



Big data -analytiikan vaikutukset yrityksen suorituskykyyn

Tietojärjestelmätieteen
kandidaatintutkielma

Laatija:
Ville Kemppainen

Ohjaaja:
FT Samuli Laato

31.07.2025

Turku

Turun yliopiston laatujärjestelmän mukaisesti tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck -järjestelmällä.

Kandidatutkielma

Oppiaine: Tietojärjestelmätiede

Tekijä: Ville Kempainen

Otsikko: Big data -analytiikan vaikutukset yrityksen suorituskykyyn

Ohjaaja: FT Samuli Laato

Sivumäärä: 40

Päivämäärä: 31.07.2025

Tämän kandidaatin tutkielman tavoitteena on selvittää, miten Big Data -analytiikka (BDA) vaikuttaa yrityksen suorituskykyyn. Tutkielma toteutettiin kirjallisuuskatsauksena, ja tutkimusongelmaa lähestyttiin kahden apututkimuskysymyksen avulla: miten yrityksen suorituskykyä on käsitelty akateemisessa kirjallisuudessa, ja miten BDA:n roolia yrityksen liiketoiminnassa on tutkittu.

Yrityksen suorituskykyä lähestyttiin erityisesti kahden teoreettisen viitekehyksen avulla: Balanced Scorecard -malli (Kaplan & Norton, 1992) sekä resurssiperusteinen näkökulma (Barney, 1991). Näiden pohjalta suorituskyvyn ulottuvuuksia tarkasteltiin muun muassa taloudellisten mittareiden, asiakasnäkökulman, sisäisten prosessien sekä innovaation ja oppimisen näkökulmista.

BDA:n osalta tutkielmassa käsiteltiin sen määritelmiä, analyysimenetelmiä ja vaikutuksia liiketoimintaan. BDA mahdollistaa laajojen ja monimuotoisten tietomassojen analysoinnin, mikä voi johtaa parempaan päätöksentekoon, tehokkuuteen, asiakasymmärrykseen sekä nopeampaan innovointiin. BDA:n käyttö on kuitenkin tehokasta vain, jos organisaatiolla on riittävät teknologiset ja organisatoriset kyvykkyydet, mukaan lukien tiedonhallinta ja oppiminen. Tutkimuksissa korostetaan, että pelkkä analytiikan käyttö ei riitä, vaan vaikutukset realisoituvat vasta, kun BDA:n hyödyntäminen on strategisesti johdettua ja osa organisaation oppivaa kulttuuria.

Tutkielma osoittaa, että BDA voi tukea yrityksen suorituskykyä monella tasolla, mutta sen vaikutukset vaihtelevat organisaation kontekstin mukaan. Tämän vuoksi jatkotutkimukselle ja käytännön sovelluksille on edelleen tarvetta.

Avainsanat: big data, big data -analytiikka, yrityksen suorituskyky

SISÄLLYS

1	JOHDANTO	6
2	YRITYKSEN SUORITUSKYKY	9
2.1	Yrityksen suorituskyvyn määrittely: historiallinen kehitys ja eri näkökulmat	9
2.2	Yrityksen suorituskyvyn keskeiset teoriat	11
2.2.1	Balanced Scorecard	11
2.2.2	Resurssiperusteinen näkökulma	14
3	BIG DATA -ANALYTIIKAN ROOLI YRITYKSEN LIIKETOIMINNASSA	18
3.1	Big data ja analytiikka	18
3.2	Big data -analytiikan menetelmät	21
3.3	Big datan vaikutukset työkuvaan ja liiketoimintaprosesseihin	22
3.4	Big data -analytiikan haasteet	23
3.5	Big datan vaikutukset yrityksen suorituskykyyn	24
3.5.1	Big data -analytiikan vaikutus yrityksen suorituskykyyn Balanced Scorecardin neljästä näkökulmasta	25
3.5.2	Big data -analytiikan vaikutus yrityksen suorituskykyyn Resurssiperusteisesta näkökulmasta	28
4	JOHTOPÄÄTÖKSET JA YHTEENVETO	32
	LÄHTEET	34

Kuvioluettelo

Kuvio 1: Balanced Scorecard (mukaillen Kaplan & Norton, 1992).	11
Kuvio 2: Resurssiperusteisen näkökulman osatekijöiden väliset suhteet (mukaillen Barney, 1991)	17
Kuvio 3: Big data resurssien luokittelu (mukaillen Gupta ym., 2016)	29

Taulukkoluetelo

Taulukko 1: Big datan V-ulottuvuudet	19
---	----

1 Johdanto

Maaailma on muuttunut teknologian kehityksen seurauksena suuntaan, jossa vuosi vuodelta voidaan luoda ja varastoida yhä enemmän dataa (Liu, 2014). IDC:n mukaan maailmassa tuotetun ja replikoidun datan määrä eli niin kutsuttu Global DataSphere oli vuonna 2018 noin 33 zettatavua, ja sen ennustettiin kasvavan jopa 175 zettatavuun vuoteen 2025 mennessä (Reinsel ym., 2018). Zettatavu on monille tuntematon termi, mutta yksi zettatavu tarkoittaa yhtä triljoonaa gigatavua. Vertailun vuoksi esimerkiksi yhden teratavun eli tuhannen gigatavun tallentaminen vastaa noin 250 000 HD-laatuista elokuvaa, joten zettatavun mittakaava on valtava. Datun suuren määrän takia sen varastoiminen vaatii usein datakeskuksia. Nykyisessä tietoympäristössä organisaatiot käsittelevät jatkuvasti kasvavia määriä monimuotoista dataa, mikä asettaa yhä suurempia vaatimuksia tiedon tallennus- ja käsittelykapasiteetille. Tämä kehitys on johtanut siihen, että datakeskuksista on tullut kriittisiä toimijoita: ne eivät enää ainoastaan varastoi dataa, vaan vastaavat myös sen hallinnasta, organisoinnista ja hyödyntämisestä. Suurten datamäärien tehokas käsittely vaatii infrastruktuuria, jossa yhdistyvät skaalautuvat tallennusratkaisut, nopea verkkoyhteys, energiatehokkuus sekä korkea luotettavuus. Datakeskusten jatkuva kehittäminen onkin välttämätöntä, jotta valtavia tietomääriä voidaan käsitellä tarkoituksenmukaisesti ja kustannustehokkaasti (Chen ym., 2014). Tämä kehitys puolestaan kiinnostaa yrityksiä, koska suuria datamassoja analysoimalla voidaan mahdollisesti ymmärtää syvällisemmin omaa liiketoimintaa ja asiakkaiden käyttäytymistä, ja näin ollen saavuttaa parempia tuloksia (Liu, 2014).

Valtava analysoitavan datan määrä on tuonut mukanaan termit big data ja big data -analytiikka. Kyseiset termit ovat olleet paljon esillä viimeisen vuosikymmenen aikana, mutta on hyvä käydä läpi, mistä ne koostuvat. Big data ei ole saanut mitään yksittäistä virallista määritelmää, mutta sen voidaan katsoa olevan kokoelma laajoja tietoaaineistoja, joita ei voida käsitellä hyödyntämällä perinteisiä tietokannan hallintatyökaluja (Zikopoulos & Eaton, 2011). Tietoaaineistojen laajuuden lisäksi kirjallisuudessa korostetaan niiden monimutkaisuutta, eli tietoa ei ole vain paljon, vaan myös monenlaista (Manyika ym., 2011).

Big data -analytiikka (BDA) taas määritellään kirjallisuudessa kokonaisvaltaiseksi lähestymistavaksi, joka kattaa datan hallinnan, käsittelyn ja analysoinnin viiden keskeisen ominaisuuden eli niin sanottujen viiden V:n suhteen. Näihin ominaisuuksiin kuuluvat datan määrä (engl. volume), monimuotoisuus (engl. variety), datan syntymisen nopeus (engl. velocity), paikkansapitävyys (engl. veracity) ja arvo (engl. value). (Wamba ym., 2015.) V-ulottuvuuksia käsitellään tarkemmin tutkielman luvussa 3.1. BDA:n tavoitteena on tuottaa toteutettavissa olevia

oivalluksia, jotka edistävät kestävä arvon luomista, suorituskyvyn mittaamista ja kilpailuetujen saavuttamista (Wamba ym., 2015). BDA hyödyntää tilastotieteen, data- ja tekstinlouhinnan, koneoppimisen sekä matemaattisen mallinnuksen menetelmiä ennustaakseen tulevia tuloksia aiempien tietojen perusteella. Sen mukanaan tuomat kyvykkyydet ovat luoneet mahdollisuuksia löytää massiivisen datamassan seasta hyödyllisiä oivalluksia, joiden avulla voidaan saavuttaa kilpailuetua markkinoilla. (Liu, 2014.) Hyötyjen realisoituminen kuitenkin perustuu vahvaan osaamiseen ja tarpeeksi suurten resurssien omaamiseen (Wamba ym., 2017). Osaava yritys voi löytää historiallisen datan seasta piileviä mahdollisuuksia ja välttää erilaisia riskejä (Liu, 2014).

BDA:n rooli herättää erityistä kiinnostusta myös siinä, miten se voi vaikuttaa yrityksen suorituskykyyn (engl. performance). Suorituskyky on keskeinen asia kaikille yrityksille niiden jokapäiväisessä liiketoiminnassa. Ilman suorituskykyistä liiketoimintaa, yrityksen tulevaisuus on epävarma. Accenturen ja General Electricin vuonna 2014 tehdyn tutkimuksen mukaan 87 % yrityksistä uskoi BDA:n muokkaavan kilpailuympäristöä heidän omalla alallaan seuraavan kolmen vuoden sisällä ja 89 % yrityksistä uskoi, että yritykset, jotka eivät ota käyttöön BDA:han perustuvaa strategiaa seuraavan vuoden sisällä, ovat vaarassa menettää markkinaosuutta ja momentumia (Akter ym., 2016). Näistä syistä on kiinnostavaa tarkastella BDA:n yhteyttä ja vaikutuksia yrityksen suorituskykyyn.

Tämän tutkielman tutkimusongelma on, miten big data -analytiikka vaikuttaa yrityksen suorituskykyyn. Tutkimusongelmaan vastaamiseksi hyödynnetään seuraavia apututkimuskysymyksiä:

1. Miten akateemisessa kirjallisuudessa on tutkittu yrityksen suorituskykyä?
2. Miten akateemisessa kirjallisuudessa on tutkittu big data -analytiikan roolia yrityksen liiketoiminnassa?

BDA on laajasti tutkittu ja hyödynnetty aihe analytiikan ammattilaisten ja yritysten päätöksentekijöiden keskuudessa. Tässä tutkielmassa näkökulma rajataan yritysten toimintaan ja erityisesti siihen, miten BDA liittyy yrityksen suorituskykyyn. Koska yrityksen suorituskyky on laaja ja monimuotoinen aihe, sen tarkastelu rajataan kahteen keskeiseen teoriaan, jotka tarjoavat sekä riittävän fokuksen että monipuoliset näkökulmat.

Tutkielman rakenne on seuraava: Luvussa 2 käsitellään yrityksen suorituskykyä, sen määritelmän kehitystä ja ulottuvuuksia aiemman tutkimustiedon pohjalta. Luku 3 syventyy BDA:han, sen hyödyntämiseen liiketoiminnassa sekä siihen liittyviin haasteisiin. Luvussa 3.5 yhdistetään lukujen

2 ja 3 näkökulmat ja tarkastellaan BDA:n vaikutusta yrityksen suorituskykyyn. Luku 4 kokoaa keskeiset havainnot ja esittää johtopäätökset kirjallisuuskatsauksen pohjalta.

2 Yrityksen suorituskyky

Kilpailuympäristön jatkuva muutos edellyttää, että yritykset kykenevät sekä luomaan arvoa asiakkaille että kehittämään liiketoimintaansa jatkuvasti. Yritykset, jotka panostavat innovointiin sekä korkean suorituskyvyn saavuttamiseen ja ylläpitämiseen, ovat usein voittajia markkinoilla. Suorituskyvyn ymmärtäminen ja seuraaminen on siis keskeistä yrityksille, jotka haluavat menestyä. (Taouab & Issor, 2019.) Myös Gavrea ym., (2011) korostavat, että yritysjohton näkökulmasta suorituskyky on keskeinen tekijä organisaation kasvun, kehityksen ja menestyksen mahdollistamisessa.

Luku 2 etenee seuraavasti. Luvussa 2.1 käsitellään suorituskyvyn määritelmän kehitystä ja sen tarkastelussa käytettyjä näkökulmia. Luvussa 2.2 esitellään kaksi teoriaa, jotka muodostavat perustan yrityksen suorituskyvyn ja BDA:n välisen yhteyden tarkastelulle myöhemmissä luvuissa.

2.1 Yrityksen suorituskyvyn määrittely: historiallinen kehitys ja eri näkökulmat

Yrityksen suorituskyky on keskeinen ja laajasti tutkittu aihe strategisen johtamisen kentällä (Taouab & Issor, 2019). Se muodostuu monista erilaisista tekijöistä ja sen tarkastelu tapahtuu jatkuvasti muuttuvassa toimintaympäristössä. Tästä syystä ei ole olemassa yhtä yksiselitteistä määritelmää, joka kattaisi yrityksen suorituskyvyn kokonaisvaltaisesti. Kattavan ymmärryksen saavuttamiseksi onkin tarpeen tarkastella suorituskykyä eri näkökulmista sekä hyödyntää eri aikoina laadittuja määritelmiä.

1950-luvulla yrityksen suorituskyky määriteltiin sen mukaan, missä määrin yritys — sosiaalisena järjestelmänä tarkasteltuna — saavutti asettamansa tavoitteet (Georgopoulos & Tannenbaum, 1957). Tänä aikana suorituskyvyn arviointi keskittyi erityisesti työhön, ihmisiin ja organisaation rakenteeseen. 1960- ja 1970-luvuilla organisaatiot alkoivat kehittää uusia lähestymistapoja suorituskykynsä arvioimiseen. Näiden vuosikymmenten aikana suorituskyky ymmärrettiin laajalti organisaation kykyä sopeutua ympäristöönsä sekä hankkia ja hyödyntää rajallisia resursseja tehokkaasti (Seashore & Yuchtman, 1967). 1980- ja 1990-luvuilla organisaation suorituskyvyn ymmärtäminen syveni entisestään, kun huomattiin, että tavoitteiden asettaminen ja tunnistaminen on aiemmin oletettua monimutkaisempaa. Johtamisajattelussa alettiin painottaa sitä, että organisaation menestys rakentuu sekä tavoitteiden saavuttamisesta eli vaikuttavuudesta että resurssien säästeliäästä käytöstä eli tehokkuudesta. Organisaatioteoriat alkoivatkin korostaa, että suorituskyky rakentuu tavoitteiden saavuttamisen ja rajallisten resurssien optimoinnin välisestä tasapainosta (Lusthaus & Adrien, 1998). 1980-luvun lopulla monet tutkijat ja konsultointiyritykset

alkoivat myös korostaa ei-taloudellisten mittareiden merkitystä yrityksen suorituskyvyn arvioinnissa (Gavrea ym., 2011).

Yrityksen suorituskykyä on myös määritelty 2000-luvulle tultaessa listaamalla sen osatekijöitä. Lebasin ja Eusken (2002) suunnitteleman listan mukaan:

- Suorituskyky on joukko taloudellisia ja ei-taloudellisia mittareita, jotka tarjoavat tietoa tavoitteiden ja tulosten saavuttamisen tasosta
- Suorituskyky on dynaaminen käsite, joka vaatii harkintaa ja tulkintaa
- Suorituskykyä voidaan havainnollistaa kausaalimallin avulla, joka kuvaa, miten nykyiset toimet voivat vaikuttaa tuleviin tuloksiin
- Suorituskyky voidaan ymmärtää eri tavoin sen mukaan, kuka arviointia tekee
- Suorituskyvyn määrittelemiseksi on tunnettava sen keskeiset ominaispiirteet kunkin vastuualueen osalta
- Yrityksen suorituskyvyn raportointia varten on kyettävä mittaamaan saavutetut tulokset

Kyseinen lista käsittelee yrityksen suorituskykyä monipuolisemmin kuin monet muut määritelmät. Lista kuvaa yrityksen suorituskyvyn dynaamisuutta, taloudellisia ja ei-taloudellisia mittareita, subjektiivisuutta sekä kykyä kvantifioida eli mitata ja ilmaista numeerisesti saatuja tuloksia. (Lebas & Euske, 2002.)

Yrityksen suorituskyvyn määritelmää tarkasteltaessa huomataan, miten monista eri näkökulmista suorituskykyä voidaan tutkia ja mitata. Erilaiset asiat kiinnostavat eri yrityksiä, joten on järkevää jaotella suorituskykyä muutamiin eri luokkiin, joita voidaan tutkia. Perinteinen ja kapea näkökulma tarkastelee yrityksen suorituskykyä hyödyntämällä yksinkertaisia tulosperusteisia taloudellisia mittareita, joiden oletetaan heijastavan yrityksen taloudellisten tavoitteiden täyttymistä. Tätä lähestymistapaa yrityksen suorituskykyyn kutsutaan taloudelliseksi suorituskyvyksi. Taloudellisia mittareita ovat esimerkiksi myynnin kasvu ja kannattavuus.

Yrityksen suorituskyvyn määrittelyssä voidaan myös käyttää kaksijakoista näkökulmaa, jossa tarkastellaan yrityksen taloudellisten mittareiden lisäksi myös operatiivisia mittareita. Tämä näkökulma tuo tarkasteluun mukaan tekijöitä, kuten tuotteiden laadun, markkinoinnin tehokkuuden ja markkinaosuuden. (Venkatraman & Ramanujam, 1986.)

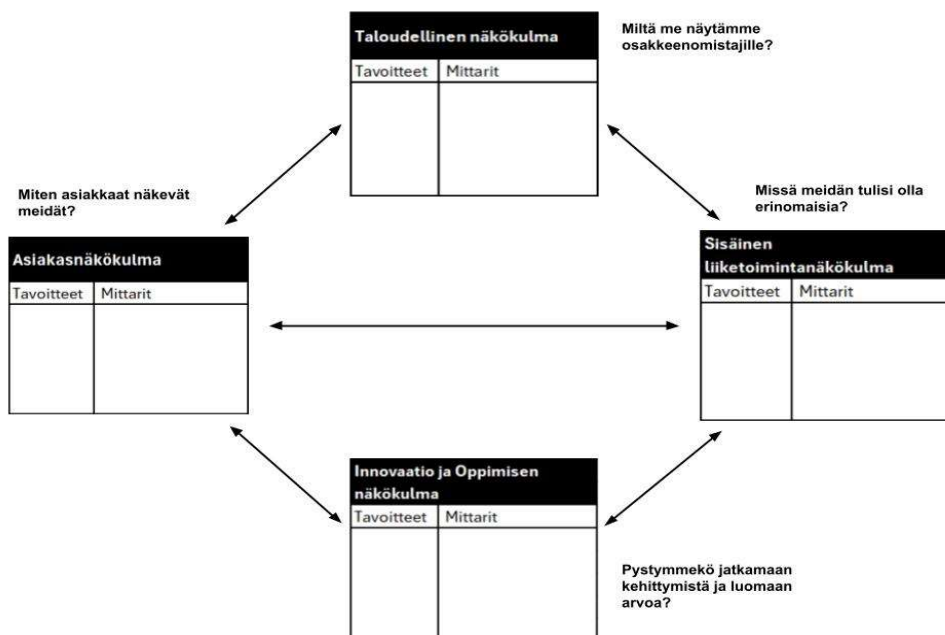
Muutama vuosi tämän kaksijakoisen määritelmän jälkeen suorituskyvyn tarkastelussa alettiin käyttää myös jakoa, jossa yrityksen suorituskykyä käsiteltiin neljästä eri näkökulmasta. Jako toi

taloudellisen suorituskyvyn seuraksi asiakasnäkökulman, sisäiset liiketoimintaprosessit sekä oppimisen ja innovaatiot. Jako näihin neljään ryhmään laajentaa tarkastelua perinteisten taloudellisten tekijöiden seuraamisesta kokonaisvaltaisempaan liiketoiminnan tarkasteluun. (Kaplan & Norton, 1992.) Kaplanin ja Nortonin vuonna 1992 esittelemää nelijakoa on hyödynnetty tutkimuksissa nykypäivään asti sen monipuolisuuden takia. Tämän takia valitsin sen yhdeksi teoriaksi, joka esitellään luvussa 2.2. Teorioiden esittelyn yhteydessä tarkastellaan tarkemmin näitä erilaisia näkökulmia ja mittareita, joita hyödynnetään yrityksen suorituskykyyn liittyvässä akateemisessa kirjallisuudessa.

2.2 Yrityksen suorituskyvyn keskeiset teoriat

2.2.1 Balanced Scorecard

1990-luvun alkupuolella kehitelty Balanced Scorecard (BSC) on yksi tunnetuimmista ja laajimmin käytetyistä strategisen suorituskyvyn mittaamisen malleista. BSC pyrkii tarjoamaan kokonaisvaltaisen näkökulman yrityksen suorituskyvystä yhdistämällä perinteisesti käytetyt taloudelliset mittarit ei-taloudellisiin tekijöihin, kuten asiakastytyväisyyteen, sisäisiin prosesseihin sekä organisaation oppimiseen ja kasvuun. BSC:n tarkoituksena on korostaa, että strategista suorituskykyä on arvioitava monipuolisesti ja tasapainoisesti, eivätkä pelkät taloudelliset mittarit riitä kuvaamaan yrityksen pitkän aikavälin menestystä. (Kaplan & Norton, 1992.)



Kuvio 1: Balanced Scorecard (mukaiillen Kaplan & Norton, 1992).

Seuraavaksi käydään läpi BSC:n kaikki neljä osa-aluetta tarkemmin. Osa-alueet ovat asiakasnäkökulma, sisäinen liiketoimintanäkökulma, innovaation ja oppimisen näkökulma sekä taloudellinen näkökulma.

Miten asiakkaat näkevät meidät? Tämä on yksi neljästä kysymyksestä, joihin yrityksen päätöksentekijöiden tulisi pystyä vastaamaan. Monet yritykset ovat hyvin keskittyneitä tuottamaan asiakkailleen arvoa, joten asiakasnäkökulman huomioon ottaminen on keskeistä suorituskyvyn tarkastelussa. BSC:n mukaan asiakkaiden huolenaiheet jakaantuvat yleensä neljään kategoriaan: aika, laatu, suorituskyky ja palvelu sekä kustannukset. Esimerkiksi laatua voidaan mitata laskemalla virheellisten tuotteiden määriä ja kysymällä asiakkaiden kokemuksia. Ajan tarkastelussa taas voidaan keskittyä esimerkiksi läpimenoajan tarkkailuun. Nykyisten tuotteiden kohdalla tämä tarkoittaa aikaa, joka kuluu asiakkaan tilauksesta siihen, kun tuote toimitetaan asiakkaalle. Jotta yritys voi saada BSC:n hyötyjä realisoitumaan, sen on tuotava esiin näihin osa-alueisiin liittyviä tavoitteita ja muunnettava nämä tavoitteet konkreettisiksi mittareiksi. BSC:n mukaan hyödynnettävän tiedon sijainti voi myös vaihdella. Joissakin tilanteissa voidaan tarvita sisäisen tiedon lisäksi myös ulkoista tietoa. Yritys voi esimerkiksi pyrkiä markkinajohtajaksi asiakastyytyväisyydessä. Kyseinen yritys voi saada tarvitsemansa tiedon ulkopuoliselta toimijalta, joka toteuttaa kyselyn asiakkaille. (Kaplan & Norton, 1992.)

Toinen kysymys, johon yrityksen on BSC:n mukaan kyettävä vastaamaan, on: missä meidän on oltava erinomaisia? Asiakasnäkökulman mittarit ovat tärkeitä, mutta ne on muunnettava mittareiksi, jotka kuvaavat sitä, mitä yrityksen on tehtävä sisäisesti vastatakseen asiakkaiden odotuksiin. Erinomainen asiakassuorituskyky perustuu loppujen lopuksi organisaation läpi kulkeviin prosesseihin, päätöksiin ja toimenpiteisiin. Yrityksen sisäisten mittareiden tulisi perustua liiketoiminnan prosesseihin, joilla on suurin vaikutus asiakkaiden tyytyväisyyteen. Esimerkiksi tekijät, jotka vaikuttavat läpimenoaikaan, laatuun, työntekijöiden osaamiseen ja tuottavuuteen voivat olla keskiössä yrityksestä riippuen. Yrityksen tulisi myös pyrkiä tunnistamaan ja mittaamaan ydinosaamisensa sekä kriittiset teknologiat, jotka ovat välttämättömiä markkinajohtajuuden säilyttämiseksi. Toinen asia, josta yrityksen tulisi päättää on, millä osaamisalueilla ja prosesseissa sen tulisi olla erinomainen. Tämän jälkeen sen tulisi määritellä kullekin alueelle mittarit, joiden avulla toimintaa voidaan seurata. Esimerkiksi toimistotuotteita valmistava yritys voi päättää, että läpimenoaikaa tulisi nopeuttaa jokin prosentuaalinen määrä. Tämän päätöksen jälkeen sen tulisi päättää, mitä mittareita sen tulisi seurata ja pyrkiä parannuksiin toiminnan alueilla, jotka vaikuttavat kyseisiin mittareihin. Esimerkiksi asiakastilausten prosessoinnin nopeuttaminen, materiaalien oston

ja saapumisen nopeuttaminen tuottajilta sekä tuotteiden valmistuksen ja kokoamisen nopeuttaminen voisivat olla keskeisiä tavoitteita ja mittareita, joita tulisi seurata. (Kaplan & Norton, 1992.)

BSC:n kolmas osio keskittyy innovaatioihin ja oppimiseen. Voimmeko jatkaa kehittymistä ja luoda arvoa? Tämä on kysymys, johon yrityksen pitäisi pystyä vastaamaan. Kilpailullinen ympäristö pakottaa yritykset tekemään jatkuvasti muutoksia liiketoiminnassaan. Yritykset voivat esimerkiksi joutua kehittämään nykyisistä tuotteistaan jatkuvasti parempia ja pyrkiä tuomaan markkinoille uusia tuotteita, jotka omaavat parempia ominaisuuksia. Yleensä yrityksen arvo sitoutuukin vahvasti sen kykyyn innovoida, kehittyä ja oppia. Kyky tuoda markkinoille uusia tuotteita, luoda asiakkaille enemmän arvoa ja parantaa jatkuvasti toiminnan tehokkuutta mahdollistavat pääsyn uusille markkinoille sekä liikevaihdon ja katteiden kasvattamisen. Toisin sanoen nämä kyvykkyydet mahdollistavat kasvun ja siten osakkeenomistajien arvon lisäämisen. (Kaplan & Norton, 1992.)

BSC:n viimeinen osio on taloudellinen näkökulma. Yrityksen on kysyttävä itseltään, miltä se näyttää osakkeenomistajille. Taloudellisen suorituskyvyn mittarit näyttävät, miten yrityksen strategia, toteutus ja toimeenpano edistävät lopulta tulosparannusta. Tyypilliset taloudelliset tavoitteet liittyvät kannattavuuteen, kasvuun ja osakkeenomistajien arvon lisäämiseen. Yhden yrityksen taloudelliset tavoitteet olivat selviytyä, menestyä ja vaurastua. Näiden tavoitteiden seuraamiseksi valittiin kullekin tavoitteelle omia mittareita, jotka auttoivat tavoitteiden saavuttamisen seuraamisessa. Selviytymistä mitattiin kassavirralla. Menestymistä arvioitiin neljännesvuosittaisella myynnin kasvulla ja liiketuloksella per liiketoimintayksikkö. Vaurastumista taas mitattiin markkinaosuuden kasvulla segmenteittäin sekä oman pääoman tuotolla. (Kaplan & Norton, 1992.)

Näiden neljän osa-alueen hyödyntäminen tekee BSC:stä monipuolisen työkalun yrityksen nykyisen toiminnan tarkasteluun ja suorituskyvyn kehittämiseksi. BSC:n toimeenpanossa on kuitenkin huomattu yksi tärkeä seikka. Sen toteuttaminen vaatii ylimmän johdon osallistumisen. Ylimmällä johdolla on kattavin näkemys yrityksen visiosta ja prioriteeteista, joten heidän panoksensa on välttämätöntä onnistumisen kannalta. Tämä huomio on merkittävä, koska useimmat olemassa olevat suorituskyvyn mittausjärjestelmät on suunniteltu ja johdettu talousasiantuntijoiden toimesta. BSC siis eroaa monista muista mittausjärjestelmistä siinä, että toteuttamiseen tarvitaan myös ylimmän johdon vahvaa osallistumista. (Kaplan & Norton, 1992.)

2.2.2 Resurssiperusteinen näkökulma

Resurssiperusteinen näkökulma (engl. Resource-based view) on 1980-luvulta nykypäivään asti hyödynnetty teoria, jossa yrityksen sisäisiä resursseja pidetään kilpailuedun lähteenä.

Resurssiperusteinen näkökulma kehittyi vastareaktion aikana hyvin suosituille ulkoisille strategioille kuten esimerkiksi Porterin viiden kilpailuvoiman mallille, jossa keskittyminen on yrityksen ulkopuolisissa tekijöissä. Monet 1980-luvulla vallitsevat kilpailuetuja tarkastelevat teoriat selittivät, että resurssiperusteinen kilpailuetu markkinoilla on vain lyhytaikaista, koska resurssit, joita yritykset hyödyntävät toteuttaakseen strategioitaan ovat erittäin liikkuvia. Tämä tarkoittaa, että yhden yrityksen saavuttama resursseihin perustuva kilpailuetu olisi lyhytkestoista, koska kilpailijat saisivat samoja resursseja käyttöönsä ennen pitkää. Resurssiperusteisen näkökulman kehittäjät vastustivat kyseistä ajattelua. Resurssiperusteinen näkökulma perustuu kahteen tärkeään oletukseen. Ensinnäkin malli perustuu oletukseen, että toimialan yritysten välillä voi olla eroja hallinnoimiensa strategisten resurssien suhteen. Toiseksi, malli olettaa, että kaikki resurssit eivät ole täysin liikkuvia yrityksiä välillä eli resurssien heterogeenisyys yritysten välillä voi olla pitkäaikaista. (Barney, 1991.)

Ennen kuin käymme Resurssiperusteisen näkökulman sisältöä tarkemmin läpi, on määriteltävä kolme eri termiä, jotka liittyvät kyseiseen teoriaan. Nämä termit ovat yrityksen resurssit, kilpailuetu ja pysyvä kilpailuetu. Perinteisesti yrityksen resursseilla on strategisessa kirjallisuudessa tarkoitettu vahvuuksia, joita yritykset voivat hyödyntää strategioiden kehittämisessä ja toteuttamisessa (Learned ym., 1969; Porter, 1981). Toinen määritelmä, joka on Resurssiperusteisen näkökulman kannalta sopivampi kertoo yrityksen resurssien kattavan kaikki ne kyvykkyydet, yrityksen ominaisuudet, varat, organisaatioprosessit, informaation ja tiedon, joita yritys hallitsee, ja jotka mahdollistavat yritykselle sellaisten strategioiden kehittämisen ja toteuttamisen, jotka parantavat sen tehokkuutta ja vaikuttavuutta (Daft, 1989). Kirjallisuudessa yrityksen resursseja on myös jaettu kolmeen kategoriaan: fyysisiin pääomaresursseihin kuten fyysisiin laitteisiin ja teknologiaan sekä raaka-aineisiin, inhimillisiin pääomaresursseihin kuten yksittäisten päätöksentekijöiden ja työntekijöiden kokemus, harkintakyky, koulutus, älykkyys, ihmissuhteet sekä oivalluskyky ja organisatorisiin pääomaresursseihin kuten yrityksen muodollinen raportointirakenne, sen muodolliset ja epämuodolliset ohjaus-, suunnittelu- ja koordinoitijärjestelmät sekä epämuodolliset suhteet organisaation eri ryhmien välillä ja yrityksen sekä sen toimintaympäristön toimijoiden välillä. (Barney, 1991.)

Yrityksellä voidaan sanoa olevan kilpailuetu, kun se toteuttaa arvonluomisen strategian, jota nykyiset tai potentiaaliset kilpailijat eivät ole toteuttamassa. Kilpailuetua voidaan pitää pysyvänä, kun muut yritykset eivät kykene saavuttamaan strategiasta saatavia hyötyjä esimerkiksi kopioimalla strategiaa. Pysyvän kilpailuedun määrittelyssä on myös keskeistä, että pysyvä kilpailuetu ei perustu kilpailuedun esiintymiseen jollakin tietyllä ajanjaksolla. Määrittelyssä useasti sanotaankin, että pysyvä kilpailuetu on kilpailuetu, joka kestää pitkän ajan. Määritelmä on ajallisesti siis hyvin suhteellinen. Tärkeämpää kilpailuedun pysyvyyden kannalta on, että kilpailuetu pysyy, vaikka sitä yritettäisiinkin kopioida kilpailijoiden toimesta. (Barney, 1991.)

Resurssiperusteisen näkökulman hyödyntämisen kannalta on pystyttävä tunnistamaan resursseja, jotka voivat potentiaalisesti edesauttaa pysyvän kilpailuedun saavuttamisessa. Resurssin on omattava neljä ominaisuutta, jotta sen avulla voidaan saavuttaa pysyvää kilpailuetua. Resurssin on oltava arvokas, harvinainen, vaikeasti jäljiteltävä eikä mikään muu korvaava resurssi saa olla strategisesti arvokas, harvinainen ja vaikeasti jäljiteltävissä. (Barney, 1991.)

Resurssit ovat arvokkaita, kun ne mahdollistavat yritykselle sen tehokkuutta ja vaikuttavuutta parantavien strategioiden kehittelyn ja toteutuksen. Perinteinen malli, joka tarkastelee yrityksen vahvuuksia, heikkouksia, mahdollisuuksia ja uhkia, kertoo, että yritys voi parantaa suorituskykyään vain sen strategioiden hyödyntäessä mahdollisuuksia tai ehkäistessä uhkia. Yrityksen ominaisuudet voivat omata muitakin piirteitä, jotka voivat tehdä niistä pysyvän kilpailuedun lähteitä, mutta näistä ominaisuuksista tulee resursseja vasta, kun ne hyödyntävät mahdollisuuksia tai ehkäisevät uhkia yrityksen ympäristössä. Resurssien arvokkuuden arvioinnissa voidaan siis hyödyntää ensin yrityksen ympäristöön keskittyviä malleja, jotta saadaan selville, mitkä ominaisuudet ovat sen arvoisia, että niitä voidaan pitää resursseina. Vasta tämän jälkeen siirrytään hyödyntämään Resurssiperusteista näkökulmaa, jonka avulla voidaan pohtia muita ominaisuuksia, joita näiltä resursseilta vaaditaan pysyvän kilpailuedun saavuttamiseksi. (Barney, 1991.)

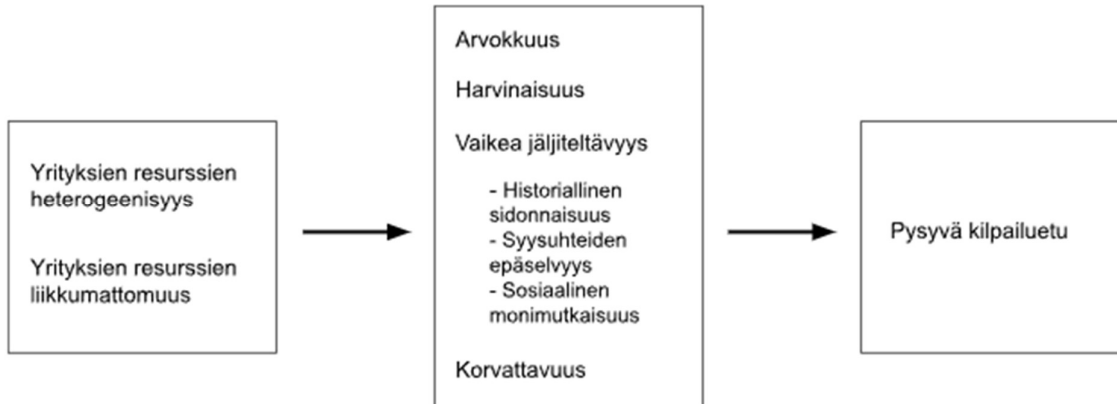
Vaikka jokin resurssi olisi arvokas, se ei tee siitä vielä pysyvän kilpailuedun tuottajaa yritykselle. Monissa tilanteissa muillakin alalla toimivilla yrityksillä on kyseistä arvokasta resurssia, joten kukaan toimija ei saa tästä arvokkaasta resurssista kilpailuetua toisiaan kohtaan. Kirjallisuuden mukaan on vaikeaa arvioida kuinka harvinainen arvokkaan resurssin tulee olla, jotta yritys voi potentiaalisesti saada kilpailuetua sen avulla. On selvää, että yrityksen omatessa ainoana toimijana jotakin resurssia, se saa vähintäänkin kilpailuetua, jos ei jopa pysyvää kilpailuetua. Toisaalta on myös mahdollista, että pienellä määrällä kilpailijoita on jotakin samaa arvokasta resurssia ja silti voidaan saada kilpailuetua markkinoilla. (Barney, 1991.)

Harvinainen ja arvokas resurssi on kuitenkin pysyvän kilpailuedun lähde ainoastaan, jos yritykset, joilla ei ole kyseistä resurssia eivät sitä myöskään kykene hankkimaan. Kyseisiä pysyvää kilpailuetua tuottavia resursseja pidetään epätäydellisesti jäljiteltävinä (engl. imperfectly imitable). Resurssi voi olla epätäydellisesti jäljiteltävissä joko yhdestä syystä tai kolmen eri syyn yhdistelmästä.

Yksi syy jäljiteltävyyden vaikeudelle on resurssin hankkimisen perustuminen yrityksen ainutlaatuisiin historiallisiin olosuhteisiin. Resurssiperusteisen näkökulman mukaan yritykset ovat luonnostaan historiallisia ja sosiaalisia toimijoita ja niiden kyky hankkia ja hyödyntää joitakin resursseja riippuu niiden ajallisesta ja paikallisesta kontekstista. Teorian mukaan yritys, joka ei hanki tällaisia resursseja oikeaan aikaan, ei välttämättä pysty hankkimaan niitä lainkaan myöhemmin. Yritys, joka sen sijaan onnistuu hankkimaan kyseisiä resursseja, voi hyödyntää niitä kehittäessään strategioita arvon luomiseksi. Toinen syy jäljiteltävyyden vaikeudelle on kausaalinen epäselvyys (engl. causal ambiguity). Tämä tarkoittaa, että joissakin tilanteissa ei ole täysin selvää linkkiä yrityksen hallitseman resurssin ja sen omaaman pysyvän kilpailuedun välillä. Epävarmuus tekee kopioimisen vaikeaksi, joten useasti kyseinen tilanne johtaa pysyvään kilpailuetuun. Teorian mukaan molempien osapuolien on oltava osittain epätietoisia siitä, mikä tai mitkä resurssit tuottavat yhdelle yritykselle pysyvää kilpailuetua. Mikäli pysyvän kilpailuedun omaava yritys ja sen työntekijät tiedostavat tarkasti, mitkä resurssit ovat edun lähteenä, voi kilpailija teoriassa kopioida edun esimerkiksi rekrytoimalla avainhenkilön, jolloin etu ei ole enää kestävää laatua. Kolmas syy jäljiteltävyyden vaikeudelle on sosiaalinen kompleksisuus tai monimutkaisuus (engl. social complexity). Koska osa resursseista perustuu sosiaaliseen kompleksisuuteen, ne voivat olla niin monimutkaisia, ettei yritys pysty vaikuttamaan niihin tai hallitsemaan niitä järjestelmällisesti. Pysyvän kilpailuedun perustuessa juuri tällaiseen monimutkaiseen ilmiöön, muiden yritysten kyky jäljitellä näitä resursseja on merkittävästi rajoittunut. Fyysinen teknologia on usein resurssi, joka on melko helposti jäljiteltävissä. Tämän fyysisen teknologian hyödyntäminen taas vaatii usein sosiaalisesti kompleksisia resursseja, joten siitä saatavat hyödyt vaihtelevat suuresti yritysten välillä. (Barney, 1991.)

Viimeinen edellytys resurssin kyvyille tuottaa pysyvää kilpailuetua on, ettei markkinoilla ole saatavilla strategisesti vastaavia resursseja, jotka ovat yleisiä tai helposti jäljiteltävissä. Kaksi resurssia ovat strategisesti vastaavia, jos kumpaakin voidaan hyödyntää erikseen saman strategian toteuttamisessa. Näin ollen, vaikka yritys käyttäisi arvokkaita, harvinaisia ja vaikeasti jäljiteltäviä resursseja, kilpailijat voivat silti heikentää kilpailuetua hyödyntämällä vaihtoehtoisia, helpommin saatavilla olevia resursseja samaan tarkoitukseen. Tällaiset korvaavat resurssit mahdollistavat

vastaavan arvonluonnin eri keinoin, mikä vähentää alkuperäisen resurssin strategista merkitystä. Siksi pysyvän kilpailuedun säilyminen edellyttää, että vastaavia korvaavia resursseja ei ole helposti saatavilla. (Barney, 1991.)



Kuvio 2: Resurssiperusteisen näkökulman osatekijöiden väliset suhteet (mukaillen Barney, 1991)

3 Big data -analytiikan rooli yrityksen liiketoiminnassa

Luvun 3 rakenne on seuraavanlainen. Luvussa 3.1 määritellään, mitä big data ja BDA ovat, miksi niitä pidetään merkittävinä nykyaikaisessa liiketoiminnassa, ja miten ne eroavat perinteisestä analytiikasta. Luvussa 3.2 esitellään keskeisiä menetelmiä, joita yritykset käyttävät BDA:n hyödyntämisessä. Luvussa 3.3 tarkastellaan BDA:n roolia ja vaikutuksia liiketoimintaan tiedon arvoketjun avulla. Luvussa 3.4 käsitellään BDA:n käyttöönottoon ja toteuttamiseen liittyviä haasteita. Luvussa 3.5 syvennyttään BDA:n ja yrityksen suorituskyvyn väliseen suhteeseen aiemman kirjallisuuden pohjalta, hyödyntäen luvussa 2.2 esiteltyjä teoreettisia viitekehyksiä analyysin tukena.

3.1 Big data ja analytiikka

BDA:n käsitteen toivat esiin Chen ym. (2012) osana liiketoimintatiedon hallinnan ja analytiikan (engl. Business Intelligence & Analytics, BI&A) kenttää. Heidän mukaansa Big data viittaa erittäin suuriin ja monimutkaisiin tietomassoihin, joiden käsittely edellyttää kehittyneitä teknologioita. BDA puolestaan tarkoittaa näiden tietojen analysointia BI&A:n menetelmin, tarjoten uusia mahdollisuuksia päätöksenteon tueksi. BDA viittaa erityisesti niihin BI&A-teknologioihin, jotka keskittyvät suurten tietomassojen analysointiin data- ja tekstilouhinnan, koneoppimisen sekä tilastollisten menetelmien avulla (Liu, 2014). Datalouhinta, jota kutsutaan myös tietokannoista tehtäväksi tiedon etsinnäksi, tarkoittaa prosessia, jossa tietokannoista pyritään löytämään piilevää, aiemmin tuntematonta ja mahdollisesti hyödyllistä tietoa – kuten sääntöjä, rajoitteita tai säännönmukaisuuksia (Frawley ym., 1992). Johdannossa mainittiin myös BDA:n määrittelyssä hyödynnettyjä V-ulottuvuuksia, joiden avulla saadaan selkeämpi kuva siitä, miten big data eroaa perinteisestä analytiikasta ja datasta. Tässä yhteydessä annettussa määritelmässä BDA:ta pidetään kokonaisvaltaisena lähestymistapana, joka kattaa datan hallinnan, käsittelyn ja analysoinnin. (Wamba ym., 2015.) BDA:n hyödyntäminen perustuu tavoitteeseen tuottaa oivalluksia, jotka edistävät liiketoimintaa ja auttavat saavuttamaan kilpailuetuja markkinoilla (Wamba ym., 2015; Liu, 2014).

Kuten W. Edwards Deming ja Peter Drucker ovat sanoneet, digitaalisen datan räjähdysmäinen kasvu on ollut merkittävää. Heidän mukaansa big datan ansiosta johtajat voivat mitata – ja siten myös tietää – huomattavasti enemmän liiketoiminnastaan ja muuntaa tämän tiedon suoraan paremmaksi päätöksenteoksi ja suorituskyvyksi. Big data ja perinteinen analytiikka ovat sidoksissa toisiinsa. Niissä on kuitenkin muutamia keskeisiä eroja. (McAfee ym., 2012.) Käsittelemme

seuraavaksi näitä keskeisiä eroja big datan V-ulottuvuuksien kautta. Alun perin käsite big data viittasi ensisijaisesti juurikin tiedon suureen määrään (engl. volume), joka syntyy uusien teknologioiden, kuten sosiaalisen median, RFID-tunnisteiden, älypuhelinien ja sensorien, käytön seurauksena. Ajan kuluessa määritelmää on laajennettu kattamaan myös datan monimuotoisuus (engl. variety) sekä nopeus (engl. velocity), jolla dataa syntyy. Myöhemmin big datan ulottuvuuksia on laajennettu edelleen sisältämään myös datan luotettavuus (engl. veracity) ja arvo (engl. value). Nämä ovat niin sanottuja big datan V-ulottuvuuksia. Termi perustuu kyseisten ulottuvuuksien englanninkielisten sanojen ensimmäiseen kirjaimen, joka on kaikissa v-kirjain. (Gupta & George, 2016.) Kirjallisuudessa puhutaan lähteestä riippuen pääsääntöisesti joko kolmesta, neljästä tai viidestä V-ulottuvuudesta. Jotkut lähteet käsittelevät vain määrää, nopeutta ja monimuotoisuutta, mutta käsiteltävän datan luotettavuus ja arvo ovat myös tärkeitä kokonaisuuden ymmärtämiseksi, joten käsittelemme tässä tutkielmassa kaikkia viittä ulottuvuutta. Alla oleva taulukko 1 tiivistää hyvin yksinkertaistettuun muotoon käsiteltävien termien sisällön.

V-ulottuvuudet:	Kuvaus:
Volume (määrä)	Kuvaa valtavaa tietomäärää, jota organisaatiot keräävät eri lähteistä, kuten transaktioista, some-kanavista, sensoreista jne.
Velocity (nopeus)	Tarkoittaa sitä, kuinka nopeasti dataa syntyy, siirretään ja analysoidaan – usein reaaliajassa.
Variety (monimuotoisuus)	Viittaa siihen, että dataa on monenlaista: rakenteista (esim. taulukot), rakenteetonta (esim. videot, tekstit) ja puolirakenteista (esim. XML).
Veracity (luotettavuus)	Kuvaa datan luotettavuutta, tarkkuutta ja laatua – onko data ristiriidatonta, puhdasta ja käyttökelpoista?
Value (arvo)	Tarkoittaa datasta saatavaa hyötyä ja liiketoiminnallista arvoa – eli miten data muunnetaan oivalluiksi ja päätöksenteon tueksi.

Taulukko 1: Big datan V-ulottuvuudet

Seuraavaksi käsittelemme näitä viittä ulottuvuutta yksityiskohtaisemmin. Jo vuonna 2012 maailmassa luotiin 2,5 eksatavua dataa joka päivä. Tämä määrä myös kaksinkertaistui noin 40 kuukauden välein. Internetin kautta kulki joka sekunti enemmän dataa kuin koko internetissä oli tallennettuna 20 vuotta aiemmin. Vuonna 2012 Walmartin arvioitiin keräävän yli 2,5 petatavua asiakastapahtumiin liittyvää dataa joka tunti. Yksi petatavu dataa vastaa noin 20 miljoonan

arkistokaapin verran tekstiä. (McAfee ym., 2012.) Walmart keräsi dataa siis noin 50 miljoonan arkistokaapin verran joka tunti. Puhutaan siis suurista määristä dataa, joita yritykset kuten Walmart voivat hyödyntää, jos niillä on hyödyntämiseen vaadittavaa osaamista ja järjestelmiä. Tsai ym. (2015) käsittelevät datan määrän kasvuun liittyviä ongelmia. Perinteisiä data-analyysimenetelmiä ei ole suunniteltu suurten ja monimutkaisten tietomassojen käsittelyyn, mikä voi aiheuttaa merkittäviä haasteita yrityksille. Tämän vuoksi niiden skaalaaminen suurivolyymiseen big dataan on käytännössä rajoittunutta, mikä voi heikentää analytiikan suorituskykyä ja estää tehokkaan päätöksenteon datan kasvun myötä.

Monissa käyttötarkoituksissa datan syntyminenopeus voi olla jopa tärkeämpää kuin sen määrä. Esimerkiksi reaaliaikainen tai lähes reaaliaikainen tiedon hyödyntäminen on joillekin yrityksille kriittinen kilpailutekijä, sillä se mahdollistaa nopeamman reagoinnin ja ketterämmän toiminnan verrattuna kilpailijoihin. Hyvä esimerkki tästä ovat Wall Streetin analyttikot – sidosryhmä, joka voi saavuttaa merkittävää kilpailuetua reaaliaikaisen datan avulla. Nopeat havainnot mahdollistavat markkinakehityksen ennakkoinnin ja päätöksenteon ennen virallisten tietojen julkaisua. (McAfee ym., 2012.) Erevelles ym. (2016) korostavat, että BDA-kyvykkyyksien avulla päätöksentekijät voivat tukea päätöksiään reaaliaikaisella datalla, mikä vähentää intuition perustuvan päätöksenteon tarvetta.

Datasta on myös tullut monimuotoisempaa. Dataa esiintyy esimerkiksi sosiaalisen median viesteinä, päivityksinä ja kuvina, matkapuhelinten GPS-tietoina sekä sensoreiden tuottamina mittaustietoina. Laitteet kuten älypuhelimet ja muut mobiililaitteet tuottavat valtavia määriä eri muodoissa esiintyvää dataa liittyen ihmisiin, toimintaan ja sijainteihin. (McAfee ym., 2012.) Perinteisesti esimerkiksi markkinoinnissa on hyödynnetty pääasiassa rakenteista dataa, kuten tietokantoihin tallennettuja ostotapahtumia tai asiakasrekistereitä. BDA:n myötä huomio on kuitenkin yhä enemmän siirtynyt myös rakenteettomaan dataan, kuten juurikin sosiaalisen median sisältöihin, kuviin ja videoihin, mikä on laajentanut datan lähteitä ja analyysin mahdollisuuksia merkittävästi. (Erevelles ym., 2016.) Monimuotoisuus on tehnyt perinteisistä tietokannoista sopimattomia big datan hallintaan ja säilyttämiseen. Lisäksi laskeneet kustannukset esimerkiksi tallennustilassa, muistissa, prosessointitehossa ja kaistanleveydessä tekevät aiemmin kalliista ja dataintensiivisistä lähestymistavoista nykyään kannattavampia taloudellisesti. (McAfee ym., 2012.)

Datan luotettavuus on tekijä, jota ei käsitellä kaikissa akateemisissa artikkeleissa, joissa tarkastellaan big datan V-ulottuvuuksia. Luotettavuus on kuitenkin monille tahoille erittäin keskeinen tai jopa välttämätön asia. Luotettavuus määritellään epävarmuuden määräksi, joka johtuu

datan epäjohdonmukaisuudesta ja puutteellisuudesta, viiveistä, epäselvyyksistä, harhaanjohtamisesta ja mallin likimääräisyyksistä. Monet tahot puhuvatkin datan tarkkuudesta, uskottavuudesta, objektiivisuudesta ja johdonmukaisuudesta tärkeinä liiketoimintaan vaikuttavina kysymyksinä. (Reimer & Madigan, 2019.) Kwon ym. (2014) painottavat, että nämä puutteelliset tiedot tai datan epätarkkuus voivat johtaa tilanteeseen, jossa yrityksen päätöksentekijät joutuvat turvautumaan vaihtoehtoisin ratkaisuihin, kuten omiin henkilökohtaisiin arvioihin, muiden toiminnan seuraamiseen ja siitä oppimiseen tai jopa sattumanvaraiseen kokeiluun. On myös selvää, että esimerkiksi terveydenhuollossa datan luotettavuus on ratkaisevan tärkeää, sillä harhaanjohtava tai virheellinen data voi johtaa vakaviin seurauksiin potilasturvallisuuden ja hoitopäätösten kannalta.

Viides ja viimeinen tässä tutkielmassa esiin tuotava big datan V-ulottuvuus on arvo. Arvo on big dataan liittyvässä kirjallisuudessa hyvin monimuotoisen ja laajan keskustelun aiheena. (Geerts & O’Leary, 2022.) Fan & Bifet (2013) ovat sitä mieltä, että big data luo arvoa, koska se antaa organisaatiolle kilpailuedun sen tuomien kyvykkyyksien takia. Heidän mukaansa big data auttaa yrityksiä vastaamaan kysymyksiin, joihin ennen ei kyetty vastaamaan. Elgendy & Elragal (2014) taas painottavat big datan luoman arvon perustuvan nopeampaan ja parempaan päätöksentekoon. Erevelles ym. (2016) korostavat, että tavoitteena on karsia epäolennaista dataa, jotta jäljelle jäävää, arvokasta tietoa voidaan hyödyntää tehokkaasti. Tämä näkökulma nostaa esiin sen, että suuri, monimuotoinen ja nopeasti syntyvä datamassa sisältää runsaasti myös käyttökeltotonta tietoa, minkä vuoksi tarvitaan osaamista ja järjestelmiä datan tehokkaaseen suodattamiseen ja puhdistamiseen. Näiden arvoon keskittyvien näkemysten ohella myös kustannusten tarkastelu on olennainen osa big datan tuottaman arvon kokonaisvaltaista arviointia. Suuren datamäärän käsittelykapasiteetin rakentaminen, erilaisten tietolähteiden hallinta, korkean nopeuden datavirtojen käsittely sekä laadukkaan datan tuottaminen voivat edellyttää merkittäviä investointeja ja johtaa kalliiden järjestelmien rakentamiseen. Big datan tuottamaa arvoa onkin tarkasteltava suhteessa niihin kustannuksiin, joita tarvittavat kyvykkyydet ja teknologiset ratkaisut edellyttävät. (Geerts & O’Leary, 2022.)

3.2 Big data -analytiikan menetelmät

BDA:ssa hyödynnetään erilaisia analyysimenetelmiä, jotka voidaan jakaa kuvailevaan, ennustavaan ja ohjaavaan analytiikkaan (LaValle ym., 2010). Kuvaileva analytiikka on yksinkertaisin muoto ja perustuu tilastollisiin menetelmiin, kuten keskiarvoihin, hajontoihin ja tapahtumien frekvensseihin. Sitä käytetään usein esikäsittelyvaiheessa piirteiden tunnistamiseen. Ennustava analytiikka

hyödyntää koneoppimismalleja – ohjattuja, ohjaamattomia ja puoliohjattuja – tulevien tapahtumien arviointiin. Ohjaava analytiikka puolestaan optimoi liiketoimintaprosesseja ennustemallien antaman palautteen perusteella, vaikka sen käyttöönotto voi olla haastavaa. Se auttaa kuitenkin ymmärtämään syy-seuraussuhteita ja sopeuttamaan prosesseja muuttuviin olosuhteisiin. (Rehman ym., 2016.)

BDA:ssa hyödynnetään tyypillisesti kolmea menetelmäryhmää: koneoppimista, tiedonlouhintaa ja tilastollisia menetelmiä. Nämä lähestymistavat tukevat erilaisten ilmiöiden tunnistamista, ennustamista ja selittämistä suurista tietomassoista. Koneoppimisen menetelmät perustuvat siihen, että järjestelmät oppivat datasta ja tekevät ennusteita tulevasta kehityksestä. Ne voivat hyödyntää valmiiksi nimettyjä aineistoja (ohjattu oppiminen) tai etsiä rakenteita ilman ennakkotietoa (ohjaamaton oppiminen). Lisäksi osa malleista mukautuu jatkuvasti uuden tiedon perusteella (puoliohjattu oppiminen). Tiedonlouhinta keskittyy datasta löytyvien piilevien rakenteiden ja yhteyksien tunnistamiseen. Tiedonlouhinnan menetelmiä käytetään esimerkiksi siihen, että voidaan luokitella havaintoja ryhmiin tai löytää usein yhdessä esiintyviä ilmiöitä, joita voidaan hyödyntää esimerkiksi asiakas- tai ostokäyttäytymisen analysoinnissa. Tilastolliset menetelmät tukevat analytiikkaa sekä kuvailemalla aineistoa että tekemällä yleistyksiä otoksesta koko populaatioon. Näitä käytetään muun muassa hajonnan, keskiarvojen ja tilastollisten erojen arviointiin. (Rehman ym., 2016.)

3.3 Big datan vaikutukset työnkuvaan ja liiketoimintaprosesseihin

BDA:n kehittyminen on muuttanut merkittävästi tapaa, jolla dataa käsitellään ja hyödynnetään organisaatioissa. Tiedon arvoketjun tarkastelu havainnollistaa muutoksen luonnetta erityisen hyvin. Tiedon arvoketju on syklinen sarja toimintoja, joiden avulla data muunnetaan ensin informaatioksi ja edelleen tiedoksi, jota hyödynnetään päätöksenteossa (Abbasi ym., 2016; Fayyad ym., 1996a, 1996b; Han ym., 2006). Perinteisessä tiedon arvoketjussa data tallennettiin rakenteisesti organisaation palvelimille, minkä jälkeen tietokanta-asiantuntijat ja ETL-kehittäjät jalostivat sitä analysoitavaan muotoon. Analytiikka pohjautui pitkälti relaatiotietokantoihin, ja analyysit tehtiin usein IT-asiantuntijoiden toimesta (Abbasi ym., 2016).

Big data on muuttanut tätä arvoketjua kolmella keskeisellä tavalla. Ensinnäkin organisaatioiden työnkuvat ja teknologiat ovat uudistuneet. Roolien, kuten data-analytikkojen, data-arkkitehtien ja datatieteilijöiden, merkitys on kasvanut (Abbasi ym., 2016). He toimivat liiketoiminnan ja teknologian rajapinnassa, jalostaen monimuotoisesta datasta käyttökelpoista tietoa (Davenport & Patil, 2012; McAfee ym., 2012).

Toiseksi, teknologiat ja prosessit ovat yhdentyneet. Tiedon jalostus tapahtuu yhä useammin automatisoiduissa putkissa, jotka mahdollistavat reaaliaikaisen, skaalautuvan ja integroidun analytiikan. Tässä kehityksessä NoSQL-teknologiat, kuten Hadoop ja Spark, sekä pilvipohjaiset ratkaisut (IaaS, PaaS, SaaS) ovat keskeisiä. Ne vastaavat sekä datan määrän että monimuotoisuuden haasteisiin, ja mahdollistavat joustavan tallennus- ja laskentakapasiteetin. (Abbasi ym., 2016.)

Kolmanneksi, päätöksenteko on siirtynyt kohti reaaliaikaisuutta ja itsepalveluperusteista analytiikkaa. Liiketoimintayksiköt voivat hyödyntää dataa ilman IT-osaston välikäsiä, mikä nopeuttaa ja ketteröittää päätöksentekoa. Tämä edellyttää osaamisen kasvattamista myös organisaation muilla tasoilla. (Abbasi ym., 2016.)

Big data on vaikuttanut myös organisaatiokulttuuriin. Organisaatiot, joissa analytiikalle on johdon tuki ja dataan perustuva päätöksenteko on vakiintunutta, ottavat analytiikan käyttöön tehokkaammin. Samalla kasvaa tarve huolelliselle tiedonhallinnalle ja tietoturvalle, erityisesti kun kyse on datan laadusta, ylläpidosta ja yksityisyydestä. (Abbasi ym., 2016.)

3.4 Big data -analytiikan haasteet

BDA:n hyödyntäminen tuo mukanaan monia haasteita. Näitä ovat muun muassa epäselvyydet siitä, mitä big data käsitteenä tarkoittaa (Hargittai, 2015), epävarmuus kerättävän datan valinnasta (Lazer ym., 2009), yksityisyyteen liittyvät huolenaiheet (Lazer ym., 2009) sekä eettiset kysymykset datan louhintaan ja käyttöön liittyen (Boyd & Crawford, 2012). Lisäksi teknologian nopea kehitys, osaamisen puute sekä infrastruktuurien kustannukset voivat muodostaa esteitä tehokkaalle hyödyntämiselle (Akerkar, 2013; Tole, 2013; Wang & Wiebe, 2016; Zicari, 2014).

Erevelles ym. (2016) korostavat, että saadakseen hyötyä big datasta, yritysten on investoitava paitsi teknologisiin ratkaisuihin, myös inhimillisiin ja organisatorisiin resursseihin. Pelkkä teknologia ei riitä: arvokkaan tiedon tunnistaminen edellyttää edelleen asiantuntijuutta, jota monista organisaatioista puuttuu (Akerkar, 2013; Zicari, 2014).

Big datan haasteet voidaan jäsentää kolmeen pääluokkaan datan elinkaaren mukaan: dataan, prosesseihin ja hallintaan liittyviin haasteisiin. Dataan liittyvät haasteet kumpuavat sen ominaisuuksista, kuten suuresta määrästä, monimuotoisuudesta, nopeudesta, laadusta ja löydettävyydestä (Jin ym., 2015). Prosessitason haasteet koskevat muun muassa datan keruuta, integrointia, muokkausta, analyysimallien valintaa ja tulosten esittämistä päätöksenteon tueksi. Hallinnolliset haasteet liittyvät yksityisyyteen, tietoturvaan, tiedon hallintaan ja eettisiin kysymyksiin, erityisesti henkilötietojen käsittelyssä. Näiden haasteiden tunnistaminen ja

ratkaiseminen on keskeistä BDA:n tehokkaalle ja vastuulliselle hyödyntämiselle liiketoiminnassa. (Akerkar, 2013; Zicari, 2014.)

BDA:n toteutus edellyttää monivaiheista datanhallintaprosessia, jonka haasteet liittyvät erityisesti datan keruuseen, siirtoon ja integrointiin. Datan keruu sisältää tietojen hankinnan, siirron ja esikäsittelyn (Chen & Lin, 2014), ja se voi tapahtua esimerkiksi sensoriteknologiaa hyödyntäen. Vaikka sensoripohjainen tiedonkeruu on yleistynyt, sen käyttöönottoa voivat rajoittaa korkeat asennus- ja ylläpitokustannukset (Cuzzocrea ym., 2011). Tähän haasteeseen on esitetty ratkaisuksi joukkoistettu tiedonkeruu, joka mahdollistaa uusien tietotyyppien keräämisen ja parantaa datan laatua ihmistoimijoiden osallistumisen kautta (Mondal ym., 2009).

Datan siirto suurina määrinä aiheuttaa teknisiä haasteita, kuten I/O-pullonkauloja (engl. input/output bottlenecks). I/O-pullonkauloilla tarkoitetaan yleensä tilanteita, joissa tiedonsiirto laitteiden välillä on liian hidasta suhteessa laskentatehoon. Näin ollen prosessori tai analytiikkaohjelmisto joutuu odottamaan, että data siirtyy muualta. Lisäksi haasteet, kuten verkon viive ja tehottomuus tiedon replikoinnissa ovat keskeisiä big datan aikakaudella (Chong & Shi, 2015).

Kolmas keskeinen tekninen haaste liittyy heterogeenisen ja rakenteettoman datan integrointiin eri lähteistä. Tämä edellyttää yhteisiä datamalleja ja alustaratkaisuja, jotka mahdollistavat tehokkaan tiedon yhdistämisen ja jatkoanalyysin. Esimerkiksi Hadoop-GIS tarjoaa tehokkaan ratkaisun paikkatietojen integrointiin MapReduce-kehikon avulla (Aji ym., 2013). Samalla sen jatkokehitys mahdollistaa tilastollisen paikkatietoanalyysin (Chen ym., 2014). Toisaalta reaaliaikaisen dataintegraation tarpeet vaativat joustavampia ratkaisuja kuin Hadoop tarjoaa. Tätä varten on kehitetty semanttisia järjestelmiä, jotka hyödyntävät automaattista skeemojen yhdistämistä ja semanttista päättelyä tapahtumapohjaisten palveluiden yhteydessä (Esposito ym., 2015).

3.5 Big data -analytiikan vaikutukset yrityksen suorituskykyyn

Aiemmissä luvuissa on tarkasteltu erikseen yrityksen suorituskykyä ja BDA:ta. Tässä luvussa tarkastellaan niiden välistä yhteyttä. Useat tutkijat pitävät BDA:ta merkittävänä muutostekijänä, joka voi parantaa yrityksen tehokkuutta ja vaikuttavuutta sekä operatiivisesti että strategisesti (Wamba ym., 2017). Esimerkiksi asiakasanalytiikan hyödyntämisen ja yrityksen suorituskyvyn välillä on havaittu positiivinen yhteys (Germann ym., 2014), ja BDA:n avulla voidaan tukea datalähtöistä strategiatyötä (Brands, 2014). BDA:sta onkin tullut yhä keskeisempi osa organisaatioiden päätöksentekoa (Hagel, 2015).

Vaikka kirjallisuudessa korostetaan usein BDA:n hyötyjä, kaikki yritykset eivät onnistu realisoimaan näitä hyötyjä käytännössä. Suorituskyvyn parantuminen ei ole automaattista, vaan siihen vaikuttavat muun muassa organisaation kyvykkyydet ja resurssit (Günther ym., 2017).

Useat esimerkit havainnollistavat BDA:n mahdollisuuksia: Amazonin tuotesuosittelun arvioidaan vaikuttavan jopa 35 % sen verkkokaupan myynnistä (Wills, 2014), ja General Electric on pyrkinyt parantamaan kaasuturbiininsa tehokkuutta ja huoltoa BDA:n avulla – tavoitteenaan jopa miljardiluokan säästöt pitkällä aikavälillä (Ward, 2014). Myös terveydenhuollossa, rahoitusallalla ja markkinoinnissa BDA:ta käytetään muun muassa kustannusten vähentämiseen ja päätöksenteon laadun parantamiseen (Liu, 2014).

Kiinnostus BDA:n hyödyntämisestä suorituskyvyn parantamisessa on kirjallisuuden mukaan laajaa eri aloilla ja myös tunnettujen suurten yritysten keskuudessa. Seuraavaksi tarkastelemme tätä yhteyttä hyödyntämällä luvussa 2 esille tuotuja suorituskyvyn teoreettisia viitekehyksiä. Ensiksi analysoimme BDA:n vaikutuksia BSC-mallin avulla, jonka jälkeen siirrymme resurssiperusteisen näkökulman tarjoamaan selitykseen yritysten välisistä eroista hyötyjen saavuttamisessa.

3.5.1 Big data -analytiikan vaikutus yrityksen suorituskykyyn Balanced Scorecardin neljästä näkökulmasta

BDA:ta voidaan tarkastella siis esimerkiksi BSC:n neljästä näkökulmasta, jotka ovat taloudellinen näkökulma, asiakasnäkökulma, sisäinen liiketoimintanäkökulma sekä innovaation ja oppimisen näkökulma. Taloudellisesta näkökulmasta BDA voi tuoda yritykselle merkittäviä hyötyjä. Yritykset, joilla on BDA-kyvykkyyksiä, pystyvät hallitsemaan ja hyödyntämään suuria määriä hinnoitteluun liittyvää dataa. Tämän ansiosta ne voivat optimoida hintojaan analysoimalla esimerkiksi asiakkaiden ostokäyttäytymistä, kysyntävaihtelua ja kilpailijoiden hinnoittelua. (Davenport, 2006.) Esimerkiksi Major League Baseball hyödyntää BDA:ta lipunhintojen dynaamiseen säätämiseen, ottaen huomioon muuttujia kuten sään, joukkueiden suoriutumisen ja sosiaalisen median aktiivisuuden. (Erevelles ym., 2016.)

BDA:n avulla yritykset voivat myös vaikuttaa kuluihinsa. Esimerkiksi terveydenhuollossa BDA:ta on hyödynnetty hukan ja petosten vähentämisessä. Tämä näkyy säästöinä kustannuksissa, mutta vaikuttaa myös luotettavuuteen ja imagoon eli positiiviset vaikutukset voivat olla suurempia kuin vain suorat säästöt hukan ja petosten vähentymisestä. (Srinivasan & Arunasalam, 2013.) Lisäksi taloudellisesta näkökulmasta tarkasteltuna BDA:lla on huomattu olevan vaikutuksia yrityksen kannattavuuteen. Akter ym. (2016) kertovat, että analytiikkakyvykkyyksien huomioiminen osana

yrityksen strategiaa vaikuttaa myönteisesti siihen, kuinka paljon kannattavuushyötyjä yritys kykenee saavuttamaan BDA:sta. Tämän selittää se, että strateginen painotus edistää analytiikan integroimista yrityksen arjen toimintamalleihin ja päätöksentekoprosesseihin, mikä puolestaan vahvistaa analytiikasta saatavia liiketoiminnallisia hyötyjä. Wamba ym. (2019) puolestaan laajentavat Akterin ym. (2016) havaintoja tarkastelemalla BDA:n laadun roolia. Heidän mukaansa analytiikan laadun huomioiminen strategisella tasolla vahvistaa analytiikan laadun ja yrityksen kannattavuuden välistä yhteyttä. Myös tässä tapauksessa selittävänä tekijänä on strateginen tuki: laadukkaasta BDA:sta on taloudellisesti merkittävämpi hyöty, jos sen hyödyntämiselle on luotu selkeä asema yrityksen strategiassa.

Kuten jo hinnoittelua käsittelevässä osiossa todettiin, BDA-kyvykkyksiä omaavat yritykset ymmärtävät asiakkaitaan paremmin kuin yritykset, joilla ei näitä kyvykkyksiä ole. Ne pystyvät analysoimaan asiakkaiden ostokäyttäytymistä erityisesti hinnoitteluun liittyen, mikä auttaa tunnistamaan, mitkä tekijät saavat asiakkaat ostamaan enemmän. Tämä syvällinen asiakasymmärrys edistää paitsi myynnin kasvattamista myös asiakkaan elinkaariarvon maksimointia. (Davenport, 2006.) BDA auttaa organisaatioita myös segmentoimaan enemmän kuin ennen. Yritykset voivat kehittää ja tarjota tuotteita ja palveluja hyvin tarkoille segmenteille, jotta asiakkaiden tarpeisiin voidaan vastata paremmin kuin ennen. Tämä vaatii toteutuakseen suuren määrän asiakasdataa, jonka hallinta puolestaan edellyttää yritykseltä BDA-kyvykkyksiä. (Manyika ym., 2011.)

BDA-kyvykkyudet voivat auttaa yrityksiä tehostamaan ja optimoimaan monia niiden sisäisistä toiminnoista (Ward, 2014). BDA-kyvykkyksiä omaavat yritykset pystyvät vaikuttamaan varastojen hallintaan. Myös ilman BDA:ta yritykset voivat havaita varastojen tyhjenemisen, mutta ne eivät kykene ennakoimaan kysynnän muutoksia yhtä tarkasti. Nämä kyvykkyudet vaativat usein kykyä kehittää ennustavia järjestelmiä, jotka perustuvat vahvaan datan hallintaan. BDA-kyvykkyudet voivat siis auttaa yritystä ylläpitämään matalampaa varaston tasoa, mikä taas vapauttaa resursseja muuhun käyttöön. (Davenport, 2006.) Yrityksen sisäisiä liiketoiminnan hyötyjä tuotiin esille jo aiemmin, kun tutkielmassa esiteltiin GE:n esimerkki. GE aikoi hyödyntää BDA:ta parantaakseen kaasuturbiinien tehokkuutta ohjelmistojen ja verkko-optimoinnin avulla. Big datan avulla GE aikoi myös tehostaa huollon aikataulutusta sekä kaasun ja sähkön järjestelmien koordinoitua. (Ward, 2014.)

Lisäksi yritykset voivat hyödyntää big dataa mainonnan tehokkuuden kehittämisessä. Esimerkiksi Sears Holdings tehosti mainonnan prosessejaan yhdistämällä omistamiensa brändien dataa yhteen paikkaan — Hadoop-klusteriin. Tämän ansiosta dataa voitiin sekä säilyttää kustannustehokkaasti

että analysoida tehokkaasti. Kun kaikki data oli keskitetysti yhdessä paikassa, Sears välttyi aikaa vievältä vaiheelta, jossa tietoa piti ensin kerätä useista eri lähteistä ja yhdistellä analysointia varten. Yrityksen tuolloisen teknologiajohtajan Phil Shelleyn mukaan laajojen mainoskampanjoiden tuottamiseen kulunut aika lyheni kahdeksasta viikosta yhteen viikkoon, ja prosessi nopeutui edelleen jatkuvasti. Sears onnistui parantamaan sisäisten prosessiensa tehokkuutta, minkä seurauksena asiakkaille voitiin tuottaa enemmän ja paremmin kohdennettua mainontaa. (McAfee ym., 2012.)

BDA-kyvykkyydet voivat vaikuttaa suorituskyykyyn myös innovaation näkökulmasta. BDA voi mahdollistaa yrityksille uusien tuotteiden ja palveluiden kehittämisen, olemassa olevien parantamisen sekä täysin uusien liiketoimintamallien luomisen. Esimerkiksi valmistajat hyödyntävät tuotteiden käytöstä kerättyä dataa seuraavan tuotesukupolven kehittämisessä ja jälkimarkkinapalveluiden innovoinnissa. (Manyika ym., 2011.) Ford Motors kerää dataa miljoonista ajoneuvoistaan ja hyödyntää sitä tunnistaakseen kehityskohteita reaaliaikaisesti. Esimerkiksi kuljettajien puhekomentojen analyysi johti mikrofonien sijoittelun parantamiseen ja automaattisen taustamelun vaimennuksen kehittämiseen. Näin Ford pystyi innovoimaan ilman perinteisten asiakastutkimusten viivettä. (Erevelles ym., 2016.) Lisäksi reaaliaikainen sijaintidata on mahdollistanut uusien paikkatietopohjaisten palveluiden syntymisen, aina navigoinnista vakuutusmaksujen hinnoitteluun ajotavan ja sijainnin perusteella. (Manyika ym., 2011.)

Organisaation oppimisen näkökulmasta yhteyttä BDA:han ei ole tutkittu ainakaan yhtä laajasti löytämässäni kirjallisuudessa kuin taloudellista näkökulmaa ja asiakasnäkökulmaa. Kirjallisuudesta voidaan kuitenkin tehdä muutamia huomioita. Garmaki ym. (2023) mukaan BDA-kyvykkyydet voivat vahvistaa organisaation oppimista muuttamalla ongelmanratkaisua, tiedonhakua ja päätöksentekoa ennakoivampaan ja kokeilevampaan suuntaan. Tämä "kysymällä oppimisen" lähestymistapa parantaa innovatiivisuutta ja tekee oppimisesta jatkuvan osan arjen työtä. BDA-kyvykkyys kehittää myös organisaation muistia ja tiedon saatavuutta, mikä tehostaa päätöksiä ja tukee yrityksen suorituskyykyä. Myös Cetindamar ym. (2021) painottavat BDA-kyvykkyyksien ja organisaation oppimisen välisen yhteyden tärkeyttä yrityksille. Heidän mukaansa BDA-kyvykkyydet tukevat organisaation oppimiskulttuurin muuntamista käytännön toimintakyvykkyyksiksi, mikä vahvistaa yrityksen dynaamisia kyvykkyyksiä ja strategista sopeutumiskykyä. BDA-kyvykkyydet toimivat siis siltana tiedon luomisen, levittämisen ja ketterän päätöksenteon välillä. Tämä mahdollistaa paremman valmistautumisen muutoksiin ja tukee pitkäaikaista suorituskyykyä.

Yhteyttä voidaan kuitenkin tarkastella myös perustuen luvussa 2 esille tuotuihin organisaation oppimisen sisältöihin. Organisaation oppiminen on prosessi, joka vaatii organisaation sisällä olevilta ihmisiltä eri muodoissa toteutettua vuorovaikutusta (Kofman & Senge, 1993). Aikaisemmin tutkielmassa esille tuodut roolit, kuten data-analyttikot ja datatieteilijät ovat siis keskiössä dataperusteisen kulttuurin ja tiedon levittämisessä organisaation sisällä. Heidän on vakuutettava muut organisaation jäsenet ja varsinkin johtoryhmä siitä, että BDA ja big data ovat tekijöitä, joiden avulla liiketoimintaa voidaan tehostaa. Toisaalta BDA-kyvykkyydet voivat myös vaikuttaa organisaation oppimiseen, koska dataa voidaan hyödyntää entistä suurempia määriä ja data voi olla hyvin monimuotoista. Organisaatio voi siis hyödyntää oppimisen prosessissa entistä enemmän dataa ja näin ollen vähentää päätöksissä intuition määrää.

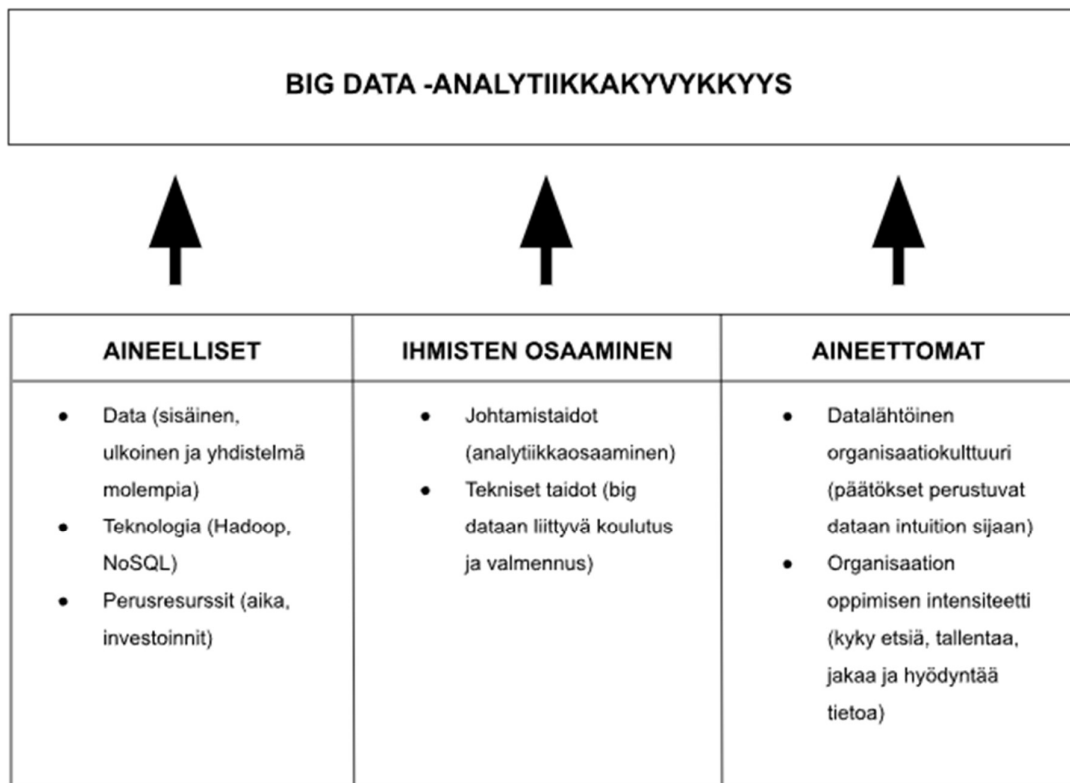
3.5.2 Big data -analytiikan vaikutus yrityksen suorituskykyyn Resurssiperusteisesta näkökulmasta

Resurssiperusteisesta näkökulmasta katsottuna BDA-kirjallisuudessa on löydetty positiivisia löydöksiä, mutta myös haasteita. Tähän mennessä on tehty useita tutkimuksia, jotka pyrkivät määrittelemään yritysten BDA-kyvykkyyden rakennuspalikoita — eli niitä resursseja, joiden varaan kyvykkyys rakentuu (Gupta & George, 2016; Kamioka & Tapanainen, 2014; McAfee ym., 2012; Mikalef ym., 2016; Wamba ym., 2017). Osa näistä tutkimuksista pohjautuu kuitenkin aikaisempaan tietojärjestelmätutkimukseen, kiinnittämättä juurikaan huomiota big data -kontekstin erityispiirteisiin (Mikalef ym., 2017).

McAfeen ym. (2012) mukaan big datan liiketoiminnallisen arvon konkretisoimiseksi on tärkeää ymmärtää sen kehittyvä ilmiö kokonaisvaltaisesti ja tunnistaa siihen liittyvät keskeiset tekijät. Nykyinen tutkimus on kuitenkin hajanaista, mikä vaikeuttaa big datan liiketoiminta-arvon arviointia. Esimerkiksi Kaisler ym. (2013) tunnistavat datan tallennuksen ja siirron tärkeiksi osa-alueiksi big datan arvon muodostumisessa. Seddon & Currie (2017) taas keskittyvät datan ominaisuuksiin liittyviin tekijöihin, kun taas Davenport & Patil (2012) korostavat big datan inhimillistä ulottuvuutta ja erityisesti datatieteilijän roolia. Nämä ovat kiistatta tärkeitä elementtejä, mutta organisaatioiden tulisi kiinnittää huomiota koko siihen resurssikirjoon, joka on tarpeen vaikeasti jäljiteltävän BDA-kyvykkyyden rakentamiseksi, sekä ymmärtää ne mekanismit ja olosuhteet, joiden kautta tämä kyvykkyys voi tuottaa liiketoiminnallista arvoa (Gupta & George, 2016).

Esille tuodut huolet tutkimusten keskittymisestä esimerkiksi geneerisesti tietojärjestelmä- tai IT-resursseihin, eikä niinkään big data -kontekstiin ovat varteenotettavia. Kirjallisuudesta löytyy

kuitenkin monia tutkimuksia, jotka käsittelevät BDA:ta ja yrityksen suorituskykyä joko suoraan tai epäsuorasti, mutta kuitenkin niin, että niitä voidaan hyödyntää BDA-kyvykkyyksien tarkastelussa. BDA-kyvykkyyksillä on kirjallisuuden mukaan monia selkeitä vaikutuksia yrityksen suorituskykyyn monesta eri näkökulmasta katsottuna. Yritykset eivät kuitenkaan saa näitä hyötyjä ilman, että niillä on tiettyjä resursseja, joiden avulla BDA:ta voidaan hyödyntää. Seuraava kuvaaja jakaa nämä resurssit kolmeen ryhmään:



Kuvio 3: Big data resurssien luokittelu (mukaillen Gupta ym., 2016)

Perustuen aiemmin mainittuun luokitteluun, aikaisemmat tutkimukset ovat korostaneet tiettyjä BDA:n osa-alueita, jotka ovat kriittisiä yrityksille. Konkreettisten resurssien osalta data, teknologia ja muut perusresurssit on tunnustettu big datan menestyksen kannalta keskeisiksi. (Mikalef ym., 2020.) Big datan määrittäviä piirteitä ovat sen määrä, monimuotoisuus ja nopeus (Chen & Zhang, 2014). Tästä huolimatta kirjallisuudessa mainitaan usein, että IT-strategit ja data-analyttikot kiinnittävät erityistä huomiota analysoimansa datan laatuun ja saatavuuteen (Brinkhues ym., 2014). Vaikka data itsessään on ydinresurssi, on yrityksille myös tärkeää, että niillä on käytössään infrastruktuuri, joka mahdollistaa datan tallentamisen, jakamisen ja analysoinnin. (Mikalef et al., 2020.) Big data edellyttää uudenlaisia teknologioita, jotka kykenevät käsittelemään suuria määriä monimuotoista ja nopeasti muuttuvaa dataa (Gupta & George, 2016). Yksi tällaisen datan keskeisistä ominaisuuksista on sen jäsentymätön muoto, mikä vaatii kehittyneitä infrastruktuuri-

investointeja, jotta datasta voidaan tuottaa merkityksellistä ja arvokasta tietoa (Ji-fan Ren ym., 2017). Perusresursseihin, kuten taloudelliseen tukeen, liittyen kirjallisuus korostaa, että big data -investoinnit vaativat usein aikaa ennen kuin ne tuottavat mitattavaa liiketoiminta-arvoa (Mikalef ym., 2017).

Ihmisten osaamisen näkökulmasta kirjallisuus tunnistaa, että big data -investoinneista saatavan arvon maksimoimiseksi tarvitaan sekä teknisiä että johtamiseen liittyviä taitoja (Pappas ym., 2018; Wamba ym., 2017). Kuten aiemmin tässä tutkielmassa jo mainittiin, kirjallisuudessa tuodaan esille myös datatieteilijän nousevan roolin merkitys big datan kontekstissa. Yksi datatieteen keskeisimmistä osa-alueista on analysoivan ajattelutavan hallitseminen, ja tämä kyvykkyys ei ole tärkeä ainoastaan datatieteilijöille, vaan koko organisaatiossa. Tämä korostuu erityisesti esihenkilö- ja johtotehtävissä toimivilla työntekijöillä. (Prescott, 2014.)

Lopuksi, aineettomien resurssien osalta kirjallisuus korostaa, että datalähtöinen kulttuuri ja organisatorinen oppiminen ovat keskeisiä tekijöitä big data -hankkeiden tehokkaassa käyttöönotossa (Mikalef & Krogstie, 2018; Vidgen ym., 2017). Yrityksissä, jotka osallistuvat big data -projekteihin, datalähtöinen kulttuuri on havaittu tärkeäksi tekijäksi niiden kokonaismenestyksen ja jatkuvuuden kannalta (LaValle ym., 2010). Big dataan liittyvän teknologian jatkuvan kehittymisen vuoksi on tärkeää, että big dataan investoivissa yrityksissä vallitsee jatkuvan oppimisen kulttuuri (Vidgen ym., 2017).

Vaikka big data -teknologiat ovat merkittävä osa liiketoimintakeskustelua, yritysten on tärkeää kiinnittää huomiota myös muihin resursseihin, jotka mahdollistavat ainutlaatuisen BDA-kyvykkyuden. Janssen ym. (2017) korostavat, että big datan tuottamaan tietoon perustuvien päätösten laatu riippuu paitsi syötteiden laadusta myös siitä, kuinka tehokkaasti data muunnetaan päätöksenteossa hyödynnettävään muotoon. Heidän mukaansa päätöksenteon laatu on vahvasti sidoksissa yrityksen kokonaisvaltaiseen BDA-kyvykkyyteen, johon kuuluvat henkilöstön osaaminen ja tietämys, yhteistyö- ja tiedonvaihtoprosessit, infrastruktuurin ja datan saatavuus sekä vakiintuneet tiedonkeruun ja -käsittelyn menetelmät. McAfee ym. (2012) painottavat datalähtöisen päätöksenteon kulttuurin merkitystä, jossa päätökset perustuvat analyysistä saatavaan tietoon, eivät pelkkään intuitioon. Vidgen ym. (2017) puolestaan huomauttavat, että datalähtöisyys ei ole pelkästään tekninen kysymys, vaan edellyttää organisaation analytiikkatoimintojen systemaattista järjestämistä sekä analytiikkakyvykkyuden ja liiketoimintastrategian yhteensovittamista. Näin ollen BDA-kyvykkyuden käsite laajentaa big datan tarkastelun yksittäisistä teknologioista

kokonaisvaltaiseen näkemykseen, jossa keskeisessä roolissa ovat kaikki ne organisatoriset resurssit, joiden avulla big datan strateginen potentiaali voidaan täysimääräisesti hyödyntää.

Näiden kaikkien osa-alueiden vaikutus yksinään ei kuitenkaan riitä rakentamaan BDA-kyvykkyyttä. Gupta & George (2016) kertovat, että yritysten on kehitettävä BDA-kyvykkyys yhdistämällä tietyt aineelliset, inhimilliset ja aineettomat resurssit. Aiemmat tutkimukset ovat korostaneet, että aineelliset resurssit ovat usein markkinoilta hankittavissa, ja siksi niiden roolia on pidetty vähemmän keskeisenä verrattuna inhimillisiin ja aineettomiin resursseihin organisaatiokyvykkyyksien rakentamisessa. Vaikka Gupta & George (2016) jakavat näkemyksen siitä, etteivät big datalle ominaiset aineelliset resurssit yksin riitä BDA-kyvykkyuden saavuttamiseen. Heidän mukaansa sama pätee myös inhimillisiin ja aineettomiin big data -resursseihin. Nimenomaan aineellisten, inhimillisten ja aineettomien resurssien ainutlaatuinen yhdistelmä mahdollistaa yrityskohtaisen analytiikkakyvykkyuden muodostumisen. Yrityksellä saattaa olla käytössään esimerkiksi laajoja datavarantoja ja tehokkaita laskentajärjestelmiä, mutta ilman riittävää johtamisosaamista ja teknistä analytiikkakompetenssia näiden resurssien tuottama potentiaalinen arvo jää hyödyntämättä. (Gupta & George, 2016.)

Kirjallisuudessa tunnistetaan suuri määrä eri resursseja ja niiden muodostamia kokonaisuuksia, jotka yhdessä muodostavat BDA-kyvykkyuden. Kirjallisuudessa ei kuitenkaan juurikaan tarkastella aihetta pysyvän kilpailuedun saavuttamisen näkökulmasta. Tämä vaikuttaa loogiselta seuraavalta keskittymisalueelta tutkimukselle, jossa BDA:ta tarkastellaan resurssiperusteisesta näkökulmasta.

4 Johtopäätökset ja yhteenveto

Tämän kandidaatintutkielman tavoitteena oli tarkastella, millaisia vaikutuksia BDA:lla voi olla yrityksen suorituskykyyn. Tutkimus toteutettiin kirjallisuuskatsauksena, ja päätutkimuskysymystä täsmennettiin kahdella apututkimuskysymyksellä: miten akateemisessa kirjallisuudessa on tutkittu yrityksen suorituskykyä, ja miten BDA:n roolia liiketoiminnassa on tutkittu.

Ensimmäisenä tarkasteltiin yrityksen suorituskykyä. Kirjallisuus osoitti, että yrityksen suorituskykyä voidaan lähestyä useista näkökulmista – perinteisesti taloudellisista mittareista, mutta myös asiakkaisiin, sisäisiin prosesseihin, innovointiin ja organisaation oppimiseen liittyvinä ulottuvuuksina. Tutkielmassa hyödynnettiin Balanced Scorecard -mallia (Kaplan & Norton, 1992) sekä resurssiperusteista näkökulmaa (Barney, 1991), joiden kautta yrityksen suorituskyvyn moniulotteisuus ja kilpailukyvyn lähteet tulivat esiin.

Tämän jälkeen käsiteltiin Big datan ja analytiikan erityispiirteitä. Big datan viisi V-ulottuvuutta erottavat sen perinteisestä datasta ja avaavat mahdollisuuksia reaaliaikaiseen, skaalautuvaan ja itsenäisesti hyödynnettävään päätöksentekoon. BDA:n analytiikkamenetelmät – kuvaileva, ennustava ja ohjaileva analytiikka – mahdollistavat syvällisen tiedon tuottamisen, joka tukee strategista ja operatiivista päätöksentekoa.

BDA:n vaikutuksia yrityksen suorituskykyyn tarkasteltiin viitekehyksien avulla. Analytiikan avulla voidaan tehostaa resurssien käyttöä, parantaa prosessien tehokkuutta ja päätöksenteon laatua sekä tukea asiakaslähtöistä kehitystä. Lisäksi BDA mahdollistaa innovaatioiden nopeamman kehittämisen, kun asiakas- ja markkinatiedon analysointi paljastaa muutostrendejä varhain. Esimerkiksi Ford Motorin tapa hyödyntää ajoneuvojen keräämää dataa tuotekehityksessä havainnollistaa BDA:n käytännön vaikutusta innovaatiotoimintaan.

Vaikka BDA:n potentiaali on laaja, sen vaikutukset eivät ole automaattisia. Useat lähteet korostavat, että vaikutusten realisoituminen riippuu yrityksen kyvykkyyksistä: osaamisesta, teknologiasta, prosesseista sekä organisaatiokulttuurista. Jos BDA jää irralliseksi työkaluksi ilman strategista johtamista, vaikutukset voivat jäädä vähäisiksi tai kohdentua väärin.

Lisäksi BDA:n käyttöönottoon liittyy haasteita, kuten datan laatu, tietoturva, osaavan henkilöstön saatavuus ja kustannukset. Yritysten on tärkeää nähdä BDA osana laajempaa digitaalista muutosta, joka vaatii uudenlaista ajattelua, rakennetta ja johtajuutta.

Tutkielma osoittaa, että BDA:lla on huomattava potentiaali parantaa yrityksen suorituskykyä, mutta sen vaikutukset eivät ole yksiselitteisiä. BDA voi tukea tehokkuutta, asiakasymmärrystä, innovaatiota ja oppimista, mutta vaikutusten toteutuminen edellyttää, että analytiikka kytetään yrityksen strategiseen toimintaan ja kulttuuriin. Organisaation on rakennettava kyvykkyyksiä, jotka mahdollistavat tiedon tehokkaan hyödyntämisen kaikilla liiketoiminnan tasoilla.

Tutkielma tarjoaa ajankohtaisen katsauksen siihen, miten BDA:ta on tutkittu ja hyödynnetty liiketoiminnassa, ja millaisilla edellytyksillä se voi tukea yrityksen suorituskykyä. Aiempaan kirjallisuuteen pohjautuva tarkastelu korostaa, että BDA ei ole vain teknologinen väline, vaan osa laajempaa organisaation kehityspolkua, joka vaatii jatkuvaa oppimista ja sopeutumista muuttuvaan toimintaympäristöön.

Lähteet

- Abbasi, A., Sarker, S., & Chiang, R. H. (2016). Big data research in information systems: Toward an inclusive research agenda. *Journal of the Association for Information Systems*, *17*(2), 3.
- Aji, A., Sun, X., Vo, H., Liu, Q., Lee, R., Zhang, X., Saltz, J., & Wang, F. (2013). *Demonstration of Hadoop-GIS: a spatial data warehousing system over MapReduce*. 528–531.
- Akerkar, R. (2013). *Big data computing*. Crc Press.
- Akter, S., Wamba, S. F., Gunasekaran, A., Dubey, R., & Childe, S. J. (2016). How to improve firm performance using big data analytics capability and business strategy alignment? *International Journal of Production Economics*, *182*, 113–131. <https://doi.org/10.1016/j.ijpe.2016.08.018>
- Barney, J. (1991). Firm Resources and Sustained Competitive Advantage. *Journal of Management*, *17*(1), 99–120. <https://doi.org/10.1177/014920639101700108>
- Boyd, D., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society*, *15*(5), 662–679.
- Brands, K. (2014). Big data and business intelligence for management accountants. *Strategic Finance (Montvale, N.J.)*, *95*(12), 64.
- Brinkhues, R., Maçada, A. C. G., & Casalinho, G. (2014). *Information Management Capabilities: Antecedents and Consequences*. AMCIS.
- Cetindamar, D., Katic, M., Burdon, S., & Gonsel, A. (2021). The Interplay among Organisational Learning Culture, Agility, Growth, and Big Data Capabilities. *Sustainability*, *13*(23). <https://doi.org/10.3390/su132313024>
- Chen, C. P., & Zhang, C.-Y. (2014). Data-intensive applications, challenges, techniques and technologies: A survey on Big Data. *Information Sciences*, *275*, 314–347.
- Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business Intelligence and Analytics: From Big Data to Big Impact. *MIS Quarterly*, *36*(4), 1165–1188. <https://doi.org/10.2307/41703503>
- Chen, M., Mao, S., & Liu, Y. (2014). Big Data: A Survey. *Mobile Networks and Applications*, *19*(2), 171–209. <https://doi.org/10.1007/s11036-013-0489-0>

- Chen, X., Vo, H., Aji, A., & Wang, F. (2014). *High performance integrated spatial big data analytics*. 11–14.
- Chen, X.-W., & Lin, X. (2014). Big data deep learning: Challenges and perspectives. *IEEE Access*, 2, 514–525.
- Chong, D., & Shi, H. (2015). Big data analytics: A literature review. *Journal of Management Analytics*, 2(3), 175–201. <https://doi.org/10.1080/23270012.2015.1082449>
- Cuzzocrea, A., Song, I.-Y., & Davis, K. C. (2011). *Analytics over large-scale multidimensional data: The big data revolution!* 101–104.
- Daft, R. L. (1989). *Organization theory and design* (3rd rd.). West.
- Davenport, T. H. (2006). Competing on analytics. *Harvard Business Review*, 84(1), 98.
- Davenport, T. H., & Patil, D. (2012). Data scientist. *Harvard Business Review*, 90(5), 70–76.
- E. Prescott, M. (2014). Big data and competitive advantage at Nielsen. *Management Decision*, 52(3), 573–601.
- Elgendy, N., & Elragal, A. (2014). Big Data Analytics: A Literature Review Paper. In P. Perner (Ed.), *Advances in Data Mining. Applications and Theoretical Aspects* (pp. 214–227). Springer International Publishing.
- Erevelles, S., Fukawa, N., & Swayne, L. (2016). Big Data consumer analytics and the transformation of marketing. *Journal of Business Research*, 69(2), 897–904. <https://doi.org/10.1016/j.jbusres.2015.07.001>
- Esposito, C., Ficco, M., Palmieri, F., & Castiglione, A. (2015). A knowledge-based platform for big data analytics based on publish/subscribe services and stream processing. *Knowledge-Based Systems*, 79, 3–17.
- Fan, W., & Bifet, A. (2013). Mining big data: Current status, and forecast to the future. *ACM SIGKDD Explorations Newsletter*, 14(2), 1–5.
- Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996a). From data mining to knowledge discovery in databases. *AI Magazine*, 17(3), 37–37.
- Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996b). The KDD process for extracting useful knowledge from volumes of data. *Communications of the ACM*, 39(11), 27–34.

- Frawley, W. J., Piatetsky-Shapiro, G., & Matheus, C. J. (1992). Knowledge discovery in databases: An overview. *AI Magazine*, 13(3), 57–57.
- Garmaki, M., Gharib, R. K., & Boughzala, I. (2023). Big data analytics capability and contribution to firm performance: The mediating effect of organizational learning on firm performance. *Journal of Enterprise Information Management*, 36(5), 1161–1184.
- Gavrea, C., Ilies, L., & Stegorean, R. (2011). Determinants of organizational performance: The case of Romania. *Management & Marketing*, 6(2).
- Geerts, G. L., & O’Leary, D. E. (2022). V-Matrix: A wave theory of value creation for big data. *International Journal of Accounting Information Systems*, 47, 100575.
- Georgopoulos, B. S., & Tannenbaum, A. S. (1957). A study of organizational effectiveness. *American Sociological Review*, 22(5), 534–540.
- Germann, F., Lilien, G. L., Fiedler, L., & Kraus, M. (2014). Do retailers benefit from deploying customer analytics? *Journal of Retailing*, 90(4), 587–593.
- Günther, W. A., Mehrizi, M. H. R., Huysman, M., & Feldberg, F. (2017). Debating big data: A literature review on realizing value from big data. *The Journal of Strategic Information Systems*, 26(3), 191–209.
- Gupta, M., & George, J. F. (2016). Toward the development of a big data analytics capability. *Big Data Commerce*, 53(8), 1049–1064. <https://doi.org/10.1016/j.im.2016.07.004>
- Hagel, J. (2015). Bringing analytics to life. *Journal of Accountancy*, 219(2), 24–26.
- Han, J., Kamber, M., & Mining, D. (2006). Concepts and techniques. *Morgan Kaufmann*, 340(1), 94104–103205.
- Hargittai, E. (2015). Is bigger always better? Potential biases of big data derived from social network sites. *The ANNALS of the American Academy of Political and Social Science*, 659(1), 63–76.
- Janssen, M., Van Der Voort, H., & Wahyudi, A. (2017). Factors influencing big data decision-making quality. *Journal of Business Research*, 70, 338–345.

- Ji-fan Ren, S., Fosso Wamba, S., Akter, S., Dubey, R., & Childe, S. J. (2017). Modelling quality dynamics, business value and firm performance in a big data analytics environment. *International Journal of Production Research*, 55(17), 5011–5026.
- Jin, X., Wah, B. W., Cheng, X., & Wang, Y. (2015). Significance and challenges of big data research. *Big Data Research*, 2(2), 59–64.
- Kaisler, S., Armour, F., Espinosa, J. A., & Money, W. (2013). *Big data: Issues and challenges moving forward*. 995–1004.
- Kamioka, T., & Tapanainen, T. (2014). Organizational Use of Big Data and Competitive Advantage-Exploration of Antecedents. *PACIS, 2014*, 18th.
- Kaplan, R., & Norton, D. (1992). The balanced scorecard—Measures that drive performance. *Harvard Business Review*, 70(1), 71–79.
- Kofman, F., & Senge, P. M. (1993). Communities of commitment: The heart of learning organizations. *Organizational Dynamics*, 22(2), 5–23.
- Kwon, O., Lee, N., & Shin, B. (2014). Data quality management, data usage experience and acquisition intention of big data analytics. *International Journal of Information Management*, 34(3), 387–394.
- LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S., & Kruschwitz, N. (2010). Big data, analytics and the path from insights to value. *MIT Sloan Management Review*.
- Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabási, A.-L., Brewer, D., Christakis, N., Contractor, N., Fowler, J., & Gutmann, M. (2009). Computational social science. *Science*, 323(5915), 721–723.
- Learned, E. P., Christensen, C. R., Andrews, K. R., & Guth, W. D. (1969). *Business policy: Text and cases*. Irwin.
- Lebas, M., & Euske, K. (2002). A conceptual and operational delineation of performance. *Business Performance Measurement: Theory and Practice*, 65, 79.
- Lusthaus, C., & Adrien, M.-H. (1998). Organizational assessment: A review of experience. *Universalia Occasional Paper*, 31(31), 1–16.

- Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A. H. (2011). *Big data: The next frontier for innovation, competition and productivity*. Technical report, McKinsey Global Institute.
- McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D., & Barton, D. (2012). Big data: The management revolution. *Harvard Business Review*, *90*(10), 60–68.
- Mikalef, P., Framnes, V. A., Danielsen, F., Krogstie, J., & Olsen, D. (2017). *Big data analytics capability: Antecedents and business value*.
- Mikalef, P., Ilias, P. O., Giannakos, M., Krogstie, J., & Lekakos, G. (2016). *Big data and strategy: A research framework*.
- Mikalef, P., & Krogstie, J. (2018). *Big Data Governance and Dynamic Capabilities: The Moderating effect of Environmental Uncertainty*. 206.
- Mikalef, P., Krogstie, J., Pappas, I. O., & Pavlou, P. (2020). Exploring the relationship between big data analytics capability and competitive performance: The mediating roles of dynamic and operational capabilities. *Information and Management*, *57*(2). Scopus.
<https://doi.org/10.1016/j.im.2019.05.004>
- Mondal, A., Madria, S. K., & Kitsuregawa, M. (2009). An economic incentive model for encouraging peer collaboration in mobile-P2P networks with support for constraint queries. *Peer-to-Peer Networking and Applications*, *2*, 230–251.
- Pappas, I. O., Mikalef, P., Giannakos, M. N., Krogstie, J., & Lekakos, G. (2018). Big data and business analytics ecosystems: Paving the way towards digital transformation and sustainable societies. *Information Systems and E-Business Management*, *16*(3), 479–491.
- Porter, M. E. (1981). The contributions of industrial organization to strategic management. *Academy of Management Review*, *6*(4), 609–620.
- Rehman, M. H., Chang, V., Batool, A., & Wah, T. Y. (2016). Big data reduction framework for value creation in sustainable enterprises. *International Journal of Information Management*, *36*(6), 917–928.

- Reimer, A. P., & Madigan, E. A. (2019). Veracity in big data: How good is good enough. *Health Informatics Journal*, 25(4), 1290–1298.
- Reinsel, D., Gantz, J., & Rydning, J. (2018). *The Digitization of the World: From Edge to Core* [White Paper]. International Data Corporation (IDC). <https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf>
- Seashore, S. E., & Yuchtman, E. (1967). Factorial analysis of organizational performance. *Administrative Science Quarterly*, 377–395.
- Seddon, J. J., & Currie, W. L. (2017). A model for unpacking big data analytics in high-frequency trading. *Journal of Business Research*, 70, 300–307.
- Srinivasan, U., & Arunasalam, B. (2013). Leveraging big data analytics to reduce healthcare costs. *IT Professional*, 15(6), 21–28.
- Taouab, O., & Issor, Z. (2019). Firm performance: Definition and measurement models. *European Scientific Journal*, 15(1), 93–106.
- Tole, A. A. (2013). Big data challenges. *Database Systems Journal*, 4(3).
- Tsai, C.-W., Lai, C.-F., Chao, H.-C., & Vasilakos, A. V. (2015). Big data analytics: A survey. *Journal of Big Data*, 2(1), 21. <https://doi.org/10.1186/s40537-015-0030-3>
- Venkatraman, N., & Ramanujam, V. (1986). Measurement of business performance in strategy research: A comparison of approaches. *Academy of Management Review*, 11(4), 801–814.
- Vidgen, R., Shaw, S., & Grant, D. B. (2017). Management challenges in creating value from business analytics. *European Journal of Operational Research*, 261(2), 626–639.
- Wamba, S. F., Akter, S., & de Bourmont, M. (2019). Quality dominant logic in big data analytics and firm performance. *Business Process Management Journal*, 25(3), 512–532.
- Wamba, S. F., Akter, S., Edwards, A., Chopin, G., & Gnanzou, D. (2015). How “big data” can make big impact: Findings from a systematic review and a longitudinal case study. *International Journal of Production Economics*, 165, 234–246. Scopus. <https://doi.org/10.1016/j.ijpe.2014.12.031>

- Wamba, S. F., Gunasekaran, A., Akter, S., Ren, S. J., Dubey, R., & Childe, S. J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research*, 70, 356–365.
<https://doi.org/10.1016/j.jbusres.2016.08.009>
- Wang, Y., & Wiebe, V. J. (2016). Big Data Analytics on the characteristic equilibrium of collective opinions in social networks. In *Big data: Concepts, methodologies, tools, and applications* (pp. 1403–1420). IGI Global.
- Ward, D. G. (2014). A guide to the strategic use of big data. *Information Management Journal*, 48(6), 45–48.
- Wills, M. J. (2014). Decisions through data: Analytics in healthcare. *Journal of Healthcare Management*, 59(4), 254–262.
- Ying Liu. (2014). Big Data and Predictive Business Analytics. *Journal of Business Forecasting*, 33(4), 40–42.
Business Source Ultimate.
- Zicari, R. V. (2014). Big data: Challenges and opportunities. *Big Data Computing*, 564, 103.
- Zikopoulos, P., & Eaton, C. (2011). *Understanding big data: Analytics for enterprise class hadoop and streaming data*.