



**TURUN
YLIOPISTO**
Kauppakorkeakoulu

Massadata-analytiikan käyttöönoton ajurit ja haasteet pk-yrityksissä

Tietojärjestelmätieteen
kandidaatintutkielma

Laatija:

Max Montin

Ohjaaja:

FT Kai Kimppa

8.12.2023

Turku

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

Kandidutkielma

Oppiaine: Tietojärjestelmätiede

Tekijä: Max Montin

Otsikko: Massadata-analytiikan käyttöönoton ajurit ja haasteet pk-yrityksissä

Ohjaaja: FT Kai Kimppa

Sivumäärä: 34 sivua

Päivämäärä: 8.12.2023

Massadata-analytiikasta on muodostunut uusi voimavara, jota yritykset haluavat hyödyntää liiketoiminnassaan. Massadata-analytiikan ansiosta yrityksissä voidaan tehdä havaintoja, jotka olisivat muuten saattaneet jäädä löytämättä. Päätöksenteon tuennan lisäksi massadata-analytiikka voi mahdollistaa muun muassa paremman asiakastuntemuksen ja lisääntyneen innovoinnin. Merkittävä osuus massadata-analytiikkaa tutkivasta kirjallisuudesta on kuitenkin melko optimistista, eikä ilmiötä ole välttämättä tarkasteltu riittävän kriittisesti. Massadata-analytiikka ei sellaisenaan takaa kilpailuetua, sillä massadatan ominaisuudet voivat muodostaa haasteita samalla tavalla kuin mahdollisuuksia.

Myös pk-yritykset ovat kiinnostuneita massadata-analytiikan hyödyntämisestä, mutta ne voivat kohdata käyttöönotossa erilaisia haasteita. Käyttöönoton haasteiden lisäksi pk-yritysten on varmistuttava siitä, että massadata-analytiikasta on ylipäättään hyötyä, järjestelmä omaksutaan ja käyttöönotto on kannattavaa. Tässä tutkielmassa selvitetään, mitkä teknologiset, organisatoriset ja ympäristölliset tekijät vaikuttavat massadata-analytiikan käyttöönottoon pk-yrityksissä. Lisäksi tarkastellaan, mitä haasteita massadata ja siihen sovellettu analytiikka voivat sellaisinaan aiheuttaa. Tutkielmassa esitetään, miten pk-yritykset voivat hyötyä massadata-analytiikasta.

Massadata-analytiikan käyttöönoton ajureiksi pk-yrityksissä tunnistettiin suhteellinen hyöty, kokeiltavuus ja tarkasteltavuus, havaitut hyödyt, ylimmän johdon tuki, organisatorinen ja taloudellinen valmius, kauppiaiden ulkopuolinen tuki sekä paikallinen sääntely. Käyttöönoton haasteiksi luokiteltiin monimutkaisuus, epävarmuus ja turvattomuus, järjestelmän omaksuminen sekä ohjelmisto- ja työmarkkinoiden ongelmat. Ylimmän johdon tuki osoittautui tärkeimmäksi käyttöönottoon vaikuttavaksi tekijäksi, ja siksi jotkin ajurit ja mahdollisuudet ovatkin riippuvaisia johdon tuesta.

Kannattavuuden arvioimiseksi pk-yritysten kannattaisi verrata havaittuja tekijöitä omiin olosuhteisiinsa ja pohtia, missä on eniten kehitettävää ja mitkä ominaisuudet ovat valmiiksi positiivisia tekijöitä. Pk-yritysten pitäisi lähestyä massadata-analytiikkaa ratkaisuna tunnistettuun ongelmaan, ei taattuna kilpailuetuna. Käyttöönottoon kannustamiseksi pk-yritysten pitäisi vähentää negatiivisia tekijöitä. Epävarmuutta ja turvattomuutta voidaan vähentää tarkasteltavuudella ja kokeiltavuudella, kun taas johdon tukea voidaan vahvistaa lisäämällä tietoisuutta saavutettavista hyödyistä. Yrityksen toimintaympäristön vakaudella voi myös olla vaikutusta siihen, minkä kategorian tekijät vaikuttavat käyttöönottoon muita enemmän.

Avainsanat: massadata, data-analytiikka, pk-yritykset

SISÄLLYS

1	Johdanto	7
2	Massadata yrityksissä	9
2.1	Massadata	9
2.2	Massadata-analytiikka	11
3	Massadata-analytiikan käyttöönotto pk-yrityksissä	16
3.1	Massadata-analytiikan käyttöönoton ajurit pk-yrityksissä	17
3.1.1	Teknologiset ajurit	17
3.1.2	Organisatoriset ajurit	19
3.1.3	Ympäristölliset ajurit	21
3.2	Massadata-analytiikan käyttöönoton haasteet pk-yrityksissä	22
3.2.1	Teknologiset haasteet	22
3.2.2	Organisatoriset haasteet	23
3.2.3	Ympäristölliset haasteet	25
4	Yhteenveto ja johtopäätökset	26
	Lähteet	30

TAULUKOT

Taulukko 1. Massadata-analytiikan käyttöönottoon vaikuttavat tekijät

27

1 Johdanto

Massadatailmiön (engl. big data) ansiosta yrityksillä on käytettävissään suurempia tietomääriä kuin koskaan aikaisemmin. Ennen massadatan aikakautta tiedon hankkiminen saatettiin kokea monessa yrityksessä työläänä ja hitaana prosessina, mutta suhteellisen lyhyessä ajassa on siirrytty tilanteeseen, jossa dataa voi olla enemmän kuin sitä ehditään käsittelemään. Yritykset saavat lähes taukoamatta dataa muun muassa asiakkaistaan, työntekijöistään ja erilaisista suorituskyvyn mittareista. Massadatan hankkiminen, käsittely ja analysointi ovat mahdollistaneet eri toimialojen yrityksille kilpailuetuja, jotka olisivat voineet jäädä aikaisemmin huomaamatta. (Müller ym., 2018; Tsai ym., 2015.) Massadata-analytiikan havaitut mahdollisuudet ovat herättäneet myös yritysjohtajien mielenkiinnon (Justy ym., 2023).

Edellä mainittujen datamäärien saatavuus ei kuitenkaan ole kaikille yrityksille taattua, eikä massadata-analytiikan käyttöönotto ole täysin vaivatonta. Esimerkiksi pienet tai keskisuuret yritykset eli pk-yritykset saattavat kärsiä resurssien ja ymmärryksen puutteesta sekä muutosvastarinnasta. Massadatalle on ominaista, että sen käsittelyyn ei voida käyttää perinteisiä liiketoiminnan hallinnan työkaluja. Sen analysoimiseksi tarvitaan järjestelmiä ja henkilöstöä, joita pk-yrityksissä ei välttämättä ole entuudestaan. (Ks. esim. Lutfi ym., 2022; Maroufkhani ym., 2020.) Käyttöönoton haasteet ja mahdollisuudet olisi siis hyvä huomioida ajoissa, jotta pk-yrityksissä voitaisiin varmistua toivottujen hyötyjen saavuttamisesta ja minimoida käyttöönoton riskejä.

Massadata-analytiikka ei myöskään takaa yrityksille välitöntä kilpailuetua, vaan sen tapauskohtaista käytettävyyttä ja hyödyllisyyttä on myös syytä harkita. Vaikka massadata-analytiikka tarjoaisikin ratkaisun johonkin ongelmaan, on mahdollista, että samaa ongelmaa varten on aikaisemmin kehitetty muita menetelmiä, jotka antaisivat saman vastauksen nopeammin ja edullisemmin. Massadata sellaisenaan ei tue yritysten toimintaa, vaan sen täytyy myös olla oikeanlaista, riittävän tarkkaa ja tarpeeksi pitkälle prosessoitua. (Fan ym., 2014; Jones, 2019.) Käyttöönoton haasteiden lisäksi yritysten kannattaisi siis arvioida saavutettavien hyötyjen käytettävyyttä, merkittävyyttä ja paikkansapitävyyttä.

Tässä kandidaatkielmassa tutustutaan massadata-analytiikan käyttöönoton tuomiin haasteisiin ja mahdollisuuksiin pk-yrityksissä sekä niiden merkitykseen käyttöönottoa

harkittaessa. Tutkielmassa selvitetään myös, mitä haasteita massadata-analytiikan hyödyntämisessä kohdataan.

Tutkimuskysymykset ovat seuraavat:

1. Mitä tekijät kannustavat pk-yrityksiä massadata-analytiikan käyttöönottoon?
 - a. Miten massadata-analytiikkaa voidaan hyödyntää pk-yrityksissä?
2. Mitä haasteita pk-yritykset kohtaavat massadata-analytiikan käyttöönotossa?
 - a. Mitä haasteita massadata ja sen analysoiminen aiheuttavat?

Tutkielmassa painotetaan pk-yritysten näkökulmaa, huomioiden esimerkiksi datan, osaamisen ja taloudellisten resurssien niukkuus suhteessa suuryrityksiin. Tutkielmassa myös selvitetään, millainen merkitys esimerkiksi muutosvastarinnalla, järjestelmän monimutkaisuudella, kilpailupaineilla, paikallisella sääntelyllä tai organisaatiokulttuurilla voi olla massadata-analytiikan käyttöönotossa.

Tutkielma sisältää kaksi osalukua. Ensimmäisessä sisältöluvussa eli luvussa 2 käsitellään massadatan ja massadata-analytiikan määritelmiä ja ominaisuuksia sekä niihin liittyviä haasteita ja mahdollisuuksia yritysten liiketoiminnassa. Luvussa selitetään, miten massadata poikkeaa muusta datasta ja miten yritykset voivat käytännössä hyödyntää massadataa. Luku 3 keskittyy massadata-analytiikkaan ja sen käyttöönoton (engl. adoption) ajureihin ja haasteisiin nimenomaan pk-yrityksissä. Luvussa 3 käytetään Tornatzky & Fleischerin (1990) esittämää innovaatioiden käyttöönoton TOE-runkoa ajureiden ja haasteiden tarkasteluun. Luvussa käsitellään myös sellaisia tekijöitä, jotka ovat uupuneet aikaisemmista TOE-runkoa käyttäneistä tutkimuksista. Luvussa 4 työ vedetään yhteen ja esitetään johtopäätökset.

2 Massadata yrityksissä

2.1 Massadata

Massadatan yksiselitteinen määrittely on vaikeaa, sillä nimitystä käytetään enemmänkin ilmiön eikä pelkän datan kuvaamiseen (Ylijoki & Porras, 2016). Massadatan ja yritysten muun datan erottaminen toisistaan on kuitenkin tärkeää, jotta voidaan tunnistaa massadatan tarjoamat mahdollisuudet ja sen analysoimiseen tarvittavat inhimilliset ja tietotekniset resurssit.

Massadatalle on annettu lukuisia eri määritelmiä, mutta monesti sen kuvaamiseen käytetään massadatan niin sanottuja v-ominaisuuksia. Yhtenä merkittävimmistä massadatan määritelmistä pidetty (Curry, 2016; Ylijoki & Porras, 2016) Laneyn (2001) kuvaus datan ominaisuuksista sisältää kolme v:tä – volume (määrä), velocity (nopeus) ja variety (valikoima). Myöhemmin määritelmää on laajennettu neljään, viiteen ja jopa seitsemään v:hen. Nämä neljä lisättyä v:tä ovat variability (vaihtelevuus), veracity (todenmukaisuus), visualization (visualisointi) ja value (arvo). (Chartier, 2016.) Massadatan v-ominaisuudet auttavat erottamaan sen tavallisesta datasta. Jos massadatalle ei olisi näitä ainutlaatuisia v-ominaisuuksia, ei yrityksillä olisi esteitä analysoida sitä perinteisillä liiketoiminta-analytiikan työkaluilla (Lu ym., 2014).

Tässä alaluvussa käsitellään alkuperäisen kolmen v:n määritelmän mukaisia ominaisuuksia, sillä niistä on saatavilla eniten tutkimusaineistoa, ja ne muodostavat suurimman osan massadatan ja massadata-analytiikan ominaisuuksiin liittyvistä haasteista ja mahdollisuuksista (ks. esim. Che ym., 2013). Lisäksi laajennetuissa määritelmässä v-ominaisuuksien valikointi saattaa vaihdella. Osa yrityksistä voi pitää esimerkiksi arvoa neljäntenä v:nä, kun taas toiset saattavat kutsua todenmukaisuutta neljänneksi v:ksi (Gandomi & Haider, 2015).

Massadatan määrä (engl. volume) on yksi sen tunnistettavimmista piirteistä. Vaikka poikkeuksellisen suuret datamäärät mahdollistavat laajemman ja tarkemman analytiikan, aiheuttavat ne myös täysin uudenlaisia haasteita. Yritysten keräämää ja käyttämää dataa on varastoitava, mikä ei ole ilmaista. Massadatan jäsentelemättömyys ja dynaaminen kasvu luovat tarpeen uudistetuille varastointiratkaisuille eli tietokannoille, jotka ovat luotettavampia, skaalautuvampia ja tehokkaampia kuin perinteiset tiedonhallintajärjestelmät. Aikaisemmin yleisessä käytössä olleet relaatiotietokannat

(engl. relational database) eivät sovellu massadatan käsittelemiseen, koska ne pakottavat datan tarkkoihin ja ennalta määriteltyihin relaatiokaavioihin. Tässä alaluvussa myöhemmin esitettävä massadatan valikoima on myös yksi syy relaatiotietokantojen soveltumattomuudelle. (Che ym., 2013; Siddiqa ym., 2017.)

Massadatan nopeudella (engl. velocity) tarkoitetaan vauhtia, jolla uutta dataa luodaan ja taltioidaan (Lu ym., 2014). Chen ym. (2013) mukaan Google esitti jo vuonna 2010, että maailmassa tuotetaan joka toinen päivä yhtä paljon dataa kuin vuoteen 2003 mennessä oli tuotettu yhteensä. Massadatan nopeus luonnollisesti tukee sen määrää, mutta samalla vaikeuttaa olennaisen tiedon keräämistä eli niin sanottua tiedonlouhintaa (engl. knowledge discovery, sittemmin data mining) (Che ym., 2013). Frawley ym. (1992) kuvaavat tiedonlouhinnan olevan ”implisiittisen, aikaisemmin tuntemattoman ja mahdollisesti hyödyllisen tiedon epätriviaalia poimintaa datasta”.¹ Larose ja Larose (2014) taas määrittelevät tiedonlouhinnan ”prosessiksi, jossa löydetään hyödyllisiä malleja ja trendejä suurista tietojoukoista”.² Nopeus tuo mukanaan uudenlaisia mahdollisuuksia, mutta voi samalla aiheuttaa haasteita pakottamalla yritykset keskittymään esimerkiksi yksityisyydensuojan toteutumiseen (Coleman ym., 2016) ja louhintatyökalujen soveltuvuuksiin eri tietokannoissa (Chen ym., 1996).

Alkuperäisen määritelmän kolmas ominaisuus on valikoima (engl. variety). Valikoima käsittää massadatan monimuotoisuuden ja uusien teknologioiden tarjoamat datatyypit, joilla voidaan laajentaa perinteistä dataa (Lu ym., 2014). Ennen yritysten keräämä data saattoi rajoittua esimerkiksi tiettyjen tuotteiden myyntimääriin eri ajankohtina (ks. esim. Che ym., 2013; Müller ym., 2018). Massadatan vaihtelevuuden ansiosta tätä analyysiä voidaan yrittää laajentaa kassajärjestelmän ulkopuolelle, esimerkiksi sosiaalisen median tai verkkokaupan liikenteeseen. Sosiaaliseen mediaan kohdistetun massadata-analytiikan onkin tunnustettu tuovan merkittävää kilpailuetua. (Dong & Yang, 2020.)

Yleisimmin käytettyjen v-ominaisuuksien lisäksi tutkijat ja yritykset ovat antaneet massadatalle muitakin määritelmiä. Microsoft (2013) määrittelee massadatan ”käsitteeksi, jota käytetään kasvavissa määrin kuvaamaan todellisen laskentatehon, eli koneoppimisen ja tekoälyn uutuuksien, soveltamista valtaviin ja usein monimutkaisiin

¹ “[...] the nontrivial extraction of implicit, previously unknown, and potentially useful information from data.”

² “Data mining is the process of discovering useful patterns and trends in large data sets.”

tietokokonaisuuksiin”.³ Ohjelmistotalo SAS (2023) sen sijaan kuvaa massadataa “dataksi, joka on niin suurta, nopeaa tai monimutkaista, että sen prosessointi perinteisillä metodeilla on vaikeaa tai mahdotonta”.⁴ Teknologiyhtiö Intel (2023) taas kuvaa massadataa ”eksatavutason tiedoksi, jota saadaan reaaliajassa erilaisina tiedostotyyppinä”, minkä analysointi ”vaatii rinnakkaisprosessointia, koneoppimista ja tekoälyä”.⁵

Vaikka massadatan määritelmät ovat heterogeenisiä, kaikilla tässä luvussa esitetyillä määritelmillä on joitakin yhteisiä piirteitä. Sekä Microsoftin, SAS:n että Intelin määritelmät tukevat aikaisemmin esitettyä kolmen v:n määritelmää siten, että nekin huomioivat massadatan suuret määrät, monimuotoisuuden ja tarpeen uudistetuille analyttisille työkaluille. Näiden määritelmien avulla on helpompi käsittää, miten massadata eroaa “tavallisesta” datasta, ja millaisia muutoksia sen analysointi voi vaatia yrityksiltä.

2.2 Massadata-analytiikka

Yritysten dataan liittyvät järjestelmät voidaan jakaa karkeasti kahteen tyyppiin: liiketoiminnallisiin järjestelmiin, jotka sisältävät muun muassa asiakkuuksienhallinnan (engl. customer relationship management, CRM) ja toimitusketjujen johtamisen (engl. supply chain management, SCM) sekä analyttisiin järjestelmiin, jotka käsittävät esimerkiksi datavarastot (engl. data warehouse) ja tiedonlouhinnan. Merkittävin ero liiketoiminnallisten ja analyttisten järjestelmien välillä on niiden käyttötarkoituksissa. Liiketoiminnalliset järjestelmät pyrkivät parantamaan olemassa olevien prosessien tehokkuutta, kun taas analyttisillä järjestelmillä yritetään tuoda esiin ainutlaatuisia innovaatioita, joilla voidaan kehittää esimerkiksi tuotteita tai palveluita. (Müller ym., 2018.) Edellä mainittujen järjestelmätyyppien erottaminen toisistaan auttaa ymmärtämään, minkälaisia resursseja yrityksillä on käytettävissään ennen massadata-analytiikan käyttöönottoa.

³ “[...] the term increasingly used to describe the process of applying serious computing power – the latest in machine learning and artificial intelligence – to seriously massive and often highly complex sets of information.”

⁴ “[...] data that is so large, fast or complex that it’s difficult or impossible to process using traditional methods.”

⁵ “[...] exabyte-scale information that arrives in real time in a wide variety of file types, from text records to sound, images, and video. Big data analytics go far beyond SQL, requiring parallel processing, machine learning, and AI.”

Ero perinteisen data-analytiikan ja massadata-analytiikan välillä on selitettävissä perinteisen data-analytiikan rajoitteilla. Data-analytiikan soveltaminen massadataan ei ole mahdollista, koska data-analytiikkaan käytettävät järjestelmät ovat skaalautumattomia, keskitettyjä, epädynaamisia ja käyttävät yhtenäisiä tietorakenteita. Käytännössä edellä mainitut ominaisuudet tarkoittavat sitä, että data-analytiikkajärjestelmien suorituskyky ei sovellu valtaviin datamääriin, järjestelmien rakenteita ei ole suunniteltu tukemaan massadatan nopeutta ja tiedon syöttämistä “lennossa”, eikä järjestelmillä ole vastausta massadatan monimuotoisuuteen ja vaihteleviin tiedostomuotoihin. (Tsai ym., 2015.)

Massadata-analytiikka voidaan jakaa kolmeen vaiheeseen: syöte, analysointi ja tuloste (engl. input, analysis, output). Syötteellä tarkoitetaan massadata-analytiikan alkuvaiheita, joissa dataa kerätään ja taltioidaan. Syötevaiheessa on tärkeää huomioida massadatan kolme v-ominaisuutta siten, että suurissa määrissä, nopeaan tahtiin ja monissa muodoissa syntyvä data on prosessoitavissa vaatimatta liikaa laskentatehoa. Analysoinnissa voidaan käyttää algoritmeja tunnistettujen ongelmien ratkaisemiseen ja koneoppimista prosessien automatisointiin tai tukemiseen. Tulostevaiheessa esitetään analysoinnin tulokset. Viimeisessä vaiheessa on tärkeää esittää tulokset oikealla tavalla, jotta tulokset ovat helposti ymmärrettävissä. Jos tuloksia ei osata tulkita oikein, ne voivat olla täysin hyödyttömiä. (Tsai ym., 2015.)

Müllerin ym. (2018) laatiman ekonometrisen analyysin mukaan massadata-analytiikka parantaa yritysten tuottavuutta ainakin IT-painotteisilla tai erityisen kilpailullisilla aloilla. Analyysissä tuottavuudella tarkoitetaan kokonaistuottavuutta (engl. multifactor productivity), joka mittaa kokonaistuotoksen ja -panosten välistä suhdetta. Yritykset ovat luonnollisesti kiinnostuneita saamaan enemmän aikaan pienemmillä kustannuksilla. Lisäksi on osoitettu, että suorituskykyyn liittyvät odotukset lisäävät massadata-analytiikan omaksumisen halukkuutta yrityksissä (Baig ym., 2023; Cabrera-Sánchez & Villarejo-Ramos, 2019; Moraes ym., 2022).

Niebel ym. (2019) osoittavat tutkimuksessaan, että massadata-analytiikalla on yhteys korkeampaan taipumukseen innovoida. Samassa tutkimuksessa tunnistettiin, että massadata-analytiikka voi vaikuttaa uusien tuotteiden ja palvelujen osuuksiin yritysten kokonaismyynneistä. Innovaatioiden on todistettu tuovan kilpailuetua myös

kansainvälisillä markkinoilla ja kasvattavan vientiä, ja uusien teknologioiden käyttöönotolla on osoitettu olevan merkitystä innovoinnissa (Azar & Ciabusch, 2017).

Suorien liiketoiminnallisten hyötyjen lisäksi massadata-analytiikka voi myös kehittää yritysten sisäistä toimintaa. Massadata-analytiikan on osoitettu muun muassa parantavan asiakastuntemusta, nopeuttavan aikaisempia prosesseja ja auttavan strategisessa päätöksenteossa (ks. esim. Marjani ym., 2017). Massadata-analytiikka voi siis tuoda kilpailuetua yrityksille, mutta etujen saavuttamiseksi yritysten on ensin saatava dataa ja hyötyä kyseisestä datasta.

Pienempien kokojensa vuoksi pk-yrityksillä voi luonnollisesti olla vähemmän dataa käytettävissä. Pk-yritykset eivät siksi voi kaikissa tapauksissa nojautua samanlaiseen dataan kuin suuryritykset, vaan ne voivat joutua etsimään uusia keinoja saadakseen sellaisia datamääriä, joita voitaisiin kutsua massadatakksi. Massadata-analytiikkaa käsittelevissä tutkimuksissa nostetaan esiin muun muassa esineiden internetin (engl. Internet of Things, IoT) (Sen ym., 2016) ja sosiaalisen median (Dong & Yang, 2020) hyödyntäminen data-aineistoissa. Esineiden internetin ja sosiaalisen median käytännössä katkeamattomat datavirrat voivat olla vastaus pk-yritysten kohtaamaan oman datan riittämättömyyteen. Datan puutteeseen vastaamisen lisäksi ulkoisen datan käyttäminen voi olla viisasta analyysien parantamiseksi ja uusien liiketoiminnallisten tutkimusten mahdollistamiseksi (Sen ym., 2016).

Esineiden internetillä on esitetty olevan lukuisia eri käyttötapauksia massadata-analytiikassa. Esineiden internetiä voidaan hyödyntää erityisesti ”älykkäissä” ympäristöissä, kuten toimitusketjuissa, maataloudessa, kaupunkiliikenteessä ja sähköverkoissa. Esineiden internetin ja massadata-analytiikan yhdistämisestä voi olla hyötyä ainakin verkkokaupoissa, älykaupungeissa, vähittäiskaupassa ja logistiikassa sekä terveydenhuollossa. Näiden osa-alueiden konkreettisiin hyötyihin lukeutuvat muun muassa ennusteiden tarkentaminen, asiakaskunnan kasvattaminen, reaaliaikainen prosessointi, kampanjoiden suunnittelu ja asiakastyytyvyyden parantaminen. (Marjani ym., 2017.) Edellä mainituilla aloilla toimivat pk-yritykset voivat hyödyntää esineiden internetin tuottamaa dataa paikatakseen datan puutetta tai laajentaakseen omalla datalla tehtyä analyysiä. Käytettävissä oleva data ei aina liity yrityksen välittömiin liiketoiminnallisiin prosesseihin, vaan sitä voidaan kerätä muun muassa tuotevarastoista tai vaikkapa sairaspöytäkirjoista (Sen ym., 2016).

Sensorien ja älyjärjestelmien lisäksi ihmiset voivat itse tuottaa massadata-analytiikkaan käytettävää dataa. Sosiaalisen median diversiteetin ja massadata-analytiikan välisen synergian on todistettu parantavan yritysten suoriutumista markkinoilla. Lisäksi pk-yritykset ovat suuryrityksiä alttiimpia tunnistamaan sosiaalisen median ja massadata-analytiikan synergian ja tuottamaan kilpailuetua siitä. Sosiaalisen median hyöty liittyy muun muassa parempaan kuluttajaymmärrykseen ja markkinointistrategioiden kehittämiseen. Pk-yritykset voivat siis hyödyntää sosiaalista mediaa massadata-analytiikassa, jos oma asiakasdata on riittämätöntä tai jos sitä halutaan tukea. (Dong & Yang, 2020.)

Massadatan tuomista mahdollisuuksista huolimatta se ei tuo yksiselitteistä kilpailuetua, joka vain odottaa löydetäksi tulemista. Kuten Sen ym. (2016) ovat esittäneet, massadataa ei pitäisi lähestyä yleishyödyllisenä työkaluna, vaan selkeällä visiolla ja tavoitteella ratkaista tunnistettuja ongelmia. Sen lisäksi, että data on löydettävä, on sitä myös käsiteltävä ja analysoitava. Samat ominaisuudet, jotka mahdollistavat massadatan liiketoiminnalliset hyödyt, voivat aiheuttaa haasteita yrityksille. Koska massadataa kerätään monesti useammasta lähteestä, datan monimutkaisuus ja mahdolliset mittausvirheet tai puuttuvat arvot voivat aiheuttaa haasteita. (Fan ym., 2014.)

Koska massadata-analytiikkaan käytettävää dataa saadaan tiedonlouhinnan kautta, koko analytiikkaprosessin tulos voi muuttua louhintavaiheen toteutustavan perusteella. Jos käytetty louhinta-algoritmi on tahallisesti tai tahattomasti puolueellinen, tekoälyavusteiset mallinnukset voivat myös olla puolueellisia. Koska tekoälyllä ei ole kontekstuaalista ymmärrystä, se ei voi itse erottaa, johtuuko otannan tulos aidosta korrelaatiosta vai kerätyn datan sattumasta. (Smith, 2020.) Esimerkiksi suuryritys Amazonin käyttämä rekryointialgoritmi saattoi tahattomasti pitää miespuolisia hakijoita naisia parempina, koska mallin kehittämisessä käytetyssä aineistossa oli enemmän miehiä kuin naisia (Reuters, 2018). Kyseisen algoritmin toiminnan todellisia syitä voidaan kuitenkin vain spekuloida, koska tapahtumasta ei ole saatavilla riittävästi tietoa. Esimerkkitapaus osoittaa silti, että algoritmit voivat olla virheellisiä ja tukea päätöksentekoa väärin perustein. Lopullinen vastuu datan perusteella tehdyistä päätöksistä on siten päättäjillä, ei data-analyttikoilla tai aineistolla.

Massadatan ominaisuuksien ja kerätyn datan laadun lisäksi on myös syytä olla kriittisiä niiden edustamia tuloksia kohtaan. Datalla voidaan esittää vain sellaisia arvoja ja lukuja,

jotka ovat esitettävissä ja kerättävissä lukuina. Massadata ei siis välttämättä anna täydellistä kuvausta nykytilasta ja kaikista muuttujista, eivätkä aikaisemmin mainitut massadatan suuret määrät tarkoita sitä, että kaikki kerätty data olisi edes käytettävissä. (Jones, 2019.) Yritysten on siten oltava tarkkoja saaduista tuloksista ja tunnistettava, millaiset huomiot ovat voineet jäädä tekemättä ja mitkä ilmiöt ovat ylipäättään tarkasteltavissa lukuina ja datana.

3 Massadata-analytiikan käyttöönotto pk-yrityksissä

Massadata-analytiikan käyttöönottoa tarkastelevat tutkimukset voidaan jakaa kahteen ryhmään: tutkimuksiin, jotka keskittyvät massadata-analytiikan käyttöönottoon yrityksen uutena voimavarana (ks. esim. Lai ym., 2018; Maroufkhani ym., 2020) ja tutkimuksiin, jotka käsittelevät uuden teknologian omaksumista käyttäjän näkökulmasta (ks. esim. Cabrera-Sánchez & Villarejo-Ramos, 2019). Selkeyden vuoksi kutsuttakoon ensimmäistä ryhmää tutkimuksia massadata-analytiikan käyttöönotoksi ja jälkimmäistä ryhmää omaksumiseksi. Kumpikin ryhmä on syytä ottaa huomioon, mutta omaksumista käsittelevien tutkimusten merkitys on huomattavasti pienempi, ja ne ovat luokiteltavissa yhdeksi käyttöönoton alaluokaksi. Teknologian omaksumisen arviointi etukäteen on pk-yritysten johtajille viisasta, mutta tässä tutkielmassa omaksumista käsitellään vain yhtenä monista massadata-analytiikan käyttöönoton tekijöistä.

Tornatzky ja Fleischer (1990) ovat esittäneet, että teknologisten innovaatioiden käyttöönottoon vaikuttavat kolme ryhmää: teknologinen (engl. technological), organisatorinen (engl. organizational) ja ympäristöllinen (engl. environmental) konteksti. Tätä innovaatioiden tarkasteluun käytettävää runkoa kutsutaan sen englanninkielisten ryhmien nimien mukaan TOE-rungoksi. Runkoa on käytetty useassa massadata-analytiikan käyttöönottoon keskittyvässä tutkimuksessa (ks. esim. Sun ym., 2018; Verma & Chaurasia, 2019), ja se on todettu muita käyttöönoton malleja paremmaksi muun muassa sen ympäristöllisen ulottuvuuden ja jatkuvan empiirisen tuen vuoksi (Oliveira & Martins, 2011).

TOE-rungon teknologisella kontekstilla tarkoitetaan yritykselle olennaisia sisäisiä ja ulkoisia teknologioita, eli nykyisiä toimintatapoja ja välineitä sekä saatavilla olevien ulkoisten teknologioiden joukkoa (Tornatzky & Fleischer, 1990). Lähdekirjallisuudessa teknologisia ajureina on tarkasteltu, miten hyvänä käyttöönotettava teknologia koetaan muihin vaihtoehtoihin nähden, kuinka teknologia sopii yhteen olemassa olevien järjestelmien kanssa ja miten käyttöönotettavan teknologian hyödyt ovat tarkasteltavissa. Teknologiset haasteet sen sijaan liittyvät järjestelmän monimutkaisuuteen sekä massadata-analytiikan aiheuttamaan epävarmuuteen ja turvallisuushuoliin. (Maroufkhani ym., 2020.)

Organisatorisella kontekstilla taas viitataan yrityksen rakenteeseen liittyviin ominaisuuksiin, kuten kokoon, johtamisrakenteisiin ja henkilöstöresurssien laatuun (Tornatzky & Fleischer, 1990). Organisatorisen kontekstin muuttujina on tutkittu johdon tuen ja organisatorisen valmiuden positiivisia vaikutuksia (ks. esim. Maroufkhani ym., 2020; Sun ym., 2018). Organisatorista kontekstia on tutkielmassa laajennettu taloudellisella valmiudella sekä järjestelmän yksilökohtaisella käyttöönotolla eli omaksumisella.

Ympäristöllistä kontekstia käytetään muun muassa toimialan, kilpailijoiden ja lainsäädännön arviointiin teknologisen innovaation käyttöönotossa (Tornatzky & Fleischer, 1990). TOE-runkoa käyttäneissä massadata-analytiikan käyttöönoton tutkimuksissa on tarkastelu kilpailijoiden ja sääntelyn vaikutuksia. Aikaisemmat TOE-runkoa käyttäneet tutkimukset eivät kuitenkaan ole huomioineet kaikkia merkittäviä ympäristöllisiä tekijöitä, kuten ohjelmisto- ja työmarkkinoiden aiheuttamia haasteita. (Ks. esim. Lutfi ym., 2022.)

3.1 Massadata-analytiikan käyttöönoton ajurit pk-yrityksissä

Tässä alaluvussa tarkastellaan kaikkia sellaisia teknologisia, organisatorisia ja ympäristöllisiä ajureita, joiden vaikutusta massadata-analytiikan käyttöönottoon on lähdeaineistossa tutkittu ja eritellään, mitkä niistä ovat osoittautuneet pk-yrityksille ominaisiksi. Lisäksi edellisen luvun näkökulmaa laajennetaan listaamalla sellaisia massadata-analytiikan käyttöönottoon kannustavia tekijöitä, jotka eivät välttämättä liity suoraan massadatan tuomiin mahdollisuuksiin.

3.1.1 Teknologiset ajurit

Massadata-analytiikan käyttöönottoa tutkivissa teoksissa teknologisia ajureina on tarkasteltu suhteellista hyötyä (engl. relative advantage), yhteensopivuutta (engl. compatibility), kokeiltavuutta (engl. trialability), tarkasteltavuutta (engl. observability) ja havaittuja hyötyjä (engl. perceived benefits) (ks. esim. Lai ym., 2018; Maroufkhani ym., 2020).

Lutfi ym. (2022) esittävät, että suhteellisella hyödyllä, eli tuotteen paremmuudella suhteessa muihin vaihtoehtoihin, on pk-yrityksissä merkittävä osuus massadata-analytiikan käyttöönottoa harkittaessa. Muutkin tutkimukset (ks. esim. Sun ym., 2018;

Verma & Chaurasia, 2019) ovat tunnistaneeet suhteellisen hyödyn olevan eräs merkittävimmistä massadata-analytiikan käyttöönoton ajureista. Jälkimmäiset tutkimukset eivät kuitenkaan keskittyneet pk-yrityksiin. Lisäksi on esitetty, ettei suhteellisella hyödyllä olisi vaikutusta massadata-analytiikan käyttöönottoon pk-yrityksissä (Maroufkhani ym., 2020). Maroufkhani ym. (2020) ovat kuitenkin selittäneet suhteellisen hyödyn poikkeuksellista merkityksettömyyttä sillä, että tutkimusaineistossa käytetyt yritykset kärsivät talouteen, henkilöstöön ja tietotekniseen infrastruktuuriin liittyvistä ongelmista. Suhteellinen hyöty ei siis sellaisenaan ole massadata-analytiikan käyttöönoton ajuri, jollei yrityksellä ole riittävästi muita valmiuksia teknologian käyttöönottoon. Suhteellista hyötyä voidaan siten tässä tutkielmassa pitää yleisenä teknologisenä ajurina, joka koskee myös pk-yrityksiä.

Yrityksen olemassa olevien tietoteknisten resurssien merkitystä ja yhteensopivuutta massadata-analytiikan käyttöönotossa on myös tutkittu, mutta niitä ei ole pidetty tilastollisesti merkittävänä tekijöinä (Sun ym., 2018; Verma & Chaurasia, 2019). Yhteensopivuuden merkityksettömyydestä ei ole kuitenkaan saavutettu yksimielisyyttä, sillä sen on esitetty vaikuttavan positiivisesti massadata-analytiikan käyttöönoton halukkuuteen toimitusketjujen johtamisessa (D. Q. Chen ym., 2015) ja verkkokaupan käyttöönottoon pk-yrityksissä (Ghobakhloo ym., 2011). Pk-yrityksiin ja massadata-analytiikkaan keskittyneet tutkimukset eivät taas ole osoittaneet yhteensopivuuden olevan merkittävä käyttöönoton tekijä (Lutfi ym., 2022; Maroufkhani ym., 2020). Yhteensopivuuden merkityksettömyyttä pk-yrityksissä on perusteltu muun muassa sillä, että pk-yritykset ovat kokonsa ansiosta sopeutuvampia ja joustavampia kuin suuryritykset, eikä pk-yrityksissä siten koeta tarvetta harkita järjestelmien yhteensopivuutta päätöksenteossa (Maroufkhani ym., 2020). Yhteensopivuus ei siis ole teknologinen ajuri, joka vaikuttaisi massadata-analytiikan käyttöönottoon pk-yrityksissä tai yrityksissä yleisesti.

Aihepiirin kirjallisuudessa kokeiltavuudella viitataan siihen, kuinka kattavasti järjestelmä on testattavissa ennen sen käyttöönottoa. Tarkasteltavuudella sen sijaan tarkoitetaan sitä, miten innovaation hyödyt ovat havaittavissa ulkopuolisille. (Maroufkhani ym., 2020.) Kokeiltavuus ja tarkasteltavuus ovat osittain samantyyllisiä ajureita, sillä niiden vaikutukset liittyvät siihen, miten hyödyt ovat tarkasteltavissa joko sisä- tai ulkopuolelta. Maroufkhani ym. (2020) ovat osoittaneet, että sekä kokeiltavuus että tarkasteltavuus vaikuttavat positiivisesti massadata-analytiikan käyttöönottoon pk-yrityksissä. Sun ym.

(2018) sijoittivat tuloksissaan kokeiltavuuden ja tarkasteltavuuden molemmat vähiten merkittäviksi massadatan omaksumiseen vaikuttaviksi tekijöiksi. Tekijöiden positiivinen vaikutus kuitenkin tunnistettiin, eikä tuloksista voida siten päätellä, ettei vaikutusta olisi lainkaan. Kokeiltavuuden ja tarkasteltavuuden merkityksistä nimenomaan massadata-analytiikan käyttöönotossa ei ole saatavilla paljon tutkimusaineistoa, mutta molempien on osoitettu vaikuttavan esimerkiksi yritysjärjestelmien (engl. enterprise systems) käyttöönottoon pk-yrityksissä (Ramdani ym., 2009). Epävarmuus ja turvattomuus ovat yleisiä massadata-analytiikan käyttöönoton esteitä nimenomaan pk-yrityksissä, ja niitä voidaan yrittää vähentää helpottamalla hyötyjen havaitsemista ja siten luomalla varmuutta onnistumisesta (Coleman ym., 2016; Maroufkhani ym., 2020). Koska muut yritykset kokevat vähemmän epävarmuutta ja turvattomuutta, eivät kokeiltavuus ja tarkasteltavuus ole siten ratkaisevia käyttöönoton tekijöitä niissä. Kokeiltavuutta ja tarkasteltavuutta voidaan siis pitää erityisesti pk-yrityksille ominaisina teknologisisina ajureina.

Käytetyssä kirjallisuudessa havaittuja hyötyjä ei ole luokiteltu omaksi muuttujakseen pk-yrityksissä. Vain Lai ym. (2018) ovat laajentaneet TOE-runkoa lisäämällä siihen havaitut hyödyt ja osoittaneet positiivisen vaikutuksen käyttöönottoon toimitusketjujen johtamisessa. Justy ym. (2023) ovat kuitenkin esittäneet yksityiskohtaisemman listan massadata-analytiikan käyttöönoton ajureista pk-yrityksissä. Listassa on esitetty strategisia ja operationaalisia ajureita, joita voidaan pitää havaittuina hyötyinä tai odotettuina suorituskyvyn parannuksina, kuten strategisen päätöksenteon tukeminen tai asiakastuntemuksen parantaminen. Pk-yritysten johtajat ovat siis havainneet massadata-analytiikan hyötyjä, ja pitävät niitä käyttöönottoon kannustavina tekijöinä. Havaitut hyödyt ovat siten teknologisia massadata-analytiikan käyttöönoton ajureita, jotka koskevat myös pk-yrityksiä.

3.1.2 Organisatoriset ajurit

Massadata-analytiikan käyttöönoton organisatorisina ajureina on tutkittu ylimmän johdon tukea (engl. top management support), organisatorista valmiutta (engl. organizational readiness) ja taloudellista valmiutta (engl. financial readiness) (ks. esim. Lai ym., 2018; Maroufkhani ym., 2020).

Useampi tutkimus esittää ylimmän johdon tuen olevan massadata-analytiikan käyttöönottoon positiivisesti vaikuttava tekijä sekä pk-yrityksissä (Justy ym., 2023; Lutfi

ym., 2022; Maroufkhani ym., 2020, 2023) että yrityksissä yleisesti (Ganeshkumar ym., 2023; Lai ym., 2018; Sun ym., 2018; Verma & Chaurasia, 2019; Youssef ym., 2022). Johdon tuella viitataan siihen, kuinka vastaanottavia ja ymmärtäväisiä johtajat ovat uutta teknologiaa kohtaan (ks. esim. Maroufkhani ym., 2020). Ylimmän johdon tukea voidaan täten pitää organisatorisena massadata-analytiikan käyttöönoton ajurina, joka koskee pk-yritysten lisäksi muitakin yritysryhmiä.

Organisatorisella valmiudella tarkoitetaan sitä, kuinka kykeneviä ja halukkaita yritykset ovat omaksumaan uusia teknologioita. Organisatorinen valmius sisältää myös yrityksen sisäisen tietoteknisen osaamisen ja infrastruktuurin. (Maroufkhani ym., 2020.) Organisatorisen valmiuden on osoitettu vaikuttavan massadata-analytiikan käyttöönottoon niin pk-yrityksissä (Maroufkhani ym., 2020) kuin yrityksissä yleisesti (Verma & Chaurasia, 2019). Organisatoriseen valmiuteen liittyen massadata-analytiikan käyttöönotossa on tutkittu myös tietoteknisen infrastruktuurin ja ominaisuuksien (engl. IT infrastructure and capabilities) vaikutusta (Lai ym., 2018). Lai ym. (2018) ovat perustelleet IT-ominaisuuksien heikkoa merkitystä kahdella syyllä: tietotekniset puutteet voidaan ratkaista ulkoistamisen kautta, tai yritysten johtajat eivät ole tarpeeksi tietoisia massadata-analytiikan tarpeista, eivätkä siten tiedä, että tämänhetkiset resurssit ovat riittämättömät. Johdon tuen vaikutus käyttöönottoon osoittautuu jälleen merkittäväksi. Organisatorinen valmius siis vaikuttaa massadata-analytiikan käyttöönottoon pk-yrityksissä, mutta tarvittaessa valmiutta voidaan parantaa esimerkiksi ulkoistamisen avulla.

Sun ym. (2018) ovat esittäneet, että käyttöönoton kustannukset ovat neljänneksi merkittävin tekijä massadata-analytiikan käyttöönotossa. Taloudelliset esteet vaikuttavat käyttöönottoon myös pk-yrityksissä (Coleman ym., 2016; Justy ym., 2023). Lai ym. (2018) yrittivät löytää positiivisen vaikutuksen taloudellisen valmiuden ja käyttöönoton halukkuuden välillä, mutta tulokset eivät tukeneet väitettä. He perustelevat merkityksettömyyttä sillä, että otannan IT-vastuuhenkilöt eivät välttämättä ole tarpeeksi tietoisia yritystensä taloudellisista tilanteista tai käyttöönotettavan teknologian kustannuksista. Koska taloudellisten esteiden on osoitettu olevan haaste massadata-analytiikan käyttöönotolle ainakin pk-yrityksissä, voidaan todeta, että taloudellinen valmius vaikuttaa niissä käyttöönottoon positiivisesti.

3.1.3 Ympäristölliset ajurit

Yrityksissä voidaan kokea myös ulkoista painetta teknologian käyttöönottoon. Kilpailupaineet (engl. competitive pressure), kauppiaiden ulkopuolinen tuki (engl. external vendor support) sekä paikallinen sääntely (engl. government regulation) voivat vaikuttaa positiivisesti käyttöönoton halukkuuteen (ks. esim. Maroufkhani ym., 2020).

Kilpailupaineet muodostavat merkittävän osuuden massadata-analytiikan käyttöönoton syistä (Sun ym., 2018; Verma & Chaurasia, 2019). Yrityksissä voidaan kokea paineita siirtyä käyttämään edistyneempää teknologiaa, jos kilpailijat ovat onnistuneet saamaan kilpailuetua kyseisellä teknologialla. Kilpailupaineiden vaikutusta massadata-analytiikan käyttöönottoon pk-yrityksissä ei ole kuitenkaan tunnistettu. (Lutfi ym., 2022; Maroufkhani ym., 2020.) Tulokset eivät ole linjassa aikaisempien tutkimusten (ks. esim. Ghobakhloo ym., 2011) kanssa, joissa esitetään kilpailun vaikuttavan ainakin verkkokaupan käyttöönottoon positiivisesti pk-yrityksissä. Koska Maroufkhanin ym. (2020) ja Lutfin ym. (2022) esittämät tulokset ovat uudempia ja keskittyvät nimenomaan massadata-analytiikan käyttöönottoon, todettakoon, että kilpailupaineet eivät ole käyttöönoton ajuri pk-yrityksissä ainakaan massadata-analytiikassa.

Kauppiaiden ulkopuolinen tuki viittaa ohjelmistokauppiaiden tai kolmansien osapuolien antamaan tukeen järjestelmän käyttöönotossa. Ulkopuolisen tuen positiivinen vaikutus massadata-analytiikan käyttöönottoon on tunnistettu pk-yrityksissä (Maroufkhani ym., 2020) ja muissakin yritysyoissa (Ganeshkumar ym., 2023). Maroufkhani ym. (2020) perustelevat ulkopuolisen tuen merkitystä sillä, että pk-yritysten johtajilla on matalampi kynnyks käyttöönottaa uusia teknologioita, jos kauppiaalta koetaan saatavan riittävästi tukea. Johdon tuen merkitys korostuu siis muissakin tekijöissä.

Paikallisen sääntelyn vaikutuksesta ei ole saavutettu selkeää yksimielisyyttä, mikä on toisaalta ymmärrettävää, koska tutkimukset on toteutettu eri ympäristöissä. Lutfi ym. (2022) ovat esittäneet, että sääntely vaikuttaa massadata-analytiikan käyttöönottoon positiivisesti etenkin sellaisissa yrityksissä, joissa taloudellisia resursseja ja teknistä osaamista ei ole riittävästi. Maroufkhani ym. (2020) ovat sen sijaan osoittaneet, ettei sääntely vaikuta massadata-analytiikan käyttöönottoon pk-yrityksissä. Lai ym. (2018) eivät löytäneet vahvaa yhteyttä sääntelyn ja käyttöönoton välillä, mutta väitettä tuettiin osittain. Sun ym. (2018) taas asettivat sääntely-ympäristön kymmenenneksi merkittävimmäksi tekijäksi 26:n massadata-analytiikan käyttöönottoon vaikuttavan

tekijän joukosta. Sääntely voi siis kannustaa yrityksiä käyttöönottamaan massadata-analytiikkaa vähentämällä epävarmuutta ja luomalla taloudellisia, teknisiä ja lainsäädännöllisiä helpotuksia yrityksille. Tulokset vaihtelevat maittain, mutta niissä maissa, joissa sääntely vaikuttaa, on vaikutus positiivista.

3.2 Massadata-analytiikan käyttöönoton haasteet pk-yrityksissä

3.2.1 Teknologiset haasteet

Kirjallisuudessa massadata-analytiikan käyttöönoton teknologisinä haasteina on tutkittu monimutkaisuutta (engl. complexity) sekä epävarmuutta (engl. uncertainty) ja turvattomuutta (engl. insecurity).

Monimutkaisuus on yksi merkittävimmistä massadata-analytiikan käyttöönottoon negatiivisesti vaikuttavista tekijöistä niin yrityksissä yleisesti (Ganeshkumar ym., 2023; Sun ym., 2018; Verma & Chaurasia, 2019) kuin pk-yrityksissäkin (Lutfi ym., 2022; Maroufkhani ym., 2020, 2023). Monimutkaisuuden vaikutuksesta ei ole kuitenkaan saavutettu selkeää yksimielisyyttä: esimerkiksi toimitusketjujen johtamisessa (Lai ym., 2018) ja vähittäiskaupassa (Youssef ym., 2022) monimutkaisuuden vaikutusta massadata-analytiikan käyttöönottoon ei ole tunnistettu. Lain ym. (2018) esittämä monimutkaisuuden merkityksettömyys saattaa johtua esimerkiksi tutkimuksessa käytettyjen yritysten teknisestä edistyneisyydestä. Monimutkaisuutta voidaan siis pitää massadata-analytiikan käyttöönoton haasteena sekä yleisellä tasolla että pk-yrityksissä. Toimialalla vaikuttaa olevan suurempi merkitys kuin yrityksen koolla, kun tarkastellaan monimutkaisuuden vaikutuksia.

Massadata-analytiikan käyttöönottoon on tunnistettu liittyvän epävarmuutta ja turvattomuuden tunnetta (Ganeshkumar ym., 2023; Maroufkhani ym., 2020; Raguseo, 2018; Sun ym., 2018). Epävarmuutta voi syntyä esimerkiksi ymmärryksen puutteesta (Coleman ym., 2016) sekä tietoturvaan ja -suojaan sekä yksityisyyteen liittyvistä huolenaiheista (Coleman ym., 2016; Maroufkhani ym., 2020; Sun ym., 2018; Youssef ym., 2022). Tulokset eivät kuitenkaan viittaa siihen, että epävarmuus ja turvattomuuteen liittyvät huolenaiheet olisivat vain pk-yrityksille ominaisia. Massadata-analytiikkaan liittyvät huolenaiheet ja epävarmuus ovat siten yleisen tason teknologisia haasteita, joita myös pk-yritykset kohtaavat.

3.2.2 Organisatoriset haasteet

Tämän tutkielman TOE-runkoa käyttäneet lähteet eivät ole tutkineet yhtäkään organisatorista haastetta massadata-analytiikan käyttöönotossa (ks. esim. Lai ym., 2018). Omaksumista käsitteleviä teoksia on kuitenkin riittävästi, ja TOE-rungon kolmesta ryhmästä omaksuminen sopii parhaiten organisatoristen tekijöiden joukkoon. Lisäksi rakenteettomammissa haastattelututkimuksissa on huomattu, että riittämättömät inhimilliset resurssit voivat olla esteenä massadata-analytiikan käyttöönotolle (Justy ym., 2023). Inhimilliset resurssit ovat kuitenkin tässä tutkielmassa osana ympäristöllisiä haasteita.

Uuden teknologian käyttöönoton haasteet eivät välttämättä lopu sen hankintaan, vaan omaksumisesta on myös varmistuttava. Teknologia tarvitsee luonnollisesti käyttäjiä, ja järjestelmän omaksuminen voi osoittautua hidasteeksi. Omaksumisen tutkimiseen käytetään usein Venkateshin ym. (2003) laatimaa UTAUT-mallia (engl. unified theory of acceptance and use of technology). UTAUT-malli on tarkoitettu nimenomaan järjestelmän omaksumisen tai hyväksymisen tarkasteluun, ja se koostuu neljästä muuttujasta: suorituskyvyn odotukset, vaivannäön odotukset, sosiaalinen vaikutus ja helpottavat olosuhteet. Tutkielman lähteissä UTAUT-mallia on laajennettu lisäämällä siihen muuttujiksi käyttövästarinta, havaitut riskit ja hinnan arvo. (Cabrera-Sánchez & Villarejo-Ramos, 2019; Moraes ym., 2022.) Lisätyistä muuttujista tässä alaluvussa tarkastellaan käyttövästarintaa alkuperäisten muuttujien lisäksi.

Suorituskyvyn odotuksilla viitataan siihen, kuinka vahvasti yksilö uskoo järjestelmän avulla saavuttavansa hyötyjä, jotka parantavat työssä suoriutumista (Venkatesh ym., 2003). Suorituskyvyn odotusten vaikutuksesta massadata-analytiikan omaksumiseen pk-yrityksissä ei ole suoraa näyttöä, mutta on osoitettu, että suorituskyvyn odotukset vaikuttavat massadata-analytiikan omaksumiseen muissa yrityksissä (Cabrera-Sánchez & Villarejo-Ramos, 2019; Moraes ym., 2022) ja lisäksi pk-yrityksissä on huomattu suorituskyvyn odotusten vaikutus muiden teknologioiden omaksumiseen (Lutfi, 2022). Voidaan siis todeta, että suorituskyvyn odotuksilla voi olla positiivinen vaikutus massadata-analytiikan omaksumiseen myös pk-yrityksissä.

Vaivannäön odotuksilla tarkoitetaan sitä, kuinka helppona järjestelmän käyttäminen nähdään (Venkatesh ym., 2003). Massadata-analytiikan omaksumisessa ei ole havaittu vaivannäön odotusten vaikutusta (Cabrera-Sánchez & Villarejo-Ramos, 2019; Moraes

ym., 2022), mutta pk-yrityksissä on havaittu vaivannäön odotusten positiivinen vaikutus muissa teknologioissa (Lutfi, 2022). Cabrera-Sánchez ja Villarejo-Ramos (2019) ovat perustelleet vähäistä merkitystä sillä, että helpottavat olosuhteet voivat lieventää odotuksia ja vaivannäön odotukset ovat siten toisesta muuttujasta riippuvaisia. Moraes ym. (2022) ovat sen sijaan perustelleet tuloksiaan käytetyn ihmisotannan nuoruudella. Vaivannäön odotuksia voidaan siten pitää positiivisesti vaikuttavana tekijänä, joka on poissuljettavissa helpottavilla olosuhteilla.

UTAUT-mallin sosiaalinen vaikutus on yksittäisen käyttäjän kokemus siitä, miten vahvasti käyttäjä uskoo muiden haluavan hänen käyttävän järjestelmää (Venkatesh ym., 2003). Sosiaalisen vaikutuksen on todistettu vaikuttavan positiivisesti massadata-analytiikan omaksumiseen (Cabrera-Sánchez & Villarejo-Ramos, 2019; Moraes ym., 2022). Sosiaalisen vaikutuksen merkitystä ei ole kuitenkaan tunnistettu pk-yrityksissä (Lutfi, 2022). Toisaalta Lutfi (2022) on esittänyt, että tulos poikkeaa aikaisemmista tutkimuksista ja johtuu tutkimuksessa tarkastellun teknologian piirteistä ja käytetystä otannasta. Sosiaalinen vaikutus todetaan siten tässä tutkielmassa pk-yrityksiinkin vaikuttavaksi omaksumisen tekijäksi.

Helpottavilla olosuhteilla viitataan siihen, kuinka helposti teknologian käyttämiseen tarvittavat resurssit ovat saatavilla (Cabrera-Sánchez & Villarejo-Ramos, 2019). Helpottavien olosuhteiden positiivinen vaikutus massadata-analytiikan omaksumiseen on todettu merkittäväksi (Cabrera-Sánchez & Villarejo-Ramos, 2019) ja merkityksettömäksi (Moraes ym., 2022). Pk-yrityksissä helpottavien olosuhteiden on tunnistettu vaikuttavan uuden teknologian omaksumiseen positiivisesti (Lutfi, 2022). Moraes ym. (2022) ovat jälleen perustelleet havaittua merkityksettömyyttä sillä, että käytetyllä otannalla oli tavallista paremmat teknologiset valmiudet. Helpottavia olosuhteita voidaan siten pitää massadata-analytiikan omaksumiseen positiivisesti vaikuttavaksi tekijäksi pk-yrityksissä ja suuremmisakin yrityksissä.

Muutos- tai käyttövästarinta on yleinen uusien teknologioiden käyttöönottoon ja omaksumiseen negatiivisesti vaikuttava tekijä (Management Association, 2019), eikä massadata-analytiikka vaikuta poikkeavan säännöstä (Cabrera-Sánchez & Villarejo-Ramos, 2019; Moraes ym., 2022; Raguseo, 2018; Verma & Chaurasia, 2019). Muutosvästarinnan negatiivinen vaikutus on havaittu myös pk-yrityksissä (Coleman ym., 2016; Justy ym., 2023). Muutos- tai käyttövästarinta voidaan todeta massadata-

analytiikan omaksumiseen negatiivisesti vaikuttavaksi tekijäksi, joka on läsnä myös pk-yrityksissä.

3.2.3 Ympäristölliset haasteet

Organisatoristen haasteiden lisäksi massadata-analytiikan käyttöönotossa ei ole tutkittu ympäristöllisiä haasteita. Muita rakenteita käyttäneet tutkimukset ovat kuitenkin osoittaneet, että pk-yritysten ympäristöstäkin löytyy haasteita muun muassa ohjelmistoihin ja työvoimaan liittyviltä markkinoilta (Coleman ym., 2016; Justy ym., 2023; Sun ym., 2018). Tässä alaluvussa esitetään kaksi uutta ympäristöllistä käyttöönoton haastetta – ohjelmistomarkkinoiden haasteet ja työvoiman saatavuus.

Coleman ym. (2016) ovat esittäneet, että pk-yrityksissä massadata-analytiikan käyttöönoton haasteiksi voivat osoittautua läpinäkymättömät ohjelmistomarkkinat, intuitiivisten ohjelmistojen puute sekä käytännöllisten ja edullisten konsultointipalveluiden heikko saatavuus. Myös Justy ym. (2023) väittävät, että markkinoilla ei ole tarpeeksi ohjelmistovälittäjiä, jotka olisivat erikoistuneita pk-yrityksiin. Riittämättömät ohjelmistomarkkinat voivat siten olla massadata-analytiikan käyttöönoton haaste pk-yrityksissä.

Koska pk-yritysten henkilöstömäärät ovat suuryrityksiä pienempiä, on niissä luonnollisesti vähemmän erikoistuneita työntekijöitä. Sekä Coleman ym. (2016) että Justy ym. (2023) ovat esittäneet, että markkinoilla olevien työntekijöiden riittämättömyys voi olla este massadata-analytiikan käyttöönotolle pk-yrityksissä. Coleman ym. (2016) esittävät työvoimaongelmat ”pullonkauloina työmarkkinoilla”, kun taas Justy ym. (2023) selvittivät, että työmarkkinoilla ei ole tarpeeksi taitavia työntekijöitä, palkkakustannukset ovat liian korkeat ja pk-yrityksissä ei ole entuudestaan data-analytiikkataitaisia työntekijöitä tai soveltuvia koulutusresursseja. Pk-yritysten oman henkilöstön riittämättömyys liittyy enemmän organisatorisiin haasteisiin, mutta ongelmaan vastaaminen on kuitenkin osa työmarkkinoita. Edellä mainittujen tulosten perusteella voidaan todeta, että työmarkkinoiden ongelmat ovat ympäristöllinen este massadata-analytiikan käyttöönotolle pk-yrityksissä.

4 Yhteenveto ja johtopäätökset

Tutkielmassa selvitettiin, mitä hyötyjä ja haasteita massadata ja massadata-analytiikka voivat tuoda yrityksille. Lisäksi selvitettiin, mitkä tekijät kannustavat pk-yrityksiä massadata-analytiikan käyttöönottoon, ja mitä haasteita pk-yritykset kohtaavat massadata-analytiikan käyttöönotossa. Massadata-analytiikan liiketoiminnallisiksi hyödyiksi tunnistettiin parempi tuottavuus, lisääntynyt innovointi, parempi asiakastuntemus, päätöksenteon tuenta ja prosessien automatisointi. Massadatan ja massadata-analytiikan haasteiksi taas osoitettiin tiedonlouhinnan mahdolliset virheet ja niiden vaikutus saatuihin tuloksiin sekä olemassa olevan datan käyttökelpoisuus ja ilmiöiden kuvaaminen datana.

Osana massadata-analytiikan haasteita tunnistettiin, että massadata-analytiikkaa pitäisi kohdella ratkaisuna selkeään ja todettuun ongelmaan, eikä taattuna kilpailuetuna. Jos massadata-analytiikka otetaan käyttöön ilman selkeää liiketoiminnallista tarkoitusta ja roolia, voivat tulokset olla vääränlaisia tai mitättömiä (ks. esim. Sen ym., 2016). Tutkielmassa esitettiin, että pk-yritysten pitäisi jo ennen käyttöönoton ajurien ja haasteiden tarkastelua pohtia, mitä massadata-analytiikalla yritetään ylipäättään saavuttaa.

Pk-yrityksissä massadata-analytiikan käyttöönoton ajurien ja haasteiden tarkasteluun käytettiin TOE-runkoa, kuten monessa muussakin käyttöönottoa tutkivassa teoksessa on ollut tapana. Tutkielmassa onnistuttiin myös laajentamaan aikaisempien tutkimusten tuloksia lisäämällä TOE-runkoon sellaisia tekijöitä, joita ei ole aikaisemmin lisätty malleihin. Lähdekirjallisuudessa taloudellista valmiutta, henkilöstöresursseja, omaksumista tai ohjelmistomarkkinoita ei ole tarkasteltu TOE-rungon vaikuttajina, mutta tässä tutkielmassa ne koettiin merkittäviksi ja siksi niiden vaikutusta tarkasteltiin. Tämän tutkielman tulokset perustuvat kuitenkin aihepiirin kirjallisuuteen, eikä lisättyjen tekijöiden väitettyä vaikutusta ole testattu empiirisesti.

Taulukko 1. Massadata-analytiikan käyttöönottoon vaikuttavat tekijät

Massadata-analytiikan käyttöönottoon vaikuttavat tekijät	
Teknologiset tekijät	
Positiivinen vaikutus yleisesti ja pk-yrityksissä	Suhteellinen hyöty Havaitut hyödyt
Positiivinen vaikutus erityisesti pk-yrityksissä	Kokeiltavuus ja tarkasteltavuus
Negatiivinen vaikutus yleisesti ja pk-yrityksissä	Monimutkaisuus Epävarmuus ja turvattomuus
Ei merkittävää vaikutusta	Yhteensopivuus
Organisatoriset tekijät	
Positiivinen vaikutus yleisesti ja pk-yrityksissä	Ylimmän johdon tuki Organisatorinen valmius Taloudellinen valmius
Negatiivinen vaikutus yleisesti ja pk-yrityksissä	Omaksuminen
Ympäristölliset tekijät	
Positiivinen vaikutus yleisesti ja pk-yrityksissä	Kauppioiden ulkopuolinen tuki Paikallinen sääntely
Negatiivinen vaikutus pk-yrityksissä	Ohjelmistomarkkinat Työmarkkinat
Ei merkittävää vaikutusta pk-yrityksissä	Kilpailupaineet

Massadata-analytiikan käyttöönoton teknologisista ajureista tutkittiin suhteellista hyötyä, yhteensopivuutta, kokeiltavuutta, tarkasteltavuutta ja havaittuja hyötyjä. Tarkastelluista ajureista suhteellinen hyöty, tarkasteltavuus, kokeiltavuus ja havaitut hyödyt todettiin pk-yritysten käyttöönottoon positiivisesti vaikuttaviksi tekijöiksi. Tekijöistä ainoastaan tarkasteltavuus ja kokeiltavuus olivat sellaisia, jotka ovat vain pk-yrityksille ominaisia. Tarkasteltavuuden ja kokeiltavuuden suurta merkitystä perusteltiin sillä, että pk-yritykset kokevat enemmän epävarmuutta ja turvattomuutta, ja hyötyjen havaitseminen voi auttaa vähentämään niitä tunteita.

Käyttöönoton organisatorisina ajureina tutkittiin ylimmän johdon tukea sekä organisatorista ja taloudellista valmiutta. Ylimmän johdon tuki osoittautui yhdeksi merkittävimmäksi käyttöönoton ajuriksi. Myös organisatorisen ja taloudellisen valmiuden vaikutus tunnistettiin. Organisatorisista ajureista yksikään ei ollut vain pk-yrityksille ominainen, vaan ne koskivat myös muunkinlaisia yrityksiä. Ylimmän johdon tuen merkitys yhdistettiin myös muihin tekijöihin ja todettiin, että riittävällä ylimmän johdon tuella muiden tekijöiden merkitystä voidaan pienentää. Joitakin tekijöitä

voidaankin käytännössä pitää johdon tukeen vaikuttavina tekijöinä käyttöönottoon vaikuttavien tekijöiden sijasta, sillä esimerkiksi epävarmuuden vähentäminen liittyy nimenomaan johtoportaan kokemaan epävarmuuteen.

Ympäristöllisinä ajureina käsiteltiin kilpailupaineita, kauppiaiden ulkopuolista tukea ja paikallista sääntelyä. Kilpailupaineiden vaikutusta massadata-analytiikan käyttöönottoon ei tunnistettu pk-yrityksissä, mutta niiden todettiin vaikuttavan muunlaisiin yrityksiin ja muihin pk-yritysten teknologioihin. Kauppiaiden ulkopuolisen tuen positiivinen vaikutus käyttöönottoon tunnistettiin, mikä tukee myös osittain kokeiltavuuden ja tarkasteltavuuden havaittuja positiivisia vaikutuksia. Paikallisen sääntelyn tulokset lähdeaineistossa vaihtelevat, mutta tutkielmassa todettiin, että jos paikallisella sääntelyllä on vaikutusta käyttöönottoon, on vaikutus positiivista.

Käyttöönoton teknologisina haasteina tutkittiin monimutkaisuutta sekä epävarmuutta ja turvattomuutta. Teknologisista haasteista kaikki tunnistettiin sellaisiksi tekijöiksi, jotka vaikuttavat niin pk-yrityksiin kuin suurempiinkin yrityksiin. Monimutkaisuuden vaikutuksen todettiin vaihtelevan todennäköisesti herkemmin toimialan kuin yrityksen koon perusteella. Epävarmuuteen ja turvattomuuteen liittyi useita huolenaiheita, muun muassa tietoturva ja yksityisyys nousivat esiin.

Organisatoristen haasteiden tarkastelussa laajennettiin olemassa olevia tutkimustuloksia esittämällä teknologian omaksuminen uudeksi haasteeksi. Omaksumisen tarkasteluun käytettiin laajennettua UTAUT-mallia, joka sisälsi suorituskyvyn ja vaivannäön odotukset, sosiaalisen vaikutuksen, helpottavat olosuhteet sekä käyttövästarinnan. Tarkastelluista tekijöistä jokainen osoittautui omaksumiseen vaikuttaviksi tekijöiksi pk-yrityksissä. Käyttövästarinta oli ainoa tekijä, jonka vaikutusta tutkittiin negatiivisena.

Ympäristöllisissä haasteissa tehtiin uusia kontribuutioita lisäämällä kaksi uutta tekijää: ohjelmistomarkkinat ja työvoiman saatavuus. Kummankin tekijän negatiivinen vaikutus massadata-analytiikan käyttöönottoon pk-yrityksissä tunnistettiin. Ohjelmistomarkkinoiden tarkastelu tekijänä koettiin tarpeelliseksi, koska muissa tekijöissä ilmeni tarve läpinäkyvyydelle ja hyötyjen tarkastelemiselle, joista molempiin ohjelmistokauppiaille on mahdollisuus vaikuttaa. Työmarkkinoiden merkitys korostuu erityisesti tämän tutkielman näkökulmassa, koska pk-yritysten olemassa olevat inhimilliset resurssit voivat olla riittämättömiä massadata-analytiikalle ja pk-yritysten on siksi harkittava työvoiman vahvistamista.

Tutkielmassa havaittuihin ajureihin ja haasteisiin vastaamiseksi pk-yritysten tulisi ensimmäiseksi ymmärtää massadata-analytiikan rajoitteet, jotta tiedetään, mitä sillä on tai ei ole mahdollista tehdä. Kaikkiin käyttöönoton tekijöihin ei ole mahdollista vaikuttaa, mutta niitä voidaan kuitenkin arvioida. Pk-yritysten on tunnistettava omat toimintaympäristönsä.

Teknologisten, organisatoristen ja ympäristöllisten tekijöiden tärkeysjärjestys voi muuttua muun muassa ympäristön epävarmuuden mukaan. Teknologisten tekijöiden merkitys korostuu kohtalaisen epävarmoissa ympäristöissä, kun taas organisatoriset ja johtamiseen liittyvät tekijät ovat merkittävämpiä erittäin epävarmoissa olosuhteissa (Mikalef ym., 2019). Organisatoristen tekijöiden vahva merkitys tässä tutkielmassa saattaa osittain selittyä sillä, että osa tutkimuksista on toteutettu maissa, joita voidaan pitää erittäin epävarmojen olosuhteiden maina, kuten Iranissa (Maroufkhani ym., 2020) ja Jordaniassa (Lutfi ym., 2022).

Jotta pk-yritykset saavat riittävää tukea massadata-analytiikan käyttöönotossa, on johtajien oltava tietoisia sen avulla saavutettavista hyödyistä. Hyötyjen tiedostamisen on esitetty olevan sellaisenaan riittävää käyttöönottoon kannustamisessa (Lutfi ym., 2022). Johtoasemien ulkopuolella olevien työntekijöiden ja ulkoisten ohjelmistokauppioiden kannattaisi siis pyrkiä lisäämään tietoisuutta ja vähentämään epävarmuutta pk-yritysten johtajien keskuudessa.

Pk-yritysten johtajien pitäisi tarkastella omaksumiseen vaikuttavia tekijöitä ja pitää huoli siitä, että yritys ei itse ole vaikeuttamassa omaksumista. Esimerkiksi helpottavat olosuhteet ovat niin merkittäviä, että niiden avulla voidaan pienentää muiden tekijöiden merkitystä (Cabrera-Sánchez & Villarejo-Ramos, 2019). Helpottaviin olosuhteisiin panostaminen olisi pk-yrityksille viisasta myös siksi, koska samalla voitaisiin parantaa organisatorista valmiutta.

Henkilöstö- ja ohjelmistomarkkinoihin on luonnollisesti vaikeaa vaikuttaa yhden yrityksen voimin, mutta kuten aikaisemmin todettiin, henkilöstö- ja ohjelmisto-ongelmat voivat olla ratkaistavissa ulkoistamisella. Pk-yritysten pitäisi siis tutustua tarjolla oleviin ulkoisiin palveluihin ja pohtia, onko kannattavampaa tehdä massadata-analytiikasta ”in-house” -ratkaisu vai viedä osa siitä yrityksen ulkopuolelle. Toisaalta esimerkiksi Colemanin ym. (2016) esittämä epävarmuus ja turvattomuus voivat vahvistua, jos tiedot eivät ole täysin yrityksen itse hallittavissa.

Lähteet

- Azar, G., & Ciabuschi, F. (2017). Organizational innovation, technological innovation, and export performance: The effects of innovation radicalness and extensiveness. *International Business Review*, 26(2), 324–336.
<https://doi.org/10.1016/j.ibusrev.2016.09.002>
- Baig, M. I., Yadegaridehkordi, E., & Nizam Bin Md Nasir, M. H. (2023). Influence of big data adoption on sustainable marketing and operation of SMEs: A hybrid approach of SEM-ANN. *Management Decision*, 61(7), 2231–2253.
<https://doi.org/10.1108/MD-06-2022-0778>
- Cabrera-Sánchez, J.-P., & Villarejo-Ramos, Á. F. (2019). FACTORS AFFECTING THE ADOPTION OF BIG DATA ANALYTICS IN COMPANIES. *Revista de Administração de Empresas*, 59(6), 415–429. <https://doi.org/10.1590/s0034-759020190607>
- Chartier, T. (2016). Vertigo Over the Seven V's of Big Data. *Journal of Corporate Accounting & Finance*, 27(3), 81–82. <https://doi.org/10.1002/jcaf.22145>
- Che, D., Safran, M., & Peng, Z. (2013). From Big Data to Big Data Mining: Challenges, Issues, and Opportunities. Teoksessa B. Hong, X. Meng, L. Chen, W. Winiwarter, & W. Song (Toim.), *Database Systems for Advanced Applications* (Vsk. 7827, ss. 1–15). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-40270-8_1
- Chen, D. Q., Preston, D. S., & Swink, M. (2015). How the Use of Big Data Analytics Affects Value Creation in Supply Chain Management. *Journal of Management Information Systems*, 32(4), 4–39.
<https://doi.org/10.1080/07421222.2015.1138364>
- Chen, M.-S., Han, J., & Yu, P. S. (1996). Data mining: An overview from a database perspective. *IEEE Transactions on Knowledge and Data Engineering*, 8(6), 866–883. <https://doi.org/10.1109/69.553155>
- Coleman, S., Göb, R., Manco, G., Pievatolo, A., Tort-Martorell, X., & Reis, M. S. (2016). How Can SMEs Benefit from Big Data? Challenges and a Path Forward. *Quality and Reliability Engineering International*, 32(6), 2151–2164.
<https://doi.org/10.1002/qre.2008>
- Curry, E. (2016). The Big Data Value Chain: Definitions, Concepts, and Theoretical Approaches. Teoksessa J. M. Cavanillas, E. Curry, & W. Wahlster (Toim.), *New*

- Horizons for a Data-Driven Economy* (ss. 29–37). Springer International Publishing. https://doi.org/10.1007/978-3-319-21569-3_3
- Dong, J. Q., & Yang, C.-H. (2020). Business value of big data analytics: A systems-theoretic approach and empirical test. *Information & Management*, 57(1), 103124. <https://doi.org/10.1016/j.im.2018.11.001>
- 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>
- Frawley, W. J., Piatetsky-Shapiro, G., & Matheus, C. J. (1992). Knowledge Discovery in Databases: An Overview. *The AI Magazine*, 13(3), 57–70. <https://doi.org/10.1609/aimag.v13i3.1011>
- Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, 35(2), 137–144. <https://doi.org/10.1016/j.ijinfomgt.2014.10.007>
- Ganeshkumar, C., Sankar, J. G., & David, A. (2023). Adoption of Big Data Analytics: Determinants and Performances Among Food Industries. *International Journal of Business Intelligence Research*, 14(1), 1–17. <https://doi.org/10.4018/IJBIR.317419>
- Ghobakhloo, M., Arias-Aranda, D., & Benitez-Amado, J. (2011). Adoption of e-commerce applications in SMEs. *Industrial Management & Data Systems*, 111(8), 1238–1269. <https://doi.org/10.1108/02635571111170785>
- Intel. (2023). *What Is Big Data, and Why Is it Important?* <https://www.intel.com/content/www/us/en/artificial-intelligence/analytics/what-is-big-data.html>
- Jones, M. (2019). What we talk about when we talk about (big) data. *The Journal of Strategic Information Systems*, 28(1), 3–16. <https://doi.org/10.1016/j.jsis.2018.10.005>
- Justy, T., Pellegrin-Boucher, E., Lescop, D., Granata, J., & Gupta, S. (2023). On the edge of Big Data: Drivers and barriers to data analytics adoption in SMEs. *Technovation*, 127, 102850. <https://doi.org/10.1016/j.technovation.2023.102850>
- Lai, Y., Sun, H., & Ren, J. (2018). Understanding the determinants of big data analytics (BDA) adoption in logistics and supply chain management: An empirical investigation. *The International Journal of Logistics Management*, 29(2), 676–703. <https://doi.org/10.1108/IJLM-06-2017-0153>

- Laney, D. (2001). 3D data management: Controlling data volume, velocity and variety. *META Group Research Note*, 6(70), 1.
- Larose, D. T., & Larose, C. D. (2014). *Discovering Knowledge in Data: An Introduction to Data Mining* (1. p.). Wiley.
<https://doi.org/10.1002/9781118874059>
- Lu, R., Zhu, H., Liu, X., Liu, J. K., & Shao, J. (2014). Toward efficient and privacy-preserving computing in big data era. *IEEE Network*, 28(4), 46–50.
<https://doi.org/10.1109/MNET.2014.6863131>
- Lutfi, A. (2022). Factors Influencing the Continuance Intention to Use Accounting Information System in Jordanian SMEs from the Perspectives of UTAUT: Top Management Support and Self-Efficacy as Predictor Factors. *Economies*, 10(4), 75. <https://doi.org/10.3390/economies10040075>
- Lutfi, A., Alsyouf, A., Almaiah, M. A., Alrawad, M., Abdo, A. A. K., Al-Khasawneh, A. L., Ibrahim, N., & Saad, M. (2022). Factors Influencing the Adoption of Big Data Analytics in the Digital Transformation Era: Case Study of Jordanian SMEs. *Sustainability*, 14(3), 1802. <https://doi.org/10.3390/su14031802>
- Management Association, I. R. (Toim.). (2019). *Cloud Security: Concepts, Methodologies, Tools, and Applications*. IGI Global.
<https://doi.org/10.4018/978-1-5225-8176-5>
- Marjani, M., Nasaruddin, F., Gani, A., Karim, A., Hashem, I., Siddiqa, A., & Yaqoob, I. (2017). Big IoT Data Analytics: Architecture, Opportunities, and Open Research Challenges. *IEEE Access*, 5, 5247–5261.
<https://doi.org/10.1109/ACCESS.2017.2689040>
- Maroufkhani, P., Iranmanesh, M., & Ghobakhloo, M. (2023). Determinants of big data analytics adoption in small and medium-sized enterprises (SMEs). *Industrial Management & Data Systems*, 123(1), 278–301. <https://doi.org/10.1108/IMDS-11-2021-0695>
- Maroufkhani, P., Tseng, M.-L., Iranmanesh, M., Ismail, W. K. W., & Khalid, H. (2020). Big data analytics adoption: Determinants and performances among small to medium-sized enterprises. *International Journal of Information Management*, 54, 102190. <https://doi.org/10.1016/j.ijinfomgt.2020.102190>
- Microsoft. (2013). The Big Bang: How the Big Data Explosion Is Changing the World. *Microsoft News Center*. <https://news.microsoft.com/2013/02/11/the-big-bang-how-the-big-data-explosion-is-changing-the-world/>

- Mikalef, P., Boura, M., Lekakos, G., & Krogstie, J. (2019). Big data analytics and firm performance: Findings from a mixed-method approach. *Journal of Business Research, 98*, 261–276. <https://doi.org/10.1016/j.jbusres.2019.01.044>
- Moraes, G. H. S. M. D., Pelegrini, G. C., De Marchi, L. P., Pinheiro, G. T., & Cappelozza, A. (2022). Antecedents of big data analytics adoption: An analysis with future managers in a developing country. *The Bottom Line, 35*(2/3), 73–89. <https://doi.org/10.1108/BL-06-2021-0068>
- Müller, O., Fay, M., & Vom Brocke, J. (2018). The Effect of Big Data and Analytics on Firm Performance: An Econometric Analysis Considering Industry Characteristics. *Journal of Management Information Systems, 35*(2), 488–509. <https://doi.org/10.1080/07421222.2018.1451955>
- Oliveira, T., & Martins, M. F. (2011). *Literature Review of Information Technology Adoption Models at Firm Level, 14*(1).
- Raguseo, E. (2018). Big data technologies: An empirical investigation on their adoption, benefits and risks for companies. *International Journal of Information Management, 38*(1), 187–195. <https://doi.org/10.1016/j.ijinfomgt.2017.07.008>
- Ramdani, B., Kawalek, P., & Lorenzo, O. (2009). Predicting SMEs' adoption of enterprise systems. *Journal of Enterprise Information Management, 22*(1/2), 10–24. <https://doi.org/10.1108/17410390910922796>
- Reuters. (2018). Amazon scraps secret AI recruiting tool that showed bias against women. *Reuters*. <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G>
- SAS. (2023). *Big Data: What it is and why it matters*. https://www.sas.com/en_us/insights/big-data/what-is-big-data.html
- Sen, D., Ozturk, M., & Vayvay, O. (2016). An Overview of Big Data for Growth in SMEs. *Procedia - Social and Behavioral Sciences, 235*, 159–167. <https://doi.org/10.1016/j.sbspro.2016.11.011>
- Siddiq, A., Karim, A., & Gani, A. (2017). Big data storage technologies: A survey. *Frontiers of Information Technology & Electronic Engineering, 18*(8), 1040–1070. <https://doi.org/10.1631/FITEE.1500441>
- Smith, G. (2020). Data mining fool's gold. *Journal of Information Technology, 35*(3), 182–194. <https://doi.org/10.1177/0268396220915600>
- Sun, S., Cegielski, C. G., Jia, L., & Hall, D. J. (2018). Understanding the Factors Affecting the Organizational Adoption of Big Data. *Journal of Computer*

Information Systems, 58(3), 193–203.

<https://doi.org/10.1080/08874417.2016.1222891>

Tornatzky, L. G., & Fleischer, M. (1990). *The processes of technological innovation* (4. print). Lexington Books.

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>

Venkatesh, Morris, Davis, & Davis. (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly*, 27(3), 425. <https://doi.org/10.2307/30036540>

Verma, S., & Chaurasia, S. (2019). Understanding the Determinants of Big Data Analytics Adoption: *Information Resources Management Journal*, 32(3), 1–26. <https://doi.org/10.4018/IRMJ.2019070101>

Ylijoki, O., & Porras, J. (2016). Perspectives to Definition of Big Data: A Mapping Study and Discussion. *Journal of Innovation Management*, 4(1), 69–91. https://doi.org/10.24840/2183-0606_004.001_0006

Youssef, M. A. E.-A., Eid, R., & Agag, G. (2022). Cross-national differences in big data analytics adoption in the retail industry. *Journal of Retailing and Consumer Services*, 64, 102827. <https://doi.org/10.1016/j.jretconser.2021.102827>