
	- Liikaa kenttiä näkyvissä	
	Ei ole tiedossa mitä täytyy täyttää tietokenttiin (Oscarissa)	
	- Ei tiedetä tarvittavia tietoja esim. Asiakasta perustettaessa (ALV-tunnus yms.)	
	Kaikkea tietoa ei ole käytettävissä	
	Dataa on niin paljon	
	- Liian paljon vaaditaan turhaa tietoa	
	Kurin palautus ja puute - sääntöjen noudattaminen ja valvonta	
	- Puutteellinen valvonta	
	Nimiketietojen muutokset suunnittelemattomia (koko, määritykset, standardit)	
	Uudet määrittelyt otettu käyttöön ilman tiedottamista	
	Sääntöjen luominen ja perehdyttäminen	
	Ei ole (ollut) selkeitä vastuita datasta	
	Ei tiedetä mitä pitää täyttää ja miten puutteellinen tieto vaikuttaa muihin	
	- Datan käyttöön ei ole ollut koulutusta	
- Puutteellinen koulutus > puutteellinen osaaminen		
- Ei ymmärretä mitä data on		
- Ei tiedetä mitä vaaditaan		
Ohjeistuksen puute		
- Asiakkaiden tietojen keruu ohjeistamatta esim. tilinumerot jne.		
- Selkeyttä täyttöohjeisiin		
- Ei tiedetä oikeaa tapaa täyttää nimiketietoja		
- Ei ole ohjeita		
- Ohjeistusta ei ole		
- Puutteellinen ohjeistus		
- Ei ole ollut tiedossa mitkä kentät tulee olla täytettynä		
- Vaatimuksia ei ole määritetty		

Roolit ja vastuut ovat onnistuneen datan hallintaohjelman selkäranka. Jos tiedon hallinnan halutaan toimivan tehokkaasti ja ydintiedon laatuun vaikuttavasti, tarvitaan organisaation kulttuurille sopivien roolien määrittely ja käyttöönotto. Vastuullisuuden virallistaminen siitä, miten ihmiset määrittelevät, tuottavat ja käyttävät dataa, vaatii rooliperusteiset käytännöt (Seiner, 2019).

Panian (2010) kuvaa omassa tutkimuksessaan organisaation itsessään olevan avainasemassa datan hallinnan onnistumisessa. Saavuttaakseen datan hallinnalle asettamansa tavoitteet, organisaatioiden täytyy määritellä datalle omistajuus, datastandardit, ja sovittujen menettelytapojen noudattamista tulee valvoa.

Kohdeyrityksen tapauksessa tulee kuitenkin huomioida, että resurssien rajallisuus ei mahdollista esim. nimikerekisterin hallinnan keskittämistä yhden tai kahden henkilön tehtäväksi. Ja juuri siitä syystä kohdeyrityksen tulee löytää datan hallintaan sellainen ratkaisu, jossa ei hajoteta nykyisiä organisaatorakenteita, eikä luoda uusia tehtä-

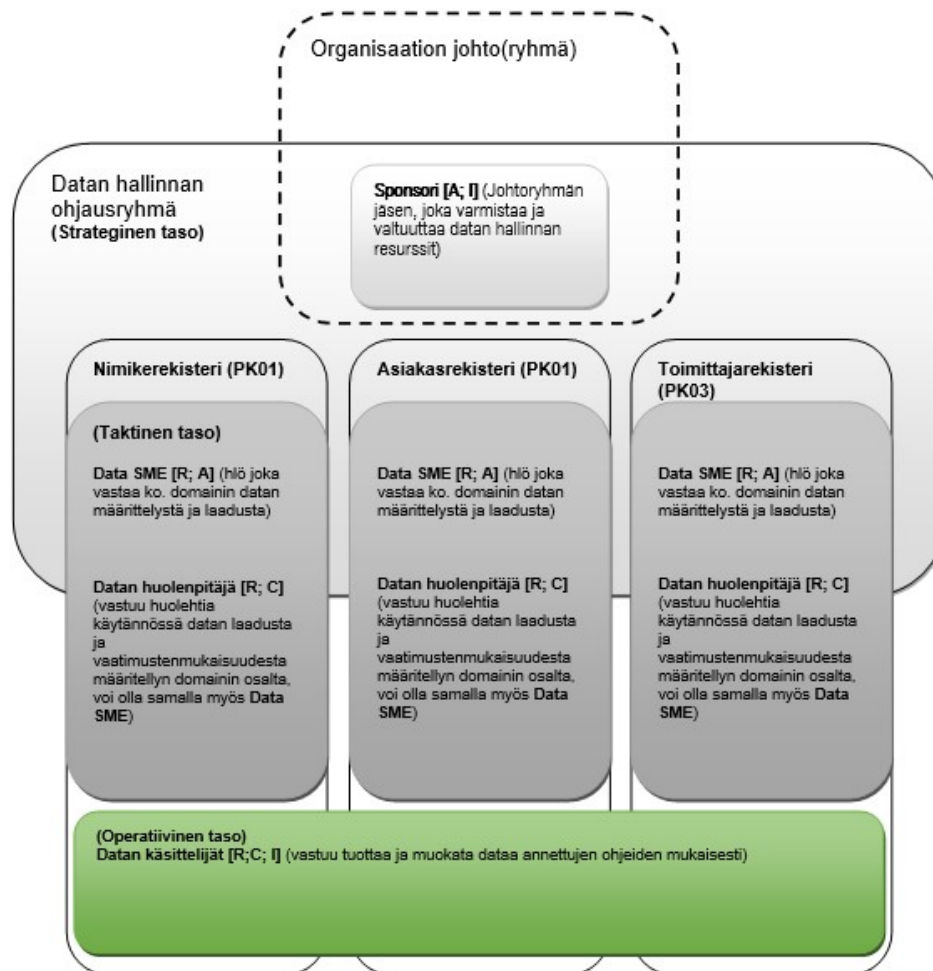
vänimikkeitä dataa hoitamaan. Ydintiedon hallinta on sisällytettävä olemassa oleviin toimintamalleihin ja ainoastaan tarvittaessa luoda uusia toimintamalleja, jotta datan hallintaa voidaan toteuttaa.

Onnistunutta datan hallintamallia ei voi vain liimata olemassa olevan organisaation päälle, vaan se täytyy huolellisesti sovittaa istumaan siihen. Jotta datan hallintamalli sopii organisaatioon, sen muodostamisessa on tärkeä huomioida yrityksen tehtävä ja tavoitteet, työyhteisön normit ja uskomukset (Väre, 2019; Weber ym., 2009). Tavoite hallintamallin luomisessa pitäisi kuitenkin olla aina se, että päätökset tehdään mahdollisimman lähellä niitä, joihin ne vaikuttavat. On myös muistettava, että mikäli asioiden päättämiseksi vaaditaan liian hidasta tai monimutkaista toimintatapaa, organisaatioon syntyy epäviralliset tavat päättää asioista (Seiner, 2014).

6.2.1 Roolit ja vastuut

Kuten teoriaosassakin monesti mainittiin, niin organisaatio, joka lähtee toteuttamaan datan hallintaa siten, ettei tarkoituksena ole harjoittaa ylhäältäpäin johdettua ”hajota ja hallitse” -tyyppistä hallintamallia, joutuu muokkaamaan datan hallinnan operatiivisen mallin omaa rakennettaan ja sen toimintatapoja ja kulttuuria vastaavaksi.

Kohdeyrityksen kohdalla se tarkoittaa organisaatorakennetta (kuviot 16), jossa tekeminen ja datan hallinta pyritään mahdollisimman paljon pitämään nykyisten roolitus-ten sisällä. Kouluttamalla vastuut datasta kaikille käyttäjille, pyritään varmistamaan se, että kaikki kokevat olevansa osa kokonaisuutta ja omalta osaltaan vaikuttamassa datan laatuun, ilman lisätyötä. Kohdeyrityksen datan hallinnan operatiivinen malli rooleineen ja vastuineen on esitelty tarkemmin liitteessä 5.



Kuvio 16. Datan hallinnan operatiivinen malli (tkA Finland Oy)

Datan hallintamallin roolien vastuut/velvollisuudet thyssenkrupp Aerospace Finland Oy:ssä:

Datan käsittelijät [R; C; I] toimivat operatiivisella tasolla ja ovat päivittäin tekemisissä datan kanssa, sitä käyttäen, ylläpitäen tai luoden.

- ovat vastuullisia datan laadusta omalla tekemisellään.
- ovat vastuullisia toimimaan annetun ohjeistuksen mukaisesti käyttäessään ja päivittäessään dataa, sekä luodessaan uutta dataa (nimikkeitä, asiakkaita ja toimittajia).
- ovat velvollisia ilmoittamaan *Datan huolenpitäjille* havaitsemistaan epäkohdista ohjeistuksessa ja/tai prosessien toimivuudessa sekä datan laadussa.

Datan huolenpitäjät [R; C] toimivat sekä operatiivisella, että taktisella tasolla ja yhdessä Data SME:n kanssa huolehtivat ydintiedolle annettujen ohjeistusten ajantasaisuudesta, kouluttamisesta sekä mittaamisesta.

- vastaavat käytössä olevan ohjeistuksen ajantasaisuudesta.
- vastaavat tehtyjen ohjeistusten jalkauttamisesta Datan käsittelijöille.
- vastaavat Datan käsittelijöiden havaitsemien epäkohtien/laaturvirheiden tutkimisesta ja viestinnästä *Data SME*:lle ja huolehtivat, että epäkohdat/virheet korjataan *Data SME*:n/*Datan hallinnan ohjausryhmän* ohjeistuksen mukaisesti.
- vastaavat käytännössä datan laadusta ja vaatimustenmukaisuudesta määritellyn domainin osalta (mittarointi).

Data SME [R; A] käytännössä sama kuin Datan huolenpitäjä, mutta on vastuussa kyseisestä datarekisteristä kokonaisuudessaan.

- vastaa ko. rekisterin datan määrittelystä, kehittämisestä ja laadusta kokonaisuudessaan.
- vastaa datan hallinnan prosessien sisällyttämisestä liiketoimintaprosesseihin määritellyn rekisterin osalta.
- vastaa soveltuvien laatumittareiden luomisesta vastuullaan olevalle datalle.
- raportoi *Sponsorille/Datan hallinnan ohjausryhmälle* oman vastuualueensa datan laadusta.

Sponsori [A; I] yrityksen johtoryhmän jäsen, joka huolehtii datan hallinnan onnistumisen edellytyksistä

- vastuussa huolehtia riittävät resurssit datan hallinnalle.

Datan hallinnan ohjausryhmä; toimielin, jossa ratkaistaan ydintietoon liittyvät ongelmat ja kehitetään datan hallintaa kokonaisuutena

- määrittelee mitä lisäohjeistuksia tarvitaan nykyisten menetelmäohjeiden lisäksi.
- vastaa viimekädessä esiin tulevien ydintiedon laatuongelmien ratkaisemisesta.
- datan hallinnan ohjausryhmän tulee kokoontua säännöllisesti katselmoimaan datan hallinnan tilannetta.

Nämä edellä luetellut roolit ja vastuut pyritään kouluttamaan kaikille kohdeyrityksen työntekijöille kevään 2020 aikana. Opinnäytetyötä kirjoittaessani koulutus on juuri aloitettu ydintiedon ja sen hallinnan kannalta keskeisimpien sidosryhmien osalta.

Kouluttamisen jälkeen tulee kaikille henkilöille (rooleille) tarjota riittävästi tukea uuden toimintamallin omaksumiseksi, eikä voida olettaa, että uudet toimintatavat juurtuisivat käytäntöön ilman seurantaa ja lisäkoulutusta.

Väre (2019) nimeää yhdeksi datan hallintamallin jalkautuksen periaatteeksi uusien roolien tukemisen niin pitkäksi aikaa kuin he sitä tarvitsevat. Eri ihmiset sisäistävät uudet asiat eri tahdissa. Kun uudet roolit ja niiden sisältämät vastuut viedään ihmisille organisaatiossa, jotkut tarvitsevat tukea roolinsa oppimiseksi kauemmin kuin toiset (Väre, 2019).

6.3 Datan laadun seuranta

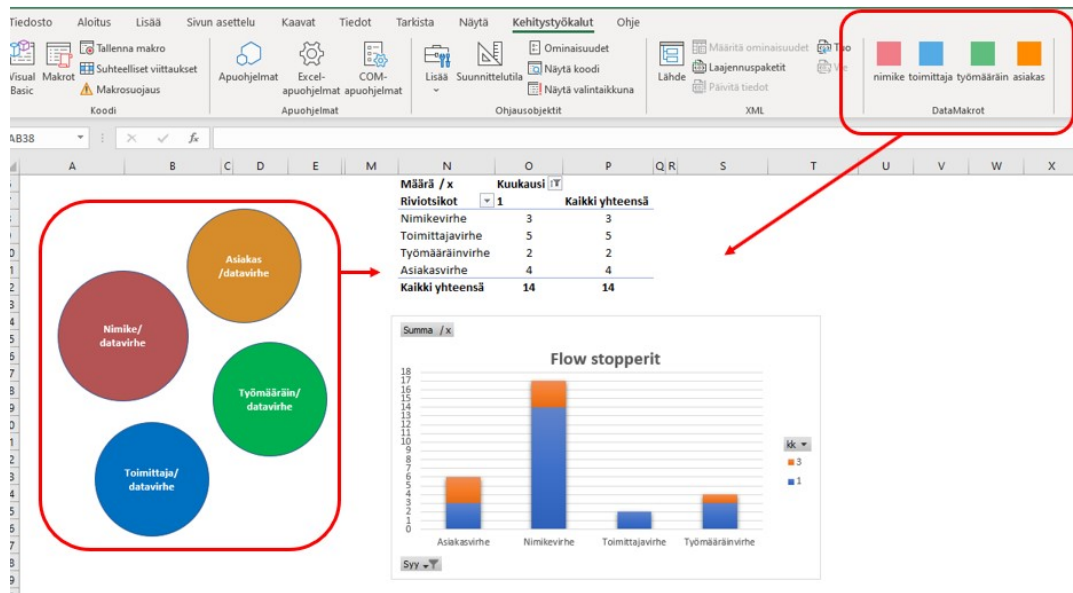
Datan hallinnan ohjausryhmän yhtenä tehtävänä uudessa datan hallinnan toimintamallissa on varmistaa ydintiedon laatu sekä määritellä ne keinot, jolla tuota laatua valvotaan. Ensivaiheessa ydintiedon laatua päätettiin lähteä seuraamaan sellaisilla mittareilla, joiden raportointi onnistuu suhteellisen pienellä panostuksella ja joiden antamaan tulokseen voidaan luottaa. Tästä syystä kaikille pakolliseksi ydintiedoksi määritellyille kentille/attribuuteille, valittiin yhdeksi mittariksi niiden täyttöaste tai tarkemmin sanottuna sisällön kattavuus. Tämä mittari kertoo yksiselitteisesti, onko pakolliseksi määritelty ydintieto olemassa vai ei.

Lisäksi sovittiin otettavaksi käyttöön tarkempi, ydintietojen muita laatuominaisuuksia seuraava mittari muutamille tärkeimmiksi katsotuille attribuuteille. Tämän mittarin seuranta tehdään otantaperusteisesti, jolloin kerran kuukaudessa 2% rekisterien tietueista tarkastetaan manuaalisesti ydintiedon määrittäjiä vastaan. Tarkemmassa seurannassa ydintiedosta tutkitaan mm. ydintietojen yhteneväisyys ja vaatimustenmukaisuus.

Näiden kahden mittarin lisäksi kohdeyrityksessä otetaan käyttöön mittari, jolla seurataan käytännön tekemisessä tapahtuvia, ydintiedon virheistä johtuvia, prosessin pysäyttäviä datan "Flow stoppeja", tilanteita, jossa jokin toiminto ei pysty suorittamaan omaa prosessiaan valmiiksi, vaan joutuu ensin etsimään ratkaisun ydintiedoissa olevan virheen poistamiseksi.

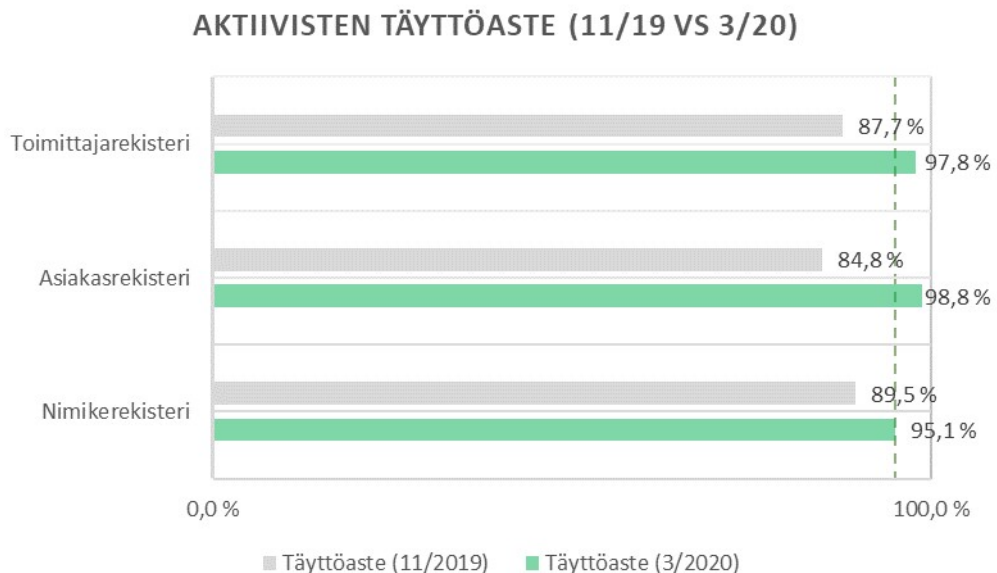
Kyseinen mittari toteutetaan Excelin avulla, jolloin mittarin raportointi voitaisiin tehdä esim. kuvion 17 mukaisilla painonapeilla, jossa ne sijaisivat vaihtoehtoisesti esim. työkalurivillä tai työkirjan taulukossa. Ideana tämän tyyppisessä "Flow stop" -mittarissa on löytää ne ydintiedon osa-alueet, jotka aiheuttavat eniten ongelmia prosessien toimintaan. Se, kuinka yksityiskohtaisia ongelmia painonappien avulla raportoidaan, on kiinni ydintiedon hallinnalle asetetuista tavoitteista. Mikäli halutaan vain yleiskuvaus datan aiheuttamista ongelmista, niin silloin raportointikriteerienkään ei tulisi olla yksityiskohtiin meneviä.

Jos taas halutaan tietää tarkasti, mikä ominaisuus esim. nimikerekisterissä aiheuttaa eniten ongelmia, tulee mittarin tarkkuutta ja siten raportoitavien ominaisuuksien määrää kasvattaa huomattavasti. Tämä puolestaan voi johtaa raportoinnin epäluotettavuuteen, jos raportointia tekevä henkilö ei tiedä todellista, prosessin pysäyttänyttä, dataominaisuutta.



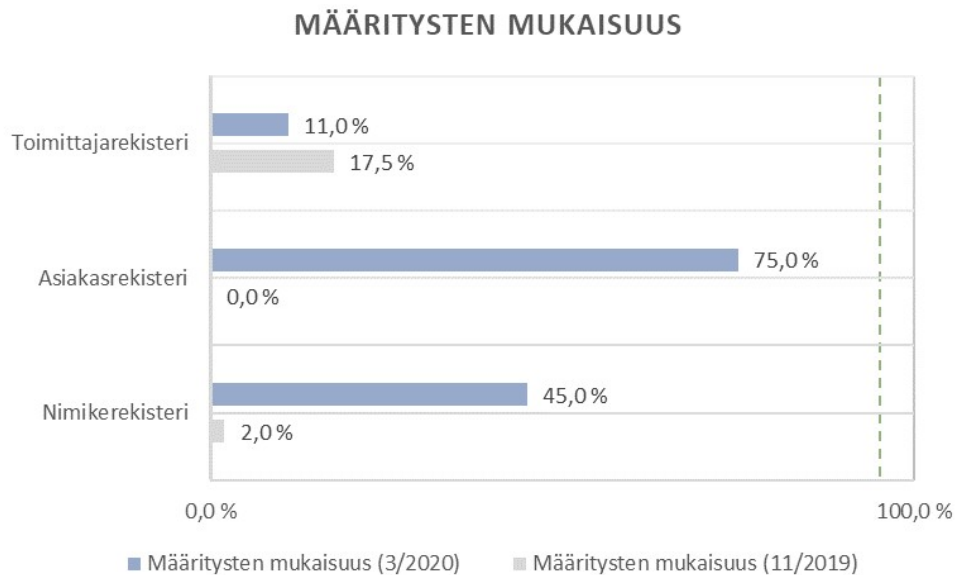
Kuvio 17. Esimerkkejä "Data Stop" nappien sijainnista excel -työkirjassa

Ydintiedoille tehtyjen määritysten jälkeen nimike-, asiakas- ja toimittajarekisterit siivottiin (syksyllä 2019) ja niiden sisältöä harmonisoitiin aktiivisten käytössä olevien tietueiden osalta. Maaliskuussa 2020 tehty ydintietojen täyttöasteen mittaus osoitti, että siivouksen avulla olikin saatu melkoisesti parannusta aikaiseksi. Nimikerekisterin täyttöaste oli 95,1%, asiakasrekisterin 98,8% ja toimittajarekisterin 97,8% (kuvio 18).



Kuvio 18. Aktiivisten tietueiden ydintietojen täyttöasteen kehitys (tkA Finland Oy)

Myös toisen, ydintietojen vaatimustenmukaisuutta seuraavan mittarin lukemat olivat parantuneet, mutta asetettuun 95% tavoitteeseen on vielä melkoisesti matkaa. Esimerkiksi toimittajarekisterin laatutaso oli jopa heikentynyt 17,5% 11%:iin (kuvio 19). Tähän on yhtenä selittävänä tekijänä kyseisen rekisterin siivoustoimenpiteiden keskeneräisyys tarkasteluhetkellä.



Kuvio 19. Ydintietojen määrittysten mukaisuus (tkA Finland Oy)

Nimikerekisteriin tehtyjen määrittysten kohdalla näyttäisi olevan tarvetta myös uudelle tarkastelulle, koska ensimmäisen määrittelykierrosten aikana oli havaittavissa jonkinlaista ”data-ahneutta” ja kukin toiminto määritteli käyttämiään attribuutteja pakollisiksi ydintiedoiksi, vaikka ei olisi välttämättä ollut tarvetta, ja joiden pakollisuuden ehdot olivat ristiriidassa käytännön tekemisen kanssa.

7 Pohdinta

7.1 Tutkimusprosessi

Suunnitelmana tätä opinnäytetyötä aloitettaessa oli selvittää mitä ydintieto ja sen hallinta kohdeyrityksessä tarkoittaa, mitä haasteita niihin liittyy, sekä millä tavalla

ydintieto saataisiin paremmin hallintaan siten, ettei pienen organisaation rakenteita jouduttaisi liiallisesti venyttämään.

Tutkimuskysymykset:

- Mitä on ydintieto kohdeyrityksessä?
- Mitä haasteita ydintietoon ja sen hallintaan liittyy?
- Minkälainen organisaatorajat ylittävä ydintiedon hallintamalli olisi kohdeyritykselle toimivin?

Kohdeyrityksessä ei tietämykseni mukaan ollut aiemmin juurikaan paneuduttu siihen, mitä ydintieto on tai mitä hyötyjä hyvin hoidetusta ydintiedosta voisi yritykselle ja sen toiminnalle olla. Kuitenkin aikana, jolloin digitalisaatio saattaa radikaalisti ja nopeastikin muuttaa toimintaympäristöjä, ei yrityksillä ole varaa jättää huomioimatta dataa elintärkeänä tekijänä, jota hyödyntämällä ja kehittämällä voidaan luoda jopa kokonaan uudenlaisia ansaintakeinoja perinteisten keinojen rinnalle.

Dataan ja data-analytiikkaan liittyvät termit, kuten AI, IoT, Blockchain jne. ovat olleet erilaisten tulevaisuuden trendejä tutkivien tutkimuslaitosten mielenkiinnon kohteena jo pitkään ja viimeaikaiset tutkimukset eivät sitä tosiasiaa ole muuttamassa, vaan päinvastoin, hyvälaatuista dataa tarvitaan jatkossa entistä enemmän.

Jotta opinnäytetyössä päästäisiin eteenpäin ja kaikkiin asetettuihin tutkimuskysymyksiin pystyttäisiin vastaamaan, tuli aivan ensimmäiseksi siis selvittää ydintiedon nykytila kohdeyrityksessä. Tämä toteutettiin kvalitatiivisen tutkimuksen eli teema-haastatteluiden avulla. Haastateltavaksi valikoitui ydintiedon eri sidosryhmien edustajia siten, että tutkimukseen saatiin mahdollisimman monipuolisesti erilaisia näkemyksiä tarkasteltavasta ilmiöstä eli ydintiedosta.

Haastatteluiden lisäksi sidosryhmien edustajia pyydettiin vastaamaan kyselyyn tiedonhallinnan kypsyydestä kohdeyrityksessä. Vastauksia kyselyyn tuli niukanlaisesti,

mikä saattoi osaltaan johtua kyseltävien asioiden vaikeasekoisuudesta tai siitä, ettei datan hallintaa koettu riittävän tärkeäksi asiaksi kyselyhetkellä.

Nykytila-analyysin jälkeen kohdeyrityksessä aloitettiin toimintatutkimus, eli kehitysprojekti, jonka tarkoituksena oli tuoda käytännön datatyöhön työkaluja, niin datastandardien kuin myös datan hallinnan kehittämisen keinoin. Kehitysprojekti toteutettiin Six Sigma DMAIC-prosessia hyödyntäen.

7.2 Tulokset

Nykytila-analyysin keskeisimpiä tuloksia oli, että ydintiedon ylätason periaatteet ovat melko hyvin hallussa, eli kohdeyrityksessä tunnistetaan ydintiedon olevan yrityksen toiminnalle tärkeää ajantasaista tietoa, jota käytetään päivittäisten prosessien suorittamiseen.

Tutkimuksen tuloksista tuli erittäin selvästi esille kuitenkin myös se tosiasia, että toiminnot ovat olleet ja ovat osin edelleenkin erittäin siiloutuneita, eikä toisten tarpeita huomioida omassa päivittäisessä tekemisessä. Moni tutkimuksen aikana haastatelluista ihmisistä kertoi aidon huolensa yhteisten pelisääntöjen puuttumisesta ydintietojen ja ylipäätään toiminnanohjausjärjestelmän datan suhteen. Ihmiset ovat saaneet elää liian pitkään ”kuin pellossa” datan suhteen, eikä nykyisen toiminnanohjausjärjestelmän pienillä muutostöillä pystyittäisi vaikuttamaan ihmisten asenteisiin ja luultuneisiin toimintatapoihin. Monet olivat myös vakaasti sitä mieltä, että ainoastaan käyttöoikeuksien tiukka rajoittaminen auttaisi saamaan ydintiedot paremmin hallintaan.

Itse olen kuitenkin varovaisen optimistinen, sen suhteen, että jos vain ymmärrettäisiin mikä datan merkitys liiketoiminnalle ja sen kehittämiseksi on nykyaikana, niin suhtautuminen siihen muuttuisi myötämielisemmäksi. Ihmisille olisi kyettävä antamaan vastaus mm. sellaisiin kysymyksiin kuin, miksi muutosta tarvitaan ja mihin muutoksella pyritään.

Toisaalta olen itsekin sitä mieltä, että tilanteessa, jossa olemassa olevaa toiminnan-ohjausjärjestelmää ei korvata täysin uudella järjestelmällä ja sitä kautta pakoteta ihmisiä uudenlaiseen ajattelu- ja toimintamalliin, on ydintietojen ja datan hallinnan kuntoon laittaminen asteen verran isompi haaste. Sillä jos päivittäinen tekeminen kuitenkin onnistuu huonollakin datalla edes jotenkin, johtaa se hyvin helposti siihen harhaluuloon, että datan laatu on riittävän hyvällä tasolla. Tällöin on helpompi pidättäytyä tutussa ja turvallisessa tavassa toimia, kuin poistua omalta mukavuusalueelta, mitä puolestaan datan hallinnan parantaminen kohdeyrityksessä ehdottomasti vaatisi.

Tällaisessa, ei pakotetussa muutostilanteessa, voisi olla hyvä tutustua ja pyrkiä seuraamaan John P Kotterin vuonna 1995 lanseeraamaa kahdeksan kohdan prosessia (Kotter, 2012), jonka avulla pyritään varmistamaan muutosprosessien onnistuminen. Kotterin prosessi juontaa juurensa aikaan, jolloin hän tutki miksi yritysten muutosohjelmat epäonnistuivat. Eräs seikka (virhe) minkä hän huomasi oli, että yrityksissä syyllistyyään liialliseen omahyväisyyteen, eikä muutokseen osallistuvat ihmiset ole valmiita poistumaan omilta mukavuusalueiltaan. Tämä johtuu Kotterin mielestä mm. siitä, että muutokseen lähdetessä ei riittävästi painoteta muutoksen kiireellisyyttä ja tärkeyttä.

Muita Kotterin havaitsemia virheitä muutosprojektien hoidossa olivat mm. riittämätön (johdon) tuki, muutokseen liittyvän kommunikaation vähyyys, ennen aikaisten voitonjuhlien järjestäminen ja tehtyjen muutosten vakiinnuttamisen epäonnistuminen.

Mutta palatakseni nykytila-analyysin tuloksiin, niin niistä kävi hyvin ilmi, että liiketointakriittiselle ydintiedolle kaivattiin määrittelyjä ja datan parissa työskenteleville yhteistä datastandardia, johon kaikki voisivat sitoutua. Tämän työn tuloksena syntyikin datastandardit kolmelle keskeiselle datarekisterille, eli asiakas-, toimittaja- ja nimikerekisterille. Datastandardien myötä saatiin myös hyvin yksityiskohtainen vastaus opinnäytetyön ensimmäiseen tutkimuskysymykseen, eli mitä ydintieto on kohdeyrityksessä.

Ydintietojen määrittelyn lisäksi oli selvää, että aiemman kaltainen tilanne, jossa kaikki saavat tehdä kaikkea, ei enää toimisi ja että ydintiedon hallinnalle pitää luoda selkeät roolit ja vastuut. Lisäksi ydintiedon laatua pitää seurata säännöllisesti ja tuoda siinä havaitut virheet esille, sekä luoda havaittujen datan laatuvirheiden käsittelylle prosessi. Tässä kohtaa on hyvä muistuttaa, että virheet ja niiden korjaaminen ovat työssä tapahtuvan oppimisen kulmakiviä.

Tätä tarvetta (ja haastetta) varten kohdeyritykselle muodostettiin datan hallinnan operatiivinen malli, johon on kirjattu kaikkien datan kanssa toimivien roolit, vastuut ja velvollisuudet. Operatiivisen mallin avulla datan hallinta tehdään läpinäkyväksi, jolloin kaikki tietävät vastuunsa ja velvollisuutensa datan suhteen sekä kuinka heidän odotetaan toimivan dataan liittyvissä ongelmatilanteissa.

Ydintiedon kehittämisessä tarvetta ja priorisointia ohjaavat liiketoiminnan ja asiakkaan tarpeet. Datan hallinnan roolien tehtävä on tunnistaa nämä tarpeet ja viedä niiden mukaisesti ydintiedon hallinnan osa-alueita eteenpäin, huomioiden kuitenkin, että koko organisaation yhteinen etu menee heidän omien liiketoiminta-alueidensa edun edelle. (Väre, 2019, Seiner 2014). Tämä voi johtaa jonkin asteiseen kipuiluun, koska monet organisaatiot ovat tottuneet toimimaan silloissa ja siirtyminen toimialarajat rikkovaan liiketoimintaan voikin tuoda mukanaan poliittisia taisteluita, mielipide-eroja ja paljon työtä (Seiner, 2014).

Täytyy myös muistaa, kuten Kotter (2012) vuonna 1995 tekemässään tutkimuksessaan huomasi, uuden toimintamallin vakiinnuttaminen pysyväksi käytännöksi tulee olemaan erittäin haasteellista ja se, että ihmiset saadaan noudattamaan ydintiedolle tehtyjä määräyksiä, vaatii vielä paljon muistuttamista ja asioiden kertausta. Tätä varten onkin erittäin tärkeää, että ydintietojen laadulle luotiin kehitysprojektin aikana mittarit, jotka auttavat tuomaan esiin ydintiedossa olevat poikkeavuudet määrittelyihin nähden.

Mittareiden avulla ongelmakohdat pystytään havaitsemaan ja siten voidaan järjestämään esim. lisäkoulutusta datan käyttäjille, tai jos ongelmat johtuvat ydintiedon virheellisistä määrittelyistä, niin siinä tapauksessa ne huomataan ottaa datan hallin-

nan ohjausryhmän tarkasteluun ja päivittää määrittelyt datastandardeihin ja toimintaohjeisiin. Ilman mittareita toiminta ydintietojen ja sen hallinnan ympärillä palautuu aikaisemman kaltaiseen hallitsemattomaan tilaan, jossa kukaan ei pysty luottamaan datan laatuun.

Opinnäytetyön tuloksena syntynyt datan hallinnan operatiivinen malli vastaa mielestäni hyvin asetettuun tutkimuskysymykseen siitä, mikä olisi kohdeyrityksen tapauksessa toimivin, organisaatorajat ylittävä ydintiedon ja datan hallintamalli.

Operatiivisen hallintamallin ja datan laadun näkyvämmän seurannan avulla, sekä kouluttamalla ydintiedoille asetetut vaatimukset ja annetut ohjeistukset kaikille datan käyttäjille lisätään myös, opinnäytetyön epävirallisen tavoitteen mukaisesti, data-tietoisuutta kohdeyrityksessä.

7.3 Tutkimuksen luotettavuus

Opinnäytetyön tuloksien perusteella kaikkiin tutkimuskysymyksiin pystyttiin vastaamaan valittujen tutkimusmenetelmien avulla. Haastatteluiden avulla pystyttiin selvittämään mitä ydintieto on, mitä haasteita ydintiedon hallintaan liittyy sekä minkälainen organisaatorajat ylittävä ydintiedon hallintamalli sopisi kohdeyritykselle parhaiten.

Toimintatutkimuksessa pystyttiin hyvinkin kattavasti käymään läpi liiketoiminnan ydintiedolle asettamat vaatimukset ja muodostamaan niistä relevantit ja ajantasaiset datastandardit. Lisäksi ydintietojen määrittelyitä tehtäessä huomioitiin mielestäni varsin hyvin se, että liiketoiminnan vaatimukset muuttuvat koko ajan ja samalla tulee myös datan määrittelyt pitää ajan tasalla ja jatkuvan seurannan alla.

Ydintiedon operatiivisen hallintamallin toimintaperiaatteeksi otettiin ydintiedon luominen kerralla oikein, mikä tukee myös edellä mainittua liiketoiminnan vaatimusten muuttumisen jatkuvaa seuraamista. Jatkossa liiketoiminnan vaatimukset pitääkin informoida aiempaa tehokkaammin sekä datan käyttäjille, että datan hallinnan oh-

jausryhmälle, jotta ohjeistusten ja määritysten ajantasaisuus pystytään varmistamaan.

Tutkimuksen teoriaosasta on tehty riittävän laaja, jotta kaikki opinnäytetyön kannalta oleellinen tieto on saatu tuotua esiin. Lähdemateriaalina on käytetty sekä alan perusteoksia, että uudempia ja enemmän tähän päivään sidottuja teoksia ja artikkeleita.

Työn ja siinä käytetyn aineiston epäluotettavimpana osana voidaan pitää sen kvalitatiivista tutkimusosuutta. Kanasen (2010) mukaan luotettavuusmittarit eivät sovellu yhteiskuntatieteisiin, jossa tutkimuskohteina on ihminen. Ihminen kun ei aina toimi samalla tavalla ja rationaalisesti, vaan voi muuttaa kantaansa äkisti ja ilman syytä. Lisäksi kvalitatiivisen tutkimusosion epäluotettavuutta lisää osaltaan myös se, että haastattelutilanteissa kirjasin ylös ainoastaan omasta mielestäni oleelliset ilmiöön liittyvät asiat. Tämän kaltainen vastausten tulkinta ja analysointi jo niitä ylös kirjattaessa saattaa aiheuttaa tulosten lopulliseen tulkintaan ja tuloksiin vääristymää.

Haastatteluiden jälkeen pyrin kuitenkin varmistamaan kirjattujen asioiden oikeellisuuden ja kattavuuden luetuttamalla puhtaaksi kirjoitetut muistiinpanot tuoreeltaan haastatelluilla. Tällä halusin varmistaa sen, että ensinnäkin kaikki olennainen tuli kirjattua ylös ja toisekseen sen, että asiat oli myös haastateltujen mielestä oikein kirjattu, eikä haastatteluiden aikana tekemäni valinnat olleet vääristäneet lopputulosta tai haastateltujen sanomisia.

7.4 Suositeltuja jatkotoimenpiteitä

Tutkimusprosessin aikana esille tulleita asioita, kuten ydintietorekisterien vanhentuneiden ja tuplanimikkeiden poistaminen/arkistointi tulee saattaa loppuun mahdollisimman nopealla aikataululla, jotta muita niihin liittyviä kehitystoimenpiteitä saadaan yrityksessä ja konsernissa eteenpäin.

Samoin on huomioitava myös se, että vaikka kohdeyrityksessä on nyt saatu ensimmäiset datastandardit tehtyä ja datan operatiivinen hallintamalli paperille, ei se suinkaan vielä anna aiheutta nostaa jalkaa pois kaasupolkimelta, vaan nyt aloitetut toimet

tulee saattaa määrätietoisesti maaliin ja vakiinnuttaa datan hallinta osaksi yrityksen normaalia toimintatapaa. Ja viimeistään, kun datan operatiivinen hallintamalli on vakiintunut käytäntöön näiden kolmen ensimmäisen datarekisterin osalta, tulee myös muut kohdeyrityksen datarekisterit ottaa datan hallinnan piiriin ja luoda niille datastandardit ja ydintiedon laadun mittarit. Tällä varmistetaan kaiken ydintiedon tasapuolinen kehittäminen.

Varmistamalla ydintietojen jatkuva kehittäminen ja laadullinen hyvä taso, varmistetaan myös yrityksen kyky vastata tuleviin muutoksiin dataan liittyvissä haasteissa ja vaatimuksissa.

Eräs haaste, joka perustuu omaan henkilökohtaiseen näkemykseeni, eikä niinkään tutkimuksen tuloksiin on, että yrityksissä työskentelevien ja ennen kaikkea yrityksiä johtavien ihmisten tulisi jatkossa aiempaa paremmin pyrkiä ajattelemaan, näkemään ja toimimaan asioiden suhteen pitkäjänteisemmin ja suunnitelmallisemmin, eikä ainoastaan keskittyä lyhyen tähtäimen voittojen kotiuttamiseen kuukausi tai vuosineljännes kerrallaan. Ruokkimalla sitä harhaa, että ainoastaan nykyhetki on tärkeä, luodaan pohjaa sille, että ihmiset priorisoivat tekemisensä lyhyen ajan tavoitteisiin peilaten.

Yritysten tulee omalla toiminnallaan varmistaa ja turvata ”pelissä” mukana pysyminen myös pitkällä aikavälillä ja siihen nimenomaan datan hallinnankin tulisi pyrkiä. Korostamalla datan merkitystä tulevaisuuden kannalta merkittävänä tekijänä autetaan ylläpitämään yrityksen olemassaolon perusedellytyksiä ja varmistamaan toiminnan kehittyminen myös jatkossa.

Voihan olla, että jatkossa tulee jälleen eteen tilanne, jossa kohdeyrityksen nykyinen toiminnanohjausjärjestelmä halutaan korvata jollain thyssenkrupp AG -konsernin yhteisellä järjestelmällä. Siinä vaiheessa olisi erittäin hyvä, jos olemassa oleva datamassa olisi jo valmiiksi mahdollisimman hyvälaatuista siirrettäväksi uuteen järjestelmään.

Lähteet

- Burbank, D., Knight, M. 2019. Raportti "Trends in Data Management." DATAVERSITY Education, LLC
- Cerwin, F. 2017. Artikkele Data Governance is Not About Governing Data." Viitattu 4.12.2019. <https://www.dataversity.net/data-governance-not-governing-data/#>.
- Cerwin, F. 2018. Artikkele "Is Master Data a Controlled Substance?" Viitattu 26.11.2019. <https://www.dataversity.net/master-data-controlled-substance>.
- Cerwin, F. 2019 Haastattelu aiheesta "Data Governance vs. Master Data Management". Viitattu 21.11.2019. <https://www.dataversity.net/data-governance-vs-master-data-management/#>.
- Cheong, L. K., Chang, V. 2007. The Need for Data Governance: A Case Study. ACIS 2007 Proceedings.
- Dama International. (ed), Mosley, M. 2010. The DAMA Guide to The Data Management Body of Knowledge (DAMA-DMBOK Guide), First Edition. Technics Publications.
- Dyché, J., Nevala, K., 2017. 10 Mistakes to avoid When launching your data governance program. SAS Institute. 2017.
- Experian Information Solutions. "The 2018 global data management benchmark report". 2018. Experian Information Solutions.
- Findwise. Mitä jokaisen kuuluu tietää EU:n uudesta tietosuojasetuksesta GDPR? Viitattu 28.11.2019. <https://findwise.com/en/gdpr-fi#1>.
- McCarthy, K. Experian Global Data Management 2019 Research, "Customer experience". 2019. Experian Information Solutions.
- Hovi, A. 2018. "Data-alan termien selitykset ja kuvaukset". Viitattu 27.9.2019. <https://www.arihovi.com/3274-2/#>.
- Institute of Governance. Artikkele " The complexity of Governance is difficult to capture in a simple definition." Viitattu 4.12.2019. <https://iog.ca/what-is-governance/>.
- Jonker, R.A., Kooistra, F.T., Cepariu, D., van Etten, J., Swartjes, S. 2011. Artikkele "Effective master data management". Viitattu 27.11.2019. <https://www.compact.nl/articles/effective-master-data-management/>.
- Kananen, J. 2010. Opinnäytetyön kirjoittamisen käytännön opas. Jyväskylän ammattikorkeakoulu.

- Kananen, J. 2014. Laadullinen tutkimus opinnäytetyönä. Jyväskylän ammattikorkeakoulu.
- Kekwaletswe, R., Lesole, T. 2016. Artikkelin ”A Framework for Improving Business Intelligence through Master Data Management”. Journal of South African Business Research.
- Kotter, J. 2012. Leading change. Boston: Harvard Business Review Press.
- Loshin, D. Master Data Management. 2009. Morgan Kaufmann Publishers.
- Maedche, A. 2010 "An ERP-centric Master Data Management Approach". AMCIS 2010 Proceedings. 384.
- Nokkala, T., Salmela, H., Toivonen, J. 2019. Tutkimus ” Data Governance in Digital Platforms”. Twenty-fifth Americas Conference on Information Systems, Cancun, 2019.
- Panian, Z. 2010. Some practical experiences in data governance. World Academy of Science, Engineering and Technology, 2010.
- Redman, T.C. 2013. Artikkelin Harvard Business Reviewissa (HBR) about ”Data’s Credibility Problem”. Viitattu 21.11.2019. <https://hbr.org/2013/12/datas-credibility-problem>.
- Sarsfield, S. 2009. The Data Governance Imperative. Cambridgeshire: IT Governance Publishing.
- Seiner, R. 2014. Non-Invasive Data Governance, The Path of Least Resistance and Greatest Success. Technics Publications.
- Seiner, R. 2018. Artikkelin ”Everybody is a Data Steward; Get Over It!”. Viitattu 4.12.2019. <https://tdan.com/everybody-is-a-data-steward-get-over-it/22637>.
- Seiner, R. 2019. Artikkelin ”Data Governance Roles and Responsibilities”. Viitattu 4.12.2019. <https://tdan.com/data-governance-roles-and-responsibilities/24774>.
- Seiner, R. 2019. Artikkelin ”Common Data Governance Challenges.” Viitattu 4.12.2019. <https://tdan.com/common-data-governance-challenges/25248>.
- Smith, H., McKeen, J. 2008. "Developments in Practice XXX: Master Data Management: Salvation or Snake Oil?" Communications of the Association for Information Systems: Vol. 23, Article 4.
- van der Staaij, A.J., de Goeij, A.S.M., van Unen, C.J.W.A., Swartjes, S., 2012. Artikkelin ” Master Data Management: dos & don’ts”. Viitattu 27.11.2019. <https://www.compact.nl/en/articles/master-data-management-dos-donts/>.
- Väre, T. 2019. Master Data. Helsinki: Alma Talent.
- Weber, K., Otto, B., Österle, H. 2009. One size does not fit all – A Contingency approach to data governance. ACM J. Data Inform. Quality 1, 1, Artikkelin 4, 2009.

Liitteet

Liite 1. Datan hallinnan kypsyysesti



The Data Governance test

This test enables you to perform a self-evaluation of your data governance programs. Testing in this way helps to focus on the thing meaningful to your organization while honestly assessing how well you address your organization's needs.

The goal of this exercise is to help you look inward to where positive aspects of your environment may be leveraged and where opportunities exist to improve governing data as a valued enterprise resource.

Before you start, please keep in mind the two questions:

1. In this day and age of increased complexities around regulatory compliance and reporting, information security, privacy, data classification, data integration, and complex transaction management, **does it make sense for us to continue governing our data as we always have or should we at least consider how to effectively formalize the way we manage these data resources?**
2. Given that competition in our industry is fierce and every company is looking for the data- and information- based edge for a competitive advantage, **does it make sense for us to continue governing our data as we always have, or should we at least consider formalizing how we manage our data resources?**

In grading (scale from 1 to 5) each of these statements for your organization, be candid when you evaluate your present situation in terms of these discipline statements.

By grading the importance of a specific topic to your business, you are able to set the priority where the efforts to improve should be targeted first.

Scoring partial is allowed e.g. 1,4 if you aren't sure whether to give 1 or 2.

Data Governance Disciplines	Score (1 to 5)	Importance to business (1 to 5)
<p>1. Risk Management/ Riskienhallinta: Hallitsemme dataan liittyviä riskejä. Organisaationi ymmärtää tarpeen sopeutua nopeasti dataan liittyviin riskeihin, ja monet näistä säännöistä tulevat organisaation ulkopuolelta. Meillä on henkilö/ henkilöstö tai neuvosto (tai kaikki edellä oleva), joka keskittyy ymmärtämään kaikki riskien tasot tietojen hallinnassa. Henkilö / henkilöstö / neuvosto välittää säännöllisesti tietoja tietoriskeistä niin, että kaikki ymmärtävät riskikäyttäytymisen ja turvallisen käytöksen suhteen tietojen käsittelyssä.</p>	1,33	4
<p>2. Data Compliance & Regulatory Control/ Tietojen yhdenmukaisuus ja sääntelyn valvonta: Organisaationa kiinnitämme paljon huomiota sääntöjen noudattamiseen ja sääntelyyn liittyviin huolenaiheisiin, jotka liittyvät tietoihin, joita keräämme, käyttämme ja jaamme osana päätöksentekoa ja liiketoimintaa. Joku on vastuussa sääntöjen dokumentoinnista ja toimittamisesta kaikille organisaation henkilöille, jotka käsittelevät näitä tietoja. Kun meitä auditoidaan, voimme osoittaa tarkastajille selvästi, että noudatamme tietoja koskevia sääntöjä.</p>	1	4
<p>3. Information Security & Data Classification/ Tietoturva ja tietojen luokittelu: Organisaationa kiinnitämme paljon huomiota kaikkien jäseneltyjen ja jäsentämättömien tietojen tietoturvaan. Meillä on tietoturvakäytäntö ja / tai jotain vastaavaa, esim. ohjeet, toimeksiannot. Olemme luottavaisia kykyymme kommunikoida, erotella ja hallita sääntöjä, jotka liittyvät erittäin luottamuksellisiin, sisäiseen käyttöön tarkoitettuihin ja julkisiin tietoihin. Ihmiset, jotka jakavat tietoja organisaatiossamme, jakavat myös tietoja koskevat dokumentoidut säännöt. Emme usko, että tietoturva on huolenaihe.</p>	2	4,67
<p>4. Metadata Management/ Metatietojen hallinta: Meillä on metatiedot tärkeimmistä hallitsemistamme tiedoista. Organisaationi tietää, mitä tietoja meillä on, missä ne sijaitsevat ja miten nämä tiedot määritellään, tuotetaan ja käytetään jaetuissa tietokannoissa ja ihmisten työasemilla. Tärkeimmistä tiedoistamme olevat tiedot ovat kaikkien tarvitsevien käytettävissä. Yhtä tärkeätä on, että olemme tunnustaneet ja sitouttaneet ne ihmiset, joilla on muodollinen vastuu metatietojen määrittelystä, tuottamisesta ja käytöstä.</p>	2	4,33
<p>5. Data Quality Management/ Tietojen laadunhallinta: Organisaatiomme keskittyy jatkuvasti tiedon laatuun. Meillä on muodolliset keinot tietojen laatuongelmien kirjaamiseksi, ja meillä on ennakoivia ja reagoivia menetelmiä ongelmien löytämiseen ja ratkaisemiseen, kun ne löydämme. Ja meillä on ihmisiä, jotka vastaavat poikkeuksien hallinnasta, arvottamisesta ja priorisoinnista. Tärkeintä on, että meillä on selkeä käsitys ydintietojen liiketoimintastandardeista, jotka helpottavat korkean laadun erottamista heikosta laadusta.</p>	1,33	4,67

Data Governance Disciplines	Score (1 to 5)	Importance to business (1 to 5)
<p>6. Business Intelligence & Data Integration/ Liiketoimintatieto ja Data integraatio: Meillä on tietovarasto ympäristö, joka hyödyntää täysin siinä olevia tietoja ja käyttää niitä täysimääräisesti. Tämä tarkoittaa, että ihmisillä on helppo pääsy tietoihin, he ymmärtävät tiedon ja auttavat meitä parantamaan jatkuvasti tiedon laatua. Ymmärrämme, että tiedonhallinnalla on tärkeä rooli tietovarastoinnin onnistumisessa tai epäonnistumisessa tiedon integroinnin yhteydessä. Ymmärrämme, että tietojen integrointi on vaikea kurinalaisuus. Mutta koska hallitsemme tietoja molemmin puolin - lähde ja kohde -, olemme tyytyväisiä liiketiedustelupalvelumme tehokkaaseen luonteeseen.</p>	1,67	4,67
<p>7. Master Data Management/ Ydintietojen hallinta: Organisaatiomme tiedostaa, että ydintietojen hallinta (MDM) on yksi tehokkaimmista ja tärkeimmistä datan aloista, joista nykyään puhutaan. Olemme tunnistanee ihmiset hallitsemaan MDM-aloitettamme ja olemme alkaneet myös tunnistamaan tekniikoita, jotka auttavat meitä hallitsemaan ja jakamaan ydin- ja referenssitietojamme. Kun täydennämme MDM-ympäristöämme, päätöksentekoa hallitaan kurinalaisesti isäntä- ja resurssijärjestelmän, metatietokomponentin sekä ydintietojen viestinnän ja saatavuuden suhteen. Olemme hyvässä asemassa ydintietoaloitteen saattamiseksi päätökseen budjetissa ja aikataulussa.</p>	2	5
<p>8. Data Governance & Data Stewardship/ Tietohallinto ja tiedonhallinta: Viimeisenä, mutta ei vähäisimpänä, meillä on tiedonhallintaohjelma, joka määrittelee selkeästi roolit ja vastuut operatiivisella, taktisella, strategisella ja tukitasolla. Ohjelmamme keskittyy hyödyntämään tietovastaavien nykyistä tietämystä tiedoista. Johtajamme, tietovastaavat, liiketoiminta- ja teknologiahenkilöt ovat omaksuneet valintamme lähestymistapaan, ja siinä käsitellään datan hallintaa ennakoivasti ja reagoivasti. Tietojenhallintaohjelmamme on ensisijainen tekijä menestyksessämme kaikilla tässä testissä luetelluilla aloilla.</p>	1,33	4

Liite 3. Asiakasrekisterin Datastandardi

Kenttä	Luokittelu	Määrittelmä	MAASTER		EHDOKSESI		Oletus	Muut säännöt (ehdollisesti pakollinen yms.)
			Data	Pakollisuus	pakollinen	Ehdollisuus		
Perustiedot								
Tunnus		tunnus	VARCHAR2[12 CHAR]		Asiakastunnus		104764	
Ryhmä		ryhmä	VARCHAR2[8 CHAR]	x			1	
ALV-tunnus		vattunnus	VARCHAR2[20 CHAR]	x	ALV tunnus		FI2345601	
Y-tunnus		ytunnus	VARCHAR2[11 CHAR]	x	Y tunnus		123456-01	
Nimi		nimi1	VARCHAR2[40 CHAR]	x	Asiakas A			Kirjoitus luten asiakkaan verkkosivulla tai muussa dokumentaatioissa
Postiosoite		katu1	VARCHAR2[40 CHAR]	x	Jalostamontie 1			
Postinumero		postiosoite	VARCHAR2[10 CHAR]	x	42300			
Kaupunki		postiosoite	VARCHAR2[40 CHAR]	x	JAMSAANKOSKI			
Maa		maakodi	VARCHAR2[3 CHAR]	x	FI			Referenssilähtöä kotimaan osalta
Maa (nimi)		maanimi	VARCHAR2[16 CHAR]	x	FINLAND			Referenssilähtöä
Yhteyshenkilö		yhteyshenkilö	VARCHAR2[30 CHAR]	x	Artti Asiakas			Kirjoitus: Etuimi Sukuimi
Puhelin		puhelin1	VARCHAR2[30 CHAR]	x	+358 40 123 4567			
Sähköposti		email	VARCHAR2[60 CHAR]	x	artti.asiakas@hotmail.com			
Tunnus		aztunnus	VARCHAR2[12 CHAR]	x	104765			
Nimi		aznimi1	VARCHAR2[40 CHAR]	x	Asiakas A Varasto			
Postiosoite		azkatu1	VARCHAR2[40 CHAR]	x	Jalostamontie1B			
Postinumero		azpostioinro	VARCHAR2[10 CHAR]	x	42300			
Kaupunki		azpostiosoite	VARCHAR2[40 CHAR]	x	JAMSAANKOSKI			Referenssilähtöä kotimaan osalta
Maa		azmaakodi	VARCHAR2[3 CHAR]	x	FI			Referenssilähtöä
Maa (nimi)		azmaanimi	VARCHAR2[16 CHAR]	x	FINLAND			Kirjoitus: Etuimi Sukuimi
Yhteyshenkilö		yhteyshenkilö	VARCHAR2[30 CHAR]	x	Ville Varastopäällikkö			
Puhelin		puhelin1	VARCHAR2[30 CHAR]	x	+358 40 123 9876			
Sähköposti		email	VARCHAR2[60 CHAR]	x	ville.varastopaallikko@hotmail.com			
Maksuehto		mekdi	VARCHAR2[3 CHAR]	x	1			
Toimitusohje		tekdi	VARCHAR2[3 CHAR]	x	2			
Toimitustapa		tskdi	VARCHAR2[3 CHAR]	x	2			
Sopimusnumero		sopimusnumero	VARCHAR2[20 CHAR]	x	251659			
Verkkolaskutusnumero		verkkolaskutusnumero	VARCHAR2[35 CHAR]	x				
Verkkolaskuoperaattori		vloger	VARCHAR2[20 CHAR]	x				
Välittäjän tunnus		kat1	VARCHAR2[20 CHAR]	x				
Laskutus								
Luottoraja		luottoraja	FLOAT					
Luottokielto		luottokielto	VARCHAR2[1 CHAR]	x				
Sähköpostilasku		spostilasku	VARCHAR2[1 CHAR]	x				
Sähköp. jakun osoite		spostilaskuemail	VARCHAR2[60 CHAR]	x	mailja.makela@maa.fi			
Perintä s. postit		spostinumerot	VARCHAR2[60 CHAR]	x				
Laskun väkioikeus1		mlasvakiotekst1	VARCHAR2[8 CHAR]	x				
Laskun väkioikeus2		mlvakiotekst2	VARCHAR2[8 CHAR]	x				
Valuutta		valuutta	VARCHAR2[3 CHAR]	x	EUR			
Luokittelu								
Kielikodi		kielikodi	VARCHAR2[3 CHAR]	x				
Myyntitili		myyntitili	VARCHAR2[8 CHAR]	x				
Myyntisaamistili		myyntisaamistili	VARCHAR2[8 CHAR]	x				

Liite 4. Toimittajarekisterin Datastandardi

Kenttä	Perustiedot	Taulunimi	Tyyppi	Määrittelmä	Dataa	Pakollisuus	pakollinen	Ehdollisuus	Esimerkki	Oletus	Muut säännöt (ehdöllisesti pakollinen yms.)
Tunnus		tunnus	VARCHAR(12 CHAR)	Toimittajätunnus	x	x			201683		
Ryhmä		ryhma	VARCHAR(28 CHAR)		x	x			2		Uusi kenttä!
Ryhmä2		ryhma2	VARCHAR(28 CHAR)	Toimittajaluokittelu	x	x			BE0534825336		Kotimaisille toimittajille
ALV-tunnus		alv-tunnus	VARCHAR(20 CHAR)	ALV-tunnus	x	x			123456-01		Sierretään kerättään Ryhmä2 (referenssilista, vähentää virheiden mahdollisuutta)
Y-tunnus		y-tunnus	VARCHAR(211 CHAR)	Y-tunnus	x	(x)	x		AERO		Sertifikaatin voimaosion päättymislähteen → siirto hyöppöön kerättään?
Vapaa-arvot		vapav1	VARCHAR(20 CHAR)	Toimittajaluokittelu	x				2020-08		Kryptotassu kuten asiakkaan verkkosivulla tai muussa dokumentaatioissa
		vapav2	VARCHAR(20 CHAR)	AERO-hyöppörynnin voimaosio	x				Precision Metals EU		
Nimi		nimi1	VARCHAR(40 CHAR)	Toimittajan nimi	x	x			Wayenborgstraat 25		
Postiosoite		katu1	VARCHAR(40 CHAR)	Toimittajan postiosoite	x	x			B-2800		
Postinumero		postinro	VARCHAR(10 CHAR)	Toimittajan postinumero	x	x			Mechelen		Referenssilistalta kotimaan osalta
Kaupunki		postios	VARCHAR(40 CHAR)	Toimittajan postitoimipaikka	x	x			BE		Referenssilistalta
Maa		maakodi	VARCHAR(23 CHAR)	Toimittajan maakoodi	x	x			BELGIUM		Kryptotassu
Maa (nimi)		maan nimi	VARCHAR(216 CHAR)	Toimittajan maa	x	x			Stef Florenty		Referenssilistasta
Yhteyshenkilö		yhteyshenkilö	VARCHAR(30 CHAR)	Toimittajan yhteyshenkilö	x	x			+32(0) 15 44 89 89		Kryptotassu: Etunimi Sukunimi
Puhelin		puhelin1	VARCHAR(230 CHAR)	Ensisijainen puhelinnumero	x	x			sales@precision-metals.be		
Sähköposti		email	VARCHAR(260 CHAR)	sähköpostiosoite	x	x					
Hyväksytyt		hyvaksyty	VARCHAR(21 CHAR)	AERO-hyöppörynnin voimaosio	x	x					Uusi kenttä!
Pvm		hyoppm	DATE		x						
Ennakkoperintärekisteröinti voimassa		egruppm	DATE		x						
Maksumerkki		mekodi	VARCHAR(23 CHAR)		x	x			1		
Toimintusielto		tekd1	VARCHAR(23 CHAR)		x	x			2		
Toimintusgaa		thkd1	VARCHAR(23 CHAR)		x	x			1		
Laskutus											
Luottokielto		luottokielto	VARCHAR(21 CHAR)		x						
Viestityyppi		viestityyppi	VARCHAR(21 CHAR)		x	(x)	x				
Jakuluista		jakolustus	VARCHAR(20 CHAR)		x						
Ulkomaan kulut		ulkokulut	VARCHAR(21 CHAR)		x						Jaeattu
Luokittelu											
Ostotili		ostotili	VARCHAR(28 CHAR)		x						
Ostovelkaili		ostovelkaili	VARCHAR(28 CHAR)		x						
Kieli		kielikodi	VARCHAR(23 CHAR)		x				EN		
Valuutta		valuutta	VARCHAR(23 CHAR)		x						
Tilauksen valitsemiseksi		otilauksenvalitsemiseksi	VARCHAR(28 CHAR)		x				OTT03		Voidaan lisätä myös Info-ikkunaan, jolloin näkyy ostotilauksen luodessa (pop-up ikkuna)
Vapaat tekstikentät		vtext1	VARCHAR(240 CHAR)		x						
		vtext2	VARCHAR(240 CHAR)		x						
		vtext3	VARCHAR(240 CHAR)		x						
		vtext4	VARCHAR(240 CHAR)		x						
Pankit (alavälilehti)											
Tilin tyyppi		acttype			x						
Pankin tilinumero		account			x						
Oletustili		defaultaccount			x						
Maksusivulieto		disabled			x						
Maksun tyyppi		btype			x						
BIC		swift			x						