

## **Keskitetyn ydintiedonhallinnan merkitys yrityksen tietojärjestelmäarkkitehtuurin näkökulmasta**

Riitta Laitinen



<b>Tekijä(t)</b> Riitta Laitinen	
<b>Koulutusohjelma</b> Tietojenkäsittelyn koulutusohjelma	
<b>Raportin/Opinnäytetyön nimi</b> Keskitetyn ydintiedonhallinnan merkitys yrityksen tieto- ja järjestelmäarkkitehtuurin näkökulmasta	<b>Sivu- ja liitesivumäärä</b> 38
<p>Opinnäytetyössä selvitetään keskitetyn ydintiedonhallinnan merkitystä yrityksen tieto- ja järjestelmäarkkitehtuurin näkökulmasta. Opinnäytetyön tavoitteena on lisätä sekä kirjoittajan omaa, että mahdollisten lukijoiden ymmärrystä ydintiedonhallinnan merkityksestä yritykselle, ja sen liiketoiminnalle. Lisäksi etsitään vastauksia kysymyksiin: onko keskitetty ydintiedonhallinta parempi kuin hajautettu ydintiedonhallinta, vai löytyykö paras ratkaisu lopulta jostain näiden kahden välimaastosta, niin sanotusta hybridimallista? Ja toisaalta, mikä on keskitetyn ydintiedonhallinnan merkitys yritykselle hyötyjen ja haittojen näkökulmasta?</p> <p>Työssä ei suunnitella, määritellä eikä mallinneta ydintiedonhallintaa tai siihen liittyviä järjestelmiä minkään yrityksen tarpeisiin tai näkökulmasta. Työssä ei käsitellä kokonaisarkkitehtuuria kokonaisuudessaan, vaan keskitytään tieto- ja järjestelmäarkkitehtuuri näkökulmiin. Tutkimus / selvitys tehdään perehtymällä alan kirjallisuuteen, julkaisuihin ja artikkeleihin.</p> <p>Työssä kerrotaan ensin, miten tieto voi merkitä dataa, informaatiota, tietämystä, ymmärrystä ja viisautta. Ydintiedoksi mielletään usein muun muassa yrityksen asiakas-, tuote-, sopimus-, ja toimittajatiedot, mutta viime kädessä ydintiedon määrittäminen riippuu siitä, mitä yritys itse mieltää omaksi ydintiedokseen. Työssä tarkistellaan myös tieto- ja järjestelmäarkkitehtuurin sisältöä, ja kuvataan niihin liittyviä keskeisiä käsitteitä.</p> <p>Työn keskiössä on ydintiedonhallinta, sen taustat, sisältö, ja suhde tietojärjestelmiin. Laatu on saanut opinnäytetyössä oman alaluvun, onhan laadun parantaminen alan asiantuntijoiden mukaan merkittävin syy yrityksen ydintiedonhallinnan määrittelylle, suunnittelulle, ja ratkaisujen kehittämiselle.</p> <p>Tietosisällön jälkeen tarkastellaan ydintietoarkkitehtuuria erilaisten mallien avulla. Mallit voidaan jakaa kolmeen luokkaan: keskitettyihin malleihin, hajautettuihin malleihin, ja hybridimalleihin. Näiden mallien avulla hahmotellaan myös keskitetyn ydintiedonhallinnan hyötyjä ja haittoja.</p> <p>Keskitetystä ydintiedonhallinnasta on yritykselle selviä hyötyjä. Näitä ovat muun muassa parantunut tuottavuus, kilpailukyky, laatu, asiakastuntemus, ja asiakaspalvelu, sekä kustannussäästöt. Haasteiksi koetaan lähinnä: joustamattomuus, hitaus, ylläpito, sekä kokonaisuuden monimutkaisuus / laajuus. Hajautettuun malliin verrattuna keskitetty ydintiedonhallinta on teoriassa selvästi parempi vaihtoehto. Sitä, voisiko keskitetyn mallin haittoja lieventää tekemällä siihen joitain perusteltuja poikkeuksia, jolloin se saisi hybridimallin piirteitä, tulee arvioida tapauskohtaisesti.</p>	
<b>Asiasanat</b> ydintieto, ydintiedon hallinta, tietoarkkitehtuuri, tietojärjestelmäarkkitehtuuri	

## Sisällys

1	Johdanto .....	1
2	Tieto ja ydintieto .....	3
2.1	Mitä on tieto? .....	3
2.2	Mitä on ydintieto? .....	4
3	Tieto- ja tietojärjestelmäarkkitehtuuri osana yrityksen kokonaisarkkitehtuuria .....	6
3.1	Tietoarkkitehtuuri .....	7
3.2	Tietojärjestelmäarkkitehtuuri .....	10
4	Yrityksen ydintiedonhallinta ja laatu .....	12
4.1	Taustaa .....	13
4.2	Ydintiedonhallinnan käytännöt .....	14
4.3	Ydintiedonhallinta ja tietojärjestelmät .....	15
4.4	Laatu kantavana ajatuksena .....	17
5	Ydintietoarkkitehtuurin vaihtoehtoiset mallit .....	19
5.1	Keskitetty malli .....	20
5.2	Hajautettu malli .....	22
5.3	Hybridimalli .....	25
6	Keskitetyn ydintiedonhallinnan hyödyt ja haitat yritykselle .....	30
6.1	Hyödyt .....	30
6.2	Haitat .....	32
7	Pohdinta .....	34
	Lähteet .....	37

# 1 Johdanto

Valitsin opinnäytetyöni aiheeksi yrityksen ydintiedonhallinnan, ja sen keskittämisen merkityksen tieto- ja tietojärjestelmäarkkitehtuurien näkökulmasta. Aiheen valinta johtuu omasta kiinnostuksestani aiheeseen, sillä olen viimeisen viidentoista vuoden aikana huomannut sen nousevan esiin eri yrityksissä ja eri yhteyksissä, tavallisimmin kehityshankkeissa. Ydintiedonhallinnan merkityksestä ja sen vaikutuksista käydään parhaillaan runsaasti keskustelua Business Intelligence (BI), ja tiedolla johtamisen teemojen yhteydessä. Aiheeseen törmätään niin ikään sähköisiä palveluita kehitettäessä, sekä laajoissa toiminnan-, talouden-, ja asiakkuudenhallinnan hankkeissa, missä järjestelmiä integroidaan, ja etsitään mahdollisuuksia kasvattaa yrityksen kilpailukykyä muun muassa kustannustehokkuuden ja paremman asiakaspalvelun kautta.

Ydintiedonhallinta on tullut itselleni tutuksi erityisesti asiakas- ja tuotetietojenhallinnan parissa työskennellessäni, toisaalta yrityksen johdon raportointia kehittäessäni. Arjen tilanteissa ydintiedossa esiintyviin virheisiin ja haasteisiin suhtaudutaan hyvin ratkaisukeskeisesti, toisin sanoen virhe etsitään ja virhe korjataan. Aika ajoin jään kuitenkin pohtimaan, mistä virheet johtuvat, ja miten niiden syntyminen voitaisiin estää. Ajatuksiin nousee malli keskitetystä ydintiedonhallinnasta. Mallista, missä yritykselle tärkeä ja yrityksen toiminoissa toistuva tieto ylläpidettäisiin joissain yhdessä paikassa, muutaman henkilön toimesta, ja jaettaisiin sieltä kaikkien käyttöön. Kuulostaa yksinkertaiselta – ehkä idealistiseltakin.

Yleisesti ottaen tuo ajatus keskitetystä ydintiedonhallinnasta lienee oikean suuntainen, mutta epäilen että tässäkin asiassa yksinkertaisuus katoaa, kun lisätään tarkasteluun muun muassa yrityksen käyttämät tiedot, tietojärjestelmät ja teknologiat, sekä liiketoiminnalliset tavoitteet, käyttäjiä unohtamatta. Mitä keskitettyä ydintiedonhallintaa suunnittelevan yrityksen olisi hyvä ymmärtää, ja onko keskitetty ydintiedonhallinta yhtään sen parempi kuin hajautettu ydintiedonhallinta? Löytyykö ratkaisu lopulta jostain näiden kahden välimaastosta, niin sanotusta hybridimallista?

Opinnäytetyö koostuu johdannon ja pohdinnan lisäksi seuraavista luvuista: tieto ja ydintieto, tieto- ja tietojärjestelmäarkkitehtuuri osana yrityksen kokonaisarkkitehtuuria, yrityksen ydintiedonhallinta ja laatu, ydintietoarkkitehtuurin vaihtoehtoiset mallit, sekä keskitetyn ydintiedonhallinnan hyödyt ja haitat yritykselle.

Johdannon jälkeen, toisessa luvussa kuvaan mitä tieto ja ydintieto ovat, ja miten ne tulisi ymmärtää. Kolmannessa, tieto- ja tietojärjestelmäarkkitehtuuri osana yrityksen kokonaisarkkitehtuuria, luvussa kuvaan neljästä kokonaisarkkitehtuurin näkökulmasta tarkemmin kahta, jotka usein nousevat esiin ydintiedonhallinnan yhteydessä. Neljäs luku, ydintiedonhallinta ja laatu, keskittyy ydintiedonhallinnan taustoihin, käytäntöihin, tietojärjestelmiin, sekä laatuun, mikä näyttäisi olevan koko ydintiedonhallinnan ydin. Viidennessä luvussa nostan esiin alan asiantuntijoiden määrittelemiä vaihtoehtoisia malleja ydintietoarkkitehtuurille, ja kuudennessa, eli viimeisessä luvussa ennen pohdintaa, koostan asiantuntijoilta saadut näkökulmat keskitetyn ydintiedonhallinnan hyödyt ja haitat yritykselle alaluvuiksi.

Opinnäytetyön tavoitteena on lisätä omaa ja mahdollisien lukijoiden ymmärrystä ydintiedonhallinnan merkityksestä yritykselle, ja sen liiketoiminnalle. Lisäksi etsitään vastauksia kysymyksiin: onko keskitetty ydintiedonhallinta parempi kuin hajautettu hallintamalli, sekä mikä on keskitetyn ydintiedonhallinnan merkitys yritykselle, hyötyjen ja haittojen näkökulmasta.

Työssä ei suunnitella, määritellä eikä mallinneta ydintiedonhallintaa tai siihen liittyviä järjestelmiä erityisesti minkään yrityksen tarpeisiin tai näkökulmasta. Työssä ei tulla käsittelemään kokonaisarkkitehtuuria kokonaisuudessaan. Toiminta-arkkitehtuuri, ja teknologiaarkkitehtuuri näkökulmat jäävät vain maininnoiksi, mutta tieto- ja tietojärjestelmäarkkitehtuuri näkökulmia tarkistellaan sen sijaan laajemmin. Keskitetyn ydintiedonhallinnan hyötyjä ja haittoja koostettaessa keskitytään kuvaamaan sitä, mitä hyötyä ja haittaa keskitetystä ydintiedonhallinnasta on yritykselle, ei niinkään enää verrata keskitettyä mallia hajautettuun tai hybridimalliin.

Opinnäytetyössä pyritään kuvaamaan ydintiedonhallintaa yrityksen tieto- ja tietojärjestelmäarkkitehtuurin näkökulmasta, sekä ymmärtämään keskitetyn ydintiedonhallinnan mahdolliset hyödyt ja haitat yritykselle. Jotta kuvaus tarjoaisi riittävästi näkökulmia, työssä esitellään myös vaihtoehtoja keskitetyille hallintamallille. Tutkimus tehdään perehtymällä alan kirjallisuuteen, julkaisuihin ja artikkeleihin.

Opinnäytetyössä esiintyvät termit ja käsitteet pyritään selittämään tarkemmin silloin kun ne mainitaan työssä ensimmäisen kerran.

## 2 Tieto ja ydintieto

### 2.1 Mitä on tieto?

Suomenkielessä sanalla tieto on monta merkitystä. Ackoffin (1999, 1) mukaan tieto voidaan määritellä / luokitella viidellä tavalla: dataksi, informaatioksi, tietämykseksi (knowledge), ymmärrykseksi (understanding) ja viisaudeksi (wisdom). Data on symboleja, numeroita tai muita faktojen ilmentymiä, ja on sinällään jokapäiväisissä liiketoiminnoissa sekä asianmukaisessa päätöksenteossa tarvittava informaation raaka-aine (Cleven & Wortmann 2010, 1). Data on kokoelma arvoja tai faktoja mitä käytetään laskemiseen, päättelyyn ja mittaamiseen. Se on kerättyä, varastoitua tai prosessoitua, mutta sitä ei ole liitetty mihinkään asiayhteyteen, josta sen merkityksen voisi päätellä. (Loshin & Powell 2003, 7.)

Informaatio on datan jalostamista joksikin hyödylliseksi, ja se vastaa kysymyksiin: kuka, mitä, missä, milloin, ja kuinka monta. Informaatio saa merkityksen sen perusteella, miten se on yhdistetty datasta. Informaatio voi näin ollen olla hyödyllistä, tai olla olematta. (Ackoff 1999, 1.) Informaatio on lopputulos kootusta ja järjestelystä datasta. Dataa on yhdistetty toisiinsa, mitä kautta sille on tullut asiayhteys ja tarkoitus. (Loshin & Powell 2003, 8.)

Tietämys on datan ja informaation hyödyntämistä, ja se vastaa kysymykseen: miten. Tietämys on prosessi, mikä perustuu muistin ja informaation yhdistämiseen. (Ackoff 1999, 1.) Tietämys on informaation ymmärtämistä, ja siitä näkemysten saamista (Loshin & Powell 2003, 8).

Ymmärrys antaa vastauksen kysymykseen: miksi. Se on kognitiivinen ja analyyttinen prosessi, missä tietämyksestä voidaan johtaa uutta tietämystä. Tietämyksen ja ymmärryksen ero on sama kuin muistamisen ja oppimisen ero. (Ackoff 1999, 1.)

Viisaus on arvioitua ymmärrystä, ymmärryksen jalostuneempi muoto, mihin liittyy muun muassa henkilön moraaliset ja eettiset käsitykset. Se auttaa meitä ymmärtämään sitä mitä emme ole aikaisemmin ymmärtäneet. (Ackoff 1999, 1.)

Entäpä sitten älykkyys? Informaatio, tietämys ja ymmärtäminen auttavat meitä olemaan tehokkaita (efficiency), mutta ei vaikuttavia (effectiveness). Älykkyys on kykyä lisätä tehokkuutta, kun taas viisaus on kykyä lisätä vaikuttavuutta. Järjestelmät ja laitteet voivat olla

älykkäitä, eli tehokkaita, mutta eivät viisaita, eli vaikuttavia. Viisaus käsittelee arvoja. Se sisältää arviointia, ja on aina sidoksissa arvioinnin tekijään. (Ackoff 1999, 1.)

Järjestelmät voidaan siis ohjelmoida käsittelemään dataa ja informaatiota, ne voidaan rakentaa ymmärtäviksi tietämuskannoiksi, mutta niistä tuskin saa luotua viisaita. Voidaankin ajatella, että tämä viides määrittely / luokka on ainoa, mikä erottaa ihmisen koneesta. (Ackoff 1999, 2.)

## 2.2 Mitä on ydintieto?

Haug & Arlbjørnin (2011, 289) mukaan yrityksen data voidaan jakaa kolmeen luokkaan: ydintietoon (master data), liiketoimintatietoon, kuten tilaukset ja laskut, sekä historiatietoon. Ydintiedon erottelu muusta datasta ei ole aina yksinkertaista datan monimerkityksellisyyden vuoksi. Tutkijat ja toiminnanharjoittajat ovat kuitenkin yhtä mieltä siitä, että ydintieto on dataa liiketoiminnan avain asioista, ja nämä tiedot / käsitteet tulisi määrittellä yksiselitteisesti ja yhdenmukaisesti yrityksen sisällä. (Otto 2012, 339.)

Yrityksissä on yleisesti tunnistettuja käsitteitä, jotka muodostavat perustan liiketoimintaprosesseille. Tällaisia käsitteitä ovat esimerkiksi: asiakas, tuote, nimike, sijainti, toimittaja, myyjä, työntekijä, projekti ja sopimus. Nämä käsitteet esiintyvät ja toistuvat yrityksen eri liiketoimintoja tukevilla järjestelmissä. Tiedot, jotka liittyvät edellä mainittuihin käsitteisiin, muodostavat: metadatoineen, attribuutteineen, määrittämineen, rooleineen, liittymineen ja luokitteluneen yrityksen ydintiedon. Ydintieto on se yrityksen kaikkein tärkein tieto, mitä liiketoimintajärjestelmät hyödyntävät, ja mihin mittarit ja raportointi perustuvat. (Loshin 2009, 6.) Ydintieto luo pohjan sekä operatiiviselle arvonn tuottamiselle, että analyyttiselle päätöksenteolle (Ofner, Straub, Otto & Oesterle 2013, 474).

Ydintiedolla viitataan siis ydinliiketoiminnan käsitteisiin, joita käytetään yrityksen eri liiketoiminta-alueiden järjestelmissä ja prosesseissa jatkuvasti. Cleven & Wortmannin (2010, 2) mukaan seuraavat ominaisuudet ovat tunnusomaisia ydintiedoille:

- Riippumaton olemassaolo, eli ne voivat esiintyä ilman muita tietoja.
- Eivät muutu usein, eli ne ovat pysyvämpiä operatiiviseen dataan verrattuna.
- Niiden määrä on tasainen esimerkiksi operatiiviseen dataan verrattuna.

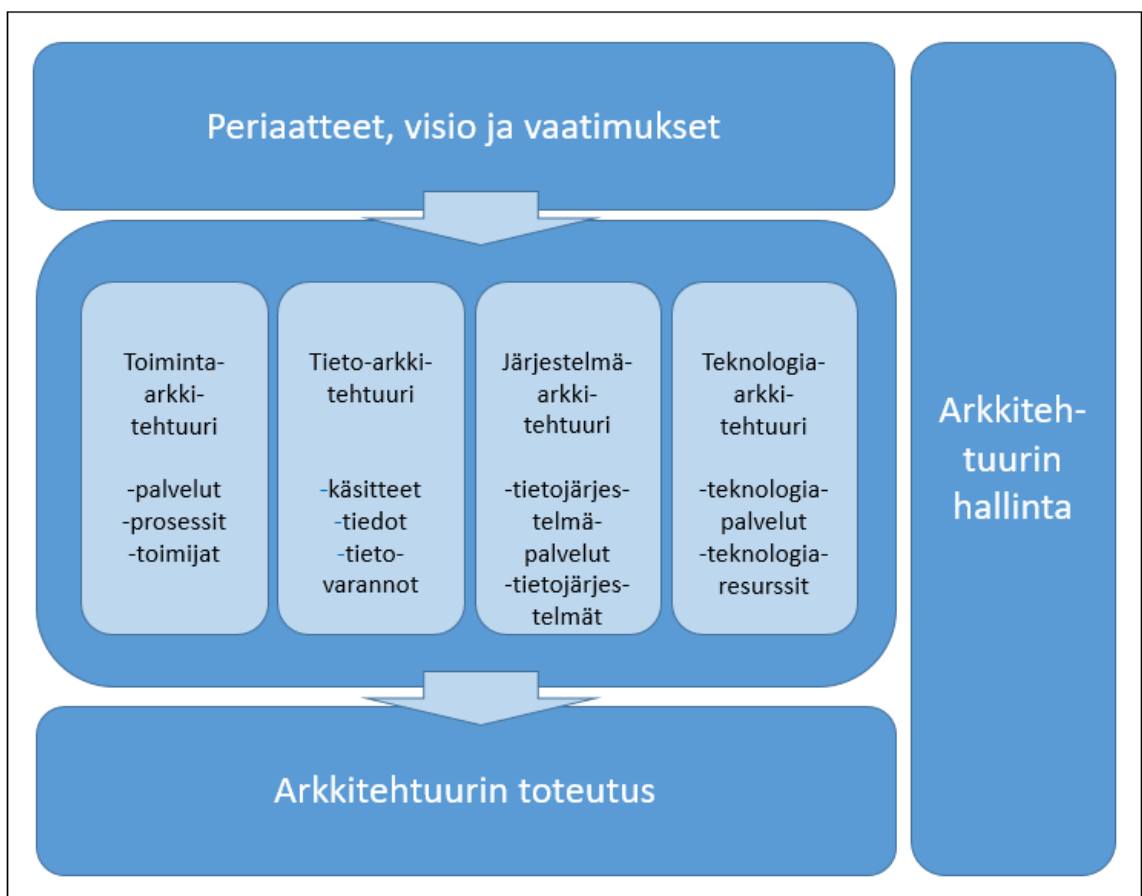
Ydintiedot edustavat yrityksen avain liiketoimintaa (Otto 2012, 338). Liiketoimintaprosessit tukeutuvat ydintietoihin, ja jos niissä on virheellisyyksiä liiketoimet eivät täytä niille asetettuja tavoitteita (Haug & Arlbjørn 2010, 289). Ydintietoja pitäisi käyttää yhdenmukaisesti

läpi yrityksen eri toimintojen. Ydintiedon tehtävä on vastata yrityksen strategisiin vaatimuksiin. (Otto 2012, 338.) Ne ovat kriittisiä liiketoiminnan kannalta, ja tulisi tästä syystä nähdä strategisena yritysvarallisuutena (Niemi & Kontra 2012, 1).



### 3 Tieto- ja tietojärjestelmäarkkitehtuuri osana yrityksen kokonaisarkkitehtuuria

JUHTA suosituksen (2017, 3) mukaan kokonaisarkkitehtuuri on toiminnan, prosessien ja palvelujen, tietojen, tietojärjestelmien ja niiden tuottamien palvelujen muodostama kokonaisuus. Kokonaisarkkitehtuurin tarkoitus on muodostaa yhteinen ymmärrys toimijoille, joiden on tarkoitus hallitusti kehittää ja muuttaa sitä. Kokonaisarkkitehtuuria hyödynnetään yrityksen strategian jalkauttamisessa ja johtamisessa, toiminnan ja palvelujen kehittämisessä, muutoshallinnassa, sekä toimintoja ja palveluja digitalisoitaessa. Se on merkittävä osa yrityksen strategiatyötä, johtamisprosessia, sekä talouden ja toiminnan suunnittelua. (JUHTA 2017, 3).



Kuva 1. Arkkitehtuurisisällön viitekehys (mukailtu JUHTA 2017, 28)

Kokonaisarkkitehtuurin tavoite on auttaa eri toimijoita ymmärtämään paremmin toistensa toimintaa, mahdollistaa yhteistyötä, sekä parantaa prosessien ja tietojärjestelmien yhteen toimivuutta. Kokonaisarkkitehtuuri tukee riskienhallintaa, tietoturvaa, sekä laadunhallintaa (JUHTA 2017, 4.), ja näitä pohdittaessa luottamuksellisuuden, eheyden ja käytettävyyden merkitys on keskeinen (JUHTA 2017, 76).

Yrityksen kokonaisarkkitehtuuri (ks. kuva 1) muodostuu toiminta-arkkitehtuurista, tietoarkkitehtuurista, tietojärjestelmäarkkitehtuurista, sekä teknologia-arkkitehtuurista (JUHTA 2017, 23).

Näistä kahta, tietoarkkitehtuuria, ja tietojärjestelmäarkkitehtuuria, kuvataan seuraavaksi tarkemmin.

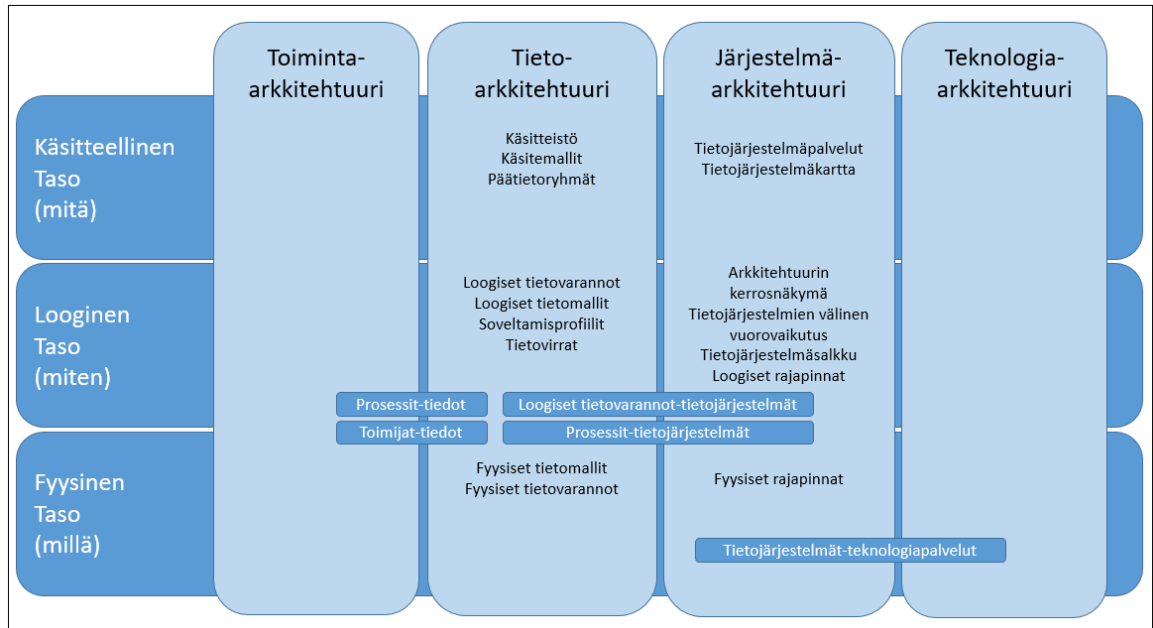
### **3.1 Tietoarkkitehtuuri**

Tietoarkkitehtuuri on yksi neljästä kokonaisarkkitehtuurin näkökulmasta. Tässä näkökulmassa keskitytään tarkastelemaan yrityksen tietotarpeita, tietopääomaa, tietojen välisiä suhteita, tietojen rakenteita, sekä tiedon organisointia ja hallintaa (JUHTA 2017, 18).

Tietoarkkitehtuurin tarkoituksena on kuvata yrityksen keskeinen tietosisältö yli organisaatio- ja tietojärjestelmärajoiden (Hovi 2009 13). Tietoarkkitehtuurin suunnittelun keskeisenä tavoitteena on tuottaa yhteinen näkemys keskeisestä tietopääomasta, sekä helpottaa tiedon keräämistä, välittämistä ja hallinnointia yritystasolla (JUHTA 2017, 18). Kun tiedot kuvataan ja määritellään riittävällä tarkkuudella, aletaan puhua yhteistä kieltä, ja vältetään usein kalliiksi tulevilta väärinkäsityksiltä (Hovi 2009, 13). Tavoitteena on tiedon helppo löydettävyys ja käytettävyys eri tarkoituksiin (JUHTA 2017, 18).

Tietoarkkitehtuuri koostuu käsitteiden määrittelystä ja standardoinnista, tietorakenteiden kuvaamisesta ja vakioinnista, tietojen hallinnasta ja varastoinnista, sekä tietojen yhteiskäyttöisyydestä yrityksen eri liiketoiminta-alueiden / osastojen välillä (JUHTA 2017, 59). Tietoarkkitehtuuri luo kokonaiskuvan yrityksen tiedoista, mahdollistaa tietovarastojen suunnittelun ja kehittämisen, edistää raportoinnin ja Business Intelligence -ratkaisujen kehittämistä, tarkentaa tarjouspyyntöjen luontia, sekä helpottaa tietojärjestelmien kehittämistä ja integrointia (Hovi 2009, 13).

Tietoarkkitehtuurin voidaan katsoa rakentuvan eri tasoista: käsitteellisestä tasosta, loogisesta tasosta, sekä fyysisestä tasosta (JUHTA 2017, 29).



Kuva 2. Arkkitehtuuritasot (mukailtu JUHTA 2017, 33)

Kuten kuvassa 2 esitetään, käsitteellinen taso vastaa kysymykseen, mitä, ja sen piiriin kuuluvat: käsitteet, käsittemallit, ja päätieto- eli ydintietoryhmät (JUHTA 2017, 29). Tieto-arkkitehtuurin käsitteellisellä tasolla keskitytään kuvaamaan mitä tietoa yrityksessä tarvitaan, käytetään, tuotetaan ja miten tiedot liittyvät muihin tietoihin. Käsitteistön selvittäminen aloitetaan esimerkiksi tarkastelemalla yrityksen, tai sen kehitettävän alueen keskeisiä prosesseja ja palveluita. Tulee selvittää mitä tietoja prosessi tai palvelu tarvitsee toimiakseen, sekä ne tiedot mitä prosessi tai palvelu tuottaa. Lähtökohdaksi voidaan prosessin tai palvelun sijaan ottaa myös yrityksen liiketoimintaa tukevat tietojärjestelmät, ja selvittää mitä tietoja ne sisältävät ja käsittelevät. (JUHTA 2017, 59.)

Käsitteistöstä tuotetaan käsitemalleja, joita voidaan kuvata eri tasoilla ja erilaisia kohde-ryhmiä silmällä pitäen. Ylätasolla käsitemalli voi olla vaikkapa kuvaus yrityksen toiminnalle kriittisestä tiedosta, kuvaten ydintietojen väliset suhteet. Tarkemmalla tasolla käsitemalli voi olla kuvaus, mitä käytetään osana tietojärjestelmän vaatimusmäärittelyä. Kaikista tarkimman tason kuvauksia käsitteiden välisistä suhteista kutsutaan tietomalleiksi. (JUHTA 2017, 60.)

Tiedoista voidaan muodostaa ydintietoryhmiä, joiden avulla on mahdollista tunnistaa tietovarantoja, lajitella, luokitella ja nimetä yrityksen tai kehitettävän alueen tietoja, sekä luoda niistä ylätasoinen tietomalli ja hallittava hierarkia. Myös ydintietoryhmät auttavat tietojen suhteiden selvittämisessä ja määrittelyssä. Ydintietoryhmille ei ole yleistä standardia, vaan ne ovat aina määritetty yrityksen, tai tarkasteltavan kohteen tietotarpeiden mukaisesti. Ydin-

tietoryhmien selvittäminen kannattaa aloittaa prosessiin tulevista ja lähtevistä tietovirroista, ja ne voidaan yhdistää niitä käyttäviin tietojärjestelmiin. Ydintietoryhmien kuvaaminen auttaa ymmärtämään mitkä ovat yrityksen keskeisimmät tiedot. (JUHTA 2017, 60.)

Tietoarkkitehtuurin käsitteellisen tason ymmärtäminen auttaa tunnistamaan ja välttämään päällekkäisyyksiä. Kuvauksia tehtäessä voidaan nähdä, että esimerkiksi toimittajatiedot ovat jo olemassa, joten niitä ei tarvitse luoda joka järjestelmään uudelleen. Keskiöön nousee tietojenhallinta ja se, että tunnistetaan kaikki erilaiset toimittajatiedon käyttötarkoitukset. (JUHTA 2017, 61.)

Tietoarkkitehtuurin looginen taso vastaa kysymykseen, miten, ja sen alueeseen kuuluvat loogiset tietovarannot, loogiset tietomallit, soveltamisprofiilit, ja tietovirrat. Kuten kuvasta 2 nähdään, looginen taso kytkeytyy toiminta-arkkitehtuurin prosessi- ja toimijatietoihin, sekä tietojärjestelmäarkkitehtuuriin tietovarantojen ja tietojärjestelmäprosessien kautta. (JUHTA 2017, 29.) Tietoarkkitehtuurin loogisella tasolla kuvataan yrityksen toiminnan kannalta keskeiset tietovarannot, sekä niiden suhteet ja käyttötarkoitukset. Loogisella tasolla selvitetään yhteentoimivuuteen liittyvät ja vaikuttavat keskeiset tekijät, sekä kiinnitetään huomiota tietojen suojaamiseen ja tietojen luokitteluun. (JUHTA 2017, 61.)

Looginen tietovaranto (tietovarasto, data warehouse) on tietojen kokoelma, mihin on koottu yrityksen toiminnan ja palveluiden kannalta olennaiset, yhteisesti hallinnoidut, tiedot. Siihen sisältyy usein eri tietojärjestelmien tietokantoja ja / tai rekistereitä. Yrityksessä on olennaista tunnistaa erilaiset tietovarannot, sillä se auttaa toimimaan kustannustehokkaasti. (JUHTA 2017, 61.) Tietovarantoon kootut tiedot indeksoidaan alkuperäisessä formaatissaan, ja data siivotaan, mikä tarkoittaa muun muassa yhdenmukaisuuden tarkistamista. Tarkistamiseen liittyy muun muassa tutkia, että data on annettu säädetyissä rajoissa, eheyden tarkistaminen, eli varmistaa, että data on kokonainen, indeksoinnin tarkistaminen, sekä audit trail -tarkistukset (Waltz 2003, 286.) Tietovarastossa data on järjesteltyssä, rakenteisessa, ja siistityssä muodossa operatiivisiin järjestelmiin verrattuna, mistä syystä se on hyvä ratkaisu tukemaan käyttäjiä tiedonhauissa, ja tietojen manipuloinnissa, mikä tapahtuu päätöksentekojärjestelmien, kuten analyysityökalujen, avulla (Choo 2002, 34).

Looginen tietomalli täydentää käsitteellisellä tasolla tehtyä käsittemallia. Tiedoille annetaan ominaisuuksia eli attribuutteja, ja niiden välisiä suhteita kuvataan tarkemmin. Jos yrityksessä on tietosisältöjä, joita siirretään ja käytetään eri tietojärjestelmissä, on hyvä liittää loogiseen tietomallin myös soveltamisprofiili. Soveltamisprofiiliin nojaten voidaan toteuttaa teknisiä tiedonsiirto-rajapintoja, sekä yksityiskohtaisempia tietomalleja tietovarantojen ja

sovelluskehityksen tietokantarakenteiden tarpeisiin. Soveltamisprofiilit ovat riippumattomia toteutuksesta ja teknologiasta, ja parhaiten niiden tekemisessä onnistutaan, kun työ tehdään yhteistyössä liiketoiminnan edustajien ja tiedonhallinnan ammattilaisten kanssa. (JUHTA 2017, 61.)

Tietotyön ammattilaiset työskentelevät usein tarkkojen tietomallien kanssa, ja ylemmän tason malli puuttuu. Ylätason mallinnus on kuitenkin tarpeen, sillä sen avulla päästään keskustelemaan asioista ja varmistutaan että puhutaan samoista asioista. Tietoarkkitehtuurin tavoite on kuvata organisaation tiedot eri tasoilla. Kuvausmenetelmänä voidaan käyttää tietokantojenkin suunnittelusta tuttua käsiteanalyysiä (ER-mallia). (Hovi 2009, 12.) Ylimmän tason malli mahtuu Hovin (2009, 12) mukaan yhdelle paperiarkille, ja toimii hyvin asian esittelyyn, sekä tietointegraatioiden suunnittelun pohjaksi, onhan siinä kokonaiskuvan yrityksen tiedoista, sekä niistä järjestelmistä, missä tiedot ovat.

Tietovirtakuvausten avulla mallinnetaan mitä tietoja kulkee esimerkiksi tietojärjestelmien, tietojärjestelmäpalveluiden, tietovarantojen, tai yrityksen eri liiketoiminta-alueiden välillä. Tietovirtakuvausten avulla hahmotetaan sekä digitaaliset, että manuaaliset tietovirrat. Ne ovat kuvauksia yrityksen tietointegraatioista, missä erottuu toisaalta integraatiopalveluiden kautta, ja toisaalta suoraan tietojärjestelmien välillä tapahtuva tiedonsiirto. Tietovirtakuvauksia kannattaa käyttää erityisesti yrityksen keskeisiä prosesseja tarkasteltaessa, jolloin voidaan selvittää yrityksen päätieto- eli ydintietoryhmiä sekä tunnistaa rajapintoja. (JUHTA 2017, 62.)

Fyysinen taso vastaa kysymykseen, millä, ja sen keskiössä ovat fyysiset tietomallit, ja fyysiset tietovarannot (JUHTA 2017, 29). Fyysisellä tasolla kuvataan, missä yrityksen tuottamat ja käyttämät tiedot fyysisesti sijaitsevat. Fyysisen tietomallin avulla ymmärretään mitä yrityksen tietoja sijaitsee minkäkin tietojärjestelmän tietokannoissa, ja mitkä ovat niiden väliset suhteet tietokantatasolla. Fyysiset tietovarannot ja rekisterit käsittävät tietokantojen kuvaukset, sekä rajapintojen tekniset määritykset. Kuvaukset auttavat ymmärtämään paremmin tietojen yhdenmukaistamiseen, ja suojaustasoihin liittyvät tarpeet. Fyysisen tason kuvauksia ja määrityksiä varten yrityksessä on tietojärjestelmäsalkku. (JUHTA 2017, 63).

### **3.2 Tietojärjestelmäarkkitehtuuri**

Tietojärjestelmäarkkitehtuuri on kokonaisarkkitehtuurin näkökulma, jossa kuvataan yrityksen keskeiset, tietoja käsittelevät sovellukset, ja sovelluskokonaisuudet (tietojärjestelmät), sovellusten väliset suhteet ja riippuvuudet, sekä näiden sovellusten keskeiset ominaisuudet.

det. Tietojärjestelmäksi kutsutaan järjestelmää, mikä tukee jonkun pysyvän tietojenkäsittelykokonaisuuden suorittamista. Tietojärjestelmä muodostuu tiedoista, käsittelysäännöistä ja -ohjeista, henkilö- ja laiteresursseista, sekä tietoa siirtävistä laitteista. Tietojärjestelmäarkkitehtuurin näkökulmassa ei ole kyse ohjelmistoarkkitehtuurista, vaan tarkoituksena on kuvata yritykselle tarpeelliset järjestelmät, ja niiden liiketoimintaa tukevat keskeiset toiminnallisuudet. (JUHTA 2017, 19.) Waltzin (2003, 303) mukaan tietojärjestelmäarkkitehtuuri tukee liiketoiminta-arkkitehtuuria teknologia-arkkitehtuurin ja teknisten standardien sallimissa rajoissa.

Kuten kuvassa 2 on mallinnettu, myös tietojärjestelmäarkkitehtuurin voidaan katsoa rakentuvan käsitteellisestä tasosta, loogisesta tasosta, sekä fyysisestä tasosta (JUHTA 2017, 29).

Käsitteellinen taso vastaa kysymykseen, mitä, ja se koostuu tietojärjestelmäpalveluista ja tietojärjestelmäkartasta (JUHTA 2017, 29). Tietojärjestelmäpalvelut ovat loppukäyttäjille suunnattuja, käyttöliittymän sisältäviä palveluita, joista hyvänä esimerkkinä on tunnistautuminen (JUHTA 2017, 63).

Looginen taso vastaa kysymykseen, miten, ja se muodostuu: arkkitehtuurin kerrosnäkömästä, tietojärjestelmien välisestä vuorovaikutuksesta, tietojärjestelmäsalkusta, ja loogisista rajapinnoista (JUHTA 2017, 29). Kuvaukset loogisten tietovarantojen ja tietojärjestelmien välisistä suhteista kertovat, mitä tietovarantoja kukin tietojärjestelmä tarvitsee (JUHTA 2017, 63). Arkkitehtuurin kerrosnäkömä on havainnollinen ja palvelukeskeinen kuvaus yrityksen tai sen osa-alueen arkkitehtuurikokonaisuudesta. Kuvaus kertoo mitkä tietojärjestelmät, tietojärjestelmäpalvelut, ja tietovarannot ovat olennaisia toiminnan ja palveluiden näkökulmasta, ja siihen voidaan lisätä myös järjestelmien hyödyntämät teknologiapalvelut, palvelukerroksia unohtamatta. Tietojärjestelmien välinen vuorovaikutuskaavio kuvaa prosessien käyttämät tietojärjestelmät, sekä näiden tietojärjestelmien väliset tietovirrat, yksityiskohtiin menemättä. (JUHTA 2017, 64.)

Yrityksen tulisi koota tiedot kaikista käyttämistään tietojärjestelmistä tietojärjestelmäsalkkuun. Tietojärjestelmäsalkussa pidetään yllä tietoa järjestelmien elinkaaresta, sekä versio-tiedoista. Salkku on hyvä apuväline muutoksenhallintaan, ja se tarjoaa hyvää aineistoa muun muassa riski- ja tietoturva-analyysille. (JUHTA 2017, 64.) Tietojärjestelmäsalkun olisi hyvä antaa kuva niin nykytilasta kuin tavoitetilastakin (JUHTA 2017, 65).

Fyysinen taso vastaa kysymykseen, millä, ja se kuvaa fyysisiä rajapintoja (JUHTA 2017, 29).

## 4 Yrityksen ydintiedonhallinta ja laatu

JUHTA suosituksen (2017, 59) mukaan hyvään tiedonhallintatapaan kuuluu tietojen saatavuudesta, ajantasaisuudesta, käytettävyydestä, tietoturvallisuudesta, eheydestä, ja harmonisoinnista huolehtiminen. Lisäksi tiedoilla ja tietovarannoilla tulisi olla nimetty vastuuhenkilö / omistaja, joka vastaa tietovarannon sisällöstä, kehittämisestä, palvelujen tuottamisesta, käyttöehdoista, sekä tietojen saatavuudesta ja tietoturvallisuudesta. (JUHTA 2017, 59.)

Ydintiedonhallinta koostuu Cleven & Wortmannin (2010, 3) mukaan viidestä osa-alueesta: ydintiedon rakenteesta, ydintiedon järjestelmäarkkitehtuurista, ydintiedon hallintamallista, ydintiedon prosesseista ja lopulta ydintiedon laadusta. Ydintiedon rakenteella tarkoitetaan tietojen määrittelyä ja niiden välisiä suhteita, ja se tukee erityisesti yhtenäistä ydintiedon käyttöä organisaatioissa (Cleven & Wortmann 2010, 3). Järjestelmäarkkitehtuuri keskittyy asianmukaiseen järjestelmämalliin, mikä tukee ydintiedon elinkaariajattelua: luontia, varastointia, käyttöä, arkistointia, ja käytöstä poistoa, sekä tietoturvallisuutta. Hallintamalli tarkoittaa sitä, että tiedonhallintaan liittyen on sovittu rooleista ja vastuista, eli siitä miten ja kenen toimesta hallinta tapahtuu. Ydintiedon prosessi on liiketoiminnan vastine järjestelmäarkkitehtuurille. On tärkeää määrittää, miten tietoja luodaan, käytetään, ylläpidetään ja arkistoidaan. Laatu on yhteinen nimittäjä muille edellä mainituille osa-alueille, onhan se lopulta yksi juurisista koko ydintiedonhallinnalle. (Cleven & Wortmann 2010, 4.)

Ydintiedonhallinta (MDM, Master data management) on Otton (2012, 338) mukaan järjestelmistä irrallinen prosessi, minkä tehtävä on määrittää, omistaa ja hallinnoida liiketoiminnan ydintietoryhmiin liittyvää dataa. Sitä koskevat edellä luetellut hyvään tiedonhallintatapaan liittyvät aktiviteetit. Ydintiedonhallinta on enemmän kuin vain järjestelmä. (Loshin 2009, 20.) Se on yhdistelmä ihmisiä, työvälineitä, menetelmiä ja sovittuja käytäntöjä, jotka yhdessä muokkaavat tulevaisuutta, kun yrityksessä etsitään lisäarvoa yrityksen tietopääomasta (data asset) (Loshin 2009, 20; Niemi & Kontra 2012, 1). Ydintieto edustaa yhtä yrityksen arvokkaimmista voimavaroista ja eduista. Ydintiedonhallinnan laiminlyönti johtaa ongelmiin kuten liiketoiminnan häiriöihin, epätarkoituksenmukaiseen / riittämättömään päätöksentekoon, ja resurssien tuhlaamiseen. (Cleven & Wortmann 2010, 3.)

Niemi & Kontran (2012, 2) mukaan hyvän (ydin)tiedonhallinnan perustaksi voidaan antaa kolme suositusta: organisoituminen, omistajuus, ja ylläpito. Organisoitumisessa keskeisessä roolissa on itse organisaatio ja sen käytännöt, tiimityöskentely, sekä johdon konkreettinen tuki. Omistajuus tarkoittaa asioihin perehtymistä ja niiden ymmärtämistä, motiivointia, sekä organisaatiossa tapahtuvaa vuorovaikutusta. Ylläpito puolestaan käsittää

standardoidut prosessit ja säännöt, keskittämisen siinä määrin, kun se katsotaan hyödylliseksi, sekä oivalluksen, että tiedonhallinta on osa päivittäistä liiketoimintaa. (Niemi & Kontra 2012, 2.) Haasteeksi koetaan usein puuttuva tieto yrityksen ydintiedon elinkaaresta. Elinkaaren aikana ydintieto määritellään, luodaan, aktivoidaan, hankitaan, varastoidaan, ylläpidetään, käytetään, korjataan ja lopulta tuhotaan (Ofner, Straub, Otto & Oesterle 2013, 473). Tärkeintä olisi toteuttaa tiedonhallinnan prosessit ja järjestelmät siten, että ne ovat mahdollisimman helppoja ja toimivia käyttää. (Niemi & Kontra 2012, 6).

Ydintiedonhallinnalla halutaan Otton (2012, 339) mukaan saavuttaa kolme tavoitetta: tarjota virallinen lähde korkealaatuiselle ydintiedolle, alentaa kustannuksia ja monimutkaisuutta standardoinnin avulla, sekä tukea Business Intelligenceä (BI) ja tietojen integrointia. Jotta nämä ja muut tavoitteet saavutetaan, pitää ymmärtää ydintiedon yhdistämistarpeet, määritellä ydintiedon lähteet ja ylläpitäjät, kuvata ja ylläpitää ydintietoarkkitehtuuria, sekä ylläpitää ydintietoa hallitusti. Ydintiedonhallinnassa on kaksi keskeistä roolia: datan hallinnoijat, jotka huolehtivat tietyn ydintietolajin laadusta, sekä datan omistaja, jolla on oikeus tehdä päätöksiä liittyen siihen, että ydintiedon määrittely ja käyttö vastaa kulloinkin liiketoiminnan tarpeisiin. Ydintiedonhallinta tulisi organisoida keskitettyyn, tai yrityksen yhteiseen yksikköön. (Otto 2012, 339.)

Ydintiedonhallinta tuo yritykselle mahdollisuuden integroida, analysoida ja hyödyntää tietovarantoa huolimatta siitä missä tieto kerätään (Cleven & Wortmann 2010, 1).

#### **4.1 Taustaa**

Alun perin ydintiedonhallinta ei ollut yrityksissä ongelma, sillä tietojenkäsittely oli keskitettyä, ja kaikki sovellukset ja datat pyörivät yhdessä järjestelmässä. Tämä oli aikaa ennen relaatiotietokantoja ja strukturoitua dataa. Tietojenkäsittely oli suoraviivaista ja datan toisteisuus aiheutti vain vähäisiä haasteita. 1980 -luvulla työryhmäverkkojen yleistyminen, työasemien määrän kasvu, sekä työasemilta pääsy erillisiin sovelluksiin, alkoi haastaa yrityksen tietojenhallintaa. Ensimmäinen syy ydintiedonhallinnalle oli tarve määritellä, järjestää, ja kuvata yrityksessä yleisesti käytetyt toimintamallit ja niihin liittyvät käsitteet. (Loshin 2009, 3.)

1990 -luvun puolivälissä siirryttiin takaisin keskitettyyn tietojenkäsittelyyn ERP, CRM, ja SCM -järjestelmähankkeiden myötä, missä hyödynnettiin keskitettyä tietovarastoa. Tieto haluttiin nähdä omaisuutena ja tuovan etua kilpailijoihin nähden. Keskitetty tietojenhallinta ja tietovarastot palvelivat myös raportointia ja analytiikkaa. Haasteen toi kuitenkin se, että tiedonhallinta keskitetyssä tietovarastossa eriytti raportit alkuperäisestä tiedonlähteestä.



Erot raporteissa haastoivat tiedon luotettavuutta, ja aiheuttivat epäluuloja tietojen oikeellisuuteen liittyen. Käyttäjät käyttivät tietoja operatiivisesta järjestelmästä, kun taas johto luki keskitetystä tietovarannosta haettuja tietoja raporttien muodossa. Ydintiedonhallinnan toivottiin takaavan datan laadun, eheyden ja luotettavuuden järjestelmien välillä. (Loshin 2009, 4.)

Nykyisin ydintiedonhallinnan pitäisi olla enemmänkin liiketoiminta ja prosessi suuntautunut kuin teknologia suuntautunutta. Tavoitteena tulisi olla yksi tietolähde korkealaatuiselle ydintiedolle, mikä jakelee synkronoitua ja yhtenäistä dataa yrityksen muille tietojärjestelmille. Jakelu edellyttää korkealaatuisia integraatioita, mikä Loshinin (2009, 14) mukaan nojaa seuraavien tekijöiden varaan:

- Kaikki yrityksessä hyödyntävät tietolähdettä / -lähteitä.
- On olemassa menetelmät ydintietojen tunnistamiseksi.
- On olemassa sovitut määrittelyt, käyttötapaukset ja -tarkoitukset, sekä ymmärrys hierarkiasta ja tietojen yhteyksistä.
- On kyky hallinnoida tiedon syntyperää, jakelua ja syöttöä muihin järjestelmiin.
- Data migraatiossa ydintiedon pohjaksi otetaan laadultaan paras data.
- Menettelytapa jolla palvelut hyödyntävät ydintietoa tehdään läpinäkyväksi.
- On tapa hallinnoida yrityksen ydintietoa, ja malli miten siihen tehdään tarvittaessa muutoksia.

1990 -luvulla yrityksissä, missä toteutettiin toiminnanohjaushankkeita, käynnisteltiin myös ydintiedonhallintaan liittyviä toimia. Monissa yrityksissä on kuitenkin vasta 2010 -luvulla aloitettu järjestelmällinen ydintiedonhallinta. Ydintiedonhallinnasta vastaava tiimi on useissa yrityksissä osa IT-organisaatiota, tai johtoryhmälle raportoivan projektitoimiston osa. Kriittisimmäksi onnistumistekijäksi on koettu ydintiedonhallinnasta vastaavan tiimin kyky toimia yhdistävänä tekijänä liiketoimintajohdon ja IT -organisaation välimaastossa. Suurimmaksi ongelmaksi on koettu yritysjohdon tuen puute niin arjen toiminnassa kuin kehittämisesäkin. (Niemi & Kontra 2012, 3.)

## **4.2 Ydintiedonhallinnan käytännöt**

Datan ylläpitoon on erilaisia käytäntöjä, jotka vaihtelevat erityisesti keskittämistason ja ulkoistuksen suhteen (Niemi & Kontra 2012, 5). Ei ole olemassa vain yhtä kaikille sopivaa mallia hallita ydintietoja (Cleven & Wortmann 2010, 1). Ydintiedonhallinnan menestystekijä onkin soveltuvan, ja hyvin toimivan hallintamekanismin luonti ja käyttöönotto (Radcliffe 2007, 1). Ydintiedon hallintaa voidaankin käytännössä tehdä keskitetysti, hajautetusti, tai jonkinlaisessa hybridimallissa. Näistä lisää myöhemmin.

Tyypillisesti yksi liiketoimintayksikkö ei ole yksin vastuussa tiedon elinkaarenhallinnasta, vaan tietynä ajanhetkenä sille on omat tarvitsijansa, ja organisaatioiden välinen ydintiedonhallinta on yleensä epäselvä. Tämä johtaa usein tiedon hallitsemattomaan prosessointiin, ja lopulta pirstaloituneeseen, ja ei paikkaansa pitävään ydintietoon yrityksessä. Ydintiedon elinkaaren ymmärtäminen, suunnittelu ja monitorointi katsotaan hyvän ydintiedonhallinnan kulmakiveksi. (Ofner, Straub, Otto & Oesterle 2013, 473.)

Cleven & Wortmannin (2010, 1) mukaan keskitettyyn ydintiedonhallintaan johtavista syistä yksi merkittävimmistä on laatuongelmat. Kun tietoa ylläpidetään eri järjestelmissä, se johtaa suureen määrään eroavuuksia muun muassa tietojen määrittelyissä, formaateissa, ja data-arvoissa, mikä puolestaan johtaa siihen, että yrityksen on lähes mahdoton hyödyntää sitä. Esimerkkejä eroavaisuuksista ovat: tuplaantuneet (dublikaatit) asiakastunnisteet, puutteelliset attribuutit, eroavuudet käytettyjen arvojen ja merkityksellisen metadatan välillä, sekä vapaatekstin käyttö tilanteissa missä se ei ole sopivaa. Tutkimusten mukaan virhetaso on jopa 30 prosentin luokkaa, ja jopa 83 prosenttia yrityksistä on kärsinyt huonosta ydintiedonlaadusta johtuvista ongelmista. (Cleven & Wortmann 2010, 1).

Niemi & Kontran (2012, 4) mukaan ydintiedon omistajuus voi vaihdella, mutta omistajuiden pitäisi olla selvästi määritelty, ja se tulisi olla esimerkiksi: myyntijohtaja, liiketoimintalueen johtaja, talousjohtaja, tai tietohallintojohtaja. Tämän henkilön tulee ymmärtää perusteellisesti tehtävän merkitys ja tavoite, onhan kysymyksessä kokonaiskuva yrityksen tiedonhallinnasta yli organisaatorajojen (Niemi & Kontra 2012, 4).

Niemi & Kontra (2012, 3) suosittelevat, että yritykseen perustettaisiin ydintiedonhallinnalle omistautunut, vähintään kahden hengen tiimi / organisaatio, joka varmistaa yhtenäisten käytäntöjen viemisen yrityksen liiketoiminnallisten ja maantieteellisten rajojen yli. Yhteisesti vahvistettavien käytäntöjen tulisi kattaa ainakin prosessit, data standardit ja laadun mittarit (Niemi & Kontra 2012, 3).

### **4.3 Ydintiedonhallinta ja tietojärjestelmät**

Harva yritys voi enää tänä päivänä pyörittää liiketoimintaansa yhden järjestelmän / sovelluksen varassa. Muun muassa: talousjärjestelmä, toiminnanohjausjärjestelmä (ERP), asiakkuudenhallintajärjestelmä (CRM), henkilöstöhallinnanjärjestelmä (HRM), ja erilaiset tietovarastot vaativat toimiakseen yrityksen ydintietoa. Koska järjestelmät eivät ole automaattisesti linkattuja keskenään, jokaisessa niistä ylläpidetään ydintietoa erikseen, jotta ne palvelisivat tarkoitustaan. Kaikilla on oma paikallinen tietokanta, ja jokaisessa on taulut ydintiedon säilytystarpeisiin. (Kashel, Bullerwell & Kent 2011, 37.) Ongelmaksi muodostuu tietojen siiloutuminen, eli eriytyminen irrallisiksi kokonaisuuksiksi. On toki järkevää, että eri

osastot ja toiminnot edistävät omaa toimintaansa ja hoitavat tarvitsemiaan tietoja, mutta ongelmaksi muodostuu se, ettei kenelläkään ole kokonaiskuvaa yrityksen tiedoista. Tällaisessa mallissa yli organisaatio- ja sovellusrajojen menevät tietomallit eivät ole kenenkään kiinnostuksen kohde. (Hovi 2009, 12.)

Ydintiedonhallintaan tarjolla olevat ratkaisut, järjestelmät ja työkalut eivät ole yksin riittävä tae ongelmattomalle ydintiedon käytölle yrityksen eri tietojärjestelmissä, vaan tarvitaan myös organisaatiotasoinen tuki ja tulkinta datalle, sekä prosessit datan harmonisoinnille, siivoukselle, luonnille, ylläpidolle ja käytöstä poistolle (Loser, Legner & Gizanis 2004, 5).

Kashel ym. (2011, 38) konkretisoi haastetta seuraavan esimerkin avulla. Jos tilauksenkäsitteilyjärjestelmä on paikka missä asiakastiedot luodaan, kuinka nämä asiakastiedot saadaan käyttöön CRM:ssä, missä mietitään uusia myyntimahdollisuuksia kyseiselle asiakkaalle? Ja entäpä jos CRM:ssä tarvitaan osittain erilaisia asiakastietoja. Toisaalta taas CRM:ssä voidaan luoda potentiaalinen asiakas, mikä muutaman kuukauden jälkeen muuttuu todelliseksi asiakkaaksi. Miten tieto saadaan CRM:stä tilausjärjestelmään? On todennäköistä että asiakas luodaan sekä tilausjärjestelmään että CRM:ään osin erilaisin tiedoin. Juuri tällaiset tilanteet aiheuttavat yritykselle vakaviakin ongelmia, sillä on vaikea saada yhtä / koko totuutta asiakkaasta. Todelliset ongelmat syntyvät ajan kuluessa. Tietojen lisäykset, muokkaukset ja poistot, tulisi tehdä kaikissa järjestelmissä samaan aikaan, mutta tätä ei ole helppoa toteuttaa käytännössä. Seurauksena on se, että ydintieto eriytyy yrityksen eri järjestelmissä. (Kashel, Bullerwell & Kent 2011, 38.)

Ydintietojärjestelmä on paikka, rekisteri, missä tiedon kriittiset attribuutit sijaitsevat, ja mistä tieto on saatavilla yrityksen käyttöön. Hyvä ydintiedonhallinta varmistaa tiedon luotettavuuden, laadukkuuden ja eheyden. (Loshin 2009, 8.) Ydintieto ei kärsi pelkästään epäyhtenäisyydestä vaan siitä, että yrityksessä on useita erillisiä sovelluksia prosessien tukena. Tästä syystä onkin tarvetta erityisille järjestelmille ja standardeille ydintiedonhallinnan tueksi läpi järjestelmien ja liiketoimintarajojen. (Loser, Legner & Gizanis 2004, 1.)

Ydintiedonhallinnan aloitteet yhdistetään usein yrityksen haluun parantaa yrityksen järjestelmiin liittyvää hallintoa ja laatua. Äkkiä katsottuna ydintiedonhallinta näyttää yritykseltä koota tieto yhteen järjestelmään, ja yritykseltä saavuttaa yksi totuus. (Loshin 2009, 13.) Yrityksen ydintiedot voivat kuitenkin sijaita yhdessä tai useammassa datalähteessä (Otto 2012, 338). Tavoitteena on järjestelmien ja menettelytapojen tehokkuus (Choo 2002, 33).

#### 4.4 Laatu kantavana ajatuksena

Ydintiedonhallinnan suorituskyvyn mittari on datan laatu (Hovi, Hervonen & Koistinen 2009, 67; Loser, Legner & Gizanis 2004, 1; Otto 2012, 338). Korkealaatuinen ydintieto on edellytys yrityksen strategisten liiketoimintatavoitteiden saavuttamiselle. Näitä tavoitteita ovat muun muassa: parempi päätöksenteko, kokonaisvaltainen asiakkuuksienhallinta, yhteensopivuus sopimusehtojen kanssa, sekä toiminnan tehokkuus. (Ofner, Straub, Otto & Oesterle 2013, 472.) Tämä on mahdollista saavuttaa, kun yrityksen kaikissa järjestelmissä on täysin sama ydintieto, mitä tulee niiden lajeihin, jäseniin ja attribuutteihin, sikäli kun järjestelmä näitä tietoja käyttää (Kashel, Bullerwell & Kent 2011, 40).

Haug & Arlbjørnin (2010, 300) mukaan suurimmat syyt huonoon ydintiedon laatuun johtuvat siitä, ettei ydintiedonhallinnan rooleja ja vastuita ole määritelty, ydintiedon oikeellisuuden varmistamista ei palkita, ydintiedonhallinnan rutiineita ei kontrolloida, työntekijöiden osaamisessa ja taidoissa on puutteita, ja ydintiedonhallinnan sovellus ei ole käyttäjäystävällinen.

Laatikainen & Niemen (2012, 1) mukaan tiedon laadun merkitys kasvaa päätöksenteossa ja liiketoimintaprosesseissa samalla kun datan / tiedon määrä kasvaa. Jollei sekä prosesseista että tiedonlähteistä pidetä huolta, tiedon laatu heikkenee nopeasti. Yksi tapa hallita laatua on valita laadullisesti parasta tietoa sisältävät järjestelmät ensisijaisiksi tietolähteiksi. Huonosta tiedon laadusta aiheutuu Suomessa yrityksille arviolta yli 10 miljardin euron suuret vahingot vuosittain. (Laatikainen & Niemi 2012, 1.)

Laatutavoitteiden toteutumista tarkastellaan usein seuraavista näkökulmista: virheettömyys (accuracy), mikä tarkoittaa sitä että tieto kuvaa reaali maailman tilannetta, täydellisyys (completeness), mikä tarkoittaa että kaikki tarvittava tieto on saatavissa, yhtenäisyys (consistency), mikä tarkoittaa sitä, että tieto on sama eri järjestelmien välillä, kahdentuminen (duplication), mikä tarkoittaa ettei samaa tietoa esiinny kahdesti, ja ajantasaisuus (timeliness), mikä tarkoittaa sitä, että tieto on saatavissa tarvittavan nopeasti (Laatikainen & Niemi 2012, 2).

Laadukasta ydintietoa tavoiteltaessa, tulee sopia yhteisistä määrittelyistä, kuvauksista ja ohjeista. Jos yhteiset, yritystason datan määrittelyt puuttuvat, on haastavaa mitata ja parantaa tiedon laatua. Yhteiset määrittelyt, kuten käsitteet ja tietomallit, ovat perusta laadukkaalle tiedolle. Vasta kun ylätason määrittelyt ovat selvät, voidaan syventyä yksityiskohtaisempiin kuvauksiin kuten kuvaaviin attribuutteihin, käsitteiden välisiin suhteisiin ja

riippuvuuksiin. Tieto tulisi kuvata ylhäältä alaspäin, jotta niille on mahdollista saada liiketoiminnan tuki ja hyväksyntä jo yrityksen ydintietomallin luontivaiheessa. (Laatikainen & Niemi 2012, 4.)

Datan laadullinen arvo konkretisoituu usein vasta tilanteessa, missä vieheellinen tieto aiheuttaa yritykselle suuria kustannuksia. Jotta ydintiedonhallinnalle saadaan johdon tuki, on tärkeää osoittaa laadukkaan tiedon tuottama arvo, mikä näyttäytyy esimerkiksi lisääntyneenä myyntinä, tehokkuuden paranemisena, tai parempana riskien hallintana. Laadun seurannassa havaitut virheellistä ja huonolaatuista tietoa tuottavat toimintamallit tulisi korjata välittömästi niiden löydettyä. Laatu olisi hyvä ottaa myös osaksi palkitsemismallia. (Laatikainen & Niemi 2012, 8).

## 5 Ydintietoarkkitehtuurin vaihtoehtoiset mallit

Ydintiedonhallinta on osa organisaation yleistä informaatorakennetta ja tietoarkkitehtuuria, mistä johtuen se pitää rakentaa perustana olevien teknologioiden päälle, ja integroida sekä uusien että jo olemassa olevien järjestelmien kanssa. Tekemisessä tulee huomioida: datamallinnus, tietojen laadunhallinta, lataukset/integroinnit ja synkronointi, liiketoimintapalvelut ja työnkulut, suorituskyky, skaalattavuus ja saatavuus, hallinnointi ja turvallisuus, sekä teknologiat ja arkkitehtuuri. Ilman laatumittareita on vaikea mitata ydintiedon laatua ja sen vaikutuksia liiketoimintaan. (Radcliffe 2007, 5-6.)



Kuva 3. Ydintietoarkkitehtuuri (mukailtu Radcliffe 2007, 3)

Kuten kuvasta 3 käy ilmi, ydintietoarkkitehtuurin mallinnus ja käyttöönotto on moniulotteinen tehtävä, missä tulee tasapainoilla eri sidosryhmien (stakeholder) intressien kanssa, ottaa huomioon tekniikan mukaan tuomat mahdollisuudet, ja huomioida eri ydintietolajien mukanaan tuomat vaatimukset (Otto 2012, 338).

Tietojärjestelmien tulee olla integroidut, jotta palveluprosesseja voitaisiin tukea parhaalla mahdollisella tavalla. Tämä on keskeinen vaatimus myös saman ydintiedon käytölle yrityksen eri osa-alueiden välillä. Yrityksessä tulee määrittellä, mikä tieto on ydintietoa, ja yhtenäistää ne läpi järjestelmien, sekä määrittää tuleeko ydintieto luoda ja ylläpitää keskitetysti vai hajautetusti. (Loser, Legner & Gizanis 2004, 2.)

Ydintietoarkkitehtuuria ei ole kattavasti määritelty. Yritykset määrittelyille löytyvät tietovarastojen (data warehouse) määrittelyn alueelta, missä ydintietoarkkitehtuurin katsotaan olevan esivaatimus standardoidulle ja täsmälliselle raportoinnille. Kenties täsmällisin määritelmä on se, että ydintietoarkkitehtuurilla viitataan tietoarkkitehtuuriin, mutta niin että keskitytään ainoastaan ydintietoihin. Ydintietoarkkitehtuuri käsittää sekä ydintietomallin että tietojärjestelmäarkkitehtuurin. (Otto 2012, 339.)

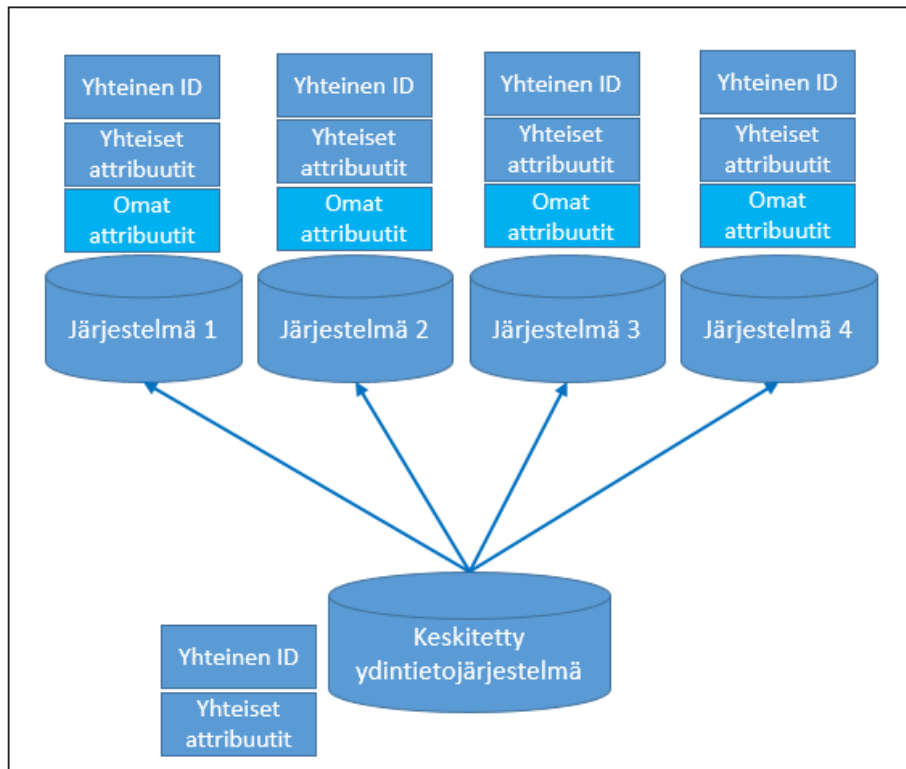
Arkkitehtuurimalliin liittyviä kysymyksiä ei Otton (2012, 345) mukaan tulisi lähestyä pelkäämään teknisistä näkökulmista, vaan sen tulisi huomioida yhtä lailla yrityksen eri organisaatiot. Ydintietoarkkitehtuuria käytetään lähestyttäessä monimutkaista asiaa, missä monimutkaisuuden aiheuttaa useiden ydintietolajien rinnakkainen käyttö, organisaatorakenne, ja organisaatioiden erilaiset käyttötapaukset. Epävarmuustekijät ydintietoarkkitehtuuriin liittyen juontuvat liiketoimintaprosesseista, erilaisista ydintietolajeista ja malleista, sekä paikallisesti ja keskitetysti hallitun ydintiedonhallinnan suhteesta. Ydintietoarkkitehtuurin ja ydintiedonhallinnan välillä on kiinteä suhde. (Otto 2012, 345.)

Niemi & Kontran (2012, 5) mukaan ei ole mahdollista antaa yhtä suositusta mikä malli sopisi kaikille. Yrityksen tulee selvittää, sopiiko sille paremmin tiedonhallinnan keskittäminen vai hajauttaminen. Tärkeää on ymmärtää myös se, että kulloinkin sopiva malli voi olla eri ydintietolajista riippuen, eli se mikä malli sopii tuotetiedonhallintaan, ei välttämättä ole paras asiakastiedonhallintaan. (Niemi & Kontra 2012, 5.)

Sekä Loser, Legner & Gizanis, että Otto, ovat mallintaneet ydintiedonhallinnan ja ydintietoarkkitehtuurin neljäksi erilaiseksi malliksi. Kashel, Bullerwell & Kent, ja Hovi, Hervonen & Koistinen käyttävät puolestaan kolmea erilaista mallia. Kaikissa näissä mallinnustavoissa on yhteistä se, että ydintiedonhallinta tapahtuu joko keskitetysti, hajautetusti tai hybridimallissa. Seuraavissa alaluvuissa näitä malleja tarkastellaan lähemmin.

## 5.1 Keskitetty malli

**Keskitetty ydintietojärjestelmä -malli** on rakennettu kuvan 4 mukaisesti, eli siten, että yhdestä keskustietojärjestelmästä tiedot toimitetaan useaan ydintietoa hyödyntävään järjestelmään. Kaikki järjestelmät käyttävät samaa ydintietoa, ja tiedon yksilöivä tunniste, esimerkiksi EAN -koodi, toimii avaimena tiedon käytölle. (Loser, Legner & Gizanis 2004, 2).



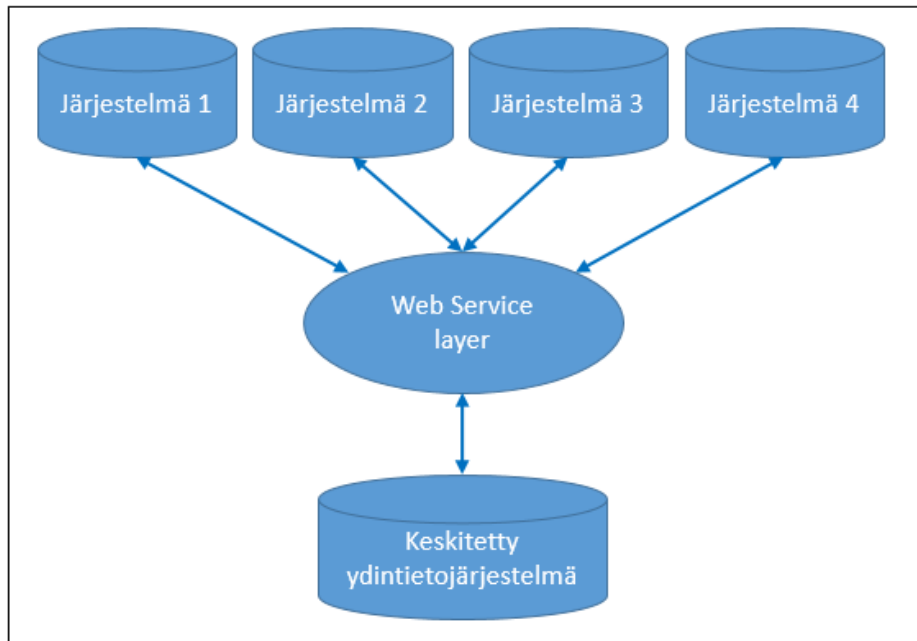
Kuva 4. Keskitetty ydintietojärjestelmä (mukailtu Loser, Legner & Gizanis 2004, 2; Otto 2012, 342)

Etu tällaisessa keskitetyssä mallissa on se, että yhteiset attribuutit luodaan aina keskitetysti, mikä takaa sen, että data on aina luotu samalla tavalla. Tämän lisäksi on mahdollista ylläpitää niin sanottuja paikallisia attribuutteja vastaanottavassa järjestelmässä itsessään. Keskusjärjestelmä toimittaa siis vain yhteiseksi määritellyn ydintiedon. Keskusjärjestelmä tai tarkoitusta varten erikseen perustettu sovellus hallinnoi avaimia. Ydintiedon välitys käynnistyy aina keskusjärjestelmästä käsin. Yleinen sääntö on, että ydintietojen ajot tapahtuvat asynkronoidusti, eli ei reaaliaikaisesti. Tämä tietojen päivittyminen viiveellä katsotaan mallin heikkoudeksi. (Loser, Legner & Gizanis 2004, 2-3).

**Vuorovaikutteisessa (transactional) -mallissa** ydintiedon attribuuttien ylläpito on MDS:n (master data server/services), eli keskitetyn ydintietojärjestelmän tehtävä. Ydintietojärjestelmä niin ikään antaa tiedoille yksilöivän tunnisteen, ja hallintaprosessi tukee ydintiedon ylläpitoa muun muassa niin, että jo tiedon luontivaiheessa tarkistetaan mahdollisten dublikaattien olemassaolo. Koska vuorovaikutteinen malli on pitkälti sama kuin keskitetty ydintietojärjestelmä malli, on ne molemmat nähtävissä kuvasta 4. Ydintietojärjestelmä vastaa ydintiedon välityksestä sitä tarvitseviin järjestelmiin. Vastaanottavat järjestelmät voivat ainoastaan lukea tietoa, mutta eivät voi tehdä siihen muutoksia. (Otto 2012, 342.)



**Vuorovaikutteinen hub (transactional hub) -malli** on Kashel, Bullerwell & Kentin (2011, 47) mukaan heidän malleistaan helpoin selittää, mutta vaikein toteuttaa. Siinä ajatus on, kuten kuvassa 5 näkyy, että on yksi ydintiedon lähde, joka palvelee kaikkia järjestelmiä. Hyötynä mallissa on se, ettei ole vaaraa dublikaateista. Haittana on puolestaan se, että yrityksessä voi olla vaikea eriyttää ydintietoa operatiivisista järjestelmistä, ja rakentaa kaikkia tarvittavia integraatioita. Vuorovaikutteinen hub nähdään usein idealistisena mallina. (Kashel, Bullerwell & Kent 2011, 48.)



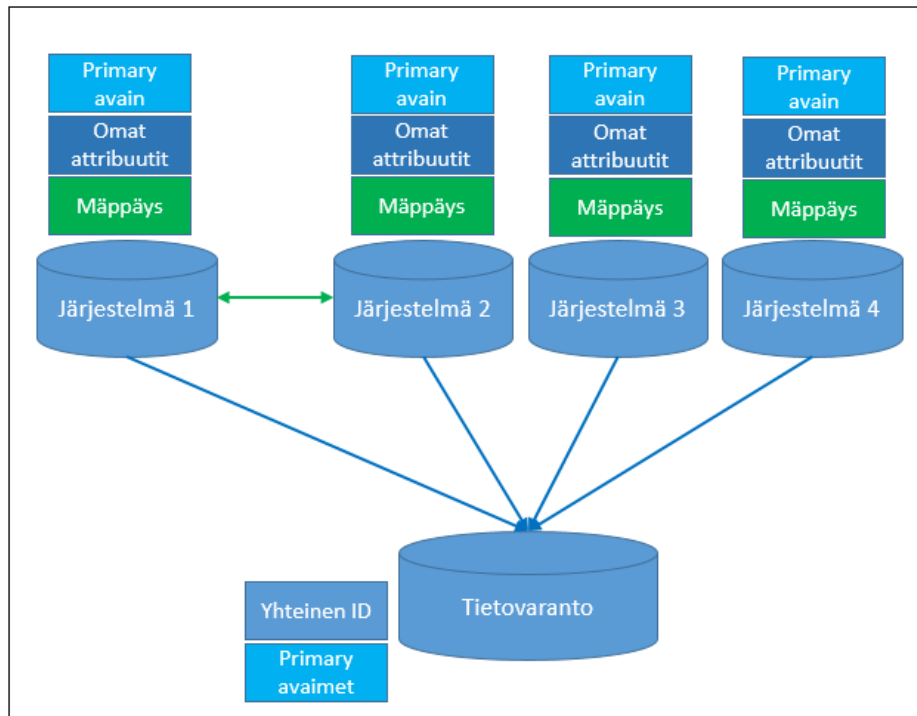
Kuva 5. Vuorovaikutteinen hub (mukailtu Kashel, Bullerwell & Kent 2011, 48)

**Keskitetty tietovaranto -mallissa** ideana on koota ja integroida yrityksen tiedot yhteen tai useampaan isoon tietokantaan, joka sisältää liiketoiminta-alueiden tietoja yhdenmukaistettuna. Yritystasoinen tietojen tarkastelu tulee tällöin mahdolliseksi. Keskitetyn tietovaraston rooli on yhdenmukaistaa operatiivisia tietoja, säilyttää niiden historia, ja toimia suurten tietomäärien varastona. Tämän tyyppinen arkkitehtuuri voi olla helposti monimutkainen, sisältäähän se monien eri alueiden tietoja. Keskitetyn tietovarannon sanotaan kuitenkin olevan yhdenmukaistettuun tietovarantoon nähden selkeämpi rakenteeltaan ja nopeampi rakentaa. (Hovi, Hervonen & Koistinen 2009, 27-28.)

## 5.2 Hajautettu malli

**Lähdetietojärjestelmä (repository) -mallissa** otetaan käyttöön keskitetty tietolähde/-varasto kaikelle tarvittavalle datalle. Sinne varastoidaan kaikki erilaiset ydintietojen luokitukset. Jos järjestelmä tarvitsee esimerkiksi asiakastietoa, se lähettää kyselyn tietovarastoon, ja saa vastauksena tiedon missä järjestelmässä asiakastieto sijaitsee, minkä jälkeen myös

kysely ohjataan suoraan sinne. Tietoja käyttävä järjestelmä huolehtii mäppäyksestä. Kuten kuvassa 6 esitetään, tässä mallissa data luodaan ja ylläpidetään hajautetusti, ja data on myös varastoitunut hajautetusti näihin järjestelmiin. Tässä mallissa ei ole datan siirtoa. Datan käyttö alkaa käyttävän järjestelmän tarpeesta. Tässä mallissa datan tunnisteet ovat yleensä eri järjestelmissä eri. Tietovarastossa on globaali avain kaikkiin datoihin, ja se hallinnoi myös primary avaimia. Muutokset ydintiedoissa ovat heti kaikkien tarvitsijoiden käytävissä. Mallia käytetään usein silloin kun kyseessä on suuret datamäärät. (Loser, Legner & Gizanis 2004, 4.)

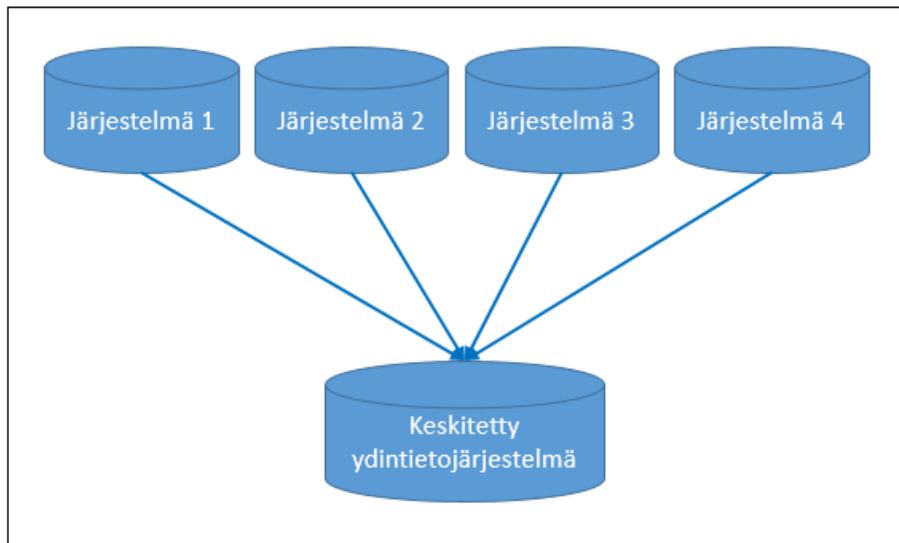


Kuva 6. Lähdetietojärjestelmä (mukailtu Loser, Legner & Gazanis 2004, 4)

Lähdetietojärjestelmä -mallissa hyödyksi katsotaan se, että järjestelmien itsenäinen asema säilyy, ja riippuvuus keskusjärjestelmään on vähäinen. Haitaksi koetaan se, että data luodaan ja ylläpidetään hajallaan eri prosesseissa. (Loser, Legner & Gizanis 2004, 5.)

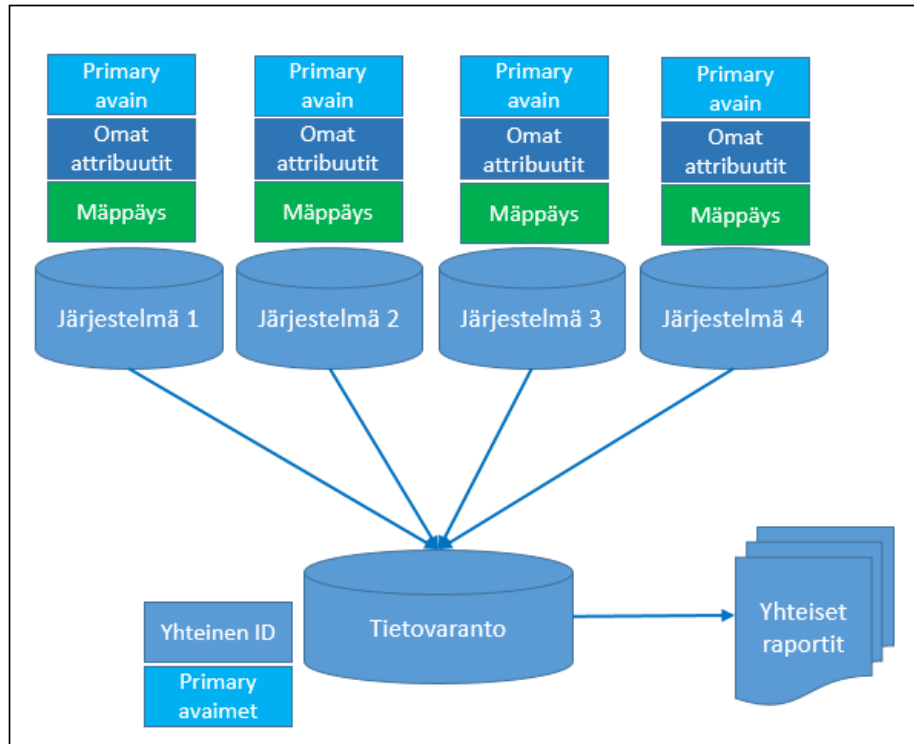
**Analyttisessä -mallissa** kaikki ydintiedonhallintaan liittyvät toiminnot (luonti, ylläpito, poisto) tapahtuvat lokaaleissa lähdejärjestelmissä (ks. kuva 7). Data käytetään ajoittain keskitetyssä ydintietojärjestelmässä, missä sille annetaan yrityksenlaajuinen yksilöivä tunniste, ja siellä on tiedot myös lokaaleista avain attribuuteista. Ydintietojärjestelmä mahdollistaa sen, että organisaatio voi hallinnoida ja ylläpitää liiketoimintadataa, mitä käytetään

liiketoimintakriittisessä päätöksenteossa. Ydintiedonhallinnan järjestelmä standardoi ja yhtenäistää yrityksen kriittiset tiedot keskitetysti. Se myös tunnistaa mahdolliset dublikaatit tiedoissa. (Otto 2012, 339.)



Kuva 7. Analyttinen malli (mukailtu Otto 2012, 342)

**Rekisteri (registry) -malli** tarkoittaa sitä, että on yksi tietokanta mikä koostaa tiedot yhteen, ja vastaa näin muodostetusta, yhtenäisestä ydintiedosta. Tätä voidaan käyttää hyödyksi muun muassa raportoinnissa. Kuten kuvassa 8 esitetään, rekisteri sisältää seuraavat tiedot: yksilöivät avaimet, lähdejärjestelmän avaimet (primary avaimet), mahdollisen metadatan sekä säännöt. Attribuutteja ei ylläpidetä rekisterissä, vaan lähdejärjestelmissä. Tuloksena on ydintietojärjestelmä, jota lähdejärjestelmät voivat ainoastaan lukea. Tämä malli ei poissulje dublikaatteiden mahdollisuutta. Epäjohdonmukaisuuksien tarkkailuun ja poistoon tarvitaan säännöt ja prosessi. Eduksi tällä mallille nähdään se, että tällainen järjestelmä on melko yksinkertainen perustaa. Haittoja ovat laajat ja hitaat, läpi tietokantojen ulottuvat kyselyt, jotka ovat välttämättömiä datan vertailussa ja epäjohdonmukaisuuksien poistossa. Ei pidä myöskään unohtaa sitä, että kyselyt pitää kirjoittaa uudelleen, kun kokonaisuuteen lisätään uusi järjestelmä. (Kashel, Bullerwell & Kent 2011, 49-50.)



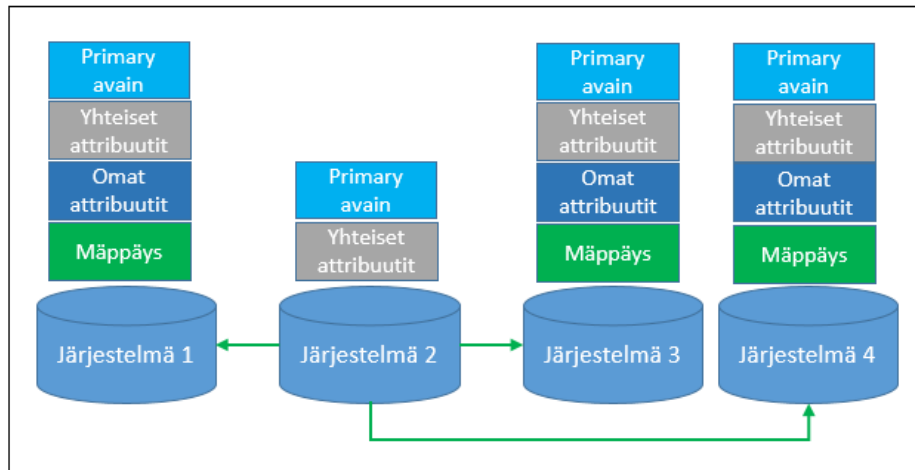
Kuva 8. Rekisteri (mukailtu Kashel, Bullerwell & Kent 2011, 50)

**Erilliset tietovarannot -maali** rakentuu Hovi, Hervonen & Koistisen (2009, 26) mukaan yhden tai useamman tietovarannon varaan. Etuna tässä mallissa on sen nopea toteutus, jolloin käyttäjille on mahdollista saada pian erilaisia raportteja. Haittana on kuitenkin erillisyyks, ja tähän malliin päädytään usein silloin kun yrityksessä ei ole aikaisemmin ollut tietostrategiaa. Samat tiedot voivat olla toistettuina useissa tietovarannoissa, mikä rasittaa operatiivista järjestelmää, ja vaikeuttaa yritystasoisena kuvan muodostamista. Erilliset varannot sopivat hyvin pieniin ympäristöihin ja organisaatioihin, missä ei ole mahdollista panostaa keskitetympään ratkaisuun. On sanottu, että keskitetyt varastot olisivat huonompia yhdenmukaistettuihin verrattuna, mutta selkeämpiä rakenteeltaan ja nopeampia rakentaa. (Hovi, Hervonen & Koistinen 2009, 26-28.)

### 5.3 Hybridimalli

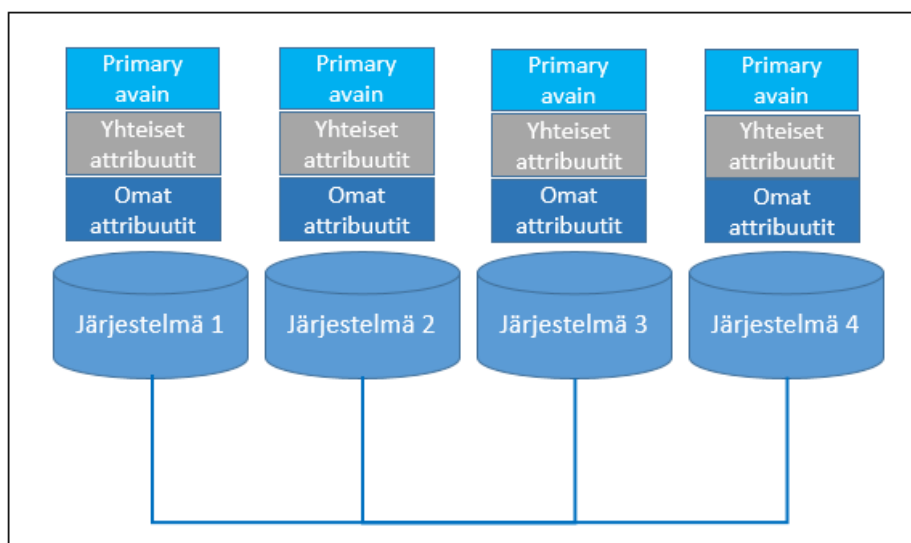
**Johtava tietojärjestelmä (Leading system) -mallissa** yksi olemassa olevista järjestelmissä nostetaan johtavaksi järjestelmäksi, kuten kuvassa 9 järjestelmä 2. Ydintieto luodaan johtavassa tietojärjestelmässä, ja se huolehtii ydintiedon jakelusta. Tässä mallissa attribuutteja ylläpidetään ja hallinnoidaan paikallisesti kaikissa järjestelmissä, eikä yhteisiä keskitetyksi hallintoituja attribuutteja ole. Tiedoilla on yksilöivät tunnisteet, mutta koska ne ovat järjestelmäkohtaisia, muodostuu tarve mäppäykselle, mikä takaa oikean tiedon välityksen järjestelmien välillä. Vain ne tiedot / attribuutit siirretään, mitä tarvitaan, ja tiedon-

siirto on myös tässä mallissa asynkronista. Tässä mallissa hyväksi katsotaan se, että vastaanottava järjestelmä säilyy itsenäisenä ja attribuutteja voi lisätä tarvittaessa jälkeempään. Integrointi on kuitenkin työläämpää määppäysten vuoksi, ja dataa pidetään järjestelmissä tarpeettomasti. (Loser, Legner & Gizanis 2004, 3.)



Kuva 9. Johtava tietojärjestelmä -malli (mukailtu Loser, Legner & Gizanis 2004, 3)

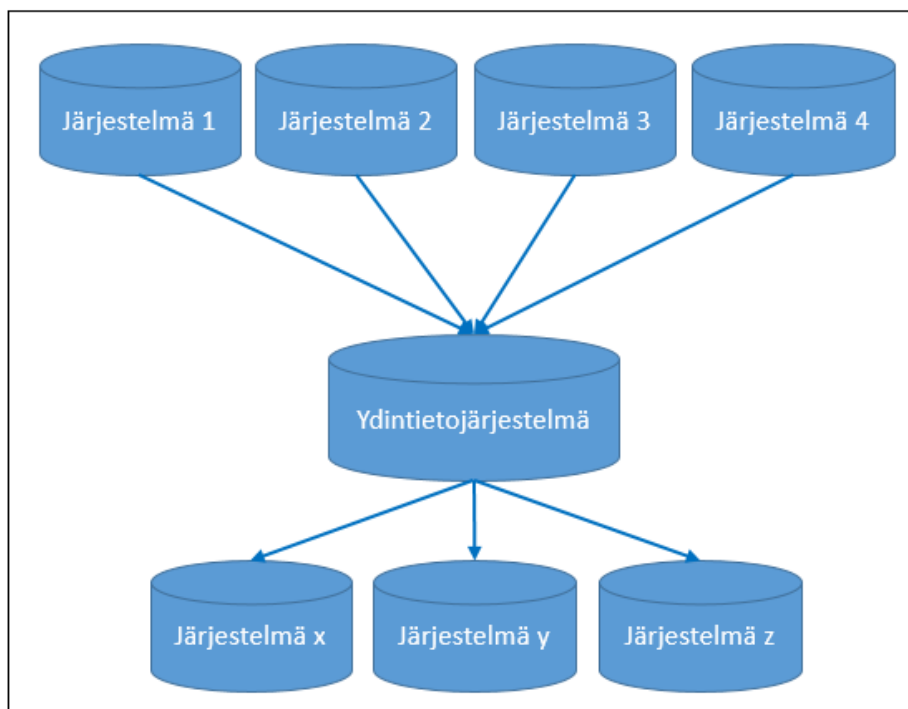
**Standardoidussa mallissa** keskiöön nousee yrityksen laajuisen standardin määrittäminen, jossa kuvataan ydintiedon rakenne, ja määritellään yleiset (niin sanotut globaalit) attribuutit, mikä tarkoittaa sitä, että niillä on sama merkitys kaikissa järjestelmissä. Integraatiokerros varmistaa, että ydintieto on rakenteeltaan sama kaikkialla, ja että yleiset attribuutit on aina luotu järjestelmissä. Koska tähän malliin ei kuulu määppäystä, on selvänä haittana dublikaattien tekemisen mahdollisuus. Esimerkiksi sama asiakas voidaan luoda useampaan järjestelmään. (Loser, Legner & Gizanis 2004, 3).



Kuva 10. Standardoitu malli (mukailtu Loser, Legner & Gizanis 2004, 4)

Kuten kuvassa 10 on esitetty, standardoidussa mallissa tietojen luonti, ja ylläpito tapahtuu hajautetusti eri järjestelmissä. Yhdessä sovitut yleiset attribuutit varmistavat sen, että kaikissa järjestelmissä ylläpidetään aina minimimäärä tietoja / attribuutteja. Tässä mallissa ei tapahdu tiedon siirtoa, eikä järjestelmien välistä eheyden tarkistusta. Selvä hyöty tässä mallissa on se, että järjestelmät pidetään erillään, ja tiedon saatavuus on hyvä. Haittana ovat mahdolliset duplikaatit sekä se, ettei yhtenäistä ydintietoa voida taata. (Loser, Legner & Gizanis 2004, 4.)

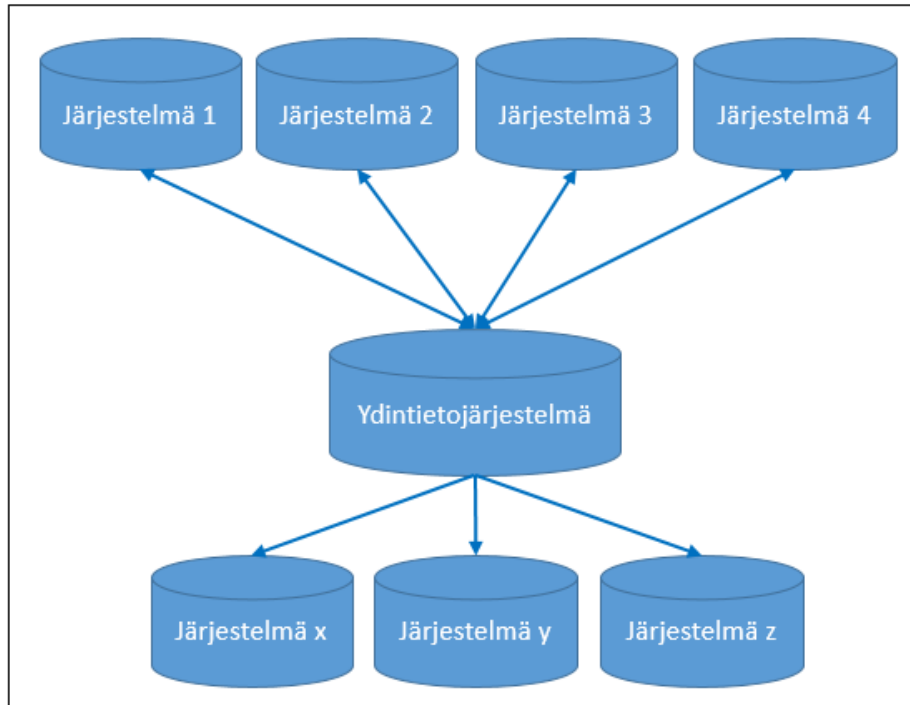
**Coexistence -malli** on analyttisen ja vuorovaikuttaisen mallin yhdistelmä (ks. kuva 11). Siinä ydintiedot ylläpidetään lokaaleissa järjestelmissä, missä myös luodaan tunnisteet ja avaimet ydintiedoille. Data vietään keskitettyyn ydintietojärjestelmään, missä sille annetaan yrityksen näkökulmasta yksilöivä tunniste. Tämän jälkeen pieni ryhmä asiantuntijoita huolehtii datan vakauttamisesta / yhtenäistämistä, minkä jälkeen tieto toimitetaan takaisin lähdejärjestelmiin, ja mikäli tarvetta, myös eteenpäin muihin kohdejärjestelmiin. (Otto 2012, 342.)



Kuva 11. Coexistence näkökulma (mukailtu Otto 2012, 342)

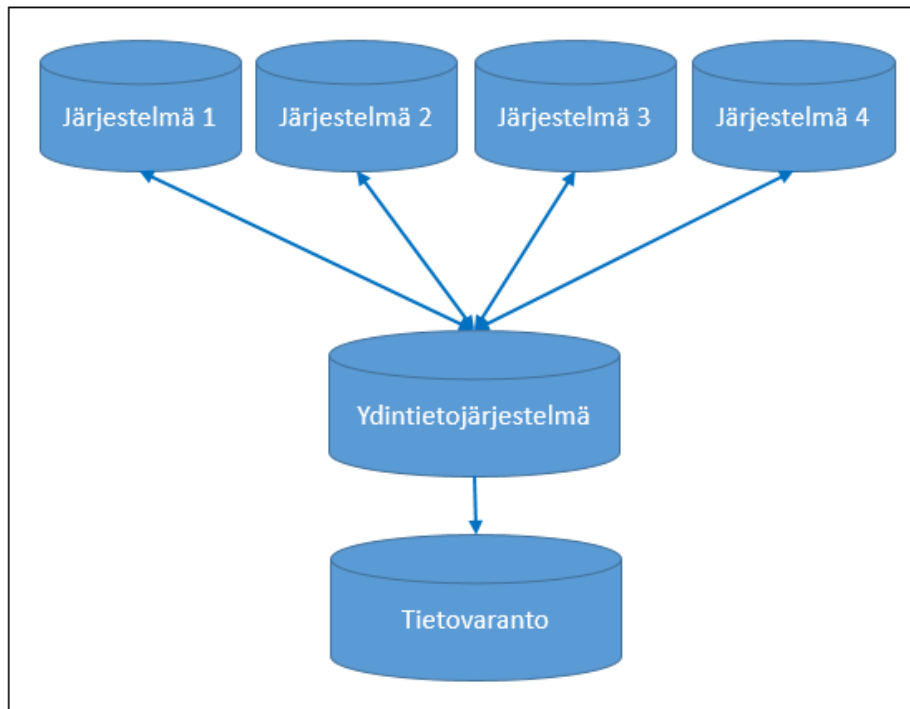
**Parallel -mallille** on tunnusomaista se, että siinä erotetaan toisistaan keskitetyn ydintiedon ylläpito (kuvassa 12 järjestelmät 1-4), keskitetty ydintietojärjestelmä, sekä lokaalit järjestelmät (kuvassa 12 järjestelmät x, y, ja z). Lokaalien järjestelmien käyttäjiä tuetaan tie-

toivintoja hallinnoimalla silloin kun luodaan ja ylläpidetään ydintietoja. Ylläpitoon liittyvät periaatteet on määritelty yhtenevästi keskitetyn ydintietojärjestelmän kanssa, joka antaa yksilöivän tunnisteiden ydintiedoille, ja tarkistaa mahdolliset dublikaatit, ja muut epä johdonmukaisuudet datassa. (Otto 2012, 342-343)



Kuva 12. Parallel näkökulma (mukailtu Otto 2012, 342)

**Hybridi -malli** on vuorovaikuttaisen hub mallin ja rekisteri mallin yhdistelmä. Siinä on erillinen tietovaranto kuten rekisterissä, mutta se sisältää myös useimmat attribuutit kuten vuorovaikutteinen hub. Kuten kuvassa 13 on esitetty, malliin kuuluu keskitetty ydintietojärjestelmä, missä ylläpidetään yksilöiviä avaimia, ja attribuutteja jotka ovat yhteiset eri järjestelmille, sekä metadataa, ja sääntöjä ylläpidolle. Liiketoiminnan järjestelmät ovat kiinni omissa tietovarannoissaan. Hybridi koostaa tiedot lokaaleista järjestelmistä, laatu varmistetaan, tiedot yhtenäistetään, ja dublikaatit poistetaan. Puhdas tieto palautetaan takaisin lokaaleihin järjestelmiin, mikä voi olla haastavaa silloin kun lokaalijärjestelmä on tällä välin jatkanut datan muokkausta. Ylläpitoprosessista, ja rooleista ja vastuista sopiminen onkin tässä mallissa erityisen tärkeää. (Kashel, Bullerwell & Kent 2011, 51-52.)



Kuva 13. Hybridi (mukailtu Kashel, Bullerwell & Kent 2011, 52)

**Yhdenmukaistetut tietovarannot -malli** käsittää yhden keskitetyn tietovarannon sijaan useita tietovarantoja. Tietovarannot rakennetaan kuitenkin yhdenmukaisiksi, jolloin osa dimensioista on yhteisiä. On huolehdittava erityisesti siitä, että järjestelmiä rakennettaessa päällekkäiset dimensiot pysyvät samoina. Keskitetyn ratkaisun puolestapuhujat sanovat yhdenmukaistettujen tietovarantojen olevan hankalasti laajennettavissa. Lisäksi katsotaan, että yhdenmukaisessa ratkaisussa tietovarannot suunnitellaan aina jonkun tietyn käyttäjäryhmän tarpeiden mukaan, jolloin niistä ei tule yhteiskäyttöisiä. Keskitettyjen tietovarantojen sanotaan olevan huonompia yhdenmukaistettuihin verrattuna. (Hovi, Hervonen & Koistinen 2009, 28.)



## **6 Keskitetyn ydintiedonhallinnan hyödyt ja haitat yritykselle**

Tähän lukuun on koottu lähdemateriaaliin pohjautuva tieto keskitetyn ydintiedonhallinnan hyödyistä ja haitoista. Koosteen tavoitteena on antaa nopeasti ja selkeästi tietoa siitä, mitkä ovat keskitetyn ydintiedonhallinnan hyödyt ja haitat yritykselle. Hyödyt ja haitat on esitetty omissa alaluvuissaan, jolloin ne voi tarvittaessa helposti kopioida vaikkapa liitteeksi erilaisiin selvityksiin tai esityksiin.

### **6.1 Hyödyt**

#### **Kustannussäästöt**

Niemi & Kontran (2012, 4) mukaan säästöjä syntyy IT-kustannuksissa, kun arkkitehtuuri yksinkertaistuu ja point-to-point integraatiot vähenevät, sekä prosessien tehostuessa.

Hovi, Hervonen & Koistisen (2009, 25) mukaan säästöjä syntyy, kun asiat tehdään vain kerran ja yhdessä paikassa, koska päällekkäinen työ aiheuttaa usein piilokustannuksia.

Loshinin (2009, 11) mukaan kustannukset alenevat, kun erilaiset lisenssi- ja ylläpitokustannukset vähenevät.

#### **Parempi tuottavuus ja kilpailukyky**

Loshinin (2009, 11) mukaan operatiivinen tehokkuus ja sitä kautta tuottavuus parantuu keskitetyn ydintiedonhallinnan vähentäessä kompleksisuutta, mikä aiheutuu saman tiedon moninkertaisesta hallinnasta ja päällekkäisestä työstä niin käyttäjätasolla kuin järjestelmätasollakin.

Loshinin (2009, 11) mukaan keskitetty ydintiedonhallinta parantaa kilpailukykyä, sillä se mahdollistaa nopeamman reagointikyvyn uusissa tilanteissa.

Loshin (2009, 1) toteaa että niissä yrityksissä, missä tietoja osataan yhdistellä ja hyödyntää operatiivisessa toiminnassa, tiedoista saadaan kilpailuetua. Tietojen yhteiskäyttö ja yhdisteltävyys ei paranna ainoastaan järjestelmien yhteentoimivuutta, se parantaa myös ihmisten yhteistoimintaa, ja tehostaa johtamista (Loshin 2009, 1).

## **Parantunut asiakaspalvelu ja asiakastyytyväisyys**

Loshinin (2009, 10) mukaan ydintiedonhallinta, missä kaikki asiakastieto on saatavissa yhdestä paikasta, niin operatiivisen toiminnan kuin analytiikankin tarpeisiin, edistää kattavaa asiakastuntemusta.

Loshin (2009, 10) toteaa että keskitetty ydintiedonhallinta mahdollistaa sen, että asiakkaan tarpeisiin voidaan vastata nopeammin ja paremmin, mikä parantaa asiakaspalvelua.

## **Parantunut raportointi ja tietojen analysointi**

Loshinin (2009, 11) mukaan keskitetty ydintiedonhallinta tarjoaa yhden datalahteen ydintiedoille, ja sitä kautta yhtenäistää raportointia

Loshin (2012, 12) kertoo keskitetyn ydintiedonhallinnan tuovan tiedot nopeasti uusienkin raporttien ja järjestelmien ulottuville, ja näin nopeuttavan tulosten saantia.

Hovi, Hervonen & Koistisen (2009, 25) mukaan tietojen raportointi ja analysointi helpottuvat ja nopeutuvat keskitetyn ydintiedonhallinnan ansiosta.

Loshinin (2009, 12) mukaan kulujen analysointi ja suunnittelu tehostuvat, kun keskitetty ydintiedonhallinta liittää datan osalta yhteen esimerkiksi tuotteet, toimittajat ja myyjät, jolloin se tukee hankintaa, myynninohjausta ja toimittajan kanssa tehtävää yhteistyötä.

## **Ketterämpi kehittäminen**

Loshinin (2009, 12) mukaan keskitetty ydintiedonhallinta yksinkertaistaa järjestelmäkehitystä, sillä ydintiedonhallintaan ei tarvita toiminnallisuuksia uusissa järjestelmissä, mutta siitä huolimatta tieto on niidenkin käytettävissä.

## **Parempi riskienhallinta**

Niemi & Kontran (2012, 4) mukaan keskitetty ydintiedonhallinta on merkittävä tekijä riskienhallinnassa ja riskien minimoinnissa, esimerkiksi tietoturvaan ja yksityisyydensuojaan liittyviä riskejä tarkasteltaessa.

Loshinin (2009, 11) mukaan keskitetty ydintiedonhallinta parantaa riskienhallintaa, sillä se poistaa haasteen mikä muodostuu mahdollisista tietojen dublikaateista, epäyhtenäisyydestä, ja puuttumisesta silloin kun riskien arviointia tehdään.

### **Parempi laatu**

Loshinin (2009, 12) mukaan keskitetty ydintiedonhallinta mahdollistaa paremman datan laadun ja sitä kautta esimerkiksi paremman sääntelyn noudattamisen. Tietojen auditointi on yksinkertaisempaa, ja sääntöjen luominen voidaan tehdä keskitetysti yhteen paikkaan, mikä tehostaa sääntelyn noudattamista (Loshin 2009, 12).

Niemi & Kontran (2012, 5) mukaan keskitetty malli pystyy takaamaan heti alusta alkaen paremman tiedon laadun. Keskittämisen kannalla olevat henkilöt ovat usein sitä mieltä, että vaikka keskitetyn ratkaisun tekemiseen voi mennä alussa enemmän aikaa, se säästelee myöhemmin, kun ei tarvitse korjata puutteita ja virheitä tiedoissa (Niemi & Kontra 2012, 5).

### **Parantunut päätöksenteko**

Loshinin (2009, 11-12) mukaan keskitetty ydintiedonhallinta mahdollistaa Business Intelligence -aktiviteetit ja erilaisten skenaarioiden tekemisen ja vertailun, parantaa luottamusta saatavissa olevaan tietoon, ja mahdollistaa näin nopeamman päätöksenteon.

## **6.2 Haitat**

### **Joustamattomuus**

Loser, Legner & Gizanis (2004, 2) toteaa, että vaikka keskitetty luonti ja ylläpito takaavat määritetyn prosessin noudattamisen, on se tämän vuoksi myös joustamattomampi.

### **Hitaus**

Niemi & Kontra (2012, 5) toteaa, että vaikka keskitetty tiedon luonti takaa paremman tiedon laadun, on kääntöpuolena mahdolliseen tiedon avaamiseen odottamiseen kuluva aika.

## **Monimutkaisuus ja laajuus**

Hovi, Hervonen & Koistisen (2009, 25) mukaan keskitetystä ydintiedonhallinnasta tulee usein monimutkainen ja laaja kokonaisuus, jolloin myös sen rakentamiseen kuluu paljon aikaa.

## **Ylläpidon haasteet**

Hovi, Hervonen & Koistisen (2009, 25) mukaan keskitetyn ratkaisun rakentaminen ja ylläpito vaativat henkilöresursseja.

Loshin (2009, 16) toteaa, että henkilöt tarvitsevat koulutusta ja perehdytystä, jotta datan hallinta onnistuu niillä tavoilla ja menetelmillä, joista on sovittu yhdessä eri organisaatioiden ja johdon kanssa.

Loshin (2009, 17) muistuttaa, että teknologia integraatioita luotaessa on ymmärrettävä, että teknologian tulee tukea toimintamallia, eikä toisinpäin.

## **Keskitetty ydintiedonhallinta aiheuttaa epäluuloja**

Loshinin (2009, 13) mukaan keskitetty ydintiedonhallinta aiheuttaa epäluuloja, koska asioita yritetään ratkaista tekniikkavetoisesti, jolloin esiin nousseet haasteet ovat johtaneet kompromisseihin, jotka eivät ole tukeneet liiketoimintaa.

Loshin (2009, 13) toteaa, että analyysien tulokset ovat lopulta vain niin hyviä kuin yrityksen kyvykyys toimia löydetyn tiedon perusteella, ja mitata toimintaa tehtyjen päätösten valossa.

## 7 Pohdinta

Opinnäytetyössä haluttiin selvittää, onko keskitetty ydintiedonhallinta parempi kuin hajautettu ydintiedonhallinta, vai löytyykö paras ratkaisu lopulta näiden kahden välimaastosta, niin sanotusta hybridimallista? Ja toisaalta, mikä on keskitetyn ydintiedonhallinnan merkitys yritykselle hyötyjen ja haittojen näkökulmasta?

Tieto- ja järjestelmäarkkitehtuurin näkökulmasta ydintiedonhallintamallit, eli ydintietoarkkitehtuurit ovat ryhmiteltävissä kolmeen luokkaan. Nämä luokat ovat: keskitetty malli, hajautettu malli, ja hybridimalli.

Keskitetyssä mallissa tietoa hallinnoidaan keskitetysti. Tämä tarkoittaa sitä, että esimerkiksi asiakastiedot luodaan ja ylläpidetään yhdessä järjestelmässä, mistä ne voidaan tarvittaessa siirtää muiden asiakastietoa tarvitsevien järjestelmien käyttöön. Hajautetussa mallissa tietoa hallinnoidaan eri järjestelmissä, ja lähtökohtaisesti kunkin järjestelmän omiin tarpeisiin. Tällöin esimerkiksi yrityksen asiakastiedot ovat näissä erillisissä järjestelmissä ylläpidettyjen tietojen summa, mihin saattaa sisältyä osittain päällekkäisiä ja osin ristiriitaisiakin tietoja. Hybridimalli on keskitetyn ja hajautetun mallin välimuoto, onhan siinä piirteitä kummastakin mallista. Siinä pyritään säilyttämään yksittäisten tietojärjestelmien itsenäisyys, ja tiedon luonnin ja hallinnan ketteryys. Jottei tiedon laadusta kuitenkaan tingittäisi, on hybridimallissa oltava mekanismeja tietojen eheyden ylläpitämiseksi ja laadun varmistamiseksi. Näitä mekanismeja ovat muun muassa keskitetty tietojen laadun ja eheyden tarkistelu ja korjaaminen, sekä eri järjestelmien välisen vuoropuhelun, ja mallin yleisen toiminnan kannalta välttämättömien mäppäysten ylläpito.

Ydintiedonhallinnan keskittäminen tuo yritykselle selkeitä hyötyjä. *Kustannussäästöjä* syntyy, kun tiedonhallinnan prosesseja tehostetaan ja keskitetään, jolloin tiedonhallintaa tehdään yhdessä paikassa ja vain kerran. Toisaalta tiedonhallinnan näkökulmasta arkkitehtuuri yksinkertaistuu, jolloin erilaiset ylläpitokustannukset vähenevät. Tiedonhallinnan keskittäminen johtaa *parempaan tuottavuuteen ja kilpailukykyyn*, mikä saavutetaan sillä, että tietoja hallitaan vain yhdessä paikassa, yksien ylläpitäjien toimesta. Kilpailukyky paranee, koska keskitetysti hallitut ydintiedot mahdollistavat yrityksen tilanteen reaaliaikaisen arvioinnin, ja tilanteisiin reagoinnin. *Asiakaspalvelu ja asiakastyytyväisyys paranevat*, sillä kuka tahansa asiakasrajapinnassa toimiva saa nähtäväkseen kattavat asiakastiedot ja voi sen perusteella keskustella asiakkaan kanssa esimerkiksi tehdyistä töistä, ja suunnitella tulevia töitä faktoihin perustuen. Kokonaiskuva asiakkaasta mahdollistaa myös potentiaalisen lisämyynnin havaitsemisen. Keskitetty ydintiedonhallinta *parantaa raportointia, tietojen analysointia, ja mahdollistaa paremman päätöksenteon*. Raportoinnin pohjana on yksi

yhteinen ydintiedon datalähde, mikä varmistaa, että raportit ja niissä olevat tiedot ovat yhtenevät, eikä esimerkiksi muuttunut tuotetieto ole jäänyt yhdessä järjestelmässä päivittämättä, aiheuttaen eroja varastotilannetta arvioitaessa. Keskitetty ydintiedonhallinta mahdollistaa tietojen analysoinnin ja erilaisten skenaarioiden tekemisen ja vertailun, mikä mahdollistaa nopeamman päätöksenteon. Keskitetty ydintiedonhallinta *mahdollistaa ketterän kehittämisen*, koska uusiin järjestelmiin ei tarvitse rakentaa ydintiedonhallintaan tarvittavia toiminnallisuuksia, mutta siitä huolimatta ydintieto on näidenkin järjestelmien käytävissä. Kun tiedot ylläpidetään yhdessä paikassa, se voidaan suojata paremmin / keskitetysti, ja näin tietoturvaan ja yksityisyydensuojaan liittyvät riskit vähenevät. *Parempi riskienhallinta* saavutetaan myös sillä, että keksitetysti hallitut tiedot pysyvät eheinä. Ydintiedonhallinnassa ja sen keskittämisessä merkittävä, jollei merkittävin hyöty on *parempi tiedonlaatu*. Kaikki edellä mainitut hyödyt edistävät omalla tavallaan myös laatua, ja keskitetyssä mallissa laatu voidaan taata jo heti alusta asti yhtenäisten sääntöjen ja käytäntöjen kautta.

Keskitetyssä ydintiedonhallinnassa on kuitenkin nähtävissä myös omat haittansa. Malli nähdään *joustamattomaksi ja hitaaksi* lähinnä siitä syystä, että keskitetty tiedon luonti ja ylläpito voidaan nähdä byrokraattisena. Se että joku jossain luo tiedon on prosessi, mihin kuluu aikaa. Keskitetty malli koetaan usein *laajaksi ja monimutkaiseksi*, minkä rakentamiseen kuluu aikaa erityisesti hankkeen alussa. Keskitetty ydintiedon *ylläpito on myös haastavaa*, sillä se vaatii keskitettyä resursointia. Ylläpitäjät tulee kouluttaa, ovathan he vastuussa yrityksen ydintiedon, yrityksen arvokkaimman tietovarannon oikeellisuudesta ja laadusta. Eipä ihme, että keskitetyn ydintiedonhallinnan paremmuutta myös epäillään. Ovatko nämä edellä mainitut haitat, se hinta mikä pitää maksaa tiedon ja toiminnan laadukkuudesta, vai voisiko näitä haittoja jotenkin välttää?

Ydintiedonhallintamalleja sekä niiden hyötyjä ja haittoja tarkasteltuani koen, että keskitetty ydintiedonhallinta, tietyn yrityksen tieto- ja arkkitehtuuriviitekehystä erillään, olisi parempi hajautettuun malliin verrattuna. Työssä esitellyissä hybridimalleissa voisi sen sijaan olla äkkiä katsottuna elementtejä, joilla puhtaasti keskitetyn ydintiedonhallinnan haittoja voitaisiin minimoida. Jousto, nopeus, yksinkertaisemmat toimintamallit, ja helpompi ylläpito voisi löytyä siten, että tietoja luotaisiin ja ylläpidettäisiin eri henkilöiden toimesta paikallisissa järjestelmissä. Mielestäni tarkempi hybridimallien tarkastelu kuitenkin osoittaa, että tiedon luontivaiheessa saavutettava jousto ja nopeus menetetään kuitenkin myöhemmin, kun tiedot pitää käyttää säännöllisin väliajoin keskitetyn tiedonhallinnan tarkistettavana ja korjattavana, jotta yhtenäisen tiedon ja laadun vaade toteutuvat. Hybridimallissakin edellytetään siis tietyn keskitetyn resurssin olemassa oloa. Mielestäni hybridimalleille tyyppisten määppäysten hallinnointi haastaa puolestaan tämän toimintamallin yksinkertaisuutta ja ylläpidon helppoutta.

Olenkin taipuvainen ajattelemaan samalla tavalla kuin Niemi & Kontran (2012, 5) mainitsemat, keskittämisen kannalla olevat henkilöt, jotka ovat sitä mieltä, että vaikka keskitetyn ratkaisun tekemiseen voi mennä alussa enemmän aikaa, se säästetään myöhemmin, kun ei tarvitse korjata puutteita ja virheitä tiedoissa. Lopulta osin samat ja osin eri seikat tekevät hybridimallista ja hajautetusta mallista mielestäni huonomman vaihtoehdon keskitetylle ydintiedonhallinnalle. Ainakin näin teoriassa.

Mutta onko keskitetty malli malleista paras myös käytännössä? Nähdäkseni mallien paremmuuden arviointi lähtee yrityksen nykytilanteen kartoittamisesta ja sen tieto- ja järjestelmäarkkitehtuurin tarkastelusta. Kuten Niemi & Kontra (2012, 5) toteavat, ei ole mahdollista antaa yhtä suositusta mikä malli sopisi kaikille, ja yrityksen tuleekin selvittää, sopiiko sille paremmin tiedonhallinnan keskittäminen vai hajauttaminen.

Työtä kirjoittaessa oli hyvä huomata, ettei keskitetyssä ydintiedonhallinnassa ja ydintiedonhallinnassa ylipäänsä ole kysymys mistään erityisen merkillisestä asiasta. Se on oikeastaan vain yleisestä tiedonhallinnasta esiin nostettu näkökulma, missä keskitytään yleisestä tiedonhallinnasta poiketen yrityksen tärkeimmäksi katsottuun tietoon, ydintietoon. Kun ydintiedosta puhutaan, tulee usein mieleen tietyt tietolajit, kuten asiakastiedot, tuotetiedon, toimittajatiedot, ja niin edelleen. Ymmärsin kuitenkin, että ydintieto on lopulta se tieto, minkä yritys kokee itselleen kaikkein arvokkaimmaksi tietovarallisuudeksi, ja se voi näin vaihdella yritysten välillä.

Olivatpa ydintiedot ja ydintiedonhallintamallit mitkä tahansa, yrityksen ydintiedonhallinta tulee nähdä koko yrityksessä eri toimintojen vuoropuheluna, ja sillä tulee olla johdon tuki. Jotta ydintiedonhallinnassa saavutetaan ne hyödyt mitä on mahdollista saavuttaa, pitää perusta sille luoda jo yrityksen strategiatyössä. On ymmärrettävä liiketoimintaprosessit missä ydintietoa käytetään, sopia tarvittavista rooleista ja vastuista, sekä tukea kokonaisuutta sopivilla ICT-ratkaisuilla.

Opinnäytetyö toimi itselleni välineenä hankkia lisää tietopääomaa ydintiedonhallinnasta, ja erilaisista tieto- ja järjestelmäarkkitehtuurimalleista siihen liittyen. Koen että minulla on nyt paremmat valmiudet suunnitella ydintietoarkkitehtuuria yrityksen tarpeisiin. Ajatus keskitetyn ydintiedonhallinnan paremmuudesta ei ollut väärä, mutta on tärkeää ymmärtää myös sen mukanaan tuomat haitat ja rajoitteet. Oli myös tärkeää oivaltaa, ettei ydintieto ole sama kaikissa yrityksissä, ja että sopiva ydintiedonhallinnan malli riippuu lopulta siitä, mikä on yrityksen tieto- ja järjestelmäarkkitehtuurin lähtötilanne.

## Lähteet

Ackoff, R. 1999. From data to wisdom. Luettavissa: <http://faculty.ung.edu/kmelton/Documents/DataWisdom.pdf>. Luettu: 21.3.2017

Choo, C. W. 2002. Information management for the intelligent organization: the art of scanning the environment. Information Today. Medford, New Jersey.

Cleven, A., & Wortmann, F. 2010. Uncovering four strategies to approach master data management. Luettavissa: <https://pdfs.semanticscholar.org/d5cb/9f4d106b02855ef77d6fc2038a99ebce6959.pdf>. Luettu: 23.4.2017.

Haug, A., & Arlbjørn, J. S. 2011. Barriers to master data quality. Journal of Enterprise Information Management, 24, 3, s. 288-303.

Hovi, A. 2009. Tietoarkkitehtuuri. Systeemityö, 2/2009, s. 12-14.

Hovi, A., Hervonen, H., & Koistinen, H. 2009. Tietovarastot ja Business Intelligence. WSOYpro. Helsinki.

JUHTA 2017. Julkisen hallinnon tietohallinnon neuvottelukunta. JHS-179 Kokonaisarkkitehtuurin suunnittelu ja kehittäminen. Luettavissa: <http://docs.jhs-suositukset.fi/jhs-suositukset/JHS179/JHS179.pdf>. Luettu: 20.3.2017.

Kashel, J., Bullerwell, M., & Kent, T. 2011. Microsoft SQL Server 2008 R2 Master Data Services. Packt Publishing. Birmingham.

Laatikainen, T., & Niemi E. 2012. Data quality johtavissa suomalaisyrityksissä: nykytila ja tarvittavat rakenteet. Luettavissa: [http://www.damafinland.fi/uploads/8/5/2/1/8521225/dama\\_dataquality\\_120430.pdf](http://www.damafinland.fi/uploads/8/5/2/1/8521225/dama_dataquality_120430.pdf). Luettu 17.4.2017.

Loser, C., Legner, C., & Gizanis, D. 2004. Master data management for collaborative service processes. Luettavissa: <https://www.alexandria.unisg.ch/65937/1/EN070.pdf>. Luettu: 23.4.2017.

Loshin, D. 2009. Master Data Management. Elsevier/Morgan Kaufmann. Burlington, Massachusetts.



Loshin, D., & Powell, R. J. 2003. Business Intelligence. Morgan Kaufmann. San Francisco.

Niemi, E., & Kontra, K. 2012. Data governance johtavissa suomalaisyrityksissä: nykytila ja yhdeksän suositusta. Luettavissa: [http://www.damafinland.fi/uploads/8/5/2/1/8521225/dama\\_datagovernance\\_120212.pdf](http://www.damafinland.fi/uploads/8/5/2/1/8521225/dama_datagovernance_120212.pdf). Luettu 20.4.2017

Ofner, M. H., Straub, K., Otto, B. & Oesterle, H. 2013. Management of the master data lifecycle: a framework for analysis. Journal of Enterprise Information Management, 26, 4, s. 472-491.

Otto, B. 2012. How to design master data architecture: Findings from a case study at Bosch. International Journal of Information Management, 32, s. 337-346.

Radcliffe, J. 2007. The Seven Building Blocks of MDM: A Framework for Success. Luettavissa: <http://www.gartnerinfo.com/mdm2/TheSevenBuildingBlocksofMDMAFrameworkforSuccess.pdf>. Luettu: 23.4.2017.

Waltz, E. 2003. Knowledge management for the intelligent enterprise. Artech House. Norwood.