

Kaisa Rantanen

**BIG DATAN HYÖDYNTÄMINEN JULKISELLA
SEKTORILLA**



JYVÄSKYLÄN YLIOPISTO
INFORMAATIOTEKNOLOGIAN TIEDEKUNTA
2018

TIIVISTELMÄ

Rantanen, Kaisa

Big datan hyödyntäminen julkisella sektorilla

Jyväskylä: Jyväskylän yliopisto, 2018, 35 s.

Tietojärjestelmätiede, kandidaatin tutkielma

Ohjaaja(t): Clements, Kati

Big data on johtavien teknologiatrendien asemassa ja yksi keskeisistä monitieteellisistä tutkimusaiheista. Sen mahdollisuudet on tunnistettu sekä julkisella että yksityisellä sektorilla, mutta yksityinen sektori on big data -analytiikan käyttöönotossa selvässä johtoasemassa. Tässä tutkimuksessa pyritään selvittämään kirjallisuuskatsauksen keinoin, kuinka julkisen sektorin erityispiirteet vaikuttavat julkisen sektorin big datan hyödyntämismahdollisuuksiin. Aihetta käsitellään vertailemalla julkisen ja yksityisen sektorin erityispiirteitä ja big datan hyödyntämiskohteita toisiinsa. Tutkimuksessa big datan hyödyntämiskohteet jaetaan kolmeen osa-alueeseen, jotka ovat päätöksenteon tukeminen, strateginen kehittäminen ja prosessien kehittäminen. Tutkimuksessa havaitaan, että julkisella sektorilla big dataa voidaan hyödyntää jokaisella osa-alueella, mutta big datan hyödyntäminen korostuu erityisesti prosessien kehittämisessä. Tutkimuksessa todettiin, että big data voidaan hyödyntää laajasti julkisella sektorilla, sillä big data -analytiikan hyödyntämiskohteita löydettiin jokaiselta tunnistetulta toimialueelta. Lisäksi tutkimuksessa käsitellään big datan hyödyntämisen keskeisiä haasteita julkisella sektorilla.

Asiasanat: big data, big data -analytiikka, julkinen sektori

ABSTRACT

Rantanen, Kaisa

Big data in public sector

Jyväskylä: University of Jyväskylä, 2018, 35 pp.

Information Systems, Bachelor's Thesis

Supervisor(s): Clements, Kati

Big data is one of the leading technology trends and it is a main subject for many multidisciplinary studies. Opportunities regarding big data have been identified in both public and private sector, yet the private sector has a clear leading role compared to the public sector in implementing big data analytics. This thesis is a literature review that seeks to examine how unique characteristics of the public sector affect the opportunities of exploiting big data in public sector. The subject is studied by comparing differences between public and private sector in general and in the context of big data applications. In this research big data applications were divided into three categories, which are decision supporting, strategic development and process development. It is discovered that in the public sector big data can be utilized in all three categories but is mainly used to improve process development. It is also found that big data could be used in various domains of the public sector since big data applications were discovered in every identified domain. Key challenges regarding big data applications in the public sector are also studied in this research.

Keywords: big data, big data analytics, public sector

TAULUKOT

TAULUKKO 1 Julkisen ja yksityisen sektorin erityispiirteet.....	15
TAULUKKO 2 Big datan hyödyntäminen julkisella ja yksityisellä sektorilla....	18

SISÄLLYS

1	JOHDANTO.....	6
2	BIG DATA.....	9
	2.1 Big datan määritelmä.....	9
	2.2 Big datan hyödyntäminen.....	11
3	JULKINEN SEKTORI.....	13
	3.1 Julkisen ja yksityisen sektorin eroavaisuudet.....	13
	3.2 Julkisen sektorin toimialueet.....	16
4	BIG DATAN HYÖDYNTÄMINEN JULKISELLA SEKTORILLA.....	17
	4.1 Päätöksenteon tukeminen.....	20
	4.2 Strateginen kehittäminen.....	21
	4.3 Prosessien kehittäminen.....	23
	4.4 Big datan hyödyntämisen haasteet.....	27
5	YHTEENVETO.....	29

1 JOHDANTO

Big data on yksi keskeisistä tutkimusaiheista sekä tietojärjestelmätieteen, johtamisen että yhteiskuntatieteen tutkimuksessa (Constantiou & Kallinikos, 2005) ja big data on noussut johtavien teknologiatrendien asemaan 2010 -luvulla (Chen, Chiang & Storey, 2012). Yleisesti big datalla tarkoitetaan sellaisia data-massoja, joiden analysointiin perinteiset työkalut eivät riitä (Chen, Mao & Liu, 2014). Big datalla on useita eri määritelmiä ja tunnistettuja määrittäviä tekijöitä. Keskeisimmäksi big dataa määrittäväksi piirteeksi on tyypillisesti mielletty datan suuruus (ks. esim. Desouza & Jacob, 2017; Gandomi & Haider, 2015; Lycett, 2013; McAfee & Brynjolfsson, 2012). Big data voidaan myös nähdä ilmiönä, jossa on kyse erilaisten datasettien yhdistelemisestä uusien analytiikkamenetelmien avulla ja jossa dataa käytetään uusissa tarkoituksissa hyväksi (Klievink, Romijn, Cunningham & de Bruijn, 2017).

Jokaisella toimialalla kerätään suuret määrät dataa (LaValle, Lesser, Shockley, Hopkins & Kruschwitz, 2011) ja big data -analytiikalle on tunnistettu useita hyödyntämiskohteita yksityisellä sektorilla. Yksityisen sektorin yritykset käyttävät big dataa sekä kustannustehokkuuden parantamiseen että liikevaihdon kasvattamiseksi maksimoidakseen yrityksen kannattavuuden (LaValle ym., 2011). Big data -analytiikalla voidaan myös esimerkiksi tukea yrityksen päätöksentekoa ja parantaa yritysten suorituskykymittausta kehittämiskohteiden tunnistamiseksi (McAfee & Brynjolfsson, 2012). On havaittu, että mitä enemmän yrityksen strategisessa kehittämisessä hyödynnetään analytiikkaa, sitä tehokkaampia ja tuottoisampia yritykset ovat (McAfee & Brynjolfsson, 2012). Big data -analytiikka voi muuttaa jopa koko yrityksen liiketoimintamallin analytiikkaan perustuvaksi (Günther, Mehrizi, Huysman & Feldberg, 2017; Lycett, 2013).

Big datan mahdollisuudet on tunnistettu yksityisen sektorin lisäksi myös julkisella sektorilla (ks. esim. Günther ym., 2017; Kim, Trimi & Chung, 2014). Vaikka big datan mahdollisuudet on tunnistettu, on kuitenkin huomattu, että big datan hyödyntäminen on julkisella sektorilla selkeästi yksityistä sektoria jäljessä (Desouza & Jacob, 2017; Klievink ym., 2017). Valtiovarainministeriön julkaisun mukaan analytiikka ja big data ovat merkittävimmät valtion tietohallinnon kehitykseen vaikuttavat tekijät viime vuosina (Hattinen ja Kiviniemi,

2017). Silti laajaa ja kokoavaa selvitystä big datan hyödyntämismahdollisuuksista julkisella sektorilla ei löydy. Tässä tutkimuksessa pyritään etsimään vastaus yhteen tutkimuskysymykseen: “Kuinka julkisen sektorin erityispiirteet vaikuttavat big datan hyödyntämiseen julkisella sektorilla?”. Tutkimuskysymystä lähestytään ensin vertailemalla julkisen ja yksityisen sektorin eroavaisuuksia yleisesti. Tämä auttaa julkisen sektorin erityispiirteiden hahmottamisessa ja tarjoaa pohjan big datan hyödyntämiskohteiden tutkimiselle erityispiirteiden näkökulmasta. Tämän jälkeen tutkitaan tunnistettujen erityispiirteiden vaikutuksia big datan hyödyntämiskohteisiin julkisella sektorilla.

Tutkimus on toteutettu kirjallisuuskatsauksena. Kirjallisuutta on haettu käyttämällä tietokantana Google Scholaria, koska se hakee julkaisut koostaen useista eri tietokannoista. Lähdemateriaalia on etsitty käyttäen sekä suomenkielisiä että englanninkielisiä hakusanoja, painottuen englanninkielisiin hakusanoihin. Tutkimuksen tärkeimmät hakusanat ovat “big data”, “big data analytics”, “julkinen sektori” ja “public sector”. Tärkeimpien hakusanojen lisäksi julkaisuja haettiin aihepiireittäin, joita etsittiin tarkentavilla hakusanoilla kuten “decision making”, “process management” ja “strategic planning”. Lähteinä on käytetty myös julkaisuja, jotka esiintyvät käytetyillä hakusanoilla löydettyjen julkaisujen lähdeluetteloista.

Lähteitä on rajattu sekä julkaisuvuoden että tekstin julkaisufoorumin mukaan. Big data -analytiikkaa käsittelevään kirjallisuuteen on valittu vuoden 2010 jälkeen julkaistua materiaalia, sillä analytiikan katsotaan nousseen tällöin johtavien teknologiatrendien asemaan (Chen ym., 2012) ja kiinnostus big dataa kohtaan on huomattu nousevan tällöin (Wamba, Akter, Edwards, Chopin & Gnanzou, 2015). Tätä rajausta on käytetty big dataa käsittelevissä kirjallisuuskatsauksissa aikaisemminkin (Günther ym., 2017). Tuoreeseen kirjallisuuteen keskittyminen mahdollistaa ilmiön nykytilanteen tarkastelun. Julkista sektoria käsittelevää kirjallisuutta ei rajattu julkaisuvuoden perusteella, sillä katsotaan, että julkisen sektorin määritelmä ja rooli yhteiskunnassa ei ole muuttunut oleellisesti viime vuosikymmenten aikana. Valittujen lähteiden julkaisukanava on arvoستeltu Julkaisufoorumin luokittelun mukaan vähintään perustason julkaisuksi. Tällä rajauksella pyritään parantamaan lähdemateriaalin akateemista uskottavuutta. Poikkeuksena tähän rajaukseen on Liikenne- ja viestintäministeriön Big datan käyttö -työryhmän julkaisema raportti (Rastas & Asp, 2014), jonka katsotaan olevan käyttökelpoinen lähde julkisen sektorin big datan hyödyntämistä arvioitaessa. Lisäksi raportin havaintoja tuetaan tieteellisten julkaisujen havainnoilla.

Tutkimus etenee noudattaen seuraavaa rakennetta. Toisessa luvussa esitellään big datan määritelmä ja big datan lähteet ja analytiikan tyypillisimmät menetelmät. Kolmas luku käsittelee julkista sektoria. Julkinen sektori pyritään kuvaamaan kattavasti ja ymmärrettävästi käsittelemällä sen eroja yksityiseen sektoriin verrattuna. Vertailun pohjalta voidaan tunnistaa julkisen sektorin erityispiirteet, joita voidaan myöhemmin käyttää big datan hyödyntämiskohteiden tutkimuksessa. Lisäksi määritellään julkisen sektorin toimialueet, jotta saadaan selkeä käsitys julkisen sektorin yhteiskunnallisesta merkityksestä. Neljän-

nessä luvussa pyritään vastaamaan tutkimuskysymykseen, ja luku esittelee big datan hyödyntämiskohteita sekä julkisella että yksityisellä sektorilla. Luvussa keskitytään erityisesti siihen, kuinka julkisen sektorin erityispiirteet vaikuttavat big data -analytiikan hyödyntämiskohteiden ilmenemiseen. Luku on jaettu kolmeen alalukuun big datan hyödyntämiskohteiden osa-alueen mukaan, joita ovat päätöksenteon tukeminen, prosessien kehittäminen ja strateginen kehittäminen. Koska big datan hyödyntämistä käsittelevässä kirjallisuudessa korostuu mahdollisuuksien lisäksi myös haasteet, luvun viimeisessä alaluvussa esitellään keskeisiä tunnistettuja hyödyntämisen haasteita. Haasteita käsitellään erityisesti julkisen sektorin näkökulmasta pohtien erityispiirteiden merkitystä haasteiden ilmenemiseen. Viimeisessä luvussa tehdään yhteenveto tutkimuksesta, jossa kerrataan tutkimuksen tärkeimmät tulokset ja esitellään tunnistetut jatkotutkimusaiheet.

2 BIG DATA

Termille big data suomenkielinen vastine on massadata. Englanninkielinen termi on kuitenkin vakiintunut käytettäväksi myös suomenkielisissä julkaisuissa (Rastas & Asp, 2014), ja tässä tutkimuksessa käytämme ilmiöstä termiä big data. Big data on terminä lähes kymmenen vuotta vanha ja tieteellisissä julkaisuissa sen määrittely ei aina ole ollut yksiselitteistä. Osa määritelmistä keskittyy big datan termin määrittämiseen, toiset taas korostavat big datan käyttökohteita ja sillä saavutettavia hyötyjä. (Gandomi & Haider, 2015) Gandomi ja Haider (2015) huomauttavat, että big datan käsitteen määrittäminen on jäänyt akateemisen kirjallisuuden piirissä vajavaiseksi samalla, kun tiedeyhteisön ulkopuolinen media on julkaissut aiheesta runsaasti kirjallisuutta. Suurimmalla osalla näistä julkaisuista pyritään markkinoimaan suurien ohjelmistoyritysten data-analytiikkaan käytettäviä järjestelmiä (Gandomi & Haider, 2015), ja selkeää ja laajamittaista kuvausta big datan tieteellisesti todistetuista hyödyistä on puutteellista. Big datalla on nähty olevan valtava määrä potentiaalia, mutta edelleen on olemassa vain rajallinen määrä tutkimusta siitä, kuinka big datan potentiaalia voidaan hyödyntää ja näin luoda arvoa organisaatioille (Günther ym., 2017). Tässä luvussa tutustutaan big datan määrittelmään ja erityisominaisuuksiin sekä eritellään big datan erilaisia lähteitä ja hyödyntämiskohteita.

2.1 Big datan määritelmä

Chen ja kollegat (2014) luonnehtivat artikkelissaan big dataa sellaisiksi datamassoiksi, joita ei pystytä käsittelemään perinteisillä analysointityökaluilla tarpeeksi tehokkaasti ja nopeasti. Jotta data luokitellaan big dataksi, tulee datan täyttää tiettyjä kriteerejä (Gandomi & Haider, 2015), joiden lukumäärästä ja laadusta on eriäviä näkemyksiä alan kirjallisuudessa. Tässä luvussa paneudutaan big datan erilaisiin määritelmiin.

Tyypillisesti big datan on hahmotettu rakentuvan kolmesta määrittävästä ominaisuudesta, jotka kuvaavat sekä big datan erityispiirteitä että big data -

analytiikkaan liittyviä haasteita. Nämä ominaisuudet ovat datan määrä (volume), datan kertymisnopeus (velocity) sekä datan monimuotoisuus (variety). Näitä kriteerejä kutsutaan big datan kolmen V:n malliksi, (Gandomi & Haider, 2015; Lycett, 2013) Datan määrän tulee olla suhteellisen valtava ollakseen big dataa ja sen määrä ilmoitetaan usein tera- ja petatavuissa. Datan kertymisnopeus on lisääntynyt viimeisten vuosien aikana uusien teknologioiden sekä uusien tietolähteiden myötä. (Gandomi & Haider, 2015) Big datan monimuotoisuudesta kertoo se, että vain pieni osa on perinteistä strukturoitua dataa (Gandomi & Haider, 2015), jolla on nimensä mukaisesti selkeä rakenne kuten taulukon sarakkeet ja rivit (Desouza & Jacob, 2017). Suurin osa big datasta onkin joko osittain strukturoitua tai strukturoimatonta dataa (Chen ym., 2014). Osittain strukturoitu data pitää sisällään strukturoitua dataa avoimempia muotoja. Strukturoimattoman data on esimerkiksi videoita, vapaata tekstiä tai kuvia. (Desouza & Jacob, 2017) Kolmen V:n mallia on käytetty eniten big dataa käsittelevässä kirjallisuudessa ja tutkimuksessa (ks. esim. Chen ym., 2014; Desouza & Jacob, 2017; Gandomi & Haider, 2015; Lycett, 2013; McAfee & Brynjolfsson, 2012) ja kyseinen määritelmä kiteyttää big datan ensimmäisenä tunnistetut erot perinteisestä datasta.

Myöhemmin big datan kolmen V:n malli on nähty riittämättömäksi ja määritelmään on lisätty neljäs määrittävä kriteeri: datan arvokkuus (value). Big datan on arvokasta vasta silloin, kun sitä hyödynnetään organisaation toiminnoissa käyttäen soveltuvia analytiikkamenetelmiä. (Lycett, 2013). Datan arvokkuuden nostaminen big datan määritelmään korostaa erityisesti sitä, mikä big datassa nähdään tärkeänä, eli liiketoiminnan hyödyn saavuttaminen analytiikalla. Viimeisimmäksi big datan tutkimuksessa on tunnistettu viides määrittävä ominaisuus: datan todenmukaisuus (veracity). Jotta big dataa voidaan käyttää liiketoiminnan tukena, tulee liikkeenjohdon voida luottaa kerättyyn tietoon (Wamba ym, 2015). Koska reaaliaikaista dataa, jollaista big data tyypillisesti on, on helppo manipuloida ja väärinkäyttöksiä hankala havaita (Desouza & Jacob, 2017), on datan todenmukaisuus tärkeää nostaa esille big datan erityispiirteisiin ja analytiikkaan liittyviin haasteisiin.

Big data voidaan määritellä myös sosio-kulttuurisena ilmiönä, joka syntyy teknologian, analyysin ja mytologian vuorovaikutuksesta. Big datassa ajatellaan ilmiönä olevan enemmän kyse asenteiden muutoksesta analyysistä kohtaan, kuin varsinaisen datan piirteistä ja niiden muutoksesta. Teknologisesta näkökulmasta big datan käsittelyyn tarvitaan maksimaalista laskentatehoa sekä tarkat algoritmit. Analyysin näkökulma ottaa huomioon sen, kuinka suurista datamääristä voidaan tunnistaa lainalaisuuksia, joiden perusteella voidaan muodostaa väitteitä ympäristöstä. Ilmiöön liittyy myös mytologiaa siitä, kuinka big datasta tehtyjen analyysien perusteella voidaan saada tarkka, totuudenmukainen ja objektiivinen kuva ympäröivästä maailmasta ja näin tehdä parempia päätöksiä. (Boyd & Crawford, 2012)

Klievink ja kollegat (2017) ovat määrittäneet big datan keskittyen sen eroavaisuuksiin perinteisestä datasta. Määritelmässään he ovat ottaneet huomioon sekä datan lähteen että sen käyttötarkoituksia. Ensimmäisenä big datan

piirteinä he mainitsevat useiden suurien datasettien yhdistelemisen. Dataa kerätään sekä ulkoisista että sisäisistä lähteistä. Toiseksi he ovat nostaneet esille sen, kuinka big datan ilmiössä korostuu strukturoidun ja strukturoimattoman datan analyysi. Kolmas erottava tekijä normaaliin dataan nähden on se, että big data kerääntyy reaaliajassa. Myös uusien analytiikkamenetelmien hyödyntäminen on nähty big datan erottavaksi tekijäksi. Viidenneksi big datan ilmiötä määrittäväksi tekijäksi on nostettu se, kuinka olemassa olevaa dataa käytetään uusissa tarkoituksissa hyväksi eli käytetään dataa johonkin muuhun, kun se on alun perin kerätty. (Klievink ym., 2017) Tämä määritelmä korostaa suurten datamäärien käytännön merkitystä organisaatioille.

2.2 Big datan hyödyntäminen

Big dataa voidaan kerätä useista eri lähteistä. Chen ja kollegat (2014) tunnistivat artikkelissaan kolme keskeisintä big datan lähdettä, jotka ovat yrityksen sisäisten järjestelmien keräämä data, esineiden internetin tuottama data sekä biolääketieteellinen data. Uusien teknologioiden myötä yritykset keräävät järjestelmillään jatkuvasti suurten määrän monimuotoista, reaaliaikaista dataa (Gandomi & Haider, 2015) muun muassa tuotteiden valmistamisesta, kaupankäynnistä (Chen ym., 2014) ja verkkosivuilla liikkumisesta tehdyistä analyysistä (McAfee & Brynjolfsson, 2012). Esineiden internetillä tarkoitetaan ilmiötä, jossa fyysisen ympäristön esineisiin, on lisätty teknologiaa, joka mahdollistaa datan keräämisen ja jakamisen internetin välityksellä (Kopetz, 2011). Nämä älykkään esineet (Kopetz, 2011) muodostavat esineiden internetin ja keräävät suuren määrän esimerkiksi paikannustietoa tai liikkuvaa kuvaa (Chen ym., 2014). Biolääketieteellistä big dataa kertyy esimerkiksi geenitutkimuksen ja kliinisen terveydenhuollon seurauksena (Chen ym., 2014). Näiden kolmen kategorian lisäksi on tunnistettu myös muita big datan lähteitä. Esimerkiksi sosiaalinen media ja mobiililaitteet keräävät suuren määrän paikkaan ja aikaan sidottua dataa käyttäjistään (McAfee & Brynjolfsson, 2012; Wamba ym, 2015).

Ennen kuin big dataa voidaan hyödyntää organisaatiossa, tulee dataa käsitellä asianmukaisesti. Big datan käsittelyprosessissa on tunnistettu kaksi päävaihetta, datan hallinta ja datan analysointi. Datan hallinta käsittää prosessissa big datan hankinnan ja keräämisen, sen muokkaamisen ja kommentoinnin sekä datan yhdistämisen, keräämisen ja muuttamisen esitettävään muotoon. Datan analysoinnin vaiheessa big datasta tehdään malleja ja analyysijä, käyttäen eri menetelmiä. Analysoinnin vaiheeseen kuuluu myös näiden mallien ja analyysien tulkinta. (Labrinidis & Jagadish, 2012)

Analyysijä voidaan tehdä big datasta useilla eri menetelmillä. Gandomi ja Haider (2015) ovat tunnistaneet artikkelissaan big datan analytiikkamenetelmiksi tekstianalytiikan, audioanalytiikan, videoanalytiikan sekä sosiaalisen median analytiikan. Tekstianalytiikan keinoin voidaan kerätä strukturoitua dataa tai tehdä tiivistelmiä suurista määristä tekstiä. Audioanalytiikkaa käytetään esimerkiksi puhelutallenteiden analysoinnissa. Videoanalytiikan avulla voidaan

automatisoida esimerkiksi turvakameroiden tallentaman materiaalin analysointia. Ongelmana videoanalytiikassa on videoiden suuri koko ja niiden käsittelemiseen vaadittavat resurssit. Sosiaalisen median analytiikkaan voidaan tehdä joko sisältöpohjaisesti tai rakennepohjaisesti. Sisältöpohjaisessa analyysissä käytetään teksti-, audio- ja videoanalytiikan menetelmiä. Rakennepohjainen analyysi keskittyy siihen, kuinka eri toimijat ovat suhteessa toisiinsa sosiaalisessa mediassa. (Gandomi & Haider, 2015) Nämä eri analytiikkamenetelmien kategoriat keskittyvät datan alkuperään sen hyödyntämiskohteiden sijaan.

Rajaraman (2016) on erotellut eri big datan analytiikkamenetelmät sen mukaan, mihin analyysijä on tarkoitus käyttää. Hänen mukaansa big datasta voidaan tehdä kuvailevaa, ennustavaa, tutkivaa tai ohjaavaa analytiikkaa. Kuvailevan analytiikan tavoitteena on kuvata historiaa tai nykytilannetta helposti ymmärrettävässä muodossa. Ennustavalla analytiikalla pyritään ennustamaan tulevaisuutta kerätyn datan perusteella. Mikäli halutaan löytää datasta uusia yhteyksiä tai syy-seuraussuhteita, käytetään big dataa hyväksi tutkivaan analytiikkaan. Ohjaava analytiikka taas toimii päätöksenteon tukena ja auttaa esimerkiksi organisaatioita optimoimaan toimintaansa. (Rajaraman, 2016)

Organisaatiot käyttävät analytiikkaa eri tavalla riippuen siitä, kuinka konkreettisesti analytiikka on otettu käyttöön organisaatiossa. LaValle ja kollegat (2011) tunnistivat organisaatioiden big data analytiikkaa tutkivassa tutkimuksessaan kolme analytiikan käyttöönoton tasoa: pyrkivät (aspirational), harjaantuneet (experienced) ja muuntautuneet (transformed) organisaatiot. Pyrkivät organisaatiot ovat analytiikan käyttöönoton alkuvaiheessa. Muuntautuneissa organisaatioissa dataohjautuvuus on omaksuttu täysin ja sitä hyödynnetään kaikessa organisaation toiminnassa. Harjaantuneet organisaatiot sijoittuvat analytiikan käyttöönotossa edellisten välille, ja analytiikka on otettu käyttöön osassa organisaation toimintaan. Analytiikan käyttöönoton taso vaikuttaa muun muassa siihen, kuinka organisaatio käyttää analytiikkaa päätöksenteon tukena, prosessien kehittämisessä ja strategisessa kehittämisessä.

3 JULKINEN SEKTORI

Julkista sektoria käsittelevää tutkimusta tarkasteltaessa vaikuttaa siltä, että sekä julkisen sektorin määritelmää että julkisen sektorin toimialueita pidetään itsensänselvyytenä. Useissa julkista sektoria käsittelevissä tieteellisissä julkaisuissa julkista sektoria tai sen vastuualueita ei ole lainkaan määritely. Yleisesti julkisella sektorilla tarkoitetaan valtion omistamaa ja hallitsemaa osaa taloudesta, jota voidaan pitää yhteiskunnan perustana (Broadbent ja Guthrie, 1992). Tässä luvussa käsitellään julkisen sektorin erityispiirteitä verrattuna yksityiseen sektoriin sekä eritellään julkisen sektorin vastuualueita yhteiskunnassa selkeän kokonaiskuvan luomiseksi julkisen sektorin toiminnasta.

3.1 Julkisen ja yksityisen sektorin eroavaisuudet

Julkinen ja yksityinen sektori eroavat toisistaan usealla toiminnan osa-alueella. Tässä alaluvussa tutustutaan sektorien erityispiirteisiin tavoitteen, rahoituksen, päätöksenteon, strategisen kehittämisen ja prosessien kehittämisen näkökulmasta. Tärkeimmät tulokset julkisen ja yksityisen sektorien eroavaisuuksista on koottu taulukkoon (Taulukko 1) tämän alaluvun loppuun.

Julkiseen sektoriin kuuluu ne organisaatiot, jotka hoitavat julkisia tehtäviä (Harisalo, Aarrevaara, Stenval & Virtanen, 2007). Julkisen sektorin toimintaa ohjaa sosiaalisen ilmapiirin kohentaminen (Lachman, 1985), valtakunnallisen järjestyksen ylläpitäminen ja kansalaisten perusoikeuksien takaaminen (Kim ym., 2014). Yksityisellä sektorilla taas keskitytään yrityksen edun ja tuoton tavoitteluun (Lachman, 1985) sekä kilpailuedun saavuttamiseen (Kim ym., 2014). Julkisen sektorin toimijat voidaan siis nähdä edustavan kansalaisten oikeuksia (Kickert, Klijn ja Koppenjan, 1997; Lachman, 1985; Perry & Wise, 1990). Myös julkisen sektorin rahoitus eroaa yksityisestä sektorista, sillä ne ovat keskushallinnon säätelemiä (Broadbent & Guthrie, 1992). Julkisen sektorin palvelut rahoitetaan verovaroin, jotka kerätään esimerkiksi ansiotyötä tekeville (Andersen, 2008). Yksityisellä sektorilla rahoitus perustuu liikevoittoon (Kim ym., 2014).

Nutt (2005) huomauttaa, että yksityisen ja julkisen sektorin välillä on havaittavissa eroja organisaatioiden päätöksentekoprosessien suhteen. Prosessien erojen ajatellaan johtuvan pääasiassa siitä, että sektoreilla on erilaiset roolit yhteiskunnassa. Julkisella sektorilla päätöksenteon tavoitteena on kansalaisten hyvinvointi ja julkisten linjausten laatiminen. (Nutt, 2005) Julkisen sektorin toiminnalla pyritään löytämään ratkaisuja suuriin valtiota koskeviin ongelmiin (Kickert ym., 1997) ja päätöksenteko liittyy usein niin sanottuihin viheliäisiin ongelmiin (Desouza & Jacob, 2017). Viheliäisillä ongelmilla tarkoitetaan sellaisia ongelmia, joihin ei ole yksiselitteistä ratkaisua. Ne voivat liittyä esimerkiksi köyhyyteen, kodittomuuteen, turvallisuuteen ja kestäväen kehityksen luomiseen. (Desouza & Jacob, 2017) Päätöksenteko yksityisellä sektorilla perustuu yrityksen ja sen sidosryhmien taloudellisen hyödyn tavoitteluun. On myös huomattu, että julkisella sektorilla päätöksentekoprosessiin osallistuu enemmän henkilöitä, kuin yksityisellä sektorilla, eikä päättävillä henkilöillä välttämättä ole substanssiosaamista juuri kyseiseltä alueelta. Tänä päivänä julkisen sektorin päätöksentekoa ajaa kaksi eri päämäärää, kustannustehokkuus ja kansalaisten hyvinvointi, mikä tekee päätöksenteosta entistä monimutkaisempaa. (Nutt, 2005) Julkisen sektorin päätöksiä tehdään demokratian valvonnassa (Harisalo ym., 2007).

Julkinen sektori eroaa yksityisen sektorin toimijoista myös strategisen kehittämisen näkökulmasta. Olsenin ja Eadien (1982, 4) mukaan julkisella sektorilla strategisella kehittämisessä tarkoitetaan niitä toimintoja, joilla pyritään luomaan keskeisiä suuntaviivoja tuleville toimille (Bryson & Roering, 1987). Näissä päätöksissä tulee ottaa huomioon perustuslailliset arvot. Yleisesti strategista kehittämistä tehdessä on otettava huomioon sekä johdon näkökulma että eri sidosryhmät. Lisäksi pyritään ottamaan huomioon mahdollisimman kattavasti organisaation ulkopuoliset uhat ja mahdollisuudet sekä arvioimaan sisäisiä vahvuuksia ja heikkouksia. (Bryson & Roering, 1987) Myös mahdollisiin kilpailijoihin kiinnitetään erityistä huomiota strategista kehittämistä tehdessä (Bryson & Roering, 1987), minkä voidaan ajatella korostuvan erityisesti yksityisellä sektorilla. Voidaan ajatella, että strategisella kehittämisessä pyritään sovittamaan organisaatio mahdollisimman hyvin toimintaympäristöönsä. Yksityisen ja julkisen sektorin strateginen kehittäminen eroaa niin ikään suunnittelun tavoitteen näkökulmasta. Yrityksen strategisessa kehittämisessä keskiössä on organisaation suorituskyky ja tehokkuus (Bryson & Roering, 1987). Myös julkisella sektorilla strateginen kehittäminen pyrkii parantamaan suorituskykyä, mutta suunnittelun keskeisimpänä lähtökohtana on kansalaiset ja kansalaisille tarjottavat palvelut, kuten liikenne tai terveydenhuolto. Strategisessa kehittämisessä voidaan keskittyä moneen eri asiaan, kuten organisaation arvoihin, tuotteiden tai palveluiden kehittämiseen, asiakassuhteisiin tai rahoitukseen. (Bryson & Roering, 1987) Julkisen sektorin strategisessa kehittämisessä korostuu erityisesti kehittämisen motiivien tarkastelu. Suunnittelussa tulee ottaa huomioon poliittinen, laillinen, eettinen ja ympäristöllinen näkökulma. (Bryson, 2018, s. 35)

Julkisella sektorilla myös prosessien hallinta eroaa yksityisestä sektorista. Liiketoimintaprosessien hallinnalla (business process management) tarkoitetaan tyypillisesti niiden prosessien dokumentointia ja kehittämistä, jossa valmistetut

tuotteet tai palvelut toimitetaan asiakkaille. Prosesseihin kuuluu esimerkiksi tuotannon suunnittelu ja toteutus, tuotteiden ja palveluiden jakelu tai teknisen tuen tarjoaminen. (Gulledge & Sommer, 2002) Harmonin (2003) mukaan liiketoimintaprosessien muutoksella tarkoitetaan prosessien uudelleensuunnittelua organisaation suorituskyvyn parantamiseksi (Stemberger & Jaklic, 2007). On huomattu, että julkisella sektorilla prosessien muutokset ovat usein hitaampia kuin yksityisellä sektorilla, sillä organisaatorakenne on usein kankeampi ja resurssit niukat. Julkisella sektorilla prosessien uudelleensuunnittelussa motiivina on suorituskyvyn parantamisen lisäksi asiakastytyväisyys ja prosessien yksinkertaistaminen asiakkaiden näkökulmasta. Asiakastytyväisyys on siis nähty tärkeänä myös julkisella sektorilla, vaikka ei pyritä yksityisen sektorin tapaan uusien asiakkaiden saamiseen tai vanhojen pitämiseen. (Stemberger & Jaklic, 2007)

Ero julkisen ja yksityisen sektorin välillä voidaan nähdä kaventuneen. Esimerkiksi Vigoda (2002) huomauttaa, että kansalaisiin tulisi suhtautua samalla tavalla kuin yksityisellä sektorilla asiakkaisiin. Palveluita tulisi kehittää heiltä kerätyn palautteen mukaan. On myös huomattu, että julkisella sektorilla kiinnitetään entistä tarkemmin huomiota kumppanuussuhteiden luomiseen sidosryhmien kanssa (Vigoda, 2002), mikä on nähty tyypillisesti yksityisen sektorin erityispiirteeksi (Kim ym., 2014). Julkisella sektorilla on nähty tarpeelliseksi omaksua yksityiseltä sektorilta myös oman toiminnan mittaaminen (Vigoda, 2002), minkä avulla organisaatioiden toimintaa on mahdollista kehittää. Yhteistyö julkisen ja yksityisen sektorin välillä on lisääntynyt huomattavasti (Vigoda, 2002), minkä voidaan ajatella tulevaisuudessa kaventavan eroa sektorien välillä entisestään eri sektorien organisaatioiden omaksuessa toimintatapoja toisiltaan.

TAULUKKO 1 Julkisen ja yksityisen sektorin erityispiirteet

	Yksityinen sektori	Julkinen sektori
Toiminnan tavoite	Tuoton tavoittelu yritykselle ja sidosryhmille (Lachman, 1985) Kilpailuedun tavoittelu (Kim ym., 2014)	Kansalaisten hyvinvointi (Nutt, 2005) ja valtion laajuisten ongelmien ratkaiseminen (Kickert ym., 1997)
Rahoitus	Liikevoitto (Kim ym., 2014)	Verovarot (Andersen, 2008)
Päätöksenteko	Taustalla yrityksen ja sidosryhmien edun tavoittelu (Nutt, 2005)	Taustalla kansalaisten hyvinvointi ja kustannustehokkuus (Nutt, 2005) Demokratian valvonta (Harisalo ym., 2007)
Strateginen kehittäminen	Suorituskyky ja tehokkuus (Bryson & Roering, 1987)	Lähtökohtana kansalaisille tarjottavat palvelut (Bryson & Roering, 1987)
Prosessien kehittäminen	Suorituskyvyn ja asiakastytyväisyyden parantaminen (Stemberger & Jaklic, 2007)	Prosessien muutokset hitaampia (Stemberger & Jaklic, 2007) Asiakastytyväisyyden korostaminen (Stemberger & Jaklic, 2007)

3.2 Julkisen sektorin toimialueet

Julkisen sektorin toimialueita on määritelty eri tavalla tutkimuksesta riippuen. Toimialueilla tarkoitetaan niitä palveluita, joita julkinen sektori tarjoaa kansalaisille. Tässä alaluvussa esitellään julkisen sektorin toimialueet lähdekirjallisuuteen perustuen. Tällä pyritään selkeyttämään julkisen sektorin roolia yhteiskunnassa.

Suomen julkinen sektorin toiminta perustuu Pohjoismaiseen hyvinvointimalliin, jossa korostuu verrattain korkealla tasolla olevat hyvinvointijärjestelmän palvelut. Näihin julkisen sektorin palveluihin kuuluu Andersenin mukaan erityisesti terveydenhuolto, koulutus ja muut sosiaalipalvelut. (Andersen, 2008). Sosiaali- ja terveystalouden sekä koulutuksen merkitystä julkisella sektorilla on korostettu myös muissa julkaisuissa. Kickert ja kollegat (1997) esittävät kirjassaan julkisen sektorin yhdeksi tärkeäksi vastuualueeksi sosiaalipalveluiden tuottamisen. Sosiaali- ja terveystalouden sekä koulutuksen tarjoaminen korostuu myös Harisalon, Aarrevaaran, Stenvallin ja Virtasen (2007) mukaan julkisen sektorin päätöksenteon osa-alueista.

Sosiaali- ja terveystalouden sekä koulutuksen tarjoamisen lisäksi julkisen sektorin toimialueeseen kuuluu valtion toiminnan mahdollistavat tekniset palvelut, jotka voidaan nähdä esimerkiksi infrastruktuurin rakentamisena ja ylläpitona (Harisalo ym., 2007). Infrastruktuurin ylläpito mielletään myös Kickertin ja kollegoiden (1997) teoksessa julkisen sektorin perinteiseksi tehtäväksi, joka on myöhemmin siirtynyt julkisen ja yksityisen sektorin yhteisvastuulle. Myös valtion turvallisuuden ja vakauden takaaminen nähdään julkisen sektorin toimijoiden vastuualueena (Kickert ym., 1997).

Kuten huomataan, julkisen sektorin toimialueella on useita eri määritelmiä. Yhteistä lähteiden määritelmille on sekä sosiaali- ja terveystalouden (ks. esim. Andersen, 2008; Flynn, 2007; Harisalo ym., 2007; Lehto, Anttonen, Haveri & Palukka, 2012) että koulutuksen korostuminen (ks. esim. Andersen, 2008; Flynn, 2007; Harisalo ym., 2007, 2007). Lisäksi turvallisuuden takaaminen (Kickert ym., 1997) sekä infrastruktuurin ylläpito ja rakentaminen (ks. esim. Flynn, 2007; Harisalo ym., 2007; Kickert ym., 1997) miellettiin julkisen sektorin keskeisiksi toimialueiksi.

4 BIG DATAN HYÖDYNTÄMINEN JULKISELLA SEKTORILLA

Valtiovarainministeriön julkaisun mukaan analytiikka ja big data ovat yhdessä mobiilipalveluiden kanssa merkittävimmät valtion tietohallinnon kehitykseen vaikuttavat tekijät (Hattinen ja Kiviniemi, 2017). Aikaisemmassa tutkimuksessa on huomattu, että big datan todellisista hyödyistä julkisella sektorilla ei ole tarpeeksi tutkittua tietoa. Sekä julkisen sektorin päättäjät että kansalaiset ovat alkaneet ideoimaan keinoja, kuinka big dataa voitaisiin hyödyntää julkisella sektorilla. On huomattu, että big data -analytiikan käyttöönotto julkisella sektorilla on alkuvaiheessa ja aihetta käsittelevästä kirjallisuudesta on pula. (Desouza & Jacob, 2017) Rajallisesta tutkimuksesta huolimatta big datan hyödyntämisen mahdollisuudet on tunnistettu ympäri maailmaa, ja eri maiden hallinnot pääasiassa Yhdysvalloissa ja Euroopassa ovat tehneet useita selvityksiä aiheesta (ks. esim. Jin, Wah, Cheng & Wang, 2015; Kim ym., 2014; Rastas & Asp, 2014). Big data on nähty tulevaisuudessa niin merkittävänä tekijänä, että on nähty tarpeelliseksi kehittää uusia analytiikkaan keskittyviä koulutusohjelmia (Rastas & Asp, 2014). On jopa arvioitu, että big datan hyödyntämisen kyvykkyydet voidaan nähdä valtion uutena vahvuuden mittarina. (Jin, ym., 2015)

Julkinen sektori kerää erityisen suuren määrän dataa päivittäin, toisin kun esimerkiksi pienet organisaatiot (Desouza & Jacob, 2017). On kuitenkin huomattu, että julkinen sektorin big datan hyödyntämisen käyttöönotto on huomattavasti jäljessä yksityisen sektorin kehitystä, eikä valtavaa määrää kerättyä dataa hyödynnetä tehokkaasti (Joseph & Johnson, 2013). Julkisella sektorilla big dataa kerätään pääasiassa kolmella eri tavalla. Dataa kertyy kansalaisten vuorovaikutuksesta verkossa, jota muodostuu esimerkiksi käyttäjien liikkumisesta julkisten organisaatioiden verkkosivuilla. Toinen big datan lähde julkisella sektorilla on sensorit, jotka voivat olla sijoitettuna julkisiin rakennuksiin, kulkuneuvoihin tai esimerkiksi julkisesti ylläpidettäville teille. Erityisen keskeisessä roolissa julkisen sektorin toimintaa mitattaessa on julkisen sektorin toimijoiden itse keräämä data omasta toiminnastaan. Big datalle tyypillisesti julkisen sektorin keräämässä datassa yhdistyy strukturoitu ja strukturoimaton data, joista osa kertyy automaattisesti ja osa kerätään pelkästään analysointia varten. Dataa kertyy entis-

tä nopeammin paperiarkistojen siirtyessä verkkoon. (Mergel, Rethemeyer & Isett, 2016)

Liikenne- ja viestintäministeriön Big datan käyttö -työryhmän selvityksessä tutkitaan sitä, kuinka big dataa voidaan hyödyntää sekä julkisella että yksityisellä sektorilla. He arvioivat, että big datalla on monia mahdollisuuksia sekä yhteiskunnallisen hyvinvoinnin kehittämisessä että taloudellisen kasvun saavuttamisessa. (Rastas & Asp, 2014) Big datan myötä on havaittu tarve ottaa analytiikka mukaan liiketoiminnan ydintoimintoihin (Davenport, Barth & Bean, 2012), kuten esimerkiksi päätöksenteon tukemiseen, prosessin tehostamiseen ja strategiseen kehittämiseen. Big datan hyödyntämiskohteet eroavat sektorien välillä (Kim ym., 2014), mutta aikaisemmassa tutkimuskirjallisuudessa ei ole laajamittaisesti määritelty, kuinka erot ilmenevät tai mistä erot johtuvat. Tässä tutkimuksessa tutkitussa kirjallisuudessa esiteltiin useita eri tyyppisiä big datan hyödyntämiskohteita, jotka jaettiin tässä tutkimuksessa kolmeen eri osa-alueeseen. Jako tehtiin tutkitussa kirjallisuudessa esiintyneiden hyödyntämiskohteiden perusteella eikä sitä ole aikaisemmin käytetty muissa tutkimuksissa. Big dataa voidaan käyttää julkisella ja yksityisellä sektorilla päätöksenteon tukemiseen, strategisen kehittämiseen ja prosessien kehittämiseen. Tärkeimmät havainnot kultakin osa-alueelta esitetään alla olevassa taulukossa (Taulukko 2). Lisäksi luvussa pohditaan sitä, kuinka julkisen sektorin tunnistetut erityispiirteet vaikuttavat big datan hyödyntämiskohteiden ilmenemiseen.

TAULUKKO 2 Big datan hyödyntäminen julkisella ja yksityisellä sektorilla

	Yksityinen sektori	Julkinen sektori
Päätöksenteon tukeminen	Tulevaisuudessa luotetaan päättäjiä enemmän analytiikkaan (McAfee & Brynjolfsson, 2012) Päätöksenteon tukeminen tai korvaaminen algoritmeilla (Wamba ym., 2015) Tehtyjen päätösten perustelu analytiikalla (Günther ym., 2017)	Kansalaisten mieltymysten paljastuminen päätöksiä tehdessä (Desouza & Jacob, 2017) Välitön palaute tehdyistä päätöksistä sosiaalisen median analyysillä (Desouza & Jacob, 2017) Ongelmakohtien tunnistaminen sosiaalisen median analyysillä (Desouza & Jacob, 2017) Big data parantaa kliinisten päätösten laatua (Wang & Hajli, 2017) Analyttikot voivat tarjota parempia analyyskejä päätöksentekijöille (Van der Voort ym., 2018)
Strategisen kehittämisen	Big data- analytiikkaa voidaan käyttää strategisessa kehittämisessä (Lycett, 2013) Yrityksen liiketoimintamallin muuttaminen (Günther ym., 2017; Lycett, 2013) Dataohjautuvat yritykset ovat 5 % tehokkaampia ja 6 % tuottoisampia (McAfee & Brynjolfsson, 2012) Dataohjautuvat yritykset tekevät	Kansalaisten segmentointi ja palveluiden kohdentaminen big data -analytiikan avulla (Rastas & asp, 2014) Läpinäkyvyyden parantaminen (Rastas & asp, 2014; Kim ym., 2014; Wamba ym., 2015)

	<p>parempia arvioita taloudellisista ja operatiivisista tuloksista (McAfee & Brynjolfsson, 2012)</p> <p>Asiakkaiden segmentointi big data -analytiikan avulla (Wamba ym., 2015)</p>	
Prosessien kehittäminen	<p>Yrityksen toimintatapojen ja tehokkuuden kehittäminen (Wamba ym., 2015)</p> <p>Uusien liiketoimintamallien ja tuotteiden kehittäminen (Wamba ym., 2015)</p> <p>Kustannustehokkuuden parantaminen (LaValle ym., 2011)</p> <p>Liikevaihdon kasvattaminen (LaValle ym., 2011)</p> <p>Asiakassuhteiden hoito (LaValle ym., 2011)</p> <p>Henkilökunnan optimointi lentokentillä (McAfee & Brynjolfsson, 2012)</p>	<p>Uusien toimintatapojen luominen koko julkiselle sektorille (Desouza & Jacob, 2017; Rastas & Asp, 2014)</p> <p>Suorituskykyvertailun kehittäminen (Rastas & Asp, 2014)</p> <p>Raporttien ja analyysien tekemiseen kulutetun ajan lyheneminen (Joseph & Johnson, 2013)</p> <p>Asiakkaiden palvelutarpeen tunnistaminen (Chatfield & Reddick, 2017)</p> <p>Lomakkeiden käsittelyn automatisointi (Joseph & Johnson, 2013)</p> <p>Veropetosten tunnistaminen (Joseph & Johnson, 2013)</p> <p>Opetuksen kehittäminen ja suunnittelu (Al Nuaimi ym., 2015; Bhat & Ahmed, 2016)</p> <p>Sairastumisen riskin alentaminen (Rastas & Asp, 2014; Wang & Hajli, 2017)</p> <p>Lyhyempi diagnosointiin käytettävä aika (Wang & Hajli, 2017)</p> <p>Big data -analytiikan avulla voidaan analysoida reaaliaikaista dataa potilaiden terveydentilasta (Al Nuaimi ym., 2015)</p> <p>Älykaupunkien rakentaminen (ks. esim. Al Nuaimi ym., 2015; Desai, 2018; Kitchin, 2014; Lowman, 2017)</p> <p>Teiden kunnan arviointi applikaation avulla (Desouza & Jacob, 2017)</p> <p>Epidemioiden ennustaminen big datan avulla (Desouza & Jacob, 2017)</p> <p>Big datan avulla voidaan tunnistaa alueita, joilla esiintyy laittomia asukkaita (Desouza & Jacob, 2017)</p> <p>Reaaliaikaisen videokuvan ja aikaisemmin kerätyn datan analysointi rikosten ennustamiseksi (Desouza & Jacob, 2017)</p> <p>Luonnonkatastrofien tehokkaampi ennustaminen (Al Nuaimi ym., 2015)</p>

4.1 Päätöksenteon tukeminen

Big datalla nähdään olevan suurta vaikutusta organisaatioiden päätöksentekoprosessiin. Tässä alaluvussa keskitytään siihen, kuinka big dataa voidaan hyödyntää päätöksentekoprosessin tukemisessa julkisella ja yksityisellä sektorilla sekä pohditaan sitä, kuinka julkisen sektorin erityispiirteet vaikuttavat big data hyödyntämiseen päätöksenteon tukemisen näkökulmasta.

McAfee ja Brynjolfsson (2012) kuvaavat artikkelissaan sitä, kuinka yksityisen sektorin päätöksentekoprosessi tulee muuttumaan big datan hyödyntämisen myötä. He uskovat, että tulevaisuudessa johtoryhmän mielipidettä useammin luotetaan siihen, millaisia analyysejä datasta voidaan tehdä. Wamban ja kollegoiden (2015) havainnot tukevat tätä. He ovat toteuttaneet kattavan kirjallisuuskatsauksen big datan hyödyntämisestä, jossa huomattiin, että noin kolmasosa tutkituista artikkeleista mainitsi päätöksenteon tukemisen tai korvaamisen algoritmeilla. Myös julkisen sektorin big datan hyödyntämistä käsittelevässä kirjallisuudessa korostuu datan vaikutukset päätöksentekoprosessiin. Van der Voort, Klievink, Arnaboldi ja Meijer (2018) ovat toteuttaneet tapaustutkimuksen, joka käsittelee julkisen sektorin big data -projekteja. Heidän tutkimuksessaan huomattiin, että big datasta tehty analyysi auttaa analyytikkoja tarjoamaan parempia analyysejä päätöksentekijöille ja vastaavasti päätöksentekijöiden on helpompi omaksua big datasta tehty analyysi verrattuna perinteisiin analyyseihin (Van der Voort ym., 2018).

Big data -analytiikalla on mahdollista tarkastella jo aikaisemmin tehtyjen päätösten vaikuttavuutta kansalaisten hyvinvointiin. On esimerkiksi huomattu, että Twitteristä kerätyn datan avulla voidaan analysoida, miten uudet päätökset ovat vaikuttaneet kansalaisten mielialaan. Lisäksi Twitteristä kerätyn big datan avulla on tunnistettu, mitkä kansalaisten tarpeet ovat jo tyydyttyneet. Analytiikan avulla voidaan myös tunnistaa ratkaisevia ongelmia kansalaisten hyvinvoinnissa. Tunnistettuja ongelmia voidaan mahdollisesti jatkossa käyttää julkisen sektorin kehityskohteiden tunnistamisessa. Laajaa tutkimusta big datan vaikuttavuudesta julkisten käytäntöjen kehittämisessä ei kuitenkaan ole vielä saatavilla. (Desouza & Jacob, 2017) Big data mahdollistaa myös kansalaisten preferenssien paljastumisen ja koostamisen (Desouza & Jacob, 2017), joita voidaan käyttää päätöksenteon tukena. Aikaisemmin on huomattu, ettei demokratia aina toimi halutulla tavalla ja big datan hyödyntäminen voisi olla ratkaisu demokratian ongelmiin (Desouza & Jacob, 2017). Big dataa voidaan käyttää päätöksenteon tukena julkisella sektorilla kansallisten päätösten lisäksi myös yksikkötasolla. Esimerkiksi terveydenhuollossa big datan on huomattu parantavan kliinisten päätösten laatua ja tarkkuutta (Wang & Hajli, 2017).

Nutt (2005) on huomannut julkisen ja yksityisen sektorin päätöksenteon eroja käsittelevässä tutkimuksessaan, että julkisen sektorin päätöksenteossa korostuu erityisesti hintaneuvottelu ja yksityisen sektorin päätöksenteossa analyysi. Havaittiin myös, että julkisen sektorin päätöksenteon valmisteluihin eli tarvittavan informaation keräämiseen käytettävä budjetti on usein tarkoin sää-

dely ja on havaittu, että vaihtoehtoisia ratkaisuja tutkitaan usein vähemmän kuin yksityisellä sektorilla (Nutt, 2005). Big data -analytiikan keinoin jo olemassa olevasta datasta voidaan tehdä erilaisia analyyskejä, joilla voidaan ennustaa esimerkiksi päätöksenteon vaikutuksia eri tekijöihin (Rajaraman, 2016). Big data -analytiikan voidaan siis ajatella vähentävän päätöksenteon valmisteluihin käytettävien resurssien tarvetta julkisella sektorilla. Tämä voisi mahdollisesti parantaa kustannustehokkuutta ilman, että päätösten laatu laskee.

Desouza ja Jacob (2017) toteavat, että big datan hyödyntäminen julkisen sektorin päätöksenteon tukena on mahdollista ja jopa suotavaa, mutta analyysit eivät voi olla ainoa keino päätöksenteon tukena. Samankaltaisia huomioita on tehty myös yksityiseltä sektorilta. Data ja siitä tehdyt analyysit ovat hyödyttömiä, ellei niistä osata etsiä oikeita asioita (McAfee & Brynjolfsson, 2012). Analyttikkojen on tärkeää kehittää strategioita datan analysoimiselle (Günther ym., 2017; McAfee & Brynjolfsson, 2012) ja tarve korostuu erityisesti erilaisten data-lähteiden hankkimisessa ja yhdistämisessä (Günther ym., 2017). Myös aikaisemmin tuntemattomat ympäristömuuttajat, kuten sääilmiöt, synnyttävät tarpeen tukeutua alan ammattilaisiin (Günther ym., 2017)

Se, missä vaiheessa big dataa käytetään päätöksentekoprosessissa, vaihtelee sen mukaan, kuinka edistynyt organisaatio on analytiikan käyttöönotossa (LaValle ym., 2011). LaValle ja kollegoiden (2011) tutkimuksessa pyrkivät ja harjaantuneet organisaatiot käyttivät big dataa pääasiassa jo tehtyjen päätösten perusteleminen datalla. Tehtyjen päätösten perusteleminen big datalla on havaittu myös muissa julkaisuissa (Günther ym., 2017). Muuntautuneet organisaatioissa koko päätöksentekoprosessi perustuu analytiikkaan (LaValle ym., 2011). Voidaan ajatella, että päätösten perustelu big data -analytiikalla ei hyödynnä dataa tehokkaasti, ja muuntautuneissa organisaatioissa käytetään big dataa tehokkaammin. Tässä kirjallisuuskatsauksessa tutkitun kirjallisuuden perusteella julkisen sektorin toimijat eivät sijoitu selkeästi millekään analytiikan käyttöönoton tasolle päätöksenteon tukemisen näkökulmasta. Big dataa käytetään jo tehtyjen päätösten vaikutusten arviointiin, mutta analytiikan avulla etsitään myös ongelmakohtia, jotka tulisi ottaa huomioon päätöksentekoprosessissa (Desouza & Jacob, 2017). Kuitenkin esimerkkejä big datan käytöstä julkisen sektorin päätöksenteon tukena löytyy vain vähän, mikä viittaa käyttöönoton olevan alkuvaiheessa.

4.2 Strateginen kehittäminen

Big dataa voidaan käyttää apuna organisaation strategisessa kehittämisessä. Tässä alaluvussa eritellään big datan hyödyntämiskohteita julkisen ja yksityisen sektorin strategisen kehittämisen näkökulmasta ja pohditaan, kuinka julkisen sektorin erityispiirteet vaikuttavat hyödyntämiskohteiden ilmenemiseen.

McAfee ja Brynjolfsson (2012) ovat tutkineet sitä, kuinka big datan hyödyntäminen vaikuttaa yrityksen menestymiseen. Tutkimuksessa mitattiin yritysten dataohjautuvuutta, jolla tarkoitetaan sitä, kuinka keskeisessä roolissa

analytiikka on yrityksen strategiassa (McAfee & Brynjolfsson, 2012). Tutkimuksessa havaittiin, että dataohjautuvuudessa parhaiten sijoittuvat yritykset olivat keskimäärin 5 % tehokkaampia ja 6 % tuottoisampi kuin muut yritykset. Tutkimuksessa toteutettiin haastattelu koskien dataohjautuvuutta sekä arvioitiin yrityksen tehokkuutta ja tuottoisuutta vuosikertomusten ja ulkopuolisten lähteiden avulla. Tutkimukseen osallistui 330 eri pohjoisamerikkalaista yritysjohtajaa (McAfee & Brynjolfsson, 2012), ja näiden tutkimustulosten voidaan olettaa olevan yleistettävissä ainakin muihin länsimaisiin yrityksiin. Lisäksi tutkimuksessa huomattiin, että mitä dataohjautuvammaksi yritykset itsensä mieltävät, sitä paremmin ne pystyvät mittaamaan taloudellisia ja operatiivisia tuloksiaan (McAfee & Brynjolfsson, 2012). Tämä voi helpottaa yrityksiä niiden tehdessä arvioita siitä, kuinka strategiaa tulisi kehittää jatkossa.

Liikenne- ja viestintäministeriön Big datan käyttö -työryhmä on tunnistanut big datan hyödyttävän julkisen sektorin strategista kehittämistä. Yhdeksi big data -analytiikan tavoitteeksi on nimetty data- ja asiakaslähtöisempi hallinto. On arvioitu, että kansalaiset voidaan segmentoida big datan avulla ja näin kohdentaa palveluita tehokkaammin. Yksilöllisemmän ja tarkemmin kohdennettujen palveluiden odotetaan parantavan kansalaistyytyväisyyttä sekä luottamusta julkisen sektorin toimintaan. (Rastas & Asp, 2014) Asiakkaiden segmentointi datan avulla on havaittu tärkeäksi big datan mahdollisuudeksi myös yksityisellä sektorilla (Wamba ym., 2015). Se, että asiakkaiden segmentointi on huomattu olevan big datan yksi hyödyntämiskohde sekä julkisella että yksityisellä sektorilla voi johtua siitä, että julkisen sektorin tulisi suhtautua yksityisen sektorin tapaan kansalaisiin (Vigoda, 2002). Segmentoinnin ja kohdennettujen palveluiden tarjoamisen voidaan ajatella täten toimivan sektorista riippumatta.

Läpinäkyvyyden edistäminen big datan avulla nousi erityisesti julkisen sektorin big datan hyödyntämistä käsittelevässä kirjallisuudessa esille. Läpinäkyvyys ei ollut merkittävä teema yksityisen sektorin big datan käyttöä tutkittaessa. Vain 14% Wamban ja kollegoiden kirjallisuuskatsauksen materiaalista mainitsi läpinäkyvyyden edistämisen yhdeksi big datan mahdollisuudeksi (Wamba ym., 2015). Muissa julkaisuissa läpinäkyvyyttä ei korostettu. Pelastuslaitosta käsittelevän tutkimuksen yksi merkittävimmistä big datan hyödyntämisen vaikutuksista oli organisaation sisäisen ja ulkoisen läpinäkyvyyden ja luottavuuden lisääntyminen (Wamba ym., 2015). Myös Rastas ja Asp (2014) korostivat raportissaan läpinäkyvyyttä esimerkiksi datan keräämisessä sekä sen analysoinnissa, joiden he odottivat parantavan kansalaisten luottamusta julkiseen hallintoon. Kimin, Trimin ja Chungin (2014) havainnot tukevat näitä havaintoja. Toiminnan läpinäkyvyyden voidaan ajatella korostuvan erityisesti julkisella sektorilla kahdesta syystä. Julkisen sektorin rahoitus perustuu verovaroihin (Andersen, 2008), ja on tärkeää, että julkinen hallinto antaa oikeellisen kuvan kansalaisten rahoittamasta toiminnastaan. Rastas ja Asp (2014) ovatkin nimenneet yhdeksi big datan hyödyntämiskohteeksi avoimemman kuvan antamisen julkisen sektorin taloudellisesta tilanteesta. Yksityisellä sektorilla ajetaan ensisijaisesti yrityksen sekä sidosryhmien etua ja kilpailuetua (Kim ym., 2014), ja voidaan ajatella, ettei yrityksen ulkopuolinen läpinäkyvyys ole keskei-

sessä asemassa. Yrityksen ja sidosryhmien välisen läpinäkyvyyden voidaan ajatella olevan tärkeää myös yksityisellä sektorilla, sillä yhteistyökumppaneilla tulee olla käsitys siitä, miten yritys toimii.

Tutkituissa lähteissä on annettu myös esimerkkejä yksityiseltä sektorilta siitä, kuinka big datan ja siitä tehdyn analyysin hyödyntäminen voi muuttaa jopa koko yrityksen liiketoimintamallin. Lycett (2013) antaa artikkelissaan esimerkin, jossa dataa on käytetty voimavarana, kun perinteisellä alalla toimiva yritys on vaihtanut liiketoimintamalliaan dataohjautuvaksi ja menestynyt tämän ansiosta. Alun perin Netflixin perusliiketoiminta muodostui DVD-levyjen postimyynistä. Nykyään yrityksen liiketoimintamalli on muuttunut suoratoistopalveluksi, jonka menestyksen arvioidaan olevan analyysien ansiota. Palvelussa sopivaa sisältöä suositellaan katsojista kerätyn datan perusteella. (Lycett, 2013) Esimerkkejä liiketoimintamallin muutoksista löytyy myös esimerkiksi vaatetusalan ja painetun median toimialoilta (Günther ym., 2017). On arvioitu, että yksityisellä sektorilla avainmenestymiseen on yrityksen johtamisen muuttaminen kokemuskeskeisestä datakeskeiseksi (McAfee & Brynjolfsson, 2012). Vastaavia laajoja liiketoimintamallin muutokseen liittyviä viittauksia ei löytynyt julkiselta sektorilta. Tämän voidaan ajatella johtuvan siitä, että julkisen sektorin toimijoiden tavoitteet ovat tehtäväkeskeisiä ja yksiköiden tehtävät ovat ennalta määrättyjä. Yksityisellä sektorilla liiketoimintamallin kokonaisvaltainen muutos on mahdollista, sillä usein tavoitellaan liikevoittoa tehtäväkeskeisyyden sijaan.

LaValleen ja kollegoiden (2011) havaintojen mukaan analytiikan käyttöönoton taso vaikuttaa siihen, kuinka organisaatio käyttää big dataa hyväkseen strategisessa kehittämisessä. Muuntautuneissa organisaatioissa, joissa analytiikka on otettu käyttöön kokonaisvaltaisesti organisaation toiminnassa, käyttävät strategisen kehittämisen perustana toimintaympäristöstä ja organisaation toiminnasta tehtyjä analyyskejä. Tutkimuksessa huomattiin, että pyrkivät organisaatiot eivät hyödynnä analyysiä strategisessa kehittämisessä, vaan keskittyvät prosessien kehittämiseen sekä päätösten perustelemiseen analyysien avulla. Harjaantuneissa organisaatioissa osa strategisesta kehittämisestä perustuu analytiikalle. (LaValle ym., 2011) Tässä kirjallisuuskatsauksessa tutkitun kirjallisuuden perusteella voidaan arvioida, että julkisen sektorin organisaatiot ovat big data -analytiikan käyttöönotossa pyrkivien tasolla. Big datan käytön mahdollisuudet julkisen sektorin strategisessa kehittämisessä on tunnistettu, mutta tutkimus big datan hyödyntämisestä strategisessa kehittämisessä puuttuu lähes kokonaan.

4.3 Prosessien kehittäminen

Big dataa voidaan käyttää apuna organisaation prosessien kehittämisessä (Lycett, 2013). Kirjallisuudessa on kuitenkin keskitytty enemmän big datan mahdollisiin hyödyntämiskohteisiin (Günther ym., 2017), ja konkreettisia esimerkkejä big datan käytöstä julkisen ja yksityisen sektorin prosessien kehittä-

misestä on rajallisesti. Tässä alaluvussa käsitellään sitä, kuinka big dataa voidaan käyttää julkiselle ja yksityisellä sektorilla organisaatioiden prosessien kehittämiseksi. Tutkittavassa kirjallisuudessa on saatettu käyttää eri termejä liiketoimintaprosessien ja jokapäiväisten toimintojen kehittämisessä, joista tässä tutkimuksessa käytetään yhtenevää nimitystä prosessien kehittäminen.

Prosesseja voidaan kehittää big datan avulla molemmilla sektoreilla. Wamban ja kollegoiden (2015) big datan hyödyntämistä käsittelevässä kirjallisuuskatsauksessa viidesosa tutkituista artikkeleista mainitsi big datan mahdolliseksi hyödyntämiskohteeksi yrityksen toimintatapojen kehittämisen sekä uusien toimintamallien ja tuotteiden kehittämisen. Desouza ja Jacob (2017) ovat arvioineet, että big dataa voitaisiin käyttää julkisella sektorilla hyväksi sekä vanhojen käytäntöjen kehittämiseen että uusien luomiseen. Myös Liikenne- ja viestintäministeriön työryhmä on tunnistanut big datan hyödyntämisen kohteiksi julkisella sektorilla prosessien ja uusien toimintatapojen kehittämisen. (Rastas & Asp, 2014)

On arvioitu, että tulevaisuudessa big datalla voidaan tehostaa suorituskykyvertailua julkisia palveluita tuottavien yksiköiden kesken. Big data mahdollistaa entistä tarkemman suorituskyvyn seurannan sekä datalähtöisemmät mittarit. On havaittu, että samankaltaisissa organisaatioissa toimii usein yhtäläiset prosessit. Yksiköiden on mahdollista oppia toistensa toiminnasta datan avulla ja kehittää yhtäläiset ja tehokkaat käytännöt prosessien suorittamiselle. (Rastas & Asp, 2014) Big data -analyysin hyödyntäminen vähentää myös raporttien ja analyysien tekemiseen kulutettavaa aikaa (Joseph & Johnson, 2013). Suorituskyvyn arviointiin ja vertailuun käytettäviä mittareita voidaan todennäköisesti käyttää jokaisessa julkisen sektorin organisaatiossa. Myös yksityisellä sektorilla big data -analytiikalla on nähty olevan suuria vaikutuksia yritysten suorituskykyyn (McAfee & Brynjolfsson, 2012). Big dataa on käytetty yksityisellä sektorilla prosessien kehittämisessä muun muassa lentokentillä. Datasta analysoimalla on tehty arvioita lentoliikenteestä, joiden perusteella esimerkiksi tarkennetaan lähtö- ja paluuaikoja ja näin optimoidaan henkilökunnan työtehtäviä lentokentillä ja näin parannetaan suorituskykyä. Teknologiyritys PASSUR on arvioinut näiden analyysien säästävän useita miljoonia vuodessa. (McAfee & Brynjolfsson, 2012)

Chatfield ja Reddick (2017) ovat tutkimuksessaan selvittäneet, kuinka Houstonin kaupungin 311-palveluita tarjoavissa organisaatioissa hyödynnetään big dataa. 311-palveluilla tarkoitetaan sellaisia julkisen sektorin palveluita, jotka eivät liity hätätilanteiden ratkaisemiseen. Tällaisia palveluita voi olla esimerkiksi sosiaalipalvelut sekä verohallinnon palvelut. Tutkimuksessa huomattiin, että big data -analytiikan käyttö paransi yksiköiden kykyä tunnistaa asiakkaiden palvelutarpeet. (Chatfield & Reddick, 2017) Julkisella sektorilla big datalla on nähty olevan mahdollisuuksia esimerkiksi lomakkeiden käsittelyn automatisointiin. On huomattu, että eri yksiköiden välisen datan jakamisen ja analysoinnin avulla voidaan nopeuttaa lomakkeiden käsittelyaikaa merkittävästi ja näin lyhentää jonoja. (Joseph & Johnson, 2013) Lomakkeiden käsittelyn automatisointia voitaisiin käyttää mahdollisesti myös yksityisellä sektorilla. Joseph ja

Johnson (2003) raportoivat big data -analytiikan olevan käytössä Yhdysvaltojen verohallinnon prosesseissa. Eri julkisten organisaatioiden datan yhdistämisen ja analysoinnin avulla verohallinto pyrkii tunnistamaan veropetoksia (Joseph & Johnson, 2013). Yksityisellä sektorilla organisaatiot keräävät verkkosivuillaan dataa siitä, kuinka kävijät liikkuvat sivulla. Tämän analyysin avulla on mahdollista parantaa asiakaskokemusta ja asiakaspalvelua. Kirjoittajat ehdottavat, että myös julkisen sektorin sivuilla voisi tehdä tällaista analyysia ja näin kehittää digitaalisten alustojen tehokkuutta. (Joseph & Johnson, 2013)

Koulutuksen ja terveydenhuollon tarjoaminen on yksi julkisen sektorin päätehtävistä (Andersen, 2008), ja big dataa voidaan hyödyntää molemmilla toimialueilla. On huomattu, että big datan avulla voidaan tehdä analyyskejä esimerkiksi oppimiseen liittyvistä puutteista, joiden avulla voidaan kehittää näihin osa-alueisiin liittyvää opetusta ja suunnittelua (Al Nuaimi, Al Neyadi, Mohamed & Al-Jaroodi, 2015). Big datalla on mahdollisuuksia parantaa sekä opetuksen laatua että oppimisen mahdollisuuksia (Bhat & Ahmed, 2016). Liikenne- ja viestintäministeriön mukaan Suomessa on kansainvälisellä mittakaavalla poikkeuksellisen laadukkaita terveystietokantoja, joita hyväksikäyttämällä voidaan tulevaisuudessa vaikuttaa terveyttä edistävasti. Sairastumisen riskiä voidaan alentaa terveydenhuollosta kerääntyvän big datan avulla (Rastas & Asp, 2014; Wang & Hajli, 2017) ja lyhentää diagnosointiin kuluva aikaa (Wang & Hajli, 2017). Big data -analytiikka mahdollistaa useiden eri terveystietokantojen (Wang & Hajli, 2017) sekä reaaliaikaisen potilasdatan analysoinnin (Al Nuaimi ym., 2015). Varsinkin terveydenhuollossa big datan hyödyntämisen voidaan katsoa vähentävän sekä kustannuksia että inhimillistä kärsimystä.

Big datalla on tunnistettu olevan monia mahdollisuuksia infrastruktuurin ylläpitoon ja kehittämiseen liittyen, joka on yksi keskeinen julkisen sektorin toimialue. Big dataa voidaan hyödyntää älykaupunkien rakentamisessa (ks. esim. Al Nuaimi ym., 2015; Desai, 2018; Kitchin, 2014; Lowman, 2017). Myös liikenne- ja viestintäministerin big dataa tutkinut työryhmä nosti yhdeksi keskeiseksi hyödyntämiskohteeksi älykkäät infrastruktuurit, joissa big datan avulla voidaan tehostaa energianjakelua tarpeeseen perustuen ja esimerkiksi mitata sisäilman laatua (Rasta & Asp, 2014). Big datan avulla voidaan myös ylläpitää infrastruktuuria. Bostonin kaupunki on esitellyt applikaation, jonka avulla voidaan analysoida julkisesti ylläpidettävien teiden kuntoa. Asukkaiden puhelimet lähettävät tietoa teiden kunnosta heidän ajomatkinsa aikana, ja tämän datan avulla voidaan tehdä arvioita siitä, mitkä tieosuudet vaativat kunnostusta. (Desouza & Jacob, 2017) Tien ylläpitoon liittyvän datan myyminen julkiselle sektorille on tunnistettu myös yksityisen sektorin toimijoiden keskuudessa (Günther ym., 2017).

Big datan hyödyntäminen julkisella sektorilla voi auttaa myös kansalaisten turvallisuuden takaamisessa. Big data -analytiikalla voidaan ennustaa esimerkiksi erilaisia luonnonkatastrofeja entistä tarkemmin (Al Nuaimi ym., 2015). New South Walesin pelastuslaitoksella big datan avulla tehdään analyyskejä historiatiedon ja reaaliaikaisen datan avulla esimerkiksi tulevista tulvista. Näiden analyysien avulla voidaan evakuoida katastrofialueen kansalaiset ennakkoon ja

keskittää resursseja oikeille alueille. (Wamba ym, 2015) On myös havaittu, että esimerkiksi sikainfluenssan kaltaisten epidemioiden ennustaminen on mahdollista yhdistelemällä aikaisemmin erillisiä tietokantoja toisiinsa (Desouza & Jacob, 2017). Big datan avulla voidaan ennustaa ja ehkäistä rikollista toimintaa. New Yorkin kaupunki on pystynyt big datan avulla määrittämään tiettyjä alueita, joilla esiintyy laittomia asukkaita ja näin puuttumaan ongelmaan tehokkaammin (Desouza & Jacob, 2017). Desouza ja Jacob (2017) raportoivat myös, että Los Angelesin poliisilaitos on pystynyt ennustamaan rikosten ilmenemistä reaaliaikaisen videokuvan ja aikaisemmin kerätyn datan analysoinnin avulla. Mikäli julkisella sektorilla datan tehokas jakaminen eri yksiköiden välillä olisi mahdollista, voisi big data -analytiikan avulla jopa ehkäistä terrorismia (Kim ym., 2014).

LaValle ja kollegat (2011) tarkastelevat tutkimuksessaan eri analytiikan käyttöönoton tasoilla olevia organisaatioita myös prosessien kehittämisen näkökulmasta. Pyrkivien tasolla olevat organisaatiot käyttävät big dataa pääasiassa kustannustehokkuuden parantamiseen ja prosessien optimointiin. Harjaantuneet organisaatiot käyttävät analytiikkaa näiden lisäksi liikevaihdon kasvattamiseksi sekä uusien palveluiden kehittämiseksi. Muuntautuneet organisaatiot käyttävät big dataa pyrkivien ja harjaantuneiden organisaatioiden tapaan, mutta lisäksi muuntautuneissa organisaatioissa keskittyvät erityisesti uusien asiakkaiden hankkimiseen ja vanhojen asiakassuhteiden ylläpitämiseen. (LaValle ym., 2011) Tässä tutkimuksessa tutkitun kirjallisuuden perusteella voidaan sanoa, että julkisella sektorilla analytiikkaa hyödynnetään pääasiassa prosessien optimointiin ja kehittämiseen eikä esimerkiksi uusien palveluiden tai asiakassuhteiden kehittämisestä löytynyt. Voidaan ajatella, että julkisen sektorin organisaatiot vastaavat parhaiten pyrkivien organisaatioiden määritelmää big data -analytiikan käyttöönotossa prosessien kehittämisen näkökulmasta.

Julkisella sektorilla big datan hyödyntämiskohteiden osa-alueista prosessien kehittäminen korostuu selkeästi päätöksenteon tukemista ja strategista kehittämistä enemmän. Prosessien kehittämisen korostumiselle voidaan tunnistaa kaksi syytä. Ensinnä sekä tässä tutkimuksessa että aikaisemmissa tutkimuksissa on huomattu julkisella sektorilla big datan käyttöönotto on alkuvaiheessa (Desouza & Jacob, 2017) ja vasta harjaantuneet organisaatiot käyttävät analytiikkaa strategisessa kehittämisessä. Toiseksi tutkitun kirjallisuuden perusteella voidaan sanoa, että julkisen sektorin strategista kehittämistä ja päätöksentekoa ohjaa yksityistä sektoria tarkemmat tavoitteet ja arvot, kuten kansalaisten hyvinvoinnin varmistaminen tai laadukkaiden palveluiden tarjoaminen. Yksityisellä sektorilla taas on mahdollista muuttaa koko liiketoimintamallia, mikäli sen katsotaan tarjoavan kilpailuetua tai taloudellista hyötyä yritykselle ja sidosryhmille. Näiden tulosten lisäksi voidaan pitää merkittävänä löytönä sitä, että big datalla nähdään olevan potentiaalia prosessien kehittämisessä koulutuksen, sosiaali- ja terveyspalveluiden, turvallisuuden sekä infrastruktuurin ylläpidon alueella eli jokaisella tunnistetulla julkisen sektorin toimialueella.

4.4 Big datan hyödyntämisen haasteet

On havaittu, että big datan hyödyntämiseen liittyvää tutkimusta löytyy erityisen vähän juuri julkiselta sektorilta (Desouza & Jacob, 2017) ja samankaltaisia havaintoja on tehty myös tätä tutkimusta tehdessä. Tätä tutkimusta tehdessä on lisäksi todettu, että julkisen sektorin big dataa käsittelevässä kirjallisuudessa korostuu myös selkeästi enemmän hyödyntämisen haasteet kuin big datan hyödyntämisen mahdollisuudet. Näistä syistä on nähty tarpeelliseksi big datan hyödyntämiskohteiden lisäksi myös big datan hyödyntämiseen liittyvien haasteiden käsitteleminen. Tässä alaluvussa esitellään keskeiset kirjallisuudessa esiintyneet big datan hyödyntämisen haasteet korostaen julkisen sektorin näkökulmaa.

Datan määrän kasvu on synnyttänyt tarpeen aikaisempia työkaluja kehittyneemmille ratkaisulle, joilla analysoida suurta määrää dataa mahdollisimman nopeasti ja tehokkaasti (Gandomi & Haider, 2015). Tällaisia tapoja ovat muun muassa datan visualisointi sekä puheen ja tekstin tunnistus sekä sosiaalisen median analyysi (LaValle ym., 2011). Wamban ja kollegoiden (2015) kirjallisuuskatsauksessa huomattiin, että kolmasosa käsitellyistä artikkeleista määritteli teknologian ja tekniikat suurimmaksi haasteeksi big datan hyödyntämislle (Wamba ym., 2015). Big datan analysointiin sopivien teknologioiden hankkiminen on suuri investointi organisaatiolle. Koska rahoitus on julkisella sektorilla peräisin verovaroista (Andersen, 2008) ja big datan hyötyjen todistaminen tieteellisesti on puutteellista (Desouza & Jacob, 2017), voi suhtautuminen uusiin teknologioihin investointiin olla erityisen varauksellista julkisella sektorilla. Uusien teknologioiden hankkimisen lisäksi julkisella sektorilla big datan hyödyntämisen keskeiseksi haasteeksi on havaittu se, että organisaatioilla ei ole selvää käsitystä siitä, kuinka big data -analytiikka tulisi käyttöönottaa (Klievink ym., 2017). Julkisen sektorin organisaatioissa muutokset ovat hitaampia ja kankeampia kuin yksityisellä sektorilla (Stemberger & Jaklic, 2007), jonka voidaan ajatella aiheuttavan uusien teknologioiden käyttöönoton hitautta julkisella sektorilla.

Yhdeksi big datan hyödyntämisen keskeiseksi haasteeksi julkisella sektorilla on tunnistettu julkisen sektorin organisaatioiden rakenne. Julkisella sektorilla yksiköt toimivat toisistaan erillisinä kokonaisuuksina, niin sanotuissa silloissa (Desouza & Jacob, 2017; Kim ym., 2014). Desouza ja Jacob (2017) arvioivat, että näiden yksiköiden välille tulisi muodostaa ryhmiä big datan hyödyntämistä edistävän strategian luomiseksi. Julkisella sektorilla toimivat tietohallintopäälliköt ovat arvioineet yhteisten järjestelmien ja datan hyödyntämiskeinojen puuttumisen olevan suurin syy siihen, minkä takia heidän organisaatioissaan ei ole aikomusta hyödyntää big dataa. (Desouza & Jacob, 2017) Myös Rastas ja Asp (2014) ovat huomioineet raportissaan, että big datan tehokas hyödyntäminen julkisella sektorilla vaatii tiivistä yhteistyötä sekä julkisen sektorin organisaatioiden että julkisen ja yksityisen sektorin välillä.

Kaikkein tärkeimpänä haasteena big datan hyödyntämisessä on nähty turvallisuuden ja yksityisyyteen liittyvät ongelmat. Desouza ja Jacob (2017) ovat nimenneet tärkeimmiksi haasteiksi big datan hallitsemisen sekä yksilön turvallisuuden. On varsin vähän tutkimusta siitä, kuinka big dataan liittyviä turvallisuuden ja yksityisyyteen liittyviä ongelmia voitaisiin ratkaista. (Desouza & Jacob, 2017) Big data -analytiikan mahdollisuus yhdistää ja löytää korrelaatioita aikaisemmin erilaisten datasettien väliltä kuuluu sen vahvuuksiin. Tämä johtaa siihen, että yksilöstä voidaan tehdä analyysejä, jotka saattavat loukata yksityisyyttä. (Desouza & Jacob, 2017) Analytiikan aiheuttamat riskit yksilön turvallisuudelle on huomioitu myös Güntherin ja kollegoiden (2017) tutkimuksessa ja he korostavat, että big datasta on mahdollista tehdä epäsoivia päätelmiä. Julkisen sektorin organisaatiot keräävät suuren määrän sensitiivistä dataa kansalaisistaan eri palveluissa, joten turvallisuuden voidaan ajatella olevan erityisen tärkeää huomioida juuri julkisen sektorin big datan hyödyntämisessä. Tämän tyyppisen big datan keräämiseen ja analysointiin tulee tulevaisuudessa kehittää oma lainsäädäntö (Kim ym., 2014).

Desouza ja Jacob (2017) esittelevät artikkelissaan, kuinka sosiaalisen median analyysejä voidaan hyväksikäyttää julkisen sektorin päätöksenteon tukena. Haasteeksi tässä kuitenkin tunnustetaan se, ettei sosiaalisen median palveluissa voida varmistua siitä, onko käyttäjätilien takana todellisia kansalaisia. On myös pohdittu sitä, kuinka kansalaisten sosioekonominen asema vaikuttaa datan luomiseen. Mikäli demokratiaa jatkossa korvattaisiin analytiikalla, voi alemmassa sosioekonomisessa asemassa olevat kansalaiset jäädä aliedustetuiksi informaatioteknologian käytön vähäisyyden takia. (Desouza & Jacob, 2017) Nämä tekijät vaikuttavat analyysien luotettavuuteen ja hankaloittaa niiden hyväksikäyttöä julkisen päätöksenteon apuvälineenä. Myös Van der Voort ja kollegat (2018) ovat tunnustaneet big datan hyödyt julkisen sektorin päätöksenteolle. Tutkimuksessa kuitenkin huomattiin, että analyttikot ja päätöksentekijät voivat ajaa omaa etuuttaan big datan avulla esimerkiksi analysoimalla vain tiettyä dataa. (Van der Voort ym., 2018) Päätöksenteon tärkeimpänä motiivina julkisella sektorilla on kansalaisten etujen ajaminen (Nutt, 2005), joten oman edun ajaminen voidaan nähdä huomattavan haitallisena erityisesti julkisella sektorilla.

5 YHTEENVETO

Tämä tutkimus toteutettiin kirjallisuuskatsauksena, joka pyrkii selvittämään kuinka big dataa voidaan hyödyntää julkisella sektorilla. Aiheeseen tutustuttiin tarkastelemalla julkisen ja yksityisen sektorin eroja, joiden perusteella määriteltiin julkisen sektorin erityispiirteet. Tutkimus vastaa yhteen tutkimuskysymykseen: ”Kuinka julkisen sektorin erityispiirteet vaikuttavat big datan hyödyntämiseen julkisella sektorilla?”. Tutkimuksen taustat ja käytetty tutkimusmenetelmä on esitelty ensimmäisessä kappaleessa eli johdannossa.

Toisessa luvussa määriteltiin big datan käsite sekä esiteltiin big datan hyödyntämistä organisaatioissa. Big data voidaan määrittää usealla eri tavalla. Määritelmässä voidaan keskittyä datan ominaisuuksiin, joita on ajateltu olevan kolme: datan määrä (volume), kertymisnopeus (velocity) ja monimuotoisuus (variety) (Gandomi & Haider, 2015; Lycett, 2013). Myöhemmin niin sanottuun kolmen V:n malliin on lisätty lisää määrittäviä kriteerejä, joita ovat datan arvokkuus (value) (Lycett, 2013) ja datan todenmukaisuus (veracity) (Wamba ym., 2015). Big data voidaan määrittää myös sosiokulttuurisena ilmiönä (Boyd & Crawford, 2012) tai big datan ja perinteisen datan eroavaisuuksien perusteella (Klievink ym., 2017). Big dataa voidaan kerätä useista eri lähteistä, joko organisaation sisäisistä järjestelmistä tai esimerkiksi sensoreista. Luvun lopussa esiteltiin lyhyesti erilaisia big datan analytiikkamenetelmiä.

Kolmannessa luvussa tutustuttiin julkisen ja yksityisen sektorin välisiin eroavaisuuksiin sekä julkisen sektorin toimialueisiin, jotta saadaan selkeä kuva julkisen sektorin toiminnasta ja erityispiirteistä. Julkinen sektori eroaa yksityisestä sektorista toiminnan tavoitteen ja rahoituksen näkökulmasta. Sektorien välillä on havaittavissa eroja myös siinä, kuinka päätöksiä tehdään sekä siinä, kuinka prosesseja ja strategiaa kehitetään. Tutkitun kirjallisuuden perusteella julkisella sektorilla on neljä keskeisintä toimialuetta eli koulutus, sosiaali- ja terveyspalvelut, turvallisuus sekä infrastruktuurin ylläpitäminen.

Neljännessä luvussa etsittiin vastausta tutkimuskysymykseen. Luvussa paneuduttiin big datan hyödyntämiseen sekä julkisella että yksityisellä sektorilla. Luvussa painotettiin erityisesti julkisen sektorin erityispiirteiden vaikutuksia big datan hyödyntämiskohteiden esiintymiseen. Lisäksi luvussa käsiteltiin big

datan hyödyntämisen keskeisiä haasteita erityisesti julkisen sektorin näkökulmasta. Tutkimuksessa saatiin kokonaiskuva siitä, miten julkisen sektorin erityispiirteet vaikuttavat hyödyntämiskohteiden esiintymiseen julkisella sektorilla. Tutkimuksessa big datan hyödyntämiskohteet jaettiin kolmeen eri osaluueeseen lähdekirjallisuudesta poimittujen osa-alueiden mukaan. Nämä osaluueet ovat päätöksenteon tukeminen, strateginen kehittäminen ja prosessien kehittäminen.

Julkisella sektorilla big dataa voidaan hyödyntää jokaisella tunnistetulla osa-alueella. Tutkitusta lähdekirjallisuudesta voidaan kuitenkin päätellä, että käyttöönotto on alkuvaiheessa kaikilla osa-alueilla. Big dataa voidaan hyödyntää julkisen sektorin päätöksenteon tukena esimerkiksi arvioimalla tehtyjen päätösten vaikutusta kansalaisiin, jolloin saadaan suora palaute tehdyistä päätöksistä (Desouza & Jacob, 2017). Lisäksi big datan avulla voidaan tunnistaa erilaisia ongelmia kansalaisten hyvinvoinnissa ja tulevaa päätöksentekoa voidaan perustaa tunnistettuihin ongelmiin (Desouza & Jacob, 2017). Big dataa voidaan käyttää strategisessa kehittämisessä erityisesti asiakaslähtöisyyden ja läpinäkyvyyden parantamiseen julkisella sektorilla. Selkeästi eniten hyödyntämiskohteita tunnistettiin prosessien kehittämisen osa-alueelta. Big datalla voidaan parantaa julkisen sektorin organisaatioiden suorituskykyä esimerkiksi suorituskykymittareiden kehittämisellä (Rastas & Asp, 2014) ja erilaisten prosessien optimoimisella. Big dataa voidaan hyödyntää myös uusien toimintatapojen kehittämisessä terveydenhuollon, koulutuksen, turvallisuuden ja infrastruktuurin ylläpitämisen toimialueilla. Tutkimuksessa huomattiin, että big datan hyödyntämisen osa-alueista julkisella sektorilla korostuu erityisesti prosessien kehittyminen, jonka ajatellaan johtuvan kahdesta eri tekijästä. Julkisella sektorilla big data -analytiikan käyttöönotto on alussa ja vasta edistyneemmät organisaatiot voivat hyödyntää analytiikkaa strategisessa kehittämisessä ja päätöksenteon tukena. Lisäksi julkisen sektorin toiminnan tavoite ohjaa yksityistä sektoria voimakkaammin päätöksentekoa ja strategista kehittämistä, minkä ajatellaan vähentävän big datan hyödyntämismahdollisuuksia näillä osa-alueilla. Lisäksi tutkimuksessa havaittiin, että big dataa voidaan käyttää laajasti julkisella sektorilla, sillä big datan sovelluskohteita löydettiin jokaiselta tunnistetulta julkisen sektorin toimialueelta. Julkisen sektorin big datan hyödyntämisen tutkimuksessa on korostunut mahdollisuuksien lisäksi haasteet ja myös näitä pohdittiin neljännessä luvussa. Uusien teknologioiden käyttöönoton ajatellaan olevan hitaampaa julkisella sektorilla rahoituksen ja muutoksen hitauden takia. Big datan käyttöönottoa vaikeuttaa myös julkisen sektorin siiloutuminen sekä sensitiivisen datan turvallisuuteen liittyvät kysymykset. Lisäksi analyysien luotettavuuden ja edustavuuden ongelmien ajateltiin korostuvan juuri julkisella sektorilla.

Tutkimusta tehdessä havaittiin, että laadukasta tutkimusta big datan hyödyntämisestä sekä julkisella että yksityisellä sektorilla on tehty hyvin vähän. Julkaisuissa keskitytään vahvasti big datan mahdollisuuksiin ja konkreettisia esimerkkejä big datan hyödyntämisestä on verrattain vähän. Tämän oletetaan johtuvan pääasiassa siitä, että big data -analytiikan käyttöönotto julkisella sek-

torilla on alussa (Desouza & Jacob, 2017). Laadukkaan tutkimuksen puutteen takia päätelmiä big datan hyödyntämisen mahdollisuuksista julkisella sektorilla joudutaan tekemään yksittäisten toimialojen kautta, koska laajamittaista tutkimusta aiheesta ei ole. Vaikka julkisen sektorin eri toimialueet hoitavat yhteiskunnassa eri tehtävää, on toimialueiden toiminnan taustalla samat vastuut ja tavoitteet. Siksi voidaan olettaa, että julkisen sektorin kaikkien toimialojen big datan sovelluskohteet ovat samankaltaisia erityisesti päätöksenteon tukemisen ja strategisen kehittämisen näkökulmasta.

Tutkimuksessa havaittiin, että big data -analytiikan käyttöönotto julkisella sektorilla on varsin alkuvaiheessa ja jatkotutkimusaiheita havaittiin useita. Eriytyisen mielenkiintoista olisi tutkia lisää sitä, millaisia big data -analytiikan hyödyntämiskohteita julkiselle sektorille on käyttöönotettu ja kuinka käyttöönotto on onnistunut. Lisäksi mielenkiintoisena tutkimusaiheena voidaan pitää sitä, millaisia hyötyä big data -analytiikalla on ollut ja onko investoinnit analytiikkaan olleet kokonaisuudessaan kannattavia.

LÄHTEET

- Al Nuaimi, E., Al Neyadi, H., Mohamed, N., & Al-Jaroodi, J. (2015). Applications of big data to smart cities. *Journal of Internet Services and Applications*, 6(1), 25.
- Andersen, T. M. (2008). Pohjoismainen malli-tulevaisuuden näkymät ja haasteet. *Yhteiskuntapolitiikka*, 73(4), 402-410.
- 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.
- Broadbent, J., & Guthrie, J. (1992). Changes in the public sector: A review of recent "alternative" accounting research. *Accounting, Auditing & Accountability Journal*, 5(2), 3-31.
- Bryson, J. M. (2018). *Strategic planning for public and nonprofit organizations: A guide to strengthening and sustaining organizational achievement*. John Wiley & Sons.
- Bryson, J. M., & Roering, W. D. (1987). Applying private-sector strategic planning in the public sector. *Journal of the American Planning Association*, 53(1), 9-22.
- Chatfield, A. T., & Reddick, C. G. (2017). Customer agility and responsiveness through big data analytics for public value creation: A case study of Houston 311 on-demand services. *Government Information Quarterly*, 25(2), 336-347.
- Chen, H., Chiang, R. H. & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS Quarterly*, 36(4), 1165-1188.
- Chen, M., Mao, S. & Liu, Y. (2014). Big data: A survey. *Mobile Networks and Applications*, 19(2), 171-209.
- Constantiou, I. D. & Kallinikos, J. (2015). New games, new rules: Big data and the changing context of strategy. *Journal of Information Technology*, 30(1), 44-57.
- Davenport, T. H., Barth, P. & Bean, R. (2012). How 'big data' is different. *MIT Sloan Management Review*, 54(1), 22-24.

- Desai, P. V. (2018). A survey on big data applications and challenges. 2018 *Second International Conference on Inventive Communication and Computational Technologies (ICICCT)* (737-740). Coimbatore: IEEE.
- Desouza, K. C. & Jacob, B. (2017). Big data in the public sector: Lessons for practitioners and scholars. *Administration & Society*, 49(7), 1043-1064.
- Flynn, N. (2007). *Public sector management*. Sage.
- Gandomi, A. & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, 35(2), 137-144.
- Gulledge Jr, T. R., & Sommer, R. A. (2002). Business process management: public sector implications. *Business process management journal*, 8(4), 364-376.
- 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.
- Harisalo, R., Aarrevaara, T., Stenvall, J., & Virtanen, P. (2007). *Julkinen toiminta-julkinen politiikka*. Tampere: Tampereen Yliopistopaino Oy.
- Harmon, P. (2003). *Business process change: a manager's guide to improving, redesigning, and automating processes*. Morgan Kaufmann.
- Hattinen, T., & Kiviniemi, M. (2016). Tietoja valtion tietohallinnosta 2015. *Valtiovarainministeriön julkaisuja 27/2016*.
- Jin, X., Wah, B. W., Cheng, X., & Wang, Y. (2015). Significance and challenges of big data research. *Big Data Research*, 2(2), 59-64.
- Joseph, R. C., & Johnson, N. A. (2013). Big data and transformational government. *IT Professional*, 15(6), 43-48.
- Kitchin, R. (2014). The real-time city? Big data and smart urbanism. *GeoJournal*, 79(1), 1-14.
- Kickert, W. J., Klijn, E. H., & Koppenjan, J. F. (Eds.). (1997). *Managing complex networks: Strategies for the public sector*. Sage.
- Kim, G. H., Trimi, S., & Chung, J. H. (2014). Big-data applications in the government sector. *Communications of the ACM*, 57(3), 78-85.
- Klievink, B., Romijn, B. J., Cunningham, S., & de Bruijn, H. (2017). Big data in the public sector: Uncertainties and readiness. *Information Systems Frontiers*, 19(2), 267-283.

- Kopetz, H. (2011). *Design principles for distributed embedded applications*. Sage.
- Labrinidis, A., & Jagadish, H. V. (2012). Challenges and opportunities with big data. *Proceedings of the VLDB Endowment*, 5(12), 2032-2033.
- Lachman, R. (1985). Public and private sector differences: CEOs' perceptions of their role environments. *Academy of Management Journal*, 28(3), 671-680.
- LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S. & Kruschwitz, N. (2011). Big data, analytics and the path from insights to value. *MIT Sloan Management Review*, 52(2), 21.
- Lehto, J., Anttonen, A., Haveri, A., & Palukka, H. (2012). *Julkisen palvelun muutos tutkimuksen kohteena*. Tampere: Tampereen Yliopistopaino Oy, 11-16.
- Lowman, M. (2017). *A practical guide to analytics for governments: Using big data for good*. Hoboken, New Jersey: Wiley.
- Lycett, M. (2013). 'Datafication': Making Sense of (Big) Data in a Complex World. *European Journal of Information Systems*, 22(4), 381-386.
- McAfee, A. & Brynjolfsson, E., (2012). Big data: The management revolution. *Harvard Business Review*, 90(10), 60-68.
- Mergel, I., Rethemeyer, R. K., & Isett, K. (2016). Big data in public affairs. *Public Administration Review*, 76(6), 928-937.
- Nutt, P. C. (2005). Comparing public and private sector decision-making practices. *Journal of Public Administration Research and Theory*, 16(2), 289-318.
- Olsen, J. B., and D. C. Eadie. (1982). *The game plan: Governance with foresight*. Washington: Council of State Planning Agencies.
- Perry, J. L., & Wise, L. R. (1990). The motivational bases of public service. *Public administration review*, 50(3), 367-373.
- Rajaraman, V. (2016). Big data analytics. *Resonance*, 21(8), 695-716.
- Rastas, T. & Asp, E. (2014) Big datan hyödyntäminen. *Liikenne- ja viestintäministeriön julkaisuja 20/2014*. Haettu 25.10.2018 osoitteesta http://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/77879/Julkaisu_20-2014.pdf
- Stemberger, M. I., & Jaklic, J. (2007). Towards E-government by business process change – A methodology for public sector. *International Journal of Information Management*, 27(4), 221-232.

- Van der Voort, H. H., Klievink, A. B., Arnaboldi, M. M., & Meijer, A. A. (2018). *Rationality and politics of algorithms. Will the promise of big data survive the dynamics of public decision making?* *Government Information Quarterly*.
- Vigoda, E. (2002). From responsiveness to collaboration: Governance, citizens, and the next generation of public administration. *Public administration review*, 62(5), 527-540.
- 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.
- Wang, Y., & Hajli, N. (2017). Exploring the path to big data analytics success in healthcare. *Journal of Business Research*, 70, 287-299.