

Emilia Oikari

**DATA-ANALYTIIKAN JA -STRATEGIAN MERKITYS  
LIIKETOIMINNAN KEHITTÄMISESSÄ**



JYVÄSKYLÄN YLIOPISTO  
INFORMAATIOTEKNOLOGIAN TIEDEKUNTA  
2023

# TIIVISTELMÄ

Oikari Emilia

Data-analytiikan ja -strategian merkitys liiketoiminnan kehittämisessä

Jyväskylä: Jyväskylän yliopisto, 2022. 33 s.

Tietojärjestelmätiede, kandidaatin tutkielma.

Ohjaaja(t): Halttunen, Veikko

Tutkimusten mukaan jo vuonna 2013 maailman ihmisistä kaksi kolmasosaa käytti jotain mobiililaitetta, jolta saatiin kerättyä merkittävä määrä dataa talteen. Tästä voidaan vain päätellä kuinka suuri osa maailman väestöstä käyttää jonkinlaista vastaavaa dataa keräävää tai tuottavaa laitetta vuonna 2022. Ihmisten käyttämistä laitteista kerätään jatkuvalla syötöllä dataa erilaisiin käyttötarpeisiin. Tätä isoa joukkoa dataa kutsutaan massadataksi (engl. Big data) ja sen tuomista mahdollisuuksista puhutaan jatkuvasti nyky-yhteiskunnassa. Tutkimuksia datasta ja sen erilaisista hyödyntämiskeinoista on tehty jo vuosikymmenien ajan, mutta liiketoiminnallista näkökulmaa on tässä aihepiirissä tutkittu jokseenkin vielä suppeasti. Data voidaan nähdä yrityksen aineettomana pääomana ja sitä useinkin pidetään organisaatiossa tärkeänä resurssina. Haaste on kuitenkin siinä, että data ei kuitenkaan itsestään tuota yritykselle tarvittavaa lisäarvoa tai kilpailuetua. Tässä tutkielmassa tarkastellaan dataa ja data-analytiikkaa etenkin liiketoiminnallisesta näkökulmasta sekä käsitellään datastrategian roolia datan tuomien haasteiden ehkäisemisessä, liiketoiminnan kehittämisessä sekä kilpailuedun tuottamisessa. Tutkielmassa pyritään luomaan mahdollisimman kokonaisvaltainen kuva data-analytiikan liiketoiminnallisesta merkityksestä sekä datastrategian roolista datan hyödyntämisessä organisaation resurssina. Tutkielman tuloksista ilmenee, että datan hyödyntämiseen liittyy merkittävässä määrin erilaisia haasteita, jonka vuoksi yritysten tulisi paremmin varautua datan hyödyntämiseen ja näin ollen kehittää omaa datapolitiikkaansa datan strategisen suunnittelun ja sen hyödyntämisen avulla.

Asiasanat: Data-analytiikka, datastrategia, massadata, ydindata, liiketoiminnan kehittäminen, liiketoimintatiedon hallinta

## ABSTRACT

Oikari, Emilia

The meaning of data-analytics and strategy in the business development

Jyväskylä: University of Jyväskylä, 2022, 33 pp.

Information Systems/ Bachelor's Thesis

Supervisor(s): Halttunen, Veikko

Research shows that already in 2013, two thirds of the world's people used a mobile device that can be used to collect and share data. We can only imagine how much of the world's population are using similar devices in 2022. There is a constant stream of data being collected from the devices people use for different purposes. This mass of data is called Big Data and its potential is constantly being discussed in today's society. Research on data and its use has been going on for decades, but the business perspective is not yet well represented in these circles. In this thesis, we will look at data and data analytics particularly from a business perspective, as well as the role of data strategy in business development and bringing competitive advantage. The aim is to create as comprehensive a picture as possible of data analytics and its importance in business. The results of the study show that there are many different challenges related to data utilization, which is why companies should prepare and plan their own data policies and data usage strategies more carefully.

Keywords: Data-analytics, Data strategy, Business development, big data, master data, business intelligence

## KUVIOT

Kuvio 1 Porterin yleisstrategiamalli (Porter, 1985).....	16
Kuvio 2: Datan hyödyntämisen prosessi (Martino ym., 2014).....	20

## TAULUKOT

Taulukko 1: Massadatan ominaisuuksista (Rijmenam, 2014, Krishnamurthy & Desouza, 2014, Sivarajah ym., 2017, Chen ym., 2013) .....	12
Taulukko 2: Datan käytön suurimmat haasteet(Chen ym., 2013; Hashem ym., 2015; Krishnamurthy & Desouza, 2014; Lycett, 2013; Manyika ym., 2011; Pehcevski, 2019; Rijmenam, 2014) .....	21
Taulukko 3: Datan hyödyntäminen päästrategioiden toteuttamisessa (Constantiou & Kallinikos, 2015; Henke ym., 2016; Manyika ym., 2011; Porter, 1985, 1991).....	25

# SISÄLLYS

TIIVISTELMÄ

ABSTRACT

KUVIOT JA TAULUKOT

1	JOHDANTO.....	6
1.1	Tutkimuskysymys .....	7
1.2	Tutkimuksen toteuttaminen.....	9
2	DATA-ANALYTIikka JA TIEDONHALLINTA .....	10
2.1	Ydindata .....	10
2.2	Massadata .....	11
2.3	Business Intelligence .....	13
3	STRATEGIA .....	14
3.1	Strategia yleisesti .....	14
3.2	Datastrategia.....	17
4	DATASTRATEGIAN ROOLI DATALÄHTÖISESSÄ LIIKETOIMINNAN KEHITTÄMISESSÄ .....	19
4.1	Datastrategian merkitys liiketoiminnalle.....	19
4.2	Esimerkkejä datastrategian hyödyistä.....	25
5	YHTEENVETO JA JATKOTUTKIMUSAIHEET .....	28
	LÄHTEET.....	31

# 1 JOHDANTO

Data-analytiikka on jo vuosien ajan ollut yksi suurimmista digitalisaation trendeistä. Datan hyödyntäminen ei kuitenkaan ole niin uusi keksintö kuin voisi luulla, sillä Rijmenamin (2014) mukaan ensimmäiset todistetut jäljet datan hyödyntämisestä yrityksen käyttöön peräytyvät kuitenkin jo yli 7000 vuoden taakse Mesopotamiaan. Data-analytiikan perimmäisenä tarkoituksena on koostaa suuria määriä dataa ja muodostaa niistä helppolukuista materiaalia organisaation hyödynnettäväksi. Kerätty data koostetaan useimmiten visuaalisesti näyttäväksi kokonaisuudeksi, jotta sen tulkinta on helppoa kaikille dataa mahdollisesti hyödyntäville osapuolille. Data-analytiikan avulla on tarkoitus luoda muuten monimutkaisesta datasta helppolukuista myös niille, jotka eivät ole raakadatan tulkinnan ammattilaisia. Data-analytiikalla voidaan siis tiivistetysti tarkoittaa kaikkea sitä, kun kerätystä datasta jalostetaan edelleen käytettävää tietoa esimerkiksi liiketoiminnan kehittämiseksi (Vincit, 2022). Data ja sen analytiikka ovat jo vuosikymmeniä olleet suuressa nosteessa. Datatieteen nopeaa kasvua onkin tukenut yksityisten tietojen lisäksi datatalouden ja digitalisoinnin kehittyminen ja räjähtävä kasvu (Cao, 2018). Data-analytiikkaa, sen eri muotoja sekä data-analytiikan käyttöönoton yleisimpiä haasteita käsitellään tarkemmin myöhemmissä luvuissa.

Vaikka data-analytiikka on nykypäivänä hyödyllinen väline yrityksen toiminnan tehostamisessa ja ohjaamisessa, ovat useat yritykset edelleen tämän osalta kehityskaaren alkupäässä. Ongelma piilee useasti siinä, että monissa organisaatioissa data-analytiikan merkitystä ja sen tuomia mahdollisuuksia tai uhkia ei ole lainkaan huomioitu yrityksen strategian toteuttamisessa. Datan hyödyntämisessä ja käsittelyssä on useita seikkoja, jotka organisaatioissa tulisi huomioida, jotta datasta saadaan kustannustehokkaasti tärkeää tietoa irti. Nykypäivänä perinteisen strategisen päätöksenteon rinnalla olisi tärkeää huomioida ja määritellä erikseen myös yrityksen datastrategia, mikäli datalla halutaan saavuttaa tuntuva vaikutusta yritykselle.

Data-analytiikasta puhuttaessa tarkoitetaan datalla yleensä massadataa (engl. Big data) tai ydindataa (engl. Master data) sekä niiden käsittelystä käytettävää termiä liiketoimintatiedon hallinta (engl. Business Intelligence) tai lyhennettä BI. Massadata on nimensä mukaisesti suuri määrä dataa, joka kehittyy ja

kasvaa jatkuvasti (Hovi, 2015). Datan jatkuva kasaantuminen tuo massadatan hyödyntämiseen haasteita, sillä saatu ja kerätty datamäärä on erittäin suurta ja sitä tulisi osata hyödyntää oikealla tavalla halutun päämäärän saavuttamiseksi. Massadataa ei useimmiten voida hyödyntää koko sen laajuudessaan, sillä dataa muodostuu mahdottoman suuria määriä jatkuvalla syötöllä. Jokainen verkon käyttäjä muodostaa huomattavan suuria määriä uutta dataa jatkuvasti ja datan muodostuminen sekä moninaistuminen on erittäin nopeaa. 66,2% koko maailman populaatiosta käytti Internetiä vuoden 2021 aikana säännöllisesti (*World Internet Users Statistics and 2022 World Population Stats*, 2022) ja tämä jatkuvasti kasvava ihmismäärä tuottaa uutta dataa jatkuvalla syötöllä. Nykypäivänä dataa kulkeutuukin verkossa jo alle sekunnissa saman verran kuin mitä sinne oli yhteensä tallennettuna 20 vuotta sitten (McAfee & Brynjolfsson, 2012).

Massadatasta poiketen ydindata on tiiviimpi joukko tietoja ja tämä datatyyppejä sisältää organisaation toiminnan kannalta tarkasteltuna ydintietoja asiakkaista (Hovi, 2015). Toisin kuin massadata, ydindata ei kehity ja muutu jatkuvasti ja näin ollen sitä voidaan hyödyntää useita vuosia organisaation käytössä ilman suurempia ylläpitoon liittyviä toimenpiteitä (Hovi, 2015). Ydindata sisältää useimmiten niitä tietueita, joita asiakkaat itse antavat ja tuottavat yritykselle. Ydindataksi luokitellaankin esimerkiksi asiakkaan yhteystiedot, joita voidaan kerätä esimerkiksi uutiskirje-tilausten yhteydessä tai yrityksen tuotteiden ja palveluiden ominaisuudet, jotka täytetään yrityksen järjestelmiin yrityksen toimesta (Väre, 2019). Näiden molempien datatyyppien tutkimiseksi tarvitaan liiketoimintatiedonhallintaa, jonka avulla kerätystä datasta voidaan muodostaa yrityksen hyväksi hyödynnettävää materiaalia.

Tässä tutkielmassa tutkitaan datastrategiaa ja sen merkitystä yrityksen liiketoiminnan kehityksessä ja sen strategisessa johtamisessa. Datastrategia perustuu data-analytiikkaan, sekä sen käyttöönottoon ja hyödyntämiseen yrityksessä myös strategisesta näkökulmasta. Tutkielmassa tarkastellaan ensin data-analytiikkaa ja sen tärkeimpiä alakäsitteitä sekä datan hyödyntämistä organisaatioissa. Tämän jälkeen tutkielmassa keskitytään yrityksen liiketoiminnan kehittämiseen strategisesta näkökulmasta. Vaikka data-analytiikan ja liiketoiminnan yhteydestä on vuosien varrella toteutettu useampiakin tutkielmia, ei datastrategian ja liiketoiminnan kehittämisen yhteydestä löydy vielä juurikaan materiaalia tai tutkimuksia. Sen vuoksi tämän kirjallisuuskatsauksen tarkoituksena on perehtyä syvällisemmin etenkin datan tuomiin liiketoiminnallisiin hyötyihin sekä datastrategian merkitykseen osana yrityksen liiketoiminnallista strategiaa.

## 1.1 Tutkimuskysymys

Tämän kirjallisuuskatsauksen on tarkoitus vastata kahteen tutkimuksessa esitettyyn tutkimuskysymykseen. Ensimmäinen tutkimuskysymys tässä tutkielmassa liittyy dataan ja sen analytiikkaan sekä näiden merkitykseen liiketoiminnan kehittämisessä. Ensimmäinen tutkimuskysymys on:

- Mikä on datan rooli liiketoiminnan kehittämisessä ja kilpailuedun luomisessa?

Tähän kysymykseen pyritään vastaamaan tutkielman toisessa luvussa. Tutkielman kolmannessa luvussa käsitellään strategiaa ja strategista johtamista, jonka myötä on syntynyt toinen tutkimuskysymys. Ensimmäisen tutkimuskysymyksen pohjalta on johdettu tutkielman toinen tutkimuskysymys:

- Voidaanko datalähtöisen liiketoiminnan kehittämisen kompastuskiviä estää datastrategian avulla?

Tähän tutkimuskysymykseen etsitään vastauksia tutkielman neljännessä luvussa.



## 1.2 Tutkimuksen toteuttaminen

Tutkimus toteutetaan laajana kirjallisuuskatsauksena. Kirjallisuuskatsauksessa tarvittavat lähteet on haettu pääosin verkosta löytyvistä tietokannoista kuten: Google Scholarista, IEEE Xplorerista, ScienceDirectistä, Scopuksesta sekä JYKDOK:ista. Hakusanoina on käytetty mm. seuraavia englanninkielisiä sanoja sekä niiden yhdistelmiä: *data strategy management, Big data in business, data strategy, data driven approach, business Intelligence, business development ja data management*. Tutkielman lähteet ja niiden laatu on pyritty pitämään mahdollisimman korkeana ja sen vuoksi jokaisen lähteen luokittelu on tarkistettu Julkaisufoorumista (<https://www.tsv.fi/julkaisufoorumi/haku.php?lang=en>). Lähteiksi on valikoitu lähtökohtaisesti vain sellaisia artikkeleita, jotka saavat julkaisufoorumissa vähintään 1-tason luokittelun. Päälähteiksi ja teorian tueksi on pyritty valitsemaan 2- tai 3-tason lähteitä, jotta tutkimuksen laatu pysyisi mahdollisimman korkeana. Joillekin tutkielmassa käytetyille lähteille ei kuitenkaan löydy luokittelua Julkaisufoorumista. Näiden lähteiden, kuten kirjojen osalta on tarkasteltu esimerkiksi Google Scholarin kautta aiempien viittausten määrää ja niiden pohjalta lähteiksi on valittu vain sellaisia teoksia, joilla on ollut mahdollisimman paljon aiempia viittauksia.

Tutkielman ensimmäisessä luvussa käsitellään tarkemmin tiedonhallintaa sekä data-analytiikkaa ja sen tärkeimpiä alakäsitteitä kuten ydindataa, massadataa sekä liiketoimintatiedon hallintaa. Luvun tarkoituksena on luoda pohja tutkittavalle aiheelle syventymällä data-analytiikan tärkeimpiin käsitteisiin. Käsitteiden läpikäynnin jälkeen suunnataan seuraavaan lukuun, jossa pohjustetaan yrityksen strategisen johtamisen näkökulmaa sekä aiheen käsitteistöä. Tämän jälkeen tutkielman neljännessä luvussa yhdistetään nämä kaksi aiemmin käsiteltyä aihealuetta yhdeksi kokonaisuudeksi niin hyvin kuin se tämän tutkielman tuloksilla on mahdollista. Lopuksi tutkielmassa käydään vielä läpi tutkielman myötä havaitut tulokset sekä pohditaan mahdollisia jatkotutkimusaiheita.

## 2 DATA-ANALYTIikka JA TIEDONHALLINTA

Tässä luvussa perehdytään data-analytiikan ja tiedonhallinnan maailmaan sekä määritellään tutkielmassa käytettäviä käsitteitä tarkemmin. Tiedonhallinta (engl. Data management) on kattokäsite muille data-analytiikkaan liittyville käsitteille. Kattokäsitteen alle kuuluvat muun muassa ydindata, massadata sekä Business Intelligence eli liiketoimintatiedon hallinta, joihin perehdytään tarkemmin tässä pääluvussa. Tämän tutkielman kirjallisuuskatsauksessa perehdytään tarkemmin massadatan maailmaan, mutta myös ydindata määritellään lyhyesti paremman kokonaiskuvan antamiseksi lukijalle. Viimeisessä alaluvussa käsitellään myös Business Intelligenceä, jonka virallinen määritelmä on suomeksi ”liiketoimintatiedon hallinta/hyödyntäminen”. Tästä käsitteestä käytetään sen suomenkielistä nimitystä ”liiketoimintatiedonhallinta” sekä lyhennettä BI, joka on yleisesti käytössä myös suomenkielisessä kirjallisuudessa.

### 2.1 Ydindata

Data-analytiikasta ja datasta puhuttaessa käsite jaetaan useimmiten kahteen alakäsitteeseen: massadataan (engl. Big data) sekä ydindataan (engl. Master data). Ydindatalla tarkoitetaan sellaisia yrityksen ydinliiketoimintaan vaikuttavia kriittisiä tietoja ja osia, joita hyödynnetään organisaation eri tasoilla ja joita ilman yrityksen toiminta ei olisi käytännössä mahdollista. Kyseessä on siis yrityksen toiminnan kannalta välttämätöntä avaindataa eli sellaisia tietoja, joita ilman yritys ei käytännössä voisi toimia lainkaan. Tällaisia tietoja ovat esimerkiksi asiakkaiden ja työntekijöiden yhteystiedot sekä tuotteiden tai osien ydintiedot ja ominaisuudet. (Loshin, 2010.)

Väreen (2019) mukaan ydindatalla on merkittävä rooli organisaatiossa ja hän kuvaileekin ydindatan vastaavan rakennuksen perustuksia ja kantavia seinä. Ydindata sijaitsee aina yrityksen ytimessä ja kaikki muu kerättävä ja analysoitava data rakennetaan tämän päälle (Väre, 2019). Toisin kuin massadata,

ydindata ei uusiudu ja kehity jatkuvalla syötöllä ja sen vuoksi kerättyä dataa voidaan hyödyntää jopa vuosia yrityksen käytössä ilman suuria datan hallinnan toimenpiteitä. (Hovi, 2015). Yksittäisiä ydindata tietueita harvoin päivitetään ja sen vuoksi tämän datatyypin ylläpitämisen voisi kuvitella olevan helppoa. Ydindatan käsittelyssä on kuitenkin omat ongelmansa. Spruitin ja Pietzkan (2015) mukaan ongelmia tuo esimerkiksi datan hajautuminen eri yksiköihin ja alueisiin sekä organisaatorakenne monine tasoineen. Näiden osalta on tärkeää huomioida datan tehokas integraatio niin, että kaikki tarvittava tieto on samassa paikassa tai helposti saavutettavissa (Spruit & Pietzka, 2015).

## 2.2 Massadata

Massadata (engl. Big data) on yksi tämän aikakauden suurimmista megatrendeistä ja sen vuoksi aiheesta on tehty viimeisen vuosikymmenen aikana hurja määrä tutkimuksia ja artikkeleita. Massadatalle on nykypäivänä niin suuri merkitys tieteeseen ja liiketoimintaan, että Sagirolu ja Sinanc (2013) jopa väittävät massadatan ja sen analysoinnin olevan nykyisen modernin tieteen ja liiketoiminnan perusta.

Massadatalle tarkoitetaan dataa, jota muodostuu huomattavan suurina määrinä lyhyessä aikajaksossa. Tällaista nopeasti monistuvaa ja muovautuvaa dataa ei voida tallentaa tavallisiin relaatiotietokantoihin. Massadataa muodostuu esimerkiksi kaikesta siitä, mitä ihminen tekee verkossa tai mobiililaitteillaan (Sagirolu & Sinanc, 2013). Jokainen lisätty kuva, kirjoitettu sähköposti, sisäänkirjautuminen, tykkäys, hakusana, sovellus sekä klikkaus muodostaa jatkuvasti uutta dataa, jota kutsutaan massadatakseksi (Zikopoulos, 2012). Rijmenam (2014) arvioi jo vuonna 2014, että kaikesta ihmisten historian aikana kerätystä datasta noin 90 % oli tuolloin kerätty viimeisen kahden vuoden aikana. Teknologian sekä sen käytön jatkuva kehittyminen ja yleistyminen vain nopeuttaa tätä uudistumisprosessia, jonka vuoksi tämäkin arvio on todennäköisesti jo vanhentunut. Tästä voidaan vain päätellä kuinka lyhyessä ajassa nykypäivänä sama määrä dataa muodostuu ja minkä vuoksi tällaisen datamäärän analysointi on niin haastavaa.

Kompleksinen sekä jatkuvasti kehittyvä ja kasaantuva data on haastavaa hyödyntää. Toisin kuin ydindata, massadatan määrä moninkertaistuu jopa minuuteissa ja tunneissa, joka muodostaa haasteita datan tehokkaaseen analysointiin ja keräämiseen (Pehcevski, 2019). Datan moninaistuminen on ollut niin vauhdikasta, että tutkijat ovat joutuneet keksimään jo uusia matemaattisia termejä kuvaamaan muodostuneen datan määrää, sillä perinteisen desimaalisysteemin ohi on menty jo vuosikausia sitten (Rijmenam, 2014). Vaikka massadata on monelle melkoisen tuore käsite, on datasta ja sen analytiikasta puhuttu jo vuosikymmeniä. McAfee ja Brynjolfsson (2012) kuvailevat massadatan ja sen analytiikan (engl. analytics) olevan hyvinkin samankaltaisia käsitteitä, sillä molempien perimmäinen tarkoitus on kääntää dataa yrityksen liiketoiminnalliseksi eduksi. Näiden kahden käsitteen välillä on kuitenkin kolme merkittävää eroavaisuutta, joita kutsutaan massadatan kolmeksi V:ksi; volyymiksi, variaatioksi sekä vauhdiksi.

Näiden kolmen tyypin ominaisuuksia vertaillaan seuraavassa taulukossa (Taulukko 1):

**Taulukko 1: Massadatan ominaisuuksista (Rijmenam, 2014, Krishnamurthy & Desouza, 2014, Sivarajah ym., 2017, Chen ym., 2013)**

	Volyyymi	Variaatio	Vauhti
Kuvaus	Kerätyn datan suuri määrä. Dataa muodostuu nopeasti suuria määriä. Esimerkiksi Walmart kerää joka tunti yli 2,5 petatavua (eng. petabyte) dataa asiakastapahtumista	Kerätyn datan monipuolisuus. Dataa muodostuu useista eri lähteistä ja eri muodoissa. Jokainen teknologian käyttäjä on kävelevä datageneraattori. Data jaetaan strukturoituun ja strukturoimattomaan dataan.	Datan muodostumisen nopeus. Monesti datan monistumisen nopeus on kerättyä datamäärää tärkeämpää. Mahdollisimman reaaliaikainen informaatio mahdollistaa ketteryyden ja kilpailuedun.
Ominaisuudet	Muodostuvan datan koko: <ul style="list-style-type: none"> <li>- Megatavu</li> <li>- Gigatavu</li> <li>- Teratavu</li> <li>- Petatavu</li> </ul>	Muodostuvan datan monipuolisuus. mm.: <ul style="list-style-type: none"> <li>- JPEG</li> <li>- MP4</li> <li>- SMS</li> <li>- XML</li> <li>- GPS</li> <li>- tykkäykset, kommentit, postaukset</li> </ul>	Datan muodostumisen nopeus: <ul style="list-style-type: none"> <li>- reaaliaikainen</li> <li>- lähes reaaliaikainen</li> <li>- määräjain</li> <li>- erissä</li> </ul>
Haasteet	Jatkuvasti muodostuvan ja lisääntyvän datan tehokas hyödyntäminen sekä muodostuneen datan yksityisyys	Monimuotoisen datan hyödyntäminen eri tarkoituksissa sekä erimuotoisten datatyypin skaalautumattomuus	Jo kerätty data vanhentuu nopeasti ja sen reaaliaikainen hyödyntäminen on haastavaa

Kuten aiemmasta taulukosta voidaan huomata, liittyy massadatan hallintaan paljon haasteita, jotka olisi hyvä tunnistaa ennakkoon ja joiden varalta yrityksellä tulisi olla jonkinlainen suunnitelma eri tilanteita varten. Dallemule ja Davenport (2017) sanovatkin artikkelissaan *"What's your data strategy?"*, että yritysten, joilla ei vielä ole data strategiaa, on alettava rakentamaan sitä mahdollisimman pian tai vaihtoehtoisesti he voivat alkaa suunnittelemaan yritykselleen exit-strategiaa. Blakelyn (2022) mukaan exit-strategialla tarkoitetaan yrityksen varasuunnitelmaa, jonka avulla henkilö voi poistua yrityksestä siinä tapauksessa, että yritys saavuttaa merkittävän voiton tai ei ole enää kannattava. Esimerkiksi yrityksen sijoittajat, omistajat, tai pääomasijoittajat voivat suunnitella irtautumisstrategioita, kun he sitoutuvat yritykseen. Exit-strategian tarkoituksena on rajoittaa tappioita tai maksimoida henkilökohtainen voitto, kun henkilö poistuu liiketoiminnasta. (Blakely, 2022.) Tämä väite kuvastaa hyvin datastrategian merkitystä nykyisessä teknologiayhteiskunnassa.

## 2.3 Business Intelligence

Massadata-analytiikka tarjoaa organisaatiolle lukuisia mahdollisuuksia parantaa sen liiketoiminnallista arvoa ja tuottavuutta. Itsessään data ei kuitenkaan tuota yritykselle lainkaan arvoa, vaan sitä täytyy pystyä hyödyntämään halutulla tavalla ylimääräisen arvon luomiseksi. Yksi tärkeimpiä massadata -analytiikan soveltamiskeinoja on liiketoimintatiedonhallinta (engl. Business Intelligence) eli BI, jonka avulla voidaan Ramin ym. (2016) mukaan muun muassa.:

- parantaa ja nopeuttaa organisaation päätöksentekokykyä
- kehittää asiakastarpeiden ymmärtämistä ja strategioiden kehittämistä uusien tuotteiden ja palvelujen lanseeraamiseksi
- tutkia uusia markkinoita ja segmenttejä
- parantaa varastojen kiertonopeutta
- nostaa henkilöstön tuottavuutta ja tehokkuutta

BI on teknologiapohjainen prosessi, jossa analysoidaan kerättyjä tietoja ja tuotetaan niistä käyttökelpoista tietoa, joka auttaa ylintä johtoa, johtajia ja työntekijöitä tekemään tietoon perustuvia liiketoimintapäätöksiä (*What Is Business Intelligence?*, 2020). BI:n avulla yritys voi hallita, visualisoida, jakaa ja analysoida dataansa paremmin ja näin ollen tehostaa datan avulla yrityksen liiketoimintaa.

Käytännössä liiketoimintatiedonhallinnalla voidaan siis tarkoittaa sitä, kun suuresta määrästä dataa muodostetaan analysoitavissa olevaa ja käyttökelpoista tietoa henkilöiden käyttöön. Liiketoimintatiedon hallintaa tarvitaankin datan hyödyntämiseksi, sillä ilman tätä prosessia, data ei ole organisaatioissa kaikkien hyödynnettävissä olevassa muodossa. Liiketoimintatiedonhallinnan ja data-analytiikan aloitteiden perimmäisenä tavoitteena on tehdä parempia liiketoimintapäätöksiä, joiden avulla organisaatiot voivat kasvattaa tulojaan, parantaa toimintansa tehokkuutta sekä muodostaa kilpailuetua kilpailijoihin nähden (*What Is Business Intelligence?*, 2020). Tutkimusten tulokset osoittavat, että liiketoimintatiedon hallinnalla on esimerkiksi suora yhteys yrityksen kansainvälistymisnopeuteen ja yhdessä organisaation ketteryuden kanssa yritys voi parantaa kansainvälistymisnopeuttaan. (Cheng ym., 2020)

## 3 STRATEGIA

Tässä luvussa perehdytään tarkemmin strategiseen liiketoiminnan johtamiseen sekä strategian merkitykseen. Ensimmäisessä alaluvussa käsitellään strategiaa ja strategista johtamista yleisellä tasolla. Toisessa alaluvussa määritellään datastrategia sekä perehdytään tarkemmin datan ja data-analytiikan strategiseen hyödyntämiseen.

### 3.1 Strategia yleisesti

Strategia voidaan nähdä tietyntylaisena karttana, oppaana tai reittinä kohti haluttua määränpäättä. Grossman (2018) määrittelee perinteisen yritysstrategian organisaation pitkän aikavälin suunnaksi ja laajuudeksi, jonka avulla saavutetaan organisaatiolle etua optimoimalla resursseja haastavassa ympäristössä ja täyttämällä näillä keinoin markkinoiden tarpeita ja sidosryhmien odotuksia. Perinteiset yrityksen strategisen suunnittelun tekniikat ja keinot keskittyvät kilpailuedun rakentamiseen ja ylläpitämiseen analysoimalla ulkoista ympäristöä ja määrittämällä yrityksen vahvuudet ja heikkoudet suhteessa siihen. Toiminnan tehokkuus ja strategia ovat molemmat välttämättömiä merkittävän kilpailuedun kannalta, joka on loppujen lopuksi jokaisen yrityksen ensisijainen tavoite ja päämäärä (Porter, 1991).

Strategisen johtamisen isänäkin pidetyn Porterin (1991) mukaan yritys voi menestyä kilpailijoitaan paremmin vain, jos se voi luoda kilpailijoihin nähden eron, jonka se voi myös säilyttää pitkällä aikavälillä eli luomalla itselleen kestävä kilpailuedun markkinoilla. Kestävän kilpailuedun myötä yritys voi tuottaa asiakkaille enemmän arvoa tai luoda vertailukelpoista arvoa alhaisemmilla kustannuksilla tai toisinaan yritys voi tehdä näitä molempia samanaikaisesti (Porter, 1991). Porterin (1985) mukaan strategista kilpailuetua voidaan hakea kolmella yleisellä päästrategiakeinolla:

**Erilaistaminen: (engl. Differentiation)**

Erilaistamisen ja eriyttämisen avulla voidaan saavuttaa merkittävä kilpailuetua. Eriyttämisellä tarkoitetaan sitä, että yritys keskittyy tuottamaan palveluita ja tuotteita, jotka eriävät jo muista markkinoilla olevista vastaavista palveluista ja tuotteista sen verran merkittävästi, että markkinakuninkuus on saavutettavissa. Yritys siis pyrkii mukauttamaan tarjontansa vastaamaan täysin markkinoiden kysyntää ja täyttää näin ollen markkina-aukon. Yrityksen ainutlaatuisuuden vuoksi yritys voi pyytää tuotteistaan korkeampaa hintaa. (Porter, 1985.)

**Kustannusjohtajuus (engl. Cost Leadership)**

Kustannusjohtajuus voidaan saavuttaa tuottamalla asiakkaille tarvittavaa arvoa kilpailijoita edullisempaan hintaan. Kun yritys ymmärtää kuluttajien kynnyshinnan ja hinta-laatu-kompromissin, voi yritys tarjota tuotteita muita edullisemmin ja näin saavuttaa kilpailujohtajuuden. (Porter, 1985.) Datalla voi olla tässä Porterin mainitsemissa keinossa merkittävä rooli, sillä datan avulla voidaan tulkita ja ennustaa näitä kustannusjohtajuuteen vaikuttavia tekijöitä huomattavasti tarkemmin.

**Fokusointi/segmentointi (engl. focus)**

Segmentoinnilla tai fokusoinnilla tarkoitetaan kolmatta Porterin yleisstrategiaa, jonka keinona pyritään keskittymään vain tiettyyn markkinaan sekä segmenttiin ja näin ollen saavuttamaan kilpailuetua luotettavuudella ja asiantuntijuudella. Tässä strategiakeinossa keskitytään kohdentamaan tiettyjä tuotteita vain tietyille kohdetyypille eli segmentille. Tämä strategia voidaan nähdä niin sanotusti kahden aiemman strategiamallin eli kustannusjohtajuuden sekä erilaistamisen yhdistelmänä eli niin sanottuna hybridistrategiana, jonka toteuttamisessa datalla on merkittävä rooli nykypäivänä. (Porter, 1985.) Seuraavassa kuviossa (Kuvio 1) on havainnoitu tarkemmin strategioita ja niiden keskittymisalueita. Fokusointi on jaettu kuviossa edelleen kahteen eri kategoriaan, sillä strategisella fokusoinnilla voidaan tavoitella niin erilaistamista kuin kustannusjohtajuuttakin. Tähän kuitenkin vaikuttaa valittu kilpailuala sekä käytettävä strategiakeino.

## Kilpailuetu

		Alhaiset kustannukset	Eriilaistaminen
		Kilpaluualan fokus	Laaja segmentti
Kapea segmentti	Kustannusjohtajuus & Fokusointi		Eriilaistaminen & Fokusointi

Kuvio 1 Porterin yleisstrategiamalli (Porter, 1985)

Tätä yritysten havittelemaa visiota ja kilpailuetua voidaan lähteä tavoittelemaan pitkän aikavälin suunnitelman eli strategian avulla. Strategiat auttavat organisaatiota määrittämään, miten visiot ja tavoitteet saavutetaan työskentelemällä perusasioiden parissa. Hyvässä liiketoimintastrategiassa otetaan huomioon olemassa olevat esteet ja resurssit. Tällaisia ovat esimerkiksi yrityksen työntekijät, omaisuus, tietotaito ja materiaalit, joiden on oltava sopuossuussa organisaation vision, mission ja yleisten tavoitteiden kanssa (Abdulwase ym., 2020). Näiden perinteisten strategiakeinojen seuraamista voidaan tehostaa datan avulla. Datan avulla saadaan merkittävä määrä tarvittavaa informaatiota markkinoiden tilanteista sekä yrityksen suunnasta. Datan käytön kanssa on kuitenkin myös paljon haasteita, eikä sen tuomat mahdollisuudet ole itsestäänselvyyksiä.



## 3.2 Datastrategia

Vaikka strategiaa käsittelevää kirjallisuutta ja tutkimuksia on vuosien varrelta kertynyt runsaasti ja useissa näistä artikkeleista korostetaan analyttisen strategian merkitystä, ei vielä ole löydetty analyttiselle dataa järjestelmällisesti hyödyntävälle strategialle yleisesti hyväksyttyä tai käytettyä määritelmää (Grossman, 2018). Grossman (2018) määrittelee datastrategian organisaation tekemiksi pitkän aikavälin päätöksiksi siitä, miten se käyttää dataa hyödyksi toteuttaessaan toimia, jotka vastaavat organisaation visioon ja tavoitteisiin. Käytännössä siis datastrategian määritelmä on hyvinkin samankaltainen Porterin (1991, 1985) ja muiden tutkijoiden määritelmään. Näillä datastrategiaan sisältyvillä päätöksillä Grossman (2018) kuitenkin tarkoittaa erityisesti organisaation analyysimahdollisuuksien valintaa ja analyttisten toimintojen, analyttisen infrastruktuurin ja analyttisten mallien integrointia organisaation tehtävien ja visioiden saavuttamiseksi (Grossman, 2018), joita ei yleisesti ole huomioitu vanhemmissa strategiamalleissa. Vaikkei yksiselitteistä määritelmää datan strategiselle johtamiselle ja hallinnalle Grossmanin (2018) mukaan olen tarkemmin vielä määriteltä, on kirjallisuudessa yleisesti hyödynnetty melko samankaltaisia määritelmiä aihealueesta ja niitä voidaankin tarvittaessa soveltaa myös datan strategiseen hyödyntämiseen.

Nykypäivänä useat yritykset ja organisaatiot hyödyntävät dataa ja sen analytiikkaa yrityksen liiketoiminnan kehittämisessä. Organisaatioiden eri tasojen johtajat luottavat yhä enemmän ja enemmän data-analytiikkaan reaaliaikaisessa päätöksenteossa sekä organisaation tulevaisuuden ohjaamisessa (Constantiou & Kallinikos, 2015). Data ja sen analytiikka ovat jo vuosia olleet keskustelun aiheena, ja sen vuoksi yritysten on nykypäivänä lähes välttämätöntä hyödyntää dataa omassa toiminnassaan, mikäli he haluavat erottua kilpailijoista sekä pysyä mukana kilpailussa. Mikäli dataa kerätään, hyödynnetään ja analysoidaan kilpailijoita paremmin, voidaan sen avulla luoda merkittävää kestävästä kilpailuetua kilpailijoihin nähden (Hagiu & Wright, 2020).

Massadatan hyödyntäminen strategisessa mielessä on kuitenkin haastavaa, sillä dataa muodostuu monille yrityksille liian nopeissa sykleissä ja liian suuria määriä kerrallaan. Useimmiten saatu data on siis jo vanhentunutta silloin, kun yritys kerkeää reagoimaan kerättyyn dataan. Mikäli yritys haluaa hyödyntää massadataa ja analytiikkaa yrityksensä strategisessa johtamisessa, tulisi yrityksen luoda erillinen datastrategia, jossa määritellään datan ja analytiikan käyttöönottoa sekä hallintaa. Tämän avulla datasta saatavat hyödyt voidaan optimoida organisaation tarpeiden mukaisiksi ja yrityksillä on jokin selkeä kohde, jota tavoitella (Pehcevski, 2019). Grossman (2018) luokittelee tutkimuksessaan viisi tärkeää kohtaa, jotka tulisi huomioida datastrategiaa suunniteltaessa:

- Organisaation tulisi käyttää analytiikkaa erottautuakseen kilpailijoista ja saadakseen sen avulla kilpailuetua
- Strategiassa tulee määritellä organisaation pitkän aikavälin analyysisuunnat
- Analyyttisten mahdollisuuksien valintaan tulisi kehittää prosessi, joka optimoi datan tuoman arvon koko organisaatiolle sekä ottaa huomioon yrityksen mahdolliset rajalliset resurssit käyttöönottoon ja mallintamiseen liittyen
- Valittujen analyyttisten keinojen tuomaa arvoa tulisi mitata ja seurata erilaisten mittareiden avulla, jotta saadaan konkreettista näyttöä onnistumisista ja epäonnistumisista
- Datastrategian myötä data tulisi nähdä ja sitä pitäisi hallita yhtenä yrityksen omaisuutena ja resurssina

Massadatan hallintaprosessin merkittävä haaste on tietojen luokittelu, mallintaminen ja kartoittaminen datatietueita tallennettaessa ja käsiteltäessä. Tämä johtuu pääasiassa datan suuresta määrästä, strukturoimattomuudesta sekä monimutkaisuudesta. Moniulotteisen ja nopeasti lisääntyvän datan haasteisiin vastaaminen vaatii uudenlaista tilastollista ajattelua sekä uusia toimintamenetelmiä, jotta dataa voidaan hyödyntää halutulla tavalla (Fan ym., 2014). Datan tehokas hallinnointi on yrityksille olennaisen tärkeää, jotta voidaan varmistaa suurten tietokokonaisuuksien joukosta louhitun ja analysoidun tiedon laatu ja näin varmistaa datan tehokas hyödyntäminen (Hashem ym., 2015).

Massadata-analytiikan aineellisten ominaisuuksien avulla yritys voi esimerkiksi automatisoida asiakastietoisien palvelutarjonnan ja inhimillismateriaaliset asiakastietoiset palvelukäytäntönsä. Lehrer ym. (2018) mukaan massadata analytiikan aineellisiin ominaisuuksiin kuuluvat muun muassa datan hankinta ja varastointi, tapahtumien tunnistaminen ja ennustaminen, käyttäytymisen tunnistaminen ja ennustaminen, sääntöpohjainen toiminta sekä datan visualisointi. Nämä analytiikan ominaisuudet voivat antaa organisaatioille mahdollisuuden uudistaa esimerkiksi arvolupauksiaan (Lehrer ym., 2018). Näiden keinojen avulla voidaan myös tuottaa yritykselle merkittävää arvoa, sillä mainiot arvolupaukset tuovat useimmiten yrityksille myös uusia asiakkaita ja näin ollen kestävä kilpailuetua. Yritykset voivat saavuttaa ketteryyttä yhdistämällä monipuolisia tietolähteitä ja ottamalla käyttöön analytiikkaa, joka havaitsee asiakkaat dynaamisessa ympäristössä sekä reagoi niihin ja tehtyihin muutoksiin (Kitchens ym., 2018). Edellä mainitut datan tuomat mahdollisuudet voidaan varmistaa datalähtöisellä strategialla ja yrityskulttuurilla, joka tähtää näiden merkittävien etujen saavuttamiseen datan avulla.

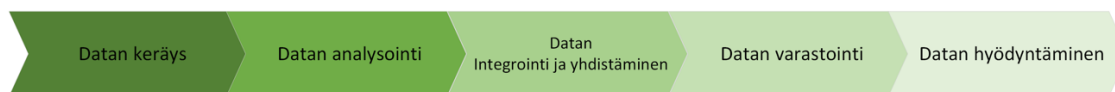
## 4 DATASTRATEGIAN ROOLI DATALÄHTÖISESSÄ LIIKETOIMINNAN KEHITTÄMISESSÄ

Tässä luvussa keskitytään toiseen tutkimuskysymykseen eli siihen voidaanko datalähtöisen liiketoiminnan kehittämisen kompastuskiviä estää datastrategian avulla. Ensimmäisessä alaluvussa kuvaillaan datastrategian merkitystä organisaatiossa ja sen mahdollisia etuja. Toisessa alaluvussa esitellään datastrategian avulla onnistuneita dataprojekteja.

### 4.1 Datastrategian merkitys liiketoiminnalle

Viime vuosikymmenen aikana yhä useammat yritykset ovat ottaneet käyttöön tietoon perustuvan lähestymistavan voidakseen tarjota kohdennetumpia palveluita, ennakoidakseen ja vähentääkseen riskejä sekä parantaakseen organisaation suorituskykyä suhteessa kilpailijoihin. On lähes kiistatonta, että voiko datan avulla tuoda yritykselle uutta arvoa ja kilpailuetua. Datan avulla kilpailuetua voidaan tuottaa esimerkiksi keräämällä ja analysoimalla asiakasdataa ja sen myötä voidaan parantaa yrityksen tarjoamia tuotteita ja palveluita sekä kohdentaa markkinointia ja mainontaa oikeille segmenteille. Datan avulla yritys voi myös löytää uusia ja tehokkaampia toimintatapoja, joiden avulla yritys voi esimerkiksi pienentää kustannuksiaan tai tehostaa tuotantoaan.

Datan tehokkaaseen hyödyntämiseen liittyy lukuisia haasteita, joita olisi mahdollista ehkäistä kattavan ja koko organisaation mukaan lukevan datastrategian avulla. Datan hyödyntäminen on monivaiheinen prosessi, jonka vuoksi siihen liittyy lukuisia eri haasteita, joita yritys ei välttämättä osaa ennakoita ilman kattavaa strategiaa. Martino ym. (2014) kuvailee datan hyödyntämisen prosessia seuraavanlaisella kuvaajalla (Kuvio 2):



Kuvio 2: Datan hyödyntämisen prosessi (Martino ym., 2014)

Ennen datan hyödyntämistä on siis jo useita eri vaiheita, joissa yrityksen on onnistuttava, jotta datasta saadaan haluttuja ominaisuuksia irti. Näitä vaiheita ovat Martinon ym. (2014) mukaan datan kerääminen, analysointi, integrointi sekä varastointi. Ensimmäisessä vaiheessa yrityksen tulee kerätä dataa joillain keinoilla. Tämän jälkeen kerätty data käydään tarkasti läpi ja analysoidaan, jonka jälkeen samankaltaiset datatietueet yhdistetään yhteen muotoon. Viimeinen vaihe ennen datan hyödyntämistä on datan varastointi, johon kuuluu datan asianmukainen varastointi, jotta tarvittava data on sellaisessa muodossa, että sen hyödyntäminen on nopeaa ja vaivatonta. (Martino ym., 2014) Datan hyödyntäminen on siis yksinkertaisestikin selitettynä monivaiheinen prosessi, jonka vuoksi moni yritys kokee haasteita datan tehokkaan hyödyntämisen kanssa. Monivaiheinen prosessi tuo myös monenlaisia haasteita, joita on koottu alta löytyvään taulukkoon (Taulukko 2).

Taulukko 2: Datan käytön suurimmat haasteet (Chen ym., 2013; Hashem ym., 2015; Krishnamurthy & Desouza, 2014; Lycett, 2013; Manyika ym., 2011; Pehcevski, 2019; Rijmenam, 2014)

<b>Datan käytön suurimmat haasteet</b>	<b>Ongelma</b>
<b>Aineiston saatavuus</b>	Data sijaitsee useissa eri tietolähteissä ja erilaisissa tietomuodoissa, jolloin sen hyödyntäminen on haasteellista tai jopa mahdotonta.
<b>Aineiston laatu</b>	Joko dataa ei ole tarpeeksi tai sitä on jopa liikaa, jonka vuoksi laatu ei korreloi kerätyn määrän kanssa. Data ei ole ajantasaista eikä sitä voida hyödyntää yrityksen eduksi.
<b>Aineiston käsittely</b>	Datan hyödyntäminen ei ole tehokasta. Data on paikoin jo vanhentunutta silloin, kun sitä käytetään. Henkilöstö ei osaa hallita dataa, jolloin siitä ei saada irti tarvittavia ominaisuuksia
<b>Datan suuri määrä</b>	Dataa kertyy niin suuria määriä, että sitä ei keretä hyödyntämään halutulla tavalla. Suuresta määrästä dataa on vaikeaa löytää tarvittavia tietoja ja kerätty data vanhenee ennen kuin siitä saadaan hyötyä yritykselle
<b>Turvallisuus ja tietosuojat</b>	Datan suojaaminen voi olla puutteellista, jolloin se voi aiheuttaa yritykselle suuria tietoturvaongelmia.
<b>Hyödyntäminen päätöksenteossa</b>	Mikäli datan käyttöön ei ole tehty tarkempaa strategista suunnitelmaa, voi datan käyttö ja hallinnointi olla puutteellista, jolloin dataa voidaan käyttää virheellisesti tai sitä voi kadota kokonaan. Henkilöstön taitotaso ei välttämättä ole tarvitulla tasolla.
<b>Teknologia infrastruktuuri</b>	Organisaation laitteet ja järjestelmät eivät riitä haluttujen toimien toteuttamiseen tai laitteet ja järjestelmät ovat vanhentuneita, jonka vuoksi niiden hyödyntäminen datan hallinnassa on ongelmallista

Moni yritys on ottanut dataa ja sen käsittelyä varten käyttöönsä tähän tarkoitukseen erikoistuneita data-analyysin käsittelyyn luotuja ohjelmia, joiden avulla organisaatiot voivat kerätä, tallentaa, hallinnoida ja analysoida suuria tietokokonaisuuksia eri lähteistä. IBM:n (2021) Tutkimusten mukaan kuitenkin jopa 60–85 % yritysten aloittamista massa data projekteista epäonnistuu vain sen vuoksi, että ennen projektin aloittamista ei ole määritelty tarvittavaa strategiaa

datan hyödyntämiseksi. Projektin epäonnistuminen johtuu esimerkiksi siitä, että projektin aikana keskitytään ainoastaan teknologiaan tai työkaluihin ja niiden tuomiin hyötyihin eikä datastrategian myötä esimerkiksi datan mahdollisiin hyötyihin liiketoiminnassa tai koko yrityksen rakenteessa. (IBM, 2021)

Data ja analytiikka ravistelevat jo nyt useita toimialoja, ja vaikutukset vain korostuvat, kun datan käyttöönotto ja hyödyntäminen saavuttavat kriittisen massan. Koneet ja teknologia kehittyvät jatkuvasti ja näin ollen ne saavat tulevaisuudessa yhä ennennäkemättömpiä kykyjä ratkaista ongelmia. Organisaatiot, jotka pystyvät valjastamaan nämä kyvyt tehokkaasti, pystyvät luomaan merkittävää arvoa ja erottautumaan muista, kun taas muut joutuvat yhä enemmän epäedulliseen asemaan kilpailukentässä. (*The age of analytics: Competing in a data-driven world* | McKinsey, 2016.) Ilman investointeja tulevaisuuden infrastruktuuriin, on riskinä se, että yritys jää kokonaan jälkeen teknologian ja datan hyödyntämisen kehityksestä, jolloin se automaattisesti heikentää omaa markkina-asemaansa.

Monissa yrityksissä dataa pidetään edelleen itsestänselvyytenä tai sen tuomia etuja ei osata huomioda, sillä datan käsittelyyn ei ole erillistä pitkän aikavälin suunnitelmaa eli strategiaa. Hagiun ja Wrightin (2020) mukaan data tuo organisaatiolle lisäarvoa vasta sitten, kun datasta saatuja tuloksia opitaan analysoimaan ja ymmärtämään niin, että niitä voidaan hyödyntää yrityksen toiminnassa. Yltä löytyvään taulukkoon onkin koottu lukuisia datan hallintaan liittyviä ongelmia, joiden ehkäiseminen voitaisiin mahdollistaa koko organisaatiotason käsittävän strategian avulla. McKinseyn (Manyika ym., 2011) tuottaman raportin mukaan esimerkiksi massadatan arvopotentiali on edelleen laajalti hyödyntämättä useissa yrityksissä. Raportin mukaan strategian ja johtamisen aiheuttamat ongelmat ovat yksi suurimmista datan käyttöön liittyvistä haasteista. Manyikan ym. tuottaman McKinseyn (2011) raportin mukaan datan avulla voidaan:

- Luoda avoimuutta ja uusia mahdollisuuksia organisaatiossa, sillä sen avulla voidaan olla laajemmin uusien potentiaalien kuten uusien toimialojen saatavilla.
- Se antaa yrityksille mahdollisuuden kokeilla rohkeammin uusia keinoja ja resursseja. Esimerkiksi datan avulla voidaan testata prosessimuutoksia ja niitä varten voidaan luoda ja analysoida suuria määriä dataa.
- Massa datan avulla voidaan luoda yksityiskohtaisempaa segmentointia asiakkaista, jonka myötä voidaan räätälöidä yrityksen resursseja ja valmistella erityisiä palveluita ja henkilökohtaisempaa markkinointia.
- Massa datan analysointi voi tukea ihmisen päätöksentekoa osoittamalla piilokorrelaatioita tai joitakin piilotettuja riskejä, joita pelkkä ihminen ei välttämättä osaa tunnistaa. Esimerkkinä tutkimuksessa mainitaan esimerkiksi vakuutusyhtiöiden riski- tai petosanalyysialgoritmit ja keinot.
- Data voi myös mahdollistaa uusia liiketoimintamalleja, tuotteita ja palveluita tai parantaa jo olemassa olevia resursseja.
- Datan avulla voidaan löytää tietoa siitä, miten tuotteita ja palveluja käytetään, jonka lisäksi sitä voidaan käyttää uusien versioiden kehittämiseen ja parantamiseen.

Henken ym. (2016) mukaan oikean vision ja strategian määrittäminen datan ja analytiikan käyttöä varten oli kaikkien osallistujien suurin este, ja yli 45 prosenttia vastaajista mainitsi sen yhdeksi kolmesta tärkeimmästä huolenaiheestaan. Tämä voidaan nähdä suurena ongelmana, koska ylimmän johdon henkilökohtaisella hyväksynnällä on suora vaikutus yrityksen sisäisiin hankkeisiin. Siirtyminen tietoon perustuvampaan päätöksentekoon ei ole niin yksinkertaista kuin esimerkiksi uuden IT-järjestelmän ostaminen, vaan se vaatii tehokasta ja esimerkillistä johtajuutta, jotta organisaatiossa voidaan saada aikaan kestävä muutos ja ottaa käyttöön kokonaan uusi tapa toimia. Johtajat, jotka ilmoittivat, että heidän yrityksensä ovat käyttäneet analytiikkaa tehokkaasti, pitivät useimmiten ylimmän johdon osallistumista tekijänä, joka on vaikuttanut eniten menestykseen. (Henke ym., 2016.)

Myös DalleMule ja Davenport (2017) tukevat McKinseyn (Manyika ym., 2011) raportin väitettä siitä, että massadatan arvopotentiaali on edelleen laajalti hyödyntämättä yrityksissä. He käsittelevät artikkelissaan laajasti datastrategian merkitystä eri toimialoilla toimiville organisaatiolle. Heidän mukaansa data ja sen strateginen hallinnointi mahdollistavat ylivoimaisen tiedonhallinnan ja analytiikan sekä antaa yritykselle välttämättömät valmiudet, jotka tukevat johdon päätöksentekoa ja viime kädessä parantavat koko yrityksen taloudellista suorituskykyä ja kilpailuasemaa.

Valtaosa yritysten keräämästä ja omistamasta datasta on strukturoimatonta ja vaikeasti analysoitavaa massadataa, jonka hallitsemiseen vaaditaan Pehcevs-kin (2019) mukaan hyvin suunniteltu datastrategia. Muutoin haasteena on se, että data vanhentuu, sen hyödyntäminen on eri tietomuotojen vuoksi on haastavaa tai data saattaa jopa pahimmassa tapauksessa kadota kokonaan. DalleMulen ja Davenportin (2017) muodostama ohjekehys määrittelee kaksi tyypillistä datastrategian ongelmaa ja helpottaa näin ollen yrityksiä selventämään ydintarkoituksen datalle sekä ohjaa heitä strategisen datan hallinnan kanssa. Yritysten tulisi datastrategiaa muodostaessaan päättää onko heidän datastrategiansa ja datan käyttönsä hyökkäävää (engl. offensive) vai puolustavaa (engl. defensive) sekä onko datan hyödyntäminen joustavaa vai kontrolloitua.

Jotta voidaan määritellä yrityksen sijainti puolustus/hyökkäys-akselilla, tulee DalleMulen ja Davenportin (2017) mukaan huomioida:

- Yrityksen kokonaiskuvan ja tavoitteen strategia
- Datan mahdollisuudet ja ominaisuudet suhteessa kilpailijoihin
- Datalla johtamisen maturiteetti
- Data-analytiikan taso ja sen budjetin koko

Yritys voi siis keskittyä joko puolustamaan dataa tai keräämään sitä. Tämä päätös tulisi kuitenkin määritellä datastrategian avulla, jotta tiedetään käytännön raamit ja rajat sekä datan hyödyntämisen suunta. (DalleMule & Davenport, 2017.)

Puolustava datastrategia perustuu tietojen suojaamiseen ja riskien minimointiin. Tähän strategiaan kuuluu aktiivinen säännösten ja lakien noudattaminen sekä seuraaminen (tietosuoja ja tietoeheys), analytiikan hyödyntäminen

petosten tunnistamisessa ja rajoittamisessa sekä järjestelmien rakentaminen varauksien estämiseksi. Puolustustoimilla varmistetaan myös yrityksen sisäisten järjestelmien kautta kulkevan tiedon eheys tunnistamalla, standardoimalla ja hallinnoimalla arvovaltaisia tietolähteitä, kuten perustavanlaatuisia asiakas- ja toimittajatietoja tai myyntitietoja, "yhden ainoan totuuden lähteen" avulla.

Hyökkäävää strategiaa hyödynnetään yleensä asiakaslähtöisten yritysten kuten myynti- ja markkinointiyritysten liiketoiminnan kehittämisessä. Sillä pyritään keskittymään tukemaan liiketoiminnan tavoitteita, kuten tulojen, kannattavuuden ja asiakaskunnan kasvattamista sekä asiakastyytyväisyyttä. Hyökkäävään strategiaan sisältyy tyypillisesti toimintoja, jotka tuottavat ymmärrystä asiakkaista ja heidän käytöksestään (data-analyysi ja mallintaminen) tai yhdistetään toisistaan poikkeavia asiakas- ja markkinatietoja johtamisen tueksi. Hyökkäävä strategia vaatii reaaliaikaisempaa datan tarkastelua sekä enemmän resursseja kuin puolustava työ, sillä hyökkäävässä strategiassa keskitytään oikeudellisiin, taloudellisiin, vaatimustenmukaisuuteen ja tietotekniikkaan liittyviin kysymyksiin.

DalleMulen ja Davenportin (2017) tuottama toimialarajat ylittävä tutkimus osoittaa, että keskimäärin vain alle puolet organisaation jäsenllystä datasta on aktiivisessa käytössä yrityksen päivittäisessä päätöksenteossa. Tutkimuksessa ilmeni, että alle 1 prosentti strukturoimattomasta datasta analysoidaan tai hyödynnetään yrityksen hyödyksi. Tämän lisäksi tutkimus osoitti, että yli 70 prosentilla työntekijöistä on pääsy tietoihin, joihin heillä ei tulisi olla pääsyä sekä 80 prosenttia analyttikoiden työajasta kuluu pelkkään oikean tiedon löytämiseen ja valmisteluun. Näiden ongelmien lisäksi datan hyödyntämiseen liittyy useita muita ongelmia organisaatiossa: tietomurrot ovat yleisiä, epärehelliset tietokokonaisuudet leviävät ja säilyvät siiloissa sekä yritysten tietotekniikka ei useinkaan vastaa sille etukäteen asetettuja vaatimuksia. Ilman strategista datan hallintaa monilla yrityksillä on haasteita datan suojelun ja hyödyntämisen kanssa (DalleMule & Davenport, 2017). Aiemmin mainitut haasteet tulisi siis huomioida yrityksen strategiassa, sillä ne ovat organisaatiolle samanlaisia riskejä kuin muutkin yleisessä yritys strategiassa huomioon otetut seikat.

Mikalef ym. (2019b) vuonna 2019 teettämä tutkimus tutkii yritysten sekä teknisiä taitoja, että johtamistaitoja massadata analytiikan mahdollisuuksien keskeisinä elementteinä. Jotta yritykset voivat realisoida liiketoiminta-arvoa tekemistään massadata -investoinneistaan, on sillä oltava jonkinlainen suunnitelma datan analysointiin ja hallintaan. Massadatan strategista arvoa tuottavat ponnistelut lähtevät organisaation strategiasta ja siten johtajilta, joilla on tarvittava osaaminen aihealueesta (Grover ym., 2018). Jotta nämä massadata -hankkeet tuottaisivat yritykselle myös myönteisiä tuloksia, on tärkeää, että organisaation silot tunnistetaan ja murretaan. Tämän lisäksi koko organisaation eri osastojen asiantuntemus ja tieto tulisi tunnistaa sekä integroida koko organisaation strategiseksi edistämiseksi. Näiden massiivisten tietokokonaisuuksien analysointi auttaa mittaamaan yritysten riskejä ja laskelmoimaan mahdollisia hyötyjä. (Fan ym., 2014.) On tärkeää huomioida, että massadata -analytiikan käyttö uusien mahdollisuuksien ja uhkien havaitsemiseksi ja niihin tarttumiseksi sekä nykyisten kyvykkyyksien muuntamiseksi edellyttää muutakin kuin vain dataa ja teknologiaa (Mikalef, Pappas, ym., 2019b). Tällaisten toimien käyttöönotossa on välttämätöntä, että



yritykset perustavat prosesseja, hallintorakenteita ja kokonaisia tiimejä, joilla on toisiaan täydentäviä datataitoja. Näiden lisäksi yritysten tulisi laatia strategiaan etenemissuunnitelma, jolla voidaan valjastaa olemassa olevat ja tulevat datavarat sekä resurssit yrityksen eduksi (Grover ym., 2018). Siksi on ratkaisevan tärkeää, että massadata -analytiikka-aloitteilla on selkeät liiketoiminnalliset tavoitteet sekä tarkka strateginen suunta, sillä ilman näitä hankkeet tulevat todennäköisesti epäonnistumaan sekä tuottavat yritykselle turhia kustannuksia.

Aiemmassa pääluvussa käsiteltyjen Porterin (1985, 1991) suosittujen päästrategiamallien toteuttamisessa voidaan myös hyödyntää tehokkaasti dataa ja sen avulla voidaan luoda organisaatiolle uusia mahdollisuuksia strategisesta johtamisen näkökulmasta. Datan avulla voidaan luoda merkittävää kilpailuetua ja se helpottaa päästrategioiden toteuttamista, joita on kuvailtu seuraavassa taulukossa (Taulukko 3):

Taulukko 3: Datan hyödyntäminen päästrategioiden toteuttamisessa (Constantiou & Kallinikos, 2015; Henke ym., 2016; Manyika ym., 2011; Porter, 1985, 1991)

Porterin Strategia	Erilaistaminen (Differentiation)	Kustannusjohtajuus (Cost Leadership)	Segmentointi (Focus)
<b>Strategian piirteet</b>	Yritys keskittyy tuottamaan palveluita ja tuotteita, jotka vastaavat täysin markkinoiden kysyntään. Tämän avulla pyritään erottumaan kilpailijoista ja täyttämään markkina-aukkoja.	Strategian mukaan yritys pyrkii tuottamaan asiakkaille tarvittavaa arvoa kilpailijoita edullisempaan hintaan.	Strategian pääpiirteiden mukaisesti keskitytään kohdentamaan tiettyjä tuotteita tai palveluita vain tietyin kriteerein valitulle kohderyhmälle
<b>Strategian tuoma kilpailuetu</b>	Yrityksen tarjoamien tuotteiden ja palveluiden ainutlaatuisuuden vuoksi yritys voi pyytää tuotteistaan kilpailijoita korkeampaa hintaa. Yritys mukauttaa tarjontansa vastaamaan markkina-aukon kysyntään, joka takaa tuotteille hyvän menekin. Muilla kilpailijoilla ei ole vastaavaa tuotetta.	Kun yritys ymmärtää kuluttajien kynnysarvon ja hinta-laatu-kompromissin, voi yritys tarjota tuotteita ja palveluita muita kilpailijoita edullisemmin ja näin saavuttaa markkinajohtajuuden kyseisessä kategoriassa.	Pyritään keskittymään vain tiettyyn markkinaan sekä valittuun segmenttiin ja näin ollen saavuttamaan kilpailuetua luotettavuudella ja asiantuntijuudella. Kysynnän ja tarjonnan kohtaaminen varmistaa arvon jatkumisen yritykselle.
<b>Datan rooli kilpailuedun tuottamisessa</b>	Datan avulla voidaan seurata markkinatilannetta ja kuluttajien ostokäyttäytymistä merkittävän tarkasti, jonka myötä yrityksen on helppoa tunnistaa markkina-aukkoja ja potentiaalisia markkinarakoja. Kysynnän tunnistaminen ja siihen vastaaminen on datan avulla helpompaa.	Datan avulla voidaan tulkita ja ennustaa näitä kustannusjohtajuuteen vaikuttavia tekijöitä huomattavasti tarkemmin.	Data mahdollistaa asiakkaiden tarkankin segmentoinnin, jonka avulla strategian toteuttaminen on entistä helpompaa

## 4.2 Esimerkkejä datastrategian hyödyistä

McAfee ja Brynjolfsson (2012) suorittivat yhdessä MIT:n kanssa tutkimuksen siitä, miten massadatan hyödyntäminen voi vaikuttaa yrityksen suorituskykyyn. Tutkimukseen osallistui 330 erisuuruista Pohjois-Amerikkalaista yritystä.

Tutkimuksen hypoteesi oli se, että datalähtöiset yritykset suoriutuvat haasteista ja liiketoiminnassaan muita paremmin. Tutkijat keräsivät tutkimuksiin osallistuvilta yrityksiltä haastatteluita heidän teknologiajohtamisestaan, organisaation johtamisesta sekä keräsivät dataa suoriutumisesta heidän vuosiraporteistaan sekä itsenäisistä tietolähteistä.

Kaikilla tutkimuksiin osallistuneilla yrityksillä ei ollut virallisesti käytössään datalähtöistä tai -ohjautuvaa päätöksen tekoa. He löysivät kuitenkin tutkimuksen myötä laajan kirjon datalähtöisiä asenteita ja lähestymistapoja kaikilla tutkimukseen osallistuneilla toimialoilla. Tutkimuksen lopputulos oli yhtäläisyys siitä, että ne yritykset, jotka laskivat itsensä data johtoiseksi yritykseksi, suoriutuivat muuta vertailuryhmää paremmin niin taloudellisesti kuin operationaalisesti. Erityisesti yritykset, jotka olivat toimialansa parhaassa kolmanneksessa tietoon ja dataan perustuvan päätöksenteon käytössä, olivat keskimäärin 5 prosenttia tuottavampia ja 6 prosenttia kannattavampia kuin muut kilpailijansa. Tämä suorituskykyero säilyi vahvana myös sen jälkeen, kun työvoiman, pääoman, ostopalvelujen ja perinteisten tietotekniikkainvestointien osuus oli otettu huomioon tuloksissa. Se oli tilastollisesti merkittävää sekä taloudellisesti tärkeää, ja se näkyi pörssi-arvojen mitattavissa olevina nousuina. (McAfee & Brynjolfsson, 2012.)

Mikalefin, Bouran ym. (2019a) tekemä tutkimus osoittaa, että massadata -analytiikka on muutakin kuin pelkkä investointi teknologiaan, valtaviin tietomäärien kerääminen ja se, että tietotekniikkaosasto voi testaila uusia analyysitekniikoita. Edellä mainittujen lisäksi tärkeitä tekijöitä, joiden avulla massadata -investoinneista saadaan liiketoiminnallista arvoa, ovat:

- sellaisten henkilöiden rekrytointi, joilla on hyvä tekninen ja hallinnollinen ymmärrys massadatasta ja sen analytiikasta
- organisaation oppimiskulttuurin edistäminen
- massadata -päätöksenteon sisällyttäminen organisaation rakenteeseen ja kulttuuriin

Näiden keinojen onnistumiseksi yrityksillä tulisi olla käytössään jonkinlaiset raamit, joita seuraamalla yritys voi tuottaa itselleen lisää arvoa datan avulla. Tämä tarkoittaa tietenkin sitä, että on otettava käyttöön lukuisia prosesseja, mikä edellyttää ylimmän johdon sitoutumista ja selkeää suunnitelmaa koko yrityksen laajuista massadata -analytiikan käyttöönottoa ja levittämistä varten. (Mikalef, Boura, ym., 2019a.) Rivien välistä voidaankin päätellä, että vaikka tutkimustulokset tukevat datan tuomaa liiketoiminnallista arvoa, olisi tärkeää, että organisaatioiden datan käyttö ja analysointi olisi rakennettu osaksi yrityksen perustuksia. Tämän vuoksi datastrategialla on tärkeä merkitys osana yrityksen strategista päätöksentekoa.

Tämän luvun tutkimuksista voidaan päätellä, että datasta ja sen analytiikasta voidaan saada merkittävää kilpailuetua ja arvoa yritykselle, mutta näiden resurssien tuomat hyödyt eivät kuitenkaan ole yrityksille itsestäänselvyksiä. Datalla on merkittävän tärkeä rooli nykyisessä yhteiskunnassa sekä tulevaisuuden liiketoiminnassa, mutta sen hallintaan ja hyödyntämiseen tulisi olla tarkka suunnitelma tai edes jonkinlaiset suuntaa antavat linjat, jotta datan

hyödyntämiseen käytetyt resurssit eivät mene yritykseltä hukkaan. Tässä tilanteessa roolinsa ottaa datastrategia, jonka avulla voidaan varmistaa, että datan hyödyntäminen on tehokasta ja palkitsevaa.

## 5 YHTEENVETO JA JATKOTUTKIMUSAIHEET

Tämän tutkielman tarkoituksena oli selvittää datan ja sen analytiikan roolia liiketoiminnan kehittämisessä sekä sitä, minkälainen rooli datastrategialla on näiden resurssien hyödyntämisessä yrityksen liiketoiminnan kehittämiseen ja kilpailuedun luomiseen. Tutkielman ensimmäisessä pääluvussa on perehdytty tarkemmin data-analytiikan pääkäsitteisiin sekä määritelty näiden resurssien tuoma arvoa sekä niiden tuomia mahdollisia haasteita yritykselle. Tämän lisäksi tutkielmassa on tutkittu datan ja sen analyysin tuottamia haasteita sekä niitä tekijöitä, joita organisaatioiden tulisi huomioida omassa datan käytössään.

Tutkimuksen alussa on käyty läpi tarkemmin vastausta ensimmäiseen tutkimuskysymykseen, jonka tarkoituksena on selvittää datan tuomat liiketoiminnalliset hyödyt. Tämän jälkeen tutkimuksessa siirryttiin tarkastelemaan toisen tutkimuskysymyksen aihepiiriä eli voidaanko datan strategisen hyödyntämisen avulla välttyä datan hyödyntämiseen liittyviltä yleisiltä ongelmilta ja haasteilta. Vaikka data on aihealueena erittäin ajankohtainen ja siitä löytyy merkittävä määrä kirjallisuutta ja tieteellisiä tutkimuksia, voidaan tämänhetkisen tutkimuksen tilanne tulkita kuitenkin jokseenkin puutteelliseksi. Datastrategiasta ja sen merkityksestä liiketoiminnalle ei löytynyt juurikaan luotettavia tuoreita tutkimuksia. Strategiasta sekä datan merkityksestä löytyi kuitenkin lukuisia tutkimuksia, mutta niissä keskityttiin näihin aihepiireihin toisistaan irrallisina tutkimuksina. Näin ollen juuri datan ja strategian yhteisvaikutuksia ei juurikaan ole tutkittu tieteellisessä kirjallisuudessa. Valtaosa aihepiirin luotettavista lähteistä oli tuotettu jo vuosia sitten, jonka vuoksi osa tuloksista on voinut jo vanhentua aihepiirin jatkuvan kehittymisen vuoksi. Tämä tuotti omalta osaltaan haasteita tutkimuksen toteuttamiseen, sillä moni lähde pohjautui jo mahdollisesti vanhentuneisiin tietolähteisiin. Uusien ja ajankohtaisten lähteiden puuttuminen on kuitenkin ymmärrettävää, sillä datan uusiutuminen ja teknologin kehittyminen on nykypäivänä merkittävän nopeaa, jonka lisäksi laadukkaiden tutkimusten julkaisemisessa voi kestää jopa vuosia. Näin ollen tieteellinen tutkimus on voinut jäädä hieman jälkeen lähivuosien data- ja teknologiakehityksen harppauksissa.

Datalla on eittämättä tärkeä rooli osana nykyistä yhteiskuntaa sekä eri organisaatioiden päätöksenteossa. Lähes jokainen voi jollain tasolla hyödyntää dataa elämässään, mutta monien yritysten kaltaisissa isoissa organisaatioissa datan

tehokas ja optimaalinen hyödyntäminen vaatii kuitenkin jonkinlaisen suunnitelman. Tämän tutkielman tulokset tukevat väitettä, jonka mukaan yrityksen on mietittävä tarkkaan, miten he haluavat hyödyntää suuria datamääriä ja kuinka he saavat tästä kokonaisuudesta irti juuri heitä ja heidän organisaatiotaan hyödyntävät tiedot. Tämän kirjallisuuskatsauksen myötä voidaan todeta, että datastrategia on varmasti hyödyllinen ja tärkeä osa yrityksen strategiaa ja datalähtöistä johtamista, mutta jatkossa tarvitaan kuitenkin vielä lisää tieteellistä tutkimusta aiheeseen liittyen. Vaikka datasta on varmasti merkittävästi hyötyä yrityksen toiminnalle, on aiheesta kerätty tietoa vielä jokseenkin keskeneräistä. Teknologia ja data eivät yksinään riitä johtamaan mitattavaan liiketoimintahyötyyn, mutta koko yrityksen laajuisen strategisen massadata-analytiikan vaikutusta suorituskykyyn ja erityisesti innovointiin ei ole vielä riittävästi tutkittu (Mikalef, Boura, ym., 2019a).

Datan ja sen analytiikan lukuisista strategisista hyödyistä huolimatta datastrategian todellisia taloudellisia seuraamuksia on tutkittu vain vähän. Etenkin datastrategian tuomia taloudellisia hyötyjä tulisi tutkia laajemmin esimerkiksi laajan ja toimialat ylittävän empiirisen tutkimuksen kautta tai muun käytännön tapaustutkimuksen kautta, jotta saataisiin ajankohtaista dataa tämänhetkisestä datan hyödyntämisestä käytännössä. Tämän myötä väitteille ja hypoteeseille voitaisiin saavuttaa vielä parempia ja ajankohtaisempia tuloksia, joiden tiimoilta datastrategian ja data-analytiikan teorioita voitaisiin päivittää nykyaikaan. Vaikka massadataa ja strategista päätöksentekoa on erikseen tutkittu huomattavan paljon, on näiden yhteinen tutkimus edelleen varsin suppeaa, jonka vuoksi aihealueesta on vaikea tehdä absoluuttisia päätelmiä. Analytiikan ja strategian yhteydestä ja niiden tuottamista hyödyistä ja eduista olisi tärkeää saada ajankohtaisempaa tietoa, jota voitaisiin hyödyntää paremmin myös käytäntöön nykyisessä alati muuttuvassa teknologiayhteiskunnassa.

Näiden aiheiden lisäksi jatkotutkimuksissa tulisi keskittyä myös datan ympäristövaikutuksiin niin sanotun green codingin näkökulmasta. Suuret ja jatkuvasti monistuvat datamäärät aiheuttavat organisaatioille ja järjestelmille merkittävän määrän kuormitusta, joka voidaan nähdä pitkällä aikavälillä suurena haasteena. Turhaa ja käyttämätöntä dataa roikkuu organisaatioiden data-altaissa, silloissa ja varastoissa merkittäviä määriä ja tällaisten tietueiden säilöminen kuormittaa ympäristöä merkittävässä määrin. Datan ja sen eettisyyden ja ympäristöystävällisyyden näkökulmasta tätä tutkimusta voitaisiin siis laajentaa vielä merkittävässä määrin. Tästä näkökulmasta olisi myös hyvä tutkia datan tuomien hyötyjen sekä haittojen välistä suhdetta tarkemmin.

Lisäksi tuloksista käy ilmi datan äärimmäisen tärkeyden ja ajankohtaisuuden lisäksi myös sen haasteellinen käsittely sekä muut analysointiin liittyvät haasteet ja resurssiongelmat. Tämä korostaa entisestään tutkimuksen tuloksia ja tuo lisää tutkimusideoita aihepiirille. Nykyinen tutkimus keskittyy pitkälti datan tuomiin hyötyihin ja etuihin, mutta datan negatiivisia vaikutuksia on tutkittu jokseenkin vähän aihepiirin ajankohtaisuutta ajatellen. Tutkimuksen myötä herääkin jatkokysymys muun muassa siitä, onko yritysten datasiilot tikittäviä aikapommeja ja voiko tästä suuresta datamäärästä muodostua yhtäkkiä vakavia ongelmia organisaatioille? Näitä aiheita olisi hyvä tutkia myös jatkossa.

Huolimatta laajemman tutkimuksen puutteesta, vastasi tämä kirjallisuuskatsaus alussa esitettyihin tutkimuskysymyksiin ja hypoteeseihin. Tutkimuksen tuottamat tulokset auttavat ymmärtämään data-analytiikan pääkäsitteitä sekä datastrategiaa käsitteenä, vaikka käsitteen virallista määritelmä onkin haastavaa löytää nykyisestä kirjallisuudesta. Tämän tutkimuksen kirjallisuuskatsauksen pohjalta voidaan jopa tulkita, että datastrategialle ei vielä ole muodostunut tarkkaa ja yhtenäistä määritelmää, jonka vuoksi lisätutkimus aiheesta olisi tarpeellista. Toisaalta esimerkiksi datalla ja sen analytiikalla on lähes niin monta määritelmää kuin on kirjoittajaa ja tutkijaakin, jonka vuoksi yksiselitteistä määritelmää datastrategialle ei välttämättä ole mahdollista saavuttaa koskaan. Yleisesti ottaen datastrategia kuitenkin jo nimensä mukaisesti kuvaa käsitettä ja sen määritelmää hyvin. Aihepiirinä data ja sen hallinta tulevat varmasti olemaan jatkossakin suuri trendi teknologian kehityksessä ja sen tutkimuksessa, joten on mahdollista, että tässä luvussa mainittuja jatkotutkimusaiheita on jo tutkittu tai tutkitaan parhailaan.

## LÄHTEET

- Abdulwase, R., Ahmed, F., Nasr, F., Alyousofi, A., & Yan, S. (2020). *The role of business strategy to create a competitive advantage in the organization*. 4, 135–138. <https://doi.org/10.15406/oajs.2020.04.00162>
- Blakely, S. (2022, helmikuuta 25). *Exit Strategies: How to Plan a Business Exit Strategy - 2022*. MasterClass. <https://www.masterclass.com/articles/exit-strategy-explained>
- Cao, L. (2018). Data Science: A Comprehensive Overview. *ACM Computing Surveys*, 50(3), 1–42. <https://doi.org/10.1145/3076253>
- Chen, J., Chen, Y., Du, X., Li, C., Lu, J., Zhao, S., & Zhou, X. (2013). Big data challenge: A data management perspective. *Frontiers of Computer Science*, 7(2), 157–164. <https://doi.org/10.1007/s11704-013-3903-7>
- Cheng, C., Zhong, H., & Cao, L. (2020). Facilitating speed of internationalization: The roles of business intelligence and organizational agility. *Journal of Business Research*, 110, 95–103. <https://doi.org/10.1016/j.jbusres.2020.01.003>
- Christensen, P., & Ackerman, B. (2021). *Define your data strategy – IBM Garage Practices*. <https://www.ibm.com/garage/method/practices/think/define-data-strategy/>
- Constantiou, I. D., & Kallinikos, J. (2015). New games, new rules: Big data and the changing context of strategy. *Journal of Information Technology*, 30(1), 44–57. <https://doi.org/10.1057/jit.2014.17>
- DalleMule, L., & Davenport, T. H. (2017, toukokuuta 1). What's Your Data Strategy? *Harvard Business Review*. <https://hbr.org/2017/05/whats-your-data-strategy>
- Fan, J., Han, F., & Liu, H. (2014). Challenges of Big Data analysis. *National Science Review*, 1(2), 293–314. <https://doi.org/10.1093/nsr/nwt032>
- Grossman, R. L. (2018). A framework for evaluating the analytic maturity of an organization. *International Journal of Information Management*, 38(1), 45–51. <https://doi.org/10.1016/j.ijinfomgt.2017.08.005>
- Grover, V., Chiang, R. H. L., Liang, T.-P., & Zhang, D. (2018). Creating Strategic Business Value from Big Data Analytics: A Research Framework. *Journal of Management Information Systems*, 35(2), 388–423. <https://doi.org/10.1080/07421222.2018.1451951>
- Hagiu, A., & Wright, J. (2020, tammikuuta 1). When Data Creates Competitive Advantage. *Harvard Business Review*. <https://hbr.org/2020/01/when-data-creates-competitive-advantage>

- Hashem, I. A. T., Yaqoob, I., Anuar, N. B., Mokhtar, S., Gani, A., & Ullah Khan, S. (2015). The rise of “big data” on cloud computing: Review and open research issues. *Information Systems, 47*, 98–115.  
<https://doi.org/10.1016/j.is.2014.07.006>
- Henke, N., Bughin, J., Chui, M., Manyika, J., Saleh, T., Wiseman, B., & Sethupathy, G. (2016, joulukuuta 7). *The age of analytics: Competing in a data-driven world* | McKinsey.  
<https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-age-of-analytics-competing-in-a-data-driven-world>
- Hovi, A. (2015). *Mitä on Master Data?* Ari Hovi.  
<https://www.arihovi.com/mita-master-data/>
- Kitchens, B., Dobolyi, D., Li, J., & Abbasi, A. (2018). Advanced Customer Analytics: Strategic Value Through Integration of Relationship-Oriented Big Data. *Journal of Management Information Systems, 35*(2), 540–574.  
<https://doi.org/10.1080/07421222.2018.1451957>
- Krishnamurthy, R., & Desouza, K. (2014). Big data analytics: The case of the social security administration. *Information Polity, 19*, 165–178.  
<https://doi.org/10.3233/IP-140337>
- Lehrer, C., Wieneke, A., vom Brocke, J., Jung, R., & Seidel, S. (2018). How Big Data Analytics Enables Service Innovation: Materiality, Affordance, and the Individualization of Service. *Journal of Management Information Systems, 35*(2), 424–460. <https://doi.org/10.1080/07421222.2018.1451953>
- Loshin, D. (2010). *Master Data Management*. Morgan Kaufmann.
- Lycett, M. (2013). ‘Datafication’: Making sense of (big) data in a complex world. *European Journal of Information Systems, 22*(4), 381–386.  
<https://doi.org/10.1057/ejis.2013.10>
- Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Hung Byers, A. (2011). *Big data: The next frontier for innovation, competition, and productivity* | McKinsey.  
<https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/big-data-the-next-frontier-for-innovation>
- Martino, B. D., Aversa, R., Cretella, G., Esposito, A., & Kołodziej, J. (2014). Big data (lost) in the cloud. *International Journal of Big Data Intelligence, 1*(1/2), 3. <https://doi.org/10.1504/IJBDI.2014.063840>
- McAfee, A., & Brynjolfsson, E. (2012). Big Data: The Management Revolution. *Harvard Business Review, 9*.
- Mikalef, P., Boura, M., Lekakos, G., & Krogstie, J. (2019). Big Data Analytics Capabilities and Innovation: The Mediating Role of Dynamic Capabilities and Moderating Effect of the Environment. *British Journal of Management, 30*(2), 272–298. <https://doi.org/10.1111/1467-8551.12343>



- Mikalef, P., Pappas, I., Krogstie, J., & Pavlou, P. (2019). Big data and business analytics: A research agenda for realizing business value. *Information & Management*.
- Pehceviski, J. (2019). *Big Data Analytics – Methods and Applications*.  
<https://web.p.ebscohost.com/ehost/ebookviewer/ebook/ZTAwMHh3d19fMjAxMzkzN19fQU41?sid=b7e97746-f6a2-43d5-9ecc-7d0c00cdf9f9@redis&vid=0&format=EB&rid=1>
- Porter, M. E. (1985). *Porter's Generic Competitive Strategies (ways of competing)*.  
<https://www.ifm.engl.cam.ac.uk/research/dstools/porters-generic-competitive-strategies/>
- Porter, M. E. (1991). Towards a dynamic theory of strategy. *Strategic Management Journal*, 12(S2), 95–117.  
<https://doi.org/10.1002/smj.4250121008>
- Ram, J., Zhang, Changyu., & Koronios, A. (2016). The Implications of Big Data Analytics on Business Intelligence: A Qualitative Study in China. *Procedia Computer Science*, 87, 221–226. <https://doi.org/10.1016/j.procs.2016.05.152>
- Rijmenam, M. V. (2014). *Think Bigger: Developing a Successful Big Data Strategy for Your Business*. AMACOM.
- Sagiroglu, S., & Sinanc, D. (2013). Big data: A review. *2013 International Conference on Collaboration Technologies and Systems (CTS)*, 42–47.  
<https://doi.org/10.1109/CTS.2013.6567202>
- Spruit, M., & Pietzka, K. (2015). MD3M: The master data management maturity model. *Computers in Human Behavior*, 51, 1068–1076.  
<https://doi.org/10.1016/j.chb.2014.09.030>
- Stedman, C. (2020, syyskuuta). *What is Business Intelligence? BI Defined*. SearchBusinessAnalytics.  
<https://www.techtarget.com/searchbusinessanalytics/definition/business-intelligence-BI>
- Väre, T. (2019). *Master data* (1.). Alma Talent.
- World Internet Users Statistics and 2022 World Population Stats*. (2021, maaliskuuta 31). <https://www.internetworldstats.com/stats.htm>
- Zikopoulos, P. (2012). *Understanding big data: Analytics for enterprise class Hadoop and streaming data*. McGraw-Hill.