

Johanna Kokkinen

**BIG DATAN REAALIAIKAINEN HYÖDYNTÄMINEN
PÄÄTÖKSENTEOSSA**



JYVÄSKYLÄN YLIOPISTO
INFORMAATIOTEKNOLOGIAN TIEDEKUNTA
2019

TIIVISTELMÄ

Kokkinen, Johanna

Big datan reaaliaikainen hyödyntäminen päätöksenteossa

Jyväskylä: Jyväskylän yliopisto, 2019, 38 s.

Tietojärjestelmätiede, kandidaatintutkielma

Ohjaaja: Seppänen, Ville

Nykymaailmassa syntyy päivittäin valtava määrä big dataa, jota organisaatiot voivat hyödyntää päätöksenteossaan. Organisaatioiden toimintaympäristö on jatkuvassa muutoksessa, joten päätöksenteon reaaliaikaisuudella on suuri rooli organisaatioiden menestyksen kannalta. Tämän systemaattisena kirjallisuuskatsauksena toteutetun tutkimuksen tarkoituksena oli selvittää big datan hyödyntämisen vaikutuksia organisaatioiden reaaliaikaiseen päätöksentekoon ja tuoda esille, millaisia mahdollisuuksia se tarjoaa eri sovellusaloilla sekä millaisia haasteita siihen liittyy. Työssä esitellään esimerkkisovelluksia liiketoiminnan, valtion ja yhteiskunnan sekä urheilun parista. Tutkimuksen keskeisin tulos on, että reaaliaikainen big data tukee päätöksentekoa ja mahdollistaa organisaatioille muun muassa tehokkuuden, tuottavuuden, suorituskyvyn ja ketteryyden parantamisen sekä uusien liiketoimintamahdollisuuksien löytämisen. Reaaliaikainen big data tuo mukanaan kuitenkin myös haasteita liittyen muun muassa teknisiin ratkaisuihin, datan määrään, nopeuteen, monimuotoisuuteen ja luotettavuuteen, organisaatiokulttuurin muutokseen, yksityisyyteen sekä eettisyyteen.

Asiasanat: big data, data-analytiikka, reaaliaikainen päätöksenteko, dataohjautuva päätöksenteko

ABSTRACT

Kokkinen, Johanna

Utilizing big data in real-time decision-making

Jyväskylä: University of Jyväskylä, 2019, 38 pp.

Information Systems, Bachelor's Thesis

Supervisor: Seppänen, Ville

In today's world, there is generated every day a huge amount of big data that organizations can utilize in their decision-making. The operating environment of organizations is constantly changing, so real-time decision-making plays a major role in organizational success. This systematic literature review attempted to investigate the impact of utilization of big data on the organization's real-time decision-making and to highlight its opportunities and challenges in different application areas. The work introduces exemplary applications from the areas of business, state and society, and sport. The most important results of this research are that real-time big data supports decision-making and enables organizations to improve their efficiency, productivity, performance and agility, as well as finding new business opportunities. The real-time big data also brings challenges, including technical solutions, data volume, velocity, variety and veracity, organizational culture change, privacy and ethics.

Keywords: big data, data-analytics, real-time decision-making, data-driven decision-making

TAULUKOT

Taulukko 1	Reaaliaikaisen big datan sovellusalueita päätöksenteossa.....	31
------------	---	----

SISÄLLYS

TIIVISTELMÄ	2
ABSTRACT	3
TAULUKOT	4
SISÄLLYS.....	5
1 JOHDANTO.....	6
2 BIG DATA JA PÄÄTÖKSENTEKO.....	9
2.1 Big data.....	9
2.2 Data-analytiikka.....	12
2.3 Dataohjautuva päätöksenteko	14
2.4 Reaalikainen päätöksenteko.....	16
3 REAALIAIKAINEN BIG DATA ERI SOVELLUSALOILLA.....	19
3.1 Sovelluksia liiketoiminnan päätöksenteossa	19
3.2 Sovelluksia valtion ja yhteiskunnan päätöksenteossa.....	22
3.3 Sovelluksia urheilun päätöksenteossa	27
4 YHTEENVETO JA JOHTOPÄÄTÖKSET	30
LÄHTEET	34

1 JOHDANTO

Nykyihmisen arkipäiväinen toiminta synnyttää valtavan määrän dataa, joka tallentuu jossain muodossa digimaailman syövereihin. Muun muassa puhelimet, verkon selaaminen, sosiaalinen media, pikaviestisovellukset, GPS-anturit ja esineiden internet (Internet of Things) tuottavat dataa osana niiden normaalia toimintaa. (McAfee & Brynjolfsson, 2012.) Tämä big dataksi kutsuttu nopeasti syntyvä, määrältään suuri ja monimuotoinen data tekee Marrin (2015) mukaan maailmasta jatkuvasti älykkäämmän. Se vaikuttaa elämän eri osa-alueilla tehden muun muassa urheilusta, asumisesta, rakkauden löytämisestä ja vanhemmuudesta älykkäämpää. Organisaatioille datan kerääminen kaikesta mahdollisesta tarjoaa jatkuvasti enemmän mahdollisuuksia hienosäätää toimintaansa parantaakseen tehokkuuttaan. Big data muuttaakin liiketoiminnan luonnetta eri aloilla aina tehdasteollisuudesta terveydenhuoltoon ja vähittäismyynnistä maatalouteen. Pelkkä datan kerääminen ei kuitenkaan riitä, vaan sen perusteella on osattava tehdä hyviä päätöksiä nopeasti. (Marr, 2015.)

Dataa siis syntyy jatkuvasti ja sitä kerätään kaikkialta, mutta sitä on osattava hyödyntää oikein ja usein jopa reaaliaikaisesti, jotta se tuottaa arvoa liiketoiminnalle (Panian, 2007). Organisaatiot tekevät päivittäin useita pieniä ja suuria päätöksiä. Niillä voi olla taloudellista merkitystä tai ne voivat vaikuttaa työntekijöiden hyvinvointiin ja sitä kautta tuottavuuteen. Päätöksillä voi olla vaikutusta myös organisaation ulkopuolisiin toimijoihin ja sidosryhmiin. Nykymaailmassa myös kilpailuedun saavuttaminen markkinoilla on yrityksille yhä haastavampaa. Yritysten toimintaympäristöön vaikuttavat niin asiakkaat, kilpailijat, yhteistyökumppanit, markkinavoimat, säännöstelyvalta kuin se, mitä maailmalla tapahtuu. Näissä jatkuvasti muuttuvissa olosuhteissa reaaliaikaisella päätöksenteolla on kriittinen rooli. (Panian, 2007.) Päätökset pitää tehdä nopeasti ja niiden on oltava oikeita. McAfeen ja Brynjolfssonin (2012) mukaan dataohjautuvat päätökset ovat parempia päätöksiä, joten big dataan perustuva päätöksenteko on kasvattanut organisaatioissa suosiotaan perinteisen intuitioon perustuvan päätöksenteon sijasta.

Powerin (2014) mukaan olemme luomassa globaalia dataohjautuvaa yhteiskuntaa, mutta tämän vaikutuksia yhteiskunnan, yksilön tai liiketoiminnan

näkökulmista ei vielä ymmärretä hyvin. Hän esittää myös, että päätöksenteon tuen ja tieteellisen tutkimuksen kannalta on tärkeää jatkuvasti tunnistaa semi-strukturoidun ja strukturoimattoman datan analysointiin liittyviä käyttötapauksia ja -esimerkkejä. Tämä tutkielma pyrkii osaltaan vastaamaan tähän selvittämällä asiaa yhteiskunnan ja liiketoiminnan näkökulmista. Näkökulmana on siis organisaatioiden päätöksenteko, yksilön näkökulma jätetään tarkastelun ulkopuolelle. Tutkielman tavoitteena on tuoda esille, millaisia mahdollisuuksia big datan reaaliaikainen hyödyntäminen päätöksenteossa tarjoaa organisaatioille eri sovellusaloilla ja toisaalta millaisia haasteita siihen liittyy. Tutkimuskysymykset ovat:

- Miten big datan hyödyntäminen vaikuttaa reaaliaikaiseen päätöksentekoon?
- Millä eri sovellusaloilla reaaliaikaista big dataa voidaan hyödyntää päätöksenteossa?
- Mitä mahdollisuuksia ja haasteita liittyy big datan reaaliaikaiseen hyödyntämiseen päätöksenteossa?

Tutkimus toteutettiin systemaattisena kirjallisuuskatsauksena hyödyntäen Templierin ja Parén (2015) esittelemää viitekehystä kirjallisuuskatsauksen laatimisesta. Tutkimus eteni viitekehysten kautta vaihetta noudattaen. Vaiheet ovat ongelman muotoilu, kirjallisuuden etsiminen, kirjallisuuden valitseminen, laadun arvioiminen, aineiston tiivistäminen sekä analyysin ja synteessin tekeminen.

Lähteenä käytettiin suurimmaksi osaksi vertaisarvioituja tieteellisiä artikkeleita, jotka haettiin pääasiassa Google Scholarin ja IEEE Xplore:n avulla käyttäen hakulauseina muun muassa seuraavia termejä: "real-time big data", "big data in real-time decision-making", "real-time analytics", "real-time sport analytics", "big data in motorsport". Tutkielmaan valittiin 50 lähdehakusanon osuvuuden ja siteerausmäärien perusteella. Lähteiden laatua ja luotettavuutta arvioitiin julkaisukanavan ja sisällön perusteella. Artikkelien lisäksi käytettiin konferenssijulkaisuja sekä muutamia kirjoja.

Tutkielma koostuu johdannosta, kahdesta sisältöluvusta sekä yhteenvedosta ja johtopäätöksistä. Ensimmäisessä sisältöluvussa käsitellään ensin big dataa käsitteenä ja esitellään sen keskeisimpiä ominaisuuksia. Sitten käydään läpi data-analytiikkaa sekä dataohjautuvaa ja reaaliaikaista päätöksentekoa vastaten tutkimuskysymykseen "Miten big datan hyödyntäminen vaikuttaa reaaliaikaiseen päätöksentekoon?". Toisessa sisältöluvussa vastataan tutkimuskysymyksiin "Millä eri sovellusaloilla reaaliaikaista big dataa voidaan hyödyntää päätöksenteossa?" ja "Mitä mahdollisuuksia ja haasteita liittyy big datan reaaliaikaiseen hyödyntämiseen päätöksenteossa?". Luvussa esitellään esimerkkejä big datan reaaliaikaisesta hyödyntämisestä eri sovellusaloilta liiketoiminnan, valtion ja yhteiskunnan sekä urheilun näkökulmista. Yhteenvedo ja johtopäätökset -luvussa kootaan yhteen big datan vaikutuksia reaaliaikaiseen päätök-

sentekoon ja esimerkkisovelluksista esiin nousseita haasteita ja mahdollisuuksia sekä esitetään jatkotutkimusaiheita.

2 BIG DATA JA PÄÄTÖKSENTEKO

Digitaalisten laitteiden yleistymisen on johtanut ennennäkemättömän datamäärän syntyyn ja kasvattanut tarvetta data-analytiikalle sekä dataohjautuvalle suunnittelulle (Gandomi & Haider, 2015). Maailmasta tulee kaiken aikaa yhtenäisempi ja välineellisempi, joten päätöksentekijöillä tulee olla yhä nopeampi pääsy jatkuvasti muuttuviin ja kasvaviin datavirtoihin. Tämä kuitenkin tekee päätöksentekijöiden työstä monimutkaisempaa, sillä on mahdotonta oppia kaikkea siinä aikaikkunassa, jossa päätökset pitäisi tehdä. Tehdäkseen hyviä päätöksiä, päätöksentekijöiden täytyykin saada selville reaaliajassa, mikä valtavassa datavirrassa on oleellista. (van der Schaar, 2014.)

Tässä luvussa määritellään big data käsitteenä ja tarkastellaan data-analytiikkaa yleisellä tasolla. Lisäksi tarkastellaan reaaliaikaisen dataohjautuvan päätöksenteon ominaispiirteitä ja reaaliaikaisen päätöstuen vaikutuksia organisaation päätöksentekoon.

2.1 Big data

Globaalin datamäärän räjähdysmäisen kasvun myötä termiä big data käytetään yleensä kuvaamaan valtavan suurien tietoa-aineistojen (Chen, Mao & Liu, 2014). Big datalle ei kuitenkaan ole olemassa yhtä yhtenäistä määritelmää. Jotkut pitävät yli kymmenen teratavua kooltaan ylittävää datamäärää big datana (Watson, 2014). Manyika ym. (2011) kuitenkin toteavat, että tietyn teratavumäärän ylittävää dataa ei voi yksiselitteisesti koon perusteella määritellä big dataksi, koska big dataksi laskettavien tietoa-aineistojen määrä muuttuu ajan teknologian edistymisen myötä ja koska sama määrä dataa voidaan tietysti sovelluksessa nähdä big dataksi, mutta toisessa ei. Chenin ym. (2014) mukaan big data eroaa perinteisestä datasta siten, että se tyypillisesti sisältää suuren määrän strukturoimatonta dataa ja vaatii enemmän reaaliaikaista analyysia. Sen käsittelyyn tarvitaan siis uudenlaisia teknologioita.

Powerin (2014) mukaan big data on markkinointitermi ja värikäs ilmaus datan keräämisessä, varastoinnissa ja hakemisessa käynnissä olevalle huomattavalle muutokselle. Se ei siis ole tekninen termi, vaan ilmiötä IT-alalla kuvaavat paremmin termit koneen tuottama data (machine data), strukturoimaton data (unstructured data), prosessidata (process data) ja uusi data (novel data).

McAfee & Brynjolfsson (2012) määrittelevät big data -ilmiön siten, että datalähteistä tulee jatkuvasti erilaisempia, moninaisempia ja dynaamisempia. Yhä useampi sidosryhmä osallistuu datan luomiseen ja analysoimiseen yhä useammilla eri laitteilla. Datan tarjoaman palautteen nopeus ja määrä suosii epäinhimillisiä toimijoita, kuten tekoälyä ja vastaavia ratkaisuja. (McAfee & Brynjolfsson., 2012.)

Vaikka big datalle ei ole olemassa yhteistä määritelmää, yleisesti sitä kuvaillaan niin sanotun kolmen V:n mallin avulla. Mallin mukaan big datalle tyypillisiä ominaisuuksia ovat sen määrä (volume), nopeus (velocity) ja monimuotoisuus (variety) (Russom, 2011). Mallia on myös laajennettu lisäämällä siihen ominaisuuksia, kuten todenmukaisuus (veracity), vaihtelevuus (variability), kompleksisuus (complexity) ja arvo (value) (Gandomi & Haider, 2015).

Määrä viittaa varastoidun datan kokoon. Dataa syntyy valtava määrä joka sekunti (Marr, 2015). Digitaalisen datan koon arvioitiin vuonna 2011 olevan yhteensä 1,8 tsettatavua, ja sen odotetaan 50-kertaistuvan vuoteen 2020 mennessä (Yadranjiaghdam, Pool & Tabrizi, 2016). Useimmiten big datan koko ilmaistaan teratavuissa, joskus myös petatavuissa (Russom, 2011). Dataa ei kuitenkaan voi määritellä big dataksi pelkästään sen koon perusteella, sillä big datan määrien kuvaukset ovat suhteellisia ja vaihtelevat useiden tekijöiden, kuten ajan ja datatyyppin mukaan. Se, mikä käsitetään isoksi tänään, ei välttämättä ole sitä tulevaisuudessa, sillä varastointikapasiteetti kasvaa ja sallii näin yhä suurempien datamäärien keräämisen. Kaksi samankokoista tiedostoa voivat vaatia erilaisia datanhallintateknologioita. (Gandomi & Haider, 2015.) Big datan kokoa voi myös määrittää laskemalla asiakirjoja, liiketapahtumia, taulukoita tai tiedostoja. Jotkut yritykset pitävät hyödyllisempänä määritellä big datan sen ajan mukaan, kuinka kauan dataa säilytetään. (Russom, 2011.)

Nopeus viittaa datan syntynopeuteen ja vauhtiin, jolla se tulisi analysoida ja ottaa käyttöön. Data liikkuu ympäri maailmaa todella nopeasti, esimerkiksi luottokorttien väärinkäyttö voidaan huomata reaaliajassa mistä päin maailmaa tahansa (Marr, 2015). Useille sovelluksille datan luonnin nopeus on tärkeämpää kuin sen määrä (McAfee & Brynjolfsson, 2012). On monia tilanteita, joissa dataa tulee hyödyntä, jos siihen perustuvaa päätöstä ei pystytä tekemään tietyssä aikaikkunassa (Yadranjiaghdam ym., 2016). Big data -teknologia tarjoaa yrityksille mahdollisuuden hyödyntää reaaliaikaisesti myös tällaista, hetkessä vanhentavaa dataa (Gandomi & Haider, 2015). Reaaliaikainen tai lähes reaaliaikainen tieto voi tehdä yrityksestä ketterämmän sen kilpailijoihin verrattuna. (McAfee & Brynjolfsson, 2012.) Toisaalta big datan suurin tekninen haaste on järjestää, hakea ja prosessoida tosiasiaa kerättyä dataa ja sisällyttää se päätöksentekoon reaaliajassa (Yadranjiaghdam ym., 2016).

Monimuotoisuus viittaa datan rakenteeseen. Big data tulee monista eri lähteistä, jotka monet ovat suhteellisen uusia, kuten sosiaalisesta mediasta, sensoreista ja matkapuhelimista, ja se voi olla muodoltaan esimerkiksi viesti, päivitys, kuva, sensorin lukema tai GPS-signaali (McAfee & Brynjolfsson, 2012). On olemassa siis hyvin erityyppistä ja -muotoista dataa. Se voi olla rakenteeltaan strukturoitua, semistrukturoitua tai strukturoimatonta. Strukturoitu data on taulukkomuotoista relaatiotietokantoihin ja taulukkolaskentaohjelmiin tallennettua dataa (Gandomi & Haider, 2015). Esimerkiksi talous- ja asiakasdata on strukturoitua. Strukturoimaton data on muodoltaan sellaista, että sitä ei voi helposti jaotella riveihin, sarakkeisiin tai kenttiin, ja sitä on vaikea analysoida perinteisillä tietokoneohjelmilla. (Marr, 2015.) Esimerkkejä tästä ovat teksti, valokuvat, ääni ja videot. Semistrukturoidun datan rakenne ei noudata tiukkoja standardeja, mutta se sisältää jonkinlaisia tunnistetietoja, joita voidaan hyödyntää analysoinnissa. Tyypillinen esimerkki tästä on XML-tekstitiedosto, joka sisältää metatietoja kuten tekijän ja päivämäärän, mutta itse teksti on rakenteeltaan strukturoimatonta. (Gandomi & Haider, 2015; Marr, 2015.)

Todenmukaisuus (veracity) viittaa datan luotettavuuteen. Labrinidisin ja Jagadishin (2012) mukaan big datan on luultu aiemmin kertovan aina totuuden, mutta tämä on kaukana todellisuudesta. Heidän mukaansa yksi big dataan liittyvistä haasteista onkin virheellisen tiedon käsittely. Big data on usein sekalaista ja sotkuista. Esimerkiksi Twitter-julkaisu voi sisältää muun muassa aiheutuniteita, lyhenteitä, kirjoitusvirheitä ja arkikielistä tekstiä. (Marr, 2015.) Myös esimerkiksi ihmisten tunteiden selvittäminen sosiaalisesta mediasta peräisin olevasta datasta on luonteeltaan epävarmaa, sillä siinä on kyse ihmisen arvioinnista. Silti tällaisesta datasta voidaan saada arvokasta tietoa, joten tarvitaan työkaluja ja analytiikkaa, jolla pystytään hallitsemaan epätasällistä ja epävarmaa tietoa. (Gandomi & Haider, 2015.)

Arvo viittaa datasta saatavaan hyötyyn. Suuressa määrässä varastoitua ja käyttämätöntä dataa piilee yleensä arvokasta tietoa (Yadranjiaghdam ym., 2016). Marrin (2015) mukaan arvo saavutetaan kuitenkin vasta data-analyysin perusteella, kun kaikki muut edellä mainitut big datan ominaisuudet on kunnolla sisäistetty. Oraclen määritelmän mukaan big data on arvotiheydeltään lähtökohtaisesti matalaa eli alkuperäisessä muodossaan datalla ei ole juuri arvoa. Korkea arvo voidaan kuitenkin saavuttaa, kun analysoidaan suuria määriä dataa. (Gandomi & Haider, 2015.)

Kompleksisuus viittaa datan moniin erilaisiin lähteisiin. Tämä luo haasteen yhdistää, sovittaa, puhdistaa ja muuntaa eri lähteistä peräisin olevaa dataa, jotta se olisi yhteensopivaa eri järjestelmien kanssa (Gandomi & Haider, 2015; Power, 2014).

Vaihtelevuus viittaa datan virtausnopeuksien vaihteluun. Datan nopeus ei ole tasaista, vaan se nousee ja laskee ajoittain. (Gandomi & Haider, 2015.)

Gandomin ja Haiderin (2015) vuonna 2012 tekemän verkkotutkimuksen mukaan 28 prosenttia big datan määritelmästä viittaa transaktiodatan massiiviseen kasvuun, 24 prosenttia uusiin big datan määrään, monimuotoisuuteen ja nopeuteen liittyviin haasteisiin vastaamaan suunniteltuihin teknologioihin, 18

prosenttia uusien datalähteiden räjähdysmäiseen kasvuun ja 19 prosenttia tarpeeseen varastoida ja arkistoida dataa säädösten ja lakien noudattamiseksi. Loput 11 prosenttia viittaavat muunlaisiin määritelmiin.

2.2 Data-analytiikka

Big datalla itsessään ei ole arvoa, vaan sitä täytyy analysoida, jotta siitä on hyötyä (Power, 2014). Big data -analytiikka (big data analytics) on Chenin, Chiangin ja Storeyn (2012) mukaan tekniikoita ja sovelluksia datalle, joka vaatii laajuutensa ja monimutkaisuutensa vuoksi ainutlaatuisia datan varastointi-, hallinta-, analyysi- ja visualisointiteknologioita. Reaaliaikaiset sovellukset perustuvat tietyllä hetkellä kerättävään dataan ja sen nopeaan analysointiin, jotta päätös tai toiminta saadaan toteutettua hyvin lyhyen ja tarkkaan määritellyn ajanjakson aikana (Yadranjiaghdam ym., 2016).

Jotta big datasta saadaan hyödyllistä tietoa, on Labrinidiksen ja Jagadishin (2012) mukaan käytävä läpi monivaiheinen prosessi, jolle big datan heterogeenisyys, epätäydellisyys, laajuus, ajantasaisuus ja yksityisyys sekä prosessin monimutkaisuus luovat haasteita. Prosessi ei ole lineaarinen, vaan aiempiin vaiheisiin joudutaan usein palaamaan, kun muutokset loppuvaiheissa aiheuttavat muutoksia myös alkupäähän. Gandomi ja Haider (2015) luokittelevat kyseisen prosessin kolme ensimmäistä vaihetta datan hallinnaksi, joka sisältää ne prosessit ja teknologiat, joilla data kerätään, varastoidaan ja valmistellaan analysoitavaksi. Kaksi viimeistä vaihetta he luokittelevat analytiikaksi, joka viittaa big datan analysointiin ja tiedon hankkimiseen käytettäviin tekniikoihin.

Ensiksi data hankitaan eli kerätään ja tallennetaan. Haasteena on, että kaikki data ei ole hyödyllistä. (Labrinidis & Jagadish, 2012.) Organisaatioiden käytössä on usein liikaa dataa, mikä aiheuttaa päätöksentekijöille kognitiivista ylikuormitusta (Merendino ym., 2018). Suuri osa datasta voidaan suodattaa, mutta suodattimet on osattava määritellä niin, ettei menetetä hyödyllistä dataa. Toinen tämän vaiheen haaste on luoda automaattisesti oikeanlaista metadatta kuvaamaan, mitä dataa on kerätty ja miten se on tallennettu ja mitattu. Metadatta voi olla ratkaisevan tärkeää analyysiprosessin loppuvaiheessa, kun analyysin tuloksia tarkastellaan ja tulkitaan. (Labrinidis & Jagadish, 2012.)

Kerätty data ei ole useinkaan analysoitavaksi sopivassa muodossa. Lähteistä poimitaan tarvittava tieto ja muutetaan se analyysiin sopivaan jäsenneltyyn muotoon. Data-analyysi on huomattavasti haastavampaa kuin paikantaminen, tunnistaminen, ymmärtäminen ja dataan viittaaminen. Kaiken tämän on tapahduttava automaattisesti, jotta analyysi olisi tehokas ja laajamittainen, joten datan on oltava tietokoneelle ymmärrettävässä muodossa. (Labrinidis & Jagadish, 2012.)

Uusien datalähteiden ja datanhallintateknologioiden avulla on mahdollista analysoida dataa eri tavoilla (Power, 2014) ja tekniikoilla (Marr, 2015; Gandomi & Haider, 2015). Yleisesti data-analytiikka jaetaan tapojen perusteella kolmeen luokkaan: kuvailevaan, ennakoivaan ja ohjailevaan (LaValle, Lesser,

Shockley, Hopkins & Kruschwitz, 2011). Yleisimmät data-analytiikkatekniikat puolestaan ovat tekstianalytiikka, puheanalytiikka sekä video- ja kuva-analytiikka (Marr, 2015).

Muun muassa liiketoimintatiedon hallinta (business intelligence) on kuvailevaa data-analytiikkaa, jossa pyritään löytämään tilastollisten työkalujen avulla historiallisesta datasta malleja ja tuloksia, joiden perusteella voidaan tehdä johtopäätöksiä tulevasta. Ennakoivassa data-analytiikassa luodaan historiallisen datan pohjalta simulaatiomalleja, jotta ymmärretään tulevaisuutta. (Power, 2014.) Se kertoo, mitä tulee tapahtumaan. Ohjaileva data-analytiikka puolestaan ehdottaa toimia. Esimerkiksi GPS-navigaattoreiden antamat ohjeet perustuvatkin siihen. (Watson, 2014.)

Ohjaileva data-analytiikka on analytiikan tyypeistä se, joka liittyy kiinteimmin datan reaaliaikaiseen analysointiin. Ohjailevassa analytiikassa käytetään suunniteltuja tilastollisia reaaliaikaisen datan analyysseja, jotka voivat käynnistää tapahtumia (Power, 2014). Siihen siis kuuluu datan käyttöä ja matemaattisia algoritmeja, jotka määräävät ja auttavat vaihtoehtoisissa päätöksissä, joihin liittyy luonteeltaan monimutkaisia tavoitteita ja vaatimuksia. Ohjaileva analytiikka sisältää monikriteeristä päätöksentekoa, optimointia ja simulointia. (Wang, Gunasekarana, Ngai & Papadopoulos, 2016.) Se voi tunnistaa optimaalisia ratkaisuja esimerkiksi niukkojen resurssien oikein kohdentamiseen. Ohjailevan analytiikan on todettu tarjoavan organisaatioille laajoja käyttömahdollisuuksia. Sitä käytetään erityisesti hinnoittelussa sellaisilla aloilla, joilla tuotteet täytyy myydä tiettyyn päivämäärään mennessä tai muuten ne tuottavat yritykselle kuluja. Tällaisia hinnoittelukohteita ovat esimerkiksi lentokoneiden istumapaikat, vuokra-autot ja hotellihuoneet. (Watson, 2014.)

Tekstianalytiikka viittaa tekniikoihin, joilla poimitaan tietoa tekstimuotoisesta datasta (Gandomi & Haider, 2015). Sen avulla voidaan tyypillisesti kategorisoida, klusteroida ja tiivistää tekstiä sekä havaita siitä käsitteitä ja mielipiteitä (Marr, 2015). Strukturoimattomasta tekstistä voidaan poimia strukturoitua dataa, kuten nimiä, ja luokitella se halutulla tavalla. Tekstiä voidaan tiivistää kahdella eri periaatteella, joko tekstin merkityksiä ymmärtämättä sen yksikköjen, kuten sanojen ja lauseiden, mukaan tai semanttisten merkitysten perusteella. Tunneanalyysissä, joka tunnetaan myös mielipiteiden louhintana, tekstistä etsitään esimerkiksi tuotteisiin, organisaatioihin, yksilöihin ja tapahtumiin liittyviä mielipiteitä. (Gandomi & Haider, 2015.) Sen perusajatuksena on luokitella teksti joko koko dokumentin tai yksittäisten kappaleiden ja lauseiden tasolla polariteetiltaan positiiviseksi, neutraaliksi tai negatiiviseksi. Tekstianalytiikka on erityisen hyödyllistä ennakoivan analytiikan kannalta. (Marr, 2015.)

Audioanalytiikka viittaa tekniikoihin, joilla poimitaan tietoa strukturoimattomasta audiodatasta. Kun audiodata sisältää puhuttua ihmisen kieltä, on kyse puheanalytiikasta. Sitä käytetään erityisesti puhelinalveluissa ja terveydenhuollossa. Teknologiat mahdollistavat puheen tunnistamisen kahdella eri periaatteella. Puhe voidaan muuttaa joko sanoiksi tai foneemeiksi eli kielen merkityksiä erottaviksi yksiköiksi, kuten tavuiksi tai kirjaimiksi, joiden perusteella järjestelmä etsii sille tallennetusta sanastosta vastineen. (Gandomi & Hai-

der, 2015.) Puheanalytiikan avulla voidaan analysoida sekä puheenaiheita että tunteita, kun huomioidaan muun muassa intonaatiot sekä puheen määrä suhteessa hiljaisuuteen (Marr, 2015).

Video- ja kuva-analytiikka sisältää monia eri tekniikoita monitoroida, analysoida ja poimia tarkoituksenmukaista tietoa videokuvista. Se on analytiikkatekniikoista tuorein ja kasvavin, sillä sen haasteena on datan suuri määrä verrattuna esimerkiksi tekstitiedostoon ja vasta big data -teknologian avulla tämä on voitu kääntää mahdollisuudeksi. (Gandomi & Haider, 2015.) Aiemmin videoista ja kuvista on voitu analysoida vain niihin liitettyjä aihetunnisteita, mutta nykyteknologia mahdollistaa myös niiden sisällön analysoinnin kohtauskohtaukselta. Videoanalytiikkaa käytetäänkin tänä päivänä monilla aloilla muun muassa kasvojen tunnistukseen, käyttäytymisen analysointiin ja muuttuvien tilanteiden tasalla pysymiseen esimerkiksi lento- ja laivaliikenteessä, voimailtoksissa ja hätäpalveluissa. (Marr, 2015.)

Lopulta päätöksentekijä on kuitenkin se, joka tulkitsee analyysin tulokset. Prosessin viimeisessä vaiheessa käyttäjien tulee ymmärtää ja varmistaa tietokoneen tuottamat tulokset, sillä tietokonejärjestelmissä voi olla vikoja, mallit sisältävät lähes aina oletuksia ja tulokset voivat perustua virheelliseen dataan. Tässä vaiheessa järjestelmän tuottama metadata siitä, miten tuloksiin on päästy, helpottaa käyttäjän työtä. (Labrinidis & Jagadish, 2012.)

2.3 Dataohjautuva päätöksenteko

Perinteisesti organisaatioissa tärkeimmät päätökset on tehty korkeassa asemassa olevien johtajien tai kalliiden ulkopuolisten konsulttien toimesta heidän kokemuksiinsa, toimintatapoihinsa ja suhteisiinsa perustuen. Tällaista päätöksentekotyylä, joka perustuu ihmisten muodostamiin mielipiteisiin siitä, mitä tulevaisuus tuo tullessaan, kutsutaan intuitioksi. Sen käyttö on järkevää silloin, kun dataa on saatavilla vain vähän, ei lainkaan digitaalisessa muodossa ja se on kallista saavuttaa. (McAfee & Brynjolfsson, 2012.) Nykyään kuitenkin big datan seurauksien kanssa kamppailu ei ole enää vain muutamien datasuuntautuneiden johtajien asia, vaan kaikkien alojen johtajat joutuvat kohtaamaan sen (Manyika ym., 2011). Tarvitaan uudenlaista päätöksentekokulttuuria (McAfee & Brynjolfsson, 2012). Tässä haasteena kuitenkin on, että päätöksentekijät kokevat usein huolta omasta valmiudestaan big datan hyödyntämiseen sekä ahdistusta siitä, millaisia vaikutuksia sillä on organisaatioon ja pystytäänkö organisaatiota muuttamaan tarpeeksi nopeasti (Merendino ym., 2018).

Dataohjautuvalla päätöksenteolla (data-driven decision-making) viitataan toimintaan, jossa päätökset perustuvat datan analysointiin puhtaan intuition sijasta (Provost & Fawcett, 2013). Dataohjautuvat päätökset voivat olla luonteeltaan erilaisia. Provost ja Fawcett (2013) erittelevät tapaustutkimuksiinsa perustuen kaksi erilaista päätöstyyppiä. Ensimmäinen näistä on päätös, joka tehdään datasta tehdyn löydön perusteella. Esimerkiksi Wal-Mart on pyrkinyt ennus-

tamaan, mitkä tuotteet ovat kysyttyjä juuri ennen hurrikaanin osumista maihin analysoimalla aiemmista hurrikaaneista kerättyä dataa. Toiseksi on olemassa erityisesti isossa mittakaavassa toistuvia päätöksiä, joita tehdessä voidaan hyötyä jo pienistä datan analysointiin perustuvista täsmällisyyksistä. Esimerkiksi televiestintäyrittysten sadoista miljoonista asiakkaista kymmenien miljoonien sopimukset erääntyvät joka kuukausi, jolloin jokaisen kohdalla todennäköisyys siirtyä lähitulevaisuudessa toisen yrityksen asiakkaaksi kasvaa. Data-analyysilla voisi potentiaalisesti saavuttaa hyvin suuren edun, jos sen avulla pystyttäisiin parantamaan arviointikykyä siitä, kuinka tuottoisaa on keskittyä kuhunkin asiakkaaseen. Sama logiikka koskee myös monia muita aloja, kuten suoramarkkinointia, online-mainontaa ja tuotesuosituksia. (Provost & Fawcett, 2013.)

Organisaatioiden päätöksiä tehdään myös yhä enemmän automaattisesti tietojärjestelmissä. Esimerkiksi kasinoiden palkkio-ohjelmat, vähittäiskauppojen suositukset ja kuluttajakohtainen mainonta verkossa ovat automatisoituja päätöksiä, jotka pystytään tekemään sekunneissa. (Provost & Fawcett, 2013.) Automaattiset päätökset sopivat parhaiten päätöksiin, jotka täytyy tehdä toistuvasti, rutiininomaisesti ja nopeasti. Päätöksentekokriteerien tulee olla hyvin strukturoituja ja kaikki huomioonotettavat tekijät täytyy ymmärtää hyvin. Vaikka päätöksentekoprosessi olisi automatisoitu, luottamukselliset, lainopilliset ja eettiset asiat vaativat kuitenkin vastuuhenkilön huolehtimaan prosessista. (Davenport & Harris, 2005.)

McAfee ja Brynjolfssonin (2012) mukaan dataohjautuvat päätökset ovat parempia päätöksiä. Alalla kuin alalla erottuvat kilpailijoistaan ne yritykset, jotka keksivät miten yhdistää alan asiantuntemus datatieteeseen. Myös LaValle ym. (2011) esittelemä tutkimus osoittaa, että suorituskyvyn ja dataohjautuvan johtamistavan välillä on selkeä yhteys, mikä ilmenee organisaation yrityksessä kasvaa, tehokkuudessa ja erottumisessa kilpailijoista. Parhaiten suoriutuvat organisaatiot hyödyntävät data-analytiikkaa päätöksissään viisi kertaa enemmän kuin heikommin suoriutuvat. Ne käyttävät data-analytiikkaa todennäköisemmin sekä suurissa että pienissä päätöksissä aina tulevaisuuden strategiasta päivittäisiin toimintoihin. (LaValle, 2011.)

McAfee ja Brynjolfssonin (2012) tekemän tutkimuksen mukaan mitä enemmän yritykset luonnehtivat itseään dataohjautuviksi, sitä paremmin he suoriutuivat taloudellisilla ja toiminnallisilla mittareilla mitattuna. Alansa eniten itsensä dataohjautuvaksi luonnehtivaan kolmanneksen kuuluvat yritykset olivat 5 % tuottavampia ja 6 % kannattavampia kilpailijoihinsa verrattuna.

Tehokas organisaatio pitää tiedon ja oleelliset päätösoikeudet samassa paikassa. Big datan aikakaudella tietoa kuitenkin luodaan ja siirretään, eikä asiantuntemus ole usein siellä missä se on ennen ollut. Organisaatioissa tulee olla ihmisiä, jotka ymmärtävät, että ongelmat pitää tuoda yhteen oikean datan kanssa sekä ihmisiä, jotka hallitsevat ongelmanratkaisutekniikat. Nokkela johtaja luo organisaatiosta niin joustavan, että se pystyy minimoimaan ”ei keksitty täällä” -efektin ja maksimoimaan toimivan yhteistyön näiden ihmisten välillä. (McAfee & Brynjolfsson, 2012.)

2.4 Reaalikainen päätöksenteko

Liiketoiminta tapahtuu reaaliaikaisesti. Elämme maailmassa, joka päivittyy joka minuutti, kun tuotamme dataa ja meillä on langaton pääsy useisiin tietolähteisiin. Tämä tekee meistä kärsimättömiä. Organisaatioilla ei siis ole varaa harjoittaa liiketoimintaansa vanhentuneiden tietojen perusteella. Asiakkaat odottavat välittömiä tuloksia, eivätkä halua odottaa vastauksia. (Panian, 2007.)

Reaaliaikaisuudelle ei ole olemassa täysin yhtenäistä määritelmää liiketoiminnan kontekstissa. Reaaliaikaisuus ei myöskään tarkoita täysin samaa kaikissa päätöksentekotilanteissa. Powerin (2011) mukaan reaaliaikainen päätöksenteko tarkoittaa, että päätöksentekoprosessin tietolähteet ovat saatavilla heti, kun ympäristössä tapahtuu muutoksia, jotka vaikuttavat näihin tietolähteisiin. Azvine, Cui ja Nauck (2005) esittävät, että reaaliaikaisuus voi tarkoittaa vaatimusta saavuttaa nollaviive prosessissa. Reaaliaikaisuudella tarkoitetaan käytännössä kuitenkin myös lähes reaaliaikaista toimintaa, sillä jonkinlaista viivettä esiintyy aina ennen kuin tilanmuutos heijastuu dataan järjestelmissä sekä ennen kuin muuttuneet tiedot ovat päättäjien saatavilla (Power, 2011). Azvinen ym. (2005) mukaan reaaliaikaisuus tarkoittaa, että järjestelmien tuottama tieto nykyhetkestä on saatavilla silloin, kun käyttäjä sitä tarvitsee. Power (2011) taas esittää, että reaaliaikaisen päätöksenteon tavoitteena on seurata tapahtumia, arvioida vaihtoehtoja sekä valita ja toimia niiden perusteella tapahtumahetkellä.

Dataohjautuvaa päätöksentekoa tukevat päätöstukijärjestelmät (decision support systems, DSS). Päätöstuki on laaja käsite, joka pitää sisällään kaikki ne tietokonejärjestelmät ja muut välineet, jotka auttavat yksilön, ryhmän ja organisaation päätöksenteossa. (Power, 2011.) Päätöstukijärjestelmät ovat kehittyneitä työkaluja, jotka käyttävät hyödykseen jatkuvasti enemmän big dataa. Niitä käytetään suunnittelemaan ja toteuttamaan yksilö- ja organisaatiotasoisia strategisia päätöksiä. (Aversa, Cabantous, & Haefliger, 2018.) Powerin (2011) mukaan Thierauf (1982) on määritellyt, että reaaliaikainen päätöstuki viittaa kaikkiin niihin järjestelmiin, jotka prosessoivat ja varastoivat dataa sillä hetkellä, kun sitä syntyy.

Panianin (2007) mukaan reaaliaikainen päätöksenteko on analyttinen prosessi, jonka avulla organisaatiot voivat automatisoida ”seuraavaksi parhaat toiminnot” tavoitteidensa ja päämääriensä perusteella. Se yhdistää reaaliaikaisen analytiikan potentiaalinen liiketoimintaprosesseihin. Tapahtumat analysoidaan niiden tapahtuessa ja suositellaan toimenpiteitä, joilla todennäköisimmin saadaan nämä tapahtumat johtamaan haluttuihin suorituskykytavoitteisiin.

Power (2011) ja Panian (2007) esittelevät reaaliaikaista päätöksentekoa tukevien järjestelmien ominaisuuksia, jotka erottavat ne perinteisistä päätöstukijärjestelmistä ja tekevät niistä ainutlaatuisen sopivia liiketoimintaprosessien hallintaan. Powerin (2011) mukaan reaaliaikaiselle päätöstuelle on ominaista valvontatoiminto, jossa sensori tai tiedonkeruulaite käy dataa läpi, käsittelee sitä ennalta määrättyjen sääntöjen perusteella ja lähettää sitten hälytyksen tekemistään havainnoista päätöksentekijöille. Toisin kuin ihmiset, automaattinen

järjestelmä voi tarkastella samanaikaisesti hyvin laajaa datamäärää määrätyltä ajanjaksolta (Panian, 2007). Data välittyy järjestelmien sekä yhden tai useamman päätöksentekijän välillä todella nopeasti. Tieto on päätöksentekijän käytettävissä heti, kun muutos tapahtuu eli se on aina ajankohtaista. (Power, 2011.) Etuna on, että ihmisen ei tarvitse olla jatkuvasti varuillaan ja suorittaa järjestelmälle kyselyitä havaitakseen nopeasti muuttuvassa datassa poikkeavuuksia, vaan analyysi tapahtuu automaattisesti (Panian, 2007). Päätöksentekijä voi hyötyä reaaliaikaisesta päätöstukijärjestelmästä myös, vaikka data ei olisikaan nopeasti muuttuvaa, sillä se mahdollistaa analyysien teon ja antaa neuvoja välittömästi käyttöhetkellä tai heti tarpeen tullen. Järjestelmän tuottama raportti voi olla muodoltaan joko raakaa tai tiivistettyä dataa. Järjestelmä voi myös muuttaa alkuperäisen datan käytettävään muotoon hyödyntäen mahdollisesti muun muassa ääntä, videoita tai tekstiä. Tämä muutos voi aloittaa tai vahvistaa transaktion tai tukea yhteistyötä. (Power, 2011.)

Mohamed & Al-Jaroodi (2014) jakavat reaaliaikaisen päätöstukisovelluksen toiminnan viiteen vaiheeseen: reaaliaikainen tapahtumien siirto, reaaliaikainen tilannekatsaus, reaaliaikainen analyysi, reaaliaikainen päätöksenteko ja reaaliaikainen reagointi. Ensimmäisessä vaiheessa tapahtumat tulee siirtää reaaliaikaisesti niiden prosessointipisteeseen ja edelleen päätöksentekoyksikköön. Tapahtumat voidaan siirtää raakoina, mutta verkkoliikenteen ja prosessointiaikojen lyhentämiseksi ne tulisi suodattaa tai yhdistää, kuitenkin vaikuttamatta negatiivisesti päätöksenteon tarkkuuteen.

Toisessa vaiheessa havaitaan reaaliaikaiset liiketoiminnassa tai operatiivisissa toiminnoissa tapahtuvat muutokset ja poikkeukset, esimerkiksi osakekurssin lasku rahoitusmarkkinoilla tai poikkeuksellisen runsas liikenne tietyllä tärkeällä alueella. Havaitsemisprosessi voi perustua liiketoimintaan tai operatiivisiin toimintoihin perustuviin sääntöihin. Ne voidaan määritellä joko ennalta tai ne voivat muuttua vuorovaikutuksessa järjestelmän analyttisten tai päätöksentekomoduulien kanssa. Ne voivat myös vaikuttaa edellisen vaiheen tapahtumansiirtoon. Haasteena on määritellä säännöt niin, että ne helpottavat oikeiden päätösten tekemistä oikeaan aikaan.

Kolmannessa vaiheessa tilannekatsauksen läpikäynyt reaaliaikainen tieto sekä varastoitu data, kuten kartat, aiemmat liiketoimintatransaktiot, tilanteet ja päätökset, analysoidaan ja niiden perusteella tarjotaan vaihtoehtoisia ratkaisuehdotuksia. Nämä ehdotukset voivat olla optimaalisia tai puolioptimaalisia riippuen käytettävissä olevasta rajallisesta aikaikkunasta sekä resursseista. Vaiheeseen liittyy paljon teknisiä haasteita, sillä varastoidun datan käsittely on usein haastavaa, varsinkin kun on kyse suurista määristä dataa, mikä vaatii datan hajauttamista pienempiin osiin. Tässä haasteena on hajautetun datan prosessointi tarpeeksi nopeasti, joten alustojen täytyy olla tehokkaita.

Neljäs vaihe on päätöksenteko. Päätösprosessi valitsee analyysin tarjoamista ratkaisuista parhaan vaihtoehdon parantamaan nykyistä liiketoimintaa tai kannattavuutta ja määrittämään sopivimmat toimet vastaamaan liiketoimintaympäristön vaatimuksiin. Päätösprosessi tarkistaa ennalta määritellyt säännöt,

voi ilmoittaa analyysin tulokset vastuuhenkilöille tai käynnistää muita prosesseja jatkotoimien suorittamiseksi.

Päätöksenteonprosessin viimeinen vaihe on päätösprosessin määrittelemän toiminnan käynnistäminen, toteuttaminen ja seuranta. Tämä prosessi on yleensä vuorovaikutuksessa muiden hajautettujen järjestelmien kanssa, ja haasteena onkin, että vuorovaikutuksen pitäisi olla luotettavaa ja reaaliaikaista. (Mohamed & Al-jaroodi, 2014.)

Big dataa hyödyntävien päätöstukijärjestelmien tutkimus on pitkään keskittynyt yllä käsiteltyihin teknologisiin aspekteihin ja big datan suunnittelun haasteisiin (Chen ym., 2012), mutta hiljattain on alettu ottaa huomioon myös organisaatioon liittyviä näkökulmia. Aversa ym. (2018) tunnistavat kolme toisiinsa liittyvää strategiseen päätöksentekoon liittyvää haastetta, jotka käyvät ilmi Formula 1 -ympäristöä tarkastelevasta tapaustutkimuksesta, mutta ovat yleistettävissä myös muiden organisaatioiden päätöksentekoon: päätöksenteon tilannesidonnainen luonne ja käyttömahdollisuudet, kognition hajanaisuus päätöksenteossa ja päätöstukijärjestelmän performatiivisuus. Organisaationäkökulma on vahvasti näkyvässä sekä päätöksenteon tilannesidonnaisessa luonteessa että kognition hajanaisuudessa.

Päätöksenteko hyödyntäen reaaliaikaista päätöstukijärjestelmää on tilannesidonnaista toimintaa, joka sijoittuu sosiaaliseen, fyysiseen ja tekniseen ympäristöön. Se pitää sisällään haasteita liittyen organisaatiokulttuuriin, kuten käyttöönotto, yhteistyö, päätöksentekoprosessi ja se, kuka tekee päätökset. Haasteita liittyy myös järjestelmien käyttömahdollisuuksiin. Järjestelmää voidaan käyttää useaan eri tarkoitukseen eikä sen tarjoaman analyysin tulkintoja ole rajattu. Käyttöympäristö kuitenkin vaikuttaa tulkintaan ja siihen, miten järjestelmää käytetään. Järjestelmä suosii, muotoilee ja kutsuu uusiin käyttötarkoituksiin, mutta myös samaan aikaan rajoittaa tiettyihin. (Aversa, ym., 2018.)

Dataohjatuissa päätöksenteossa päätöksentekijöiden kognitio on ajallisesti, sosiaalisesti ja aineellisesti jakautunut. Päätöksentekijät toisaalta tukeutuvat päätöstukijärjestelmään mutta toisaalta myös toisiinsa ryhmänä. Siksi dataohjatuissa päätöksenteon haasteena on ihmisen ja tietojärjestelmän välinen vuorovaikutus, mikä korostuu reaaliaikaisessa päätöksenteossa paineiden ja aikarajoitteiden alla. (Aversa, 2018.)

Aversan ym. (2018) mukaan päätöksentekoa big dataa hyödyntävien järjestelmien avulla voidaan pitää performatiivisena. Big dataa hyödyntävä päätöstukijärjestelmä sisältää monimutkaisia malleja, jotka toimivat välikkappaleina teorian ja datan välillä ja edustavat yksinkertaistettua tapaa ilmaista todellisuutta. Ne kokoavat ja strukturoivat sisään syötettyä dataa määrättyllä tavalla. (Aversa ym., 2018.) Performatiivisuutta voi jaotella sen mukaan, miten mallit lopulta vaikuttavat päätöksentekoon. Malleja voidaan esimerkiksi vain hyödyntää päätöksentekokäytännöissä tai ne voivat merkittävästi muuttaa päätöksentekoprosesseja. (MacKenzie, Muniesa & Siu, 2007, Aversan ym., 2018 mukaan.)

3 REAALIAIKAINEN BIG DATA ERI SOVELLUSALOILLA

Reaaliaikaista big dataa hyödynnetään tänä päivänä lukuisilla eri sovellusaloilla. Se tarjoaa monenlaisia taloudellisia ja käytännöllisiä mahdollisuuksia niin yrityksille, valtioille kuin yksilöillekin. Reaaliaikaista big dataa hyödyntävien järjestelmien menestyksenkäs suunnittelu, toteutus ja toiminta voi mahdollistaa muun muassa tuhansien ihmisten hengen pelastamisen, elämänlaadun parantumisen, riskien vähenemisen ja kannattavuuden parantumisen (Mohamed & Al-jaroodi, 2014). Toisaalta siihen liittyy myös moninaisia haasteita, jotka täytyy ottaa huomioon. Seuraavaksi tarkastellaan kirjallisuudesta esiin nousevia reaaliaikaisen data-analytiikan esimerkkisovelluksia organisaatioiden päätöksenteossa liiketoiminnan, valtion ja yhteiskunnan sekä urheilun aloilla.

3.1 Sovelluksia liiketoiminnan päätöksenteossa

Reaaliaikainen big data tarjoaa liiketoiminnalle monenlaisia mahdollisuuksia. Panianin (2007) mukaan reaaliaikainen päätöstuki tarjoaa yrityksille mahdollisuuden olla ketterämpiä, kilpailukykyisempiä, nopeammin reagoivia, luotettavampia, myöntyväisempiä ja kannattavampia. Sen avulla voidaan muun muassa parantaa markkinointia, kerätä sosiaalisesta mediasta arvokasta tietoa, valvoa rahoitusmarkkinoita, käydä kauppaa ja parantaa tietoturvaa (Mohamed & Al-jaroodi, 2014; Yadransjiaghdam ym. 2016). Reaaliaikainen päätöstuki ei kuitenkaan ainoastaan paranna yritysten tehokkuutta, vaan luo kokonaan uusia sovellusaloja sekä tapoja harjoittaa liiketoimintaa. (Panian, 2007.)

Azvinen ym. (2005) mukaan datan analysointi tuotteiden ja palveluiden markkinatrendien ennustamiseksi sekä yritysten liiketoimintajärjestelmien suorituskyvyn parantamiseksi on aina ollut osa kilpailukykyisen liiketoiminnan harjoittamista, mutta nykyään sekä analysoinnin että sen tuloksiin reagoimisen on tapahduttava reaaliajassa. Kirjallisuudesta esiin nousevien reaaliaikaisten data-analytiikkamenetelmien määrä liiketoiminnan hallinnan (business intelli-

gence) alalla kertookin omaa kieltään siitä, että se on yksi tärkeimmistä big data -analytiikan sovellusaloista ja tarvitsee reaaliaikaisia sovelluksia (Yadranjiaghdam ym., 2016). Tälle on kaksi pääasiallista syytä: yritysten toimintaympäristö muuttuu jatkuvasti ja teknologia kehittyy (Azvine ym., 2005). Perinteiset liiketoiminnan hallintajärjestelmät ovat yleensä olleet alakohtaisia ja käsitelleet ainoastaan strukturoitua yrityksen sisäistä liiketoimintadataa. Ne eivät myöskään ole kyenneet käsittelemään suuria määriä dataa. Nykypäivänä yritysten tapahtumadatan määrä on kasvanut huomasti, ja big data on mahdollistanut myös yrityksen ulkopuolella olevan datan analysoinnin ja hyödyntämisen. Datamäärä on siis kasvanut huomasti, mikä vaatii uudenlaista teknologiaa.

Vera-Baquero, Colomo-Palacios ja Molloy (2016) esittelevät ratkaisun, joka seuraa reaaliaikaisesti liiketoiminnan prosesseja, tapahtumia ja olosuhteiden muutoksia analysoimalla big dataa. Liiketoiminnan prosessien ymmärtäminen ja parantaminen on tärkeää, sillä toimimattomat prosessit voivat aiheuttaa muun muassa viivästyksiä, huonoa palvelua ja arvokkaiden resurssien hyödyntämättä jättämistä, mikä johtaa korkeisiin kustannuksiin sekä tulojen ja asiakasyytyväisyyden ja -uskollisuuden menetykseen. (Vera-Baquero ym., 2016.) Näiltä ongelmilta voidaan välttyä, kun toimimattomat prosessit havaitaan reaaliaikaisen analytiikan avulla. Tähän liittyy kuitenkin myös haasteita. Vera-Baqueron ym. (2016) mukaan yksi reaaliaikaisten liiketoiminnan hallintajärjestelmien suurimmista haasteista on hajautetun tiedon valvominen ja mittaaminen reaaliajassa. Data kulkee useiden heterogeenisten järjestelmien, kuten yrityksen toiminnanohjaus- tai asiakashallintajärjestelmien kautta ja sitä varastoidaan hyvin suuria määriä sekä keskitettyihin että hajautettuihin tietovarastoihin, joten data on hyvin levällään.

Toinen yrityksille tärkeä ja kasvava reaaliaikaisen data-analytiikan sovellusala on sosiaalinen media. Facebookin ja Twitterin kaltaisten sivustojen analysointi voi osoittautua varsin hyödylliseksi johtopäätösten ja ennusteiden tekemisessä maailman tietyillä alueilla tietyllä hetkellä tapahtuvista aktiviteeteista (Bifet, 2013). Ihmiset julkaisevat sosiaalisessa mediassa joka hetki sisältöjä lähettämällä valokuvia ja videoita, päivittämällä tiloja, kertomalla tapahtumia, jakamalla uutisia ja viestimällä mielipiteistään ja kiinnostuksenkohteistaan. Analysoimalla sosiaalisen median sisältöjä voidaan ennustaa ja havaita tärkeitä tapahtumia, trendejä ja jopa hätätilanteita, kun keskustelu ja jaot tietyn aiheen ympärillä yhtäkkiä lisääntyvät lyhyessä ajassa. (Nguyen & Jung, 2017.) Sosiaalinen media tarjoaa arvokasta tietoa asiakkaiden mielipiteistä ja siihen pystytään vastaamaan välittömästi. Esimerkiksi Starbucks hyödynsi sosiaalisen median reaaliaikaista analysointia lanseeratessaan uuden kahvituotteen. He olivat huolissaan siitä, että asiakkaat pitäisivät uuden tuotteen makua liian voimakkaana, joten he alkoivat heti aamulla analysoida blogeja, Twitteriä ja kahvikeskusteluforumeja nähdäkseen asiakkaiden reaktiot. Aamupäivään mennessä he olivat saaneet selville, että mausta kyllä pidettiin, mutta hintaa pidettiin liian korkeana. Tähän reagoitiin heti ja iltaan mennessä kaikki negatiiviset kommentit olivat hävinneet. Perinteisin keinoin alhainen myynti ja sen perimmäinen syy olisi havaittu vasta muutamien viikkojen kuluttua. (Watson, 2014.)

Myös sosiaalisen median palvelut hyödyntävät reaaliaikaista data-analytiikkaa omassa kehitystyössään. Esimerkiksi Facebook hyödynsi sitä viestisovellusta kehittäessään (Borthakur ym., 2011, Yadranjiaghdamin ym. mukaan, 2016) ja Twitter käyttää sitä hakuehdotuksissaan ja oikeinkirjoituksen korjaamisessa (Mishne, Dalton, Li, Sharma & Lin, 2013).

Muun muassa kauppaketjumarkkinat, valuutanvaihto- ja korkokurssit sekä raaka-aineiden hinnat synnyttävät joka sekunti valtavan määrän taloudellista dataa osake- ja optiomarkkinoille. Tämä data ei ole pelkästään suurta vaan myös dynaamista, joten taloudellisten mahdollisuuksien ja uhkien löytämiseksi havaitsemisprosessin tulee olla hyvin nopea. Siksi viime aikoina on kehitetty automatisoituja kaupankäyntijärjestelmiä sekä reaaliaikaisia petosten havaitsemisjärjestelmiä, joiden avulla yritykset voivat ennustaa arvopapereiden hintojen nousua tai laskua ennen kuin muutos todella tapahtuu, havaita ja pysäyttää laitton toiminta, kuten markkinoiden manipulointi, salaiset hintasopimukset ja väärin tietoihin perustuvat sisäpiirikaupat, mikä parantaa markkinoiden tehokkuutta ja turvaa sijoittajien liiketoimet. (Mohamed & Al-jaroodi, 2014.)

Reaaliaikaista big dataa hyödynnetään nykyään myös toimitusketjujen hallinnassa, logistiikassa sekä tuotannossa. Sen avulla voidaan esimerkiksi reaaliaikaisesti seurata varastotilannetta ja vähentää hävikkiä. Toimitusaikoja voidaan lyhentää, kun kuljetuksessa huomioidaan reaaliaikaiset sää- ja liikenneolosuhteet sekä kuljettajan luonteenpiirteet. Tuotannossa reaaliaikaisen analytiikan avulla voidaan valvoa tehokkaammin tuottavuutta ja tehtaisiin sijoitetut sensorit valvovat työturvallisuutta. (Waller & Fawcett, 2013.)

Yritysten ensisijainen mainos- ja markkinointikanava on jo vuosituhannen alusta ollut internet. Mobiililaitteiden suosion ennennäkemätön kasvu on luonut uusia mahdollisuuksia myös mobiilimarkkinointiin (Deng, Gao & Vuppapapati, 2015). Reaaliaikaista analytiikkaa hyödynnetään sekä parantamaan kohdennettua markkinointia käyttäjädataa analysoimalla että selvittämään, mitä kuluttajat haluavat kaupallisilla verkkosivustoilla todella nähdä (Deng & Gao, 2015). Esimerkiksi vähittäiskaupan alalla voidaan luoda asiakkaalle reaaliaikaisia yksilöllisiä tarjouksia, kun mobiililaitteista ja -sovelluksista peräisin olevaa dataa, kuten paikkatietoja, demografisia tietoja ja aiempaa ostokäyttäytymistä, analysoidaan reaaliaikaisesti (Power, 2014). Reaaliaikaisen analytiikan avulla voidaan myös muun muassa optimoida hinnat asiakkaan tai ajankohdan mukaan.

Deng ja Gao (2015) esittelevät mainostajille ja julkaisijoille suunnatun reaaliaikaisen ja staattisen palvelun, joka auttaa päättämään muun muassa milloin, miten ja missä kannattaa mainostaa. Palvelun toiminta edellyttää big datan analysointia, asiakaskäyttäytymismallien löytämistä sekä innovatiivista mainos-suositteleväjärjestelmä- ja trendien ennakoimista. Sen avulla mainostajat ja julkaisijat voivat vähentää kustannuksiaan ja parantaa tehokkuuttaan.

Reaaliaikaisen big datan mukanaan tuomia mahdollisuuksia mobiilimarkkinoinnille ovat muun muassa reaaliaikainen asiakkaiden sitoutuminen, asiakaskokemuksen parantaminen, brändiuskollisuus, tulojen kasvu ja asiakastytyväisyyden lisääminen. Haasteita puolestaan aiheuttaa se, että mobiililaitteissa

on erilaisia teknologioita ja käyttöjärjestelmiä. Myös niiden käyttäjät eroavat toisistaan niin mieltymyksiltään, käyttäytymiseltään kuin sosiaalisuudeltaan. Tämän seurauksena mobiililaitteisiin liittyvä tietoaineiston määrä kasvaa huomattavasti, ja asiakkaat tarkoituksenmukaisella tavalla sitouttaakseen yritysten täytyy analysoida valtavia ja monimuotoisia tietoaineistoja. Haasteet liittyvätkin mobiilidatan tyyppiin (strukturoituun tai strukturoimattomaan dataan), datan eksklusiivisuuteen, yksityisyyteen ja dataan liittyvään kontekstiin sekä oikeanlaisen mainontaa edistävän tiedon louhimiseen käyttäjätiedosta. (Deng ym., 2015.)

Reaaliaikainen big data tarjoaa yrityksille mahdollisuuden keksiä ja kehittää uusia tuotteita, palveluja ja liiketoimintamalleja. Esimerkiksi Netflix oli alun perin DVD-postimyyntiyritys, mutta tänä päivänä se on suoratoistopalvelu, jonka tarkoitus on ennakoita ja tarjota sitten katsojille reaaliajassa juuri sitä, mitä he haluavat nähdä. Big data -analytiikka on osa sen jokaista liiketoiminnan osa-aluetta. Sitä hyödynnetään muun muassa personalisoinnissa, viestinnässä, sisällöntuotannossa ja -jakelussa sekä laiteoptimoinnissa. (Marr, 2016.)

3.2 Sovelluksia valtion ja yhteiskunnan päätöksenteossa

Kirjallisuudesta nousee esiin useita esimerkkejä siitä, kuinka reaaliaikaista big dataa voidaan hyödyntää eri sovellusaloilla valtion ja yhteiskunnan etujen hyväksi. Valtioiden infrastruktuuria voidaan jatkuvasti rakentaa älykkäämmäksi. Moni kaupunginjohto hyödyntää jo nykyään reaaliaikaista analytiikkaa hallitakseen kaupungin toimintaa ja sääntelyä. Tulevaisuuden älykaupungeissa on kaikkialla sensoreita, digitaalisia laitteita ja infrastruktuuria, jotka tuottavat big dataa. Big data mahdollistaa reaaliaikaisen analyysin kaupungin elämästä sekä uusia kaupunginhallinnon muotoja ja tarjoaa aineksia visioida ja toteuttaa tehokkaita, kestäviä, kilpailukykyisiä, tuottavia, avoimia ja läpinäkyviä kaupunkoja. Useat paikallishallinnot muun muassa käyttävät hallintajärjestelmiä pitääkseen kirjaa ihmisten sitoutumisesta julkisiin palveluihin ja seuratakseen onko henkilöstö käsitellyt ongelmat. (Kitchin, 2014.)

Valtion sähköverkosta voidaan kunnostaa älykäs asentamalla sähköntuotantoon, -siirtoon, -jakeluun sekä kuluttajien päähän sensoreita, jotka tuottavat reaaliaikaista dataa muun muassa toimittajien ja kuluttajien käyttäytymisestä, nykyisestä sähköntuotannosta, kulutuksesta ja vioista. Älykäs sähköverkko parantaa sähkön tuotannon ja jakelun tehokkuutta, luotettavuutta, taloudellisuutta ja kestävyyttä, auttaa toteuttamaan dynaamisia hinnoittelumalleja ja tarjoaa kuluttajille lähes reaaliaikaista tietoa energiankulutuksesta. Tämä auttaa heitä hallitsemaan sähkönkulutustaan sekä tarpeidensa että hintojen mukaan. Esimerkiksi astianpesukone kannattaa pestä, kun hinnat ovat alhaisemmat. (Mohamed & Al-jaroodi, 2014.)

Liikenteessä on tänä päivänä paljon sensoreita, joiden tuottaman datan perusteella voidaan seurata reaaliaikaisesti liikenneneruuhkia suurissa kaupungeissa ja vilkkaila teillä. Sensoreita on sekä tieverkossa että ajoneuvoissa. Esimer-

kiksi valvontakamerat, tienpintaan asennetut sensorit sekä GPS-järjestelmät ja nopeusmittarit voidaan yhdistää erilaisiin viestinnän mahdollistaviin teknologioihin kuten wifiin, bluetoothiin tai satelliittiviestintään, ja näin pystytään reaaliaikaisesti seuraamaan muun muassa ajoneuvojen sijainteja, keskinopeuksia, kuljettajien ajokäyttäytymistä sekä tieolosuhteita. (Mohamed & Al-jaroodi, 2014.) Kun poliisi voi valvoa kameroita ja live-lokeja tapahtumista reaaliaikaisesti, voidaan resurssit tehokkaasti ohjata oikeisiin paikkoihin (Kitchin, 2014). Ruuhkiin ja onnettomuuksiin vaikuttavien tekijöiden reaaliaikainen tunnistaminen ja analysointi voi käynnistää turvallisuusvaroitukset ajoissa (Shi & Abdel-Aty, 2015), mikä parantaa liikenneturvallisuutta. Lisäksi datan perusteella pystytään tarjoamaan reaaliaikainen palvelu, joka kertoo käyttäjälle liikenneolosuhteiden mukaan nopeimman reitin mihin tahansa kohteeseen (Biem ym., 2010). Palvelu on hyödyllinen kenelle tahansa, mutta erityisen tärkeä se on hätäajoneuvoille, jotta apu saadaan mahdollisimman nopeasti oikeaan paikkaan. Reaaliaikaisella tiedolla voidaan myös kehittää kuljetuspalveluja. Tavaroiden toimitusaika pystytään minimoimaan ja ihmisille voidaan tarjota tietoa energiatehokkaimmasta tavasta matkustaa. (Mohamed & Al-jaroodi, 2014.)

Onnettomuuksien ja vaaratilanteiden todennäköisyys kasvaa suurissa tapahtumissa, joissa on paikalla massiivinen määrä ihmisiä. Reaaliaikainen data helpottaa viranomaisten päätöksentekoa ja väkijoukon hallintaa. Liikenteen- sekä mobiililaitteiden sijainnin seurannan perusteella voidaan tällaisissa tilanteissa parantaa yleistä turvallisuutta avaamalla lisää pysäköintialueita, ohjaamalla liikennettä, sulkemalla joitakin katuja helpottamaan jalankulkijoiden liikkuvuutta tai lisäämällä poliisin läsnäoloa tietyillä alueilla. Onnettomuuden satuttuessa viranomaiset kykenevät paremmin tunnistamaan onnettomuuspaikan ja suunnittelemaan välittömästi vastausstrategioita, kuten hätäajoneuvojen reitit, alueen ihmisten evakuoinnin ja tarvittavien tukipalvelujen tarjoamisen ajoissa. Reaaliaikainen päätöksentekojärjestelmästä on hyötyä myös muunlaisissa hätätilanteissa kuten maanjäristyksissä, tulivuoren purkauksissa, tulvissa, sodissa ja terrori-iskuissa, joissa kärsineitä on pystyttävä auttamaan heti. Päätöksentekojärjestelmä kerää mahdollisimman paljon tietoa tilanteen kehittymisestä ja käytettävissä olevista resursseista ja ehdottaa niiden perusteella hätätilatoimia nopeasti. (Mohamed & Al-jaroodi, 2014.)

Reaaliaikainen data ei auta ainoastaan onnettomuuden satuttua, vaan se voi myös varoittaa luonnonkatastrofeista ennalta, minkä ansiosta on mahdollista pelastaa tuhansia ihmishenkiä. Tällaisilla varhaisvaroitujärjestelmillä voidaan havaita esimerkiksi, minne tsunami iskee tai ennustaa metsäpalojen laajenemista. Varhaisvaroitujärjestelmät käsittelevät reaaliaikaisesti kerättyä dataa maaperässä ja veden alla olevista sensoreista sekä etäsensoreista, jotka on asennettu miehittämättömiin ilma-aluksiin ja satelliitteihin ja jotka sisältävät muun muassa säätietoja, maantieteellisiä tai alueellisia karttoja. (Mohamed & Al-jaroodi, 2014.)

Ympäristöön asennetuista sensoreista saatavaa reaaliaikaista dataa käytetään myös muihin tarkoituksiin. Koko kaupungin kattavasta sensoriverkosta dataa kokoamalla voidaan mitata muun muassa ilmansaastumista, vesitasoja

tai seismistä toimintaa (Kitchin, 2014). Esimerkiksi Japanissa sensoreilla mitataan maaperän terveyttä. Yritykset yhdistävät kerätyn datan säätietoihin ja tarjoavat tiedon avulla palvelun maanviljelijöille. (Marr, 2015.) Yadranjiaghdamin ym. (2016) mukaan yksi reaaliaikaisen data-analytiikan tärkeimmistä sovellusalueista onkin ympäristön tilan analyysi. Sillä pyritään muun muassa ennustamaan energiankulutusta ja sen vaikutuksia ympäristöön, seuraamaan jätehuollon toimintaa, paikantamaan moottoriteiden kohtia, joissa CO₂-päästöt ovat yleisimpiä, parantamaan veden laatua ja ennustamaan muun muassa säätä sekä kasvihuonepäästöjä.

Viime aikoina data-analytiikan avulla saadut tiedot on yhä useammissa kaupungeissa koottu yhteen keskittymään. Esimerkiksi Rio de Janeirossa, New Yorkissa, Dublinissa, Santanderissa ja Lontoossa on kehitetty järjestelmiä ja keskuksia, jotka kokoavat eri laitoksista, kuten julkisista palveluista ja liikenteestä, hätäpalveluista ja sääkeskuksista, peräisin olevat tiedot yhteen paikkaan. Näin eri saadaan reaaliajassa yksityiskohtaista tietoa eri järjestelmien tuottaman datan suhteista toisiinsa. Tietojen avulla voidaan tarkastella ja ennakoida eri näkökulmista kaupungin elämää sekä siinä tapahtuvia muutoksia, mikä tukee kaupunkilaisten päätöksentekoa sekä jokapäiväistä elämää ja tuo uusia visioita kaupungin jatkuvaan kehittämiseen. Osa näiden järjestelmien tuottamasta tiedosta on avointa kaikille, joten myös tavalliset asukkaat voivat esimerkiksi puhelinsovelluksen tai nettisivuston kautta päästä käsiksi muun muassa reaaliaikaiseen tietoon säästä, julkisen liikenteen viivästyksistä, yhteiskäyttöön tarkoitettujen polkupyörien saatavuudesta, joen vedenpinnan tasosta, sähkön kysynnästä, osakemarkkinoista, kaupungin Twitter-trendeistä tai jopa onnellisuuden tasosta. (Kitchin, 2014.)

Yhteiskunnan turvallisuutta voidaan parantaa reaaliaikaistamalla perinteisesti äärimmäisen hitaasta tunnistusprosessista kärsivät valvontajärjestelmät (Yadranjiaghdam ym., 2016). Baig ja Jabeen (2016) esittelevät tarkemmin opiskelijoiden käyttäytymistä seuraavan analytiikkajärjestelmän, jonka avulla voidaan arvioida kuinka todennäköisesti opiskelijat altistuvat terroristisille ideologioille. Järjestelmä hyödyntää dataa perinteisistä tietokannoista (esim. kurssisuoritusket), henkilökohtaista dataa (esim. puheluhistoria), digitaalista jalanjälkeä (esim. aktiivisuus yliopiston verkossa) sekä yliopiston hallinnoimia datanlähteitä (esim. videovalvonta ja autojen liikkumisen seuranta pysäköintialueella). Haasteena järjestelmässä on, että henkilökohtaisen datan hallussapito on laitonta ja opiskelijoiden liikkeiden seuraaminen on eettisesti kyseenalaista, mikäli ei ole hyvin vahvaa epäilystä opiskelijan terroristisille ideologioille altistumisesta. (Baig & Jabeen, 2016.) Kyseinen järjestelmä on esimerkki siitä, miten reaaliaikaisen analytiikan avulla voidaan tehostaa valvontaa. Myös Hua, Jiang ja Feng (2014) esittelevät perinteisiä valvontajärjestelmiä tukevan menetelmän, jonka avulla voidaan reaaliaikaisia kuvia tutkimalla löytää tungoksessa eksynyt lapsi. Ihmiset ottavat kaiken aikaa kuvia ympäristöstä ja julkaisevat niitä sosiaalisessa mediassa. Näitä kuvia analysoimalla saadaan selville, onko lapsi ollut lähistöllä. Haasteena on, että analysoitavien kuvien tulee olla julkisia tai järjestelmä ei pysty niitä analysoimaan. (Hua ym., 2014.)

Vaikka reaaliaikainen big data tarjoaa selvästi paljon mahdollisuuksia kaupungeille, Kitchin (2014) nostaa esiin myös big datan mukanaan tuomia haasteita kaupunkien näkökulmasta. Yksi näistä on älykaupungin tietoturvan haavoittuvuus. Kuten edellä mainituista esimerkeistäkin voi todeta, yhteiskunnan toiminta perustuu tänä päivänä yhä enemmän verkkopohjaisille tietojärjestelmille ja on riippuvainen ohjelmistoista. Siksi sen toiminta ja turvallisuus on myös jatkuvasti uhattuna erilaisten haittaohjelmien ja epäilyttävien tiedostojen takia. Jos esimerkiksi metrojärjestelmän hallintaohjelmisto kaatuu, junat eivät kulje (Kitchin, 2014). Kyberrikosten määrä on viime vuosina kasvanut nopeasti, sillä rikolliset etsivät jatkuvasti uusia tapoja kiertää turvallisuusratkaisuja päätökseen laittomasti käsiksi järjestelmiin ja verkkoihin. Reaaliaikaisen analytiikan avulla voidaan kuitenkin myös vastata tähän. Mahmood ja Afzal (2013) tarkastelevat artikkelissaan turvallisuusanalytiikkaa, jolla seurataan reaaliaikaisia tietovirtoja ja pystytään näin havaitsemaan ja torjumaan kyberhyökkäyksiä. Turvallisuusanalytiikkaa sovelletaan reaaliaikaisesti esimerkiksi verkkoliikenteessä, verkkomaksuliikenteessä, verkkopalvelimilla, datalähteissä ja käyttäjätunnusten väärinkäytösten havaitsemisessa.

Muut Kitchinin (2014) esiin nostamat haasteet liittyvät urbaanin big datan politiikkaan, teknokraattiseen hallintotapaan, teknologian lukkiutumiseen sekä kaupunkien panoptisuuteen. Vaikka data on neutraalisti ja objektiivisesti teknisten rajoitteiden puitteissa kerättyjä totuudenmukaisia kuvia ja lukemia maailmasta ilman minkäänlaista politiikkaa tai agendaa, se ei koskaan ole irrallaan ihmisistä, jotka sitä keräävät ja analysoivat. Sitä käytetään tiettyjen tavoitteiden saavuttamiseen, mikä etenkin älykaupunkien kontekstissa luo big dataan liittyvän poliittisen haasteen. On tärkeää ymmärtää, mitä arvoja ja agendoja big datan analysointimenetelmiin liittyy ja kenen etua ne palvelevat. Tahto hallita ja säädellä kaupunkia tiedon ja analytiikkajärjestelmien kautta edistää teknokraattista hallintotapaa, jossa kaikkia kaupungin osa-alueita täytyy voida mitata, valvoa ja käsitellä teknisinä ongelmina ja toisaalta ratkaisuina. Haasteena on, että tällainen hallintotapa jättää huomiotta muun muassa kulttuurin, politiikan, toimintaperiaatteet, hallinnon ja kaupunkielämää muokkaavan pääoman. Teknologiset ratkaisut eivät yksinään ratkaise kaupunkien syvään juurtuneita rakenteellisia ongelmia, sillä ne eivät puutu niiden perimmäisiin syihin. Ne pikemminkin mahdollistavat vain ongelmien ilmentymien tehokkaamman hallinnan. Big dataan perustuva päätöksenteko myös tarjoaa kaupungin päättäjille puolustuksen sellaisia päätöksiä vastaan, jotka herättävät eettisiä tai vastuullisia kysymyksiä, sillä se antaa heille mahdollisuuden perustella kaiken datan avulla. Haaste kaupungin näkökulmasta on myös, että teknologian ollessa kaupunkisuunnittelussa suuressa roolissa yritykset ottavat vallan ja toimivat omien etujensa mukaan. Maailman suurimmat ohjelmisto- ja laitevalmistajat pitävät kaupunkisuunnittelua suurina, pitkäaikaisina ja potentiaalisina markkinoina tuotteilleen. Tämä johtaa monopoliasemien syntymiseen ja teknologiseen lukkiutumiseen, kun kaupunki yhdistyy pitkällä aikavälillä vain tiettyihin teknologisiin ratkaisuihin. Panoptisuudella viitataan yhteiskunnan lisääntyneeseen valvontakulttuuriin, sillä nykyään yksilön toimia, vuorovaikutusta ja transaktioita

voidaan jäljittää hyvin yksityiskohtaisesti. Järjestelmillä pyritään tehokkaampaan hallintoon, mikä voi uhata yksilön oikeutta yksityiselämään, luottamuksellisuuteen ja sananvapauteen. Dataa käsitellään paljon yritysten tietojärjestelmissä, ja varsinkin tästä syystä on tärkeää valvoa, ettei tietoja käytetä väärin, sillä muuten kansalaiset vastustavat reaaliaikaista analytiikkaa. (Kitchin, 2014.)

Edellä mainittujen valtion infrastruktuuriin liittyvien sovellusalueiden lisäksi reaaliaikaista dataa hyödynnetään myös esimerkiksi valtioiden sotilaallisessa päätöksenteossa. Sodan voittamisen avain on oikeiden päätösten tekeminen nopeasti, joten tietoa tilanteesta pitää kerätä reaaliaikaisesti. Esimerkiksi GPS-teknologioiden, satelliittien, ilma-alusten sensoreiden ja muiden havainnointitekniikoiden tuottaman datan avulla on mahdollista seurata reaaliaikaisesti omien resurssien, kuten satojen erilaisten laitteistojen, ajoneuvojen, rakenteiden, viestintäjärjestelmien sekä sotilaiden sijaintia sekä vastustajan liikkeitä. (Mohamed & Al-jaroodi, 2014.)

Myös terveydenhuollon käytäntöjä pystytään jatkuvasti parantamaan reaaliaikaisen datan avulla, mikä lisää yhteiskunnan hyvinvointia. Kirjallisuudesta löytyy lukuisia esimerkkejä reaaliaikaisen analytiikan mahdollisuuksista terveydenhuollon kehittämiseen. Tan, Liun & Nkabinden (2016) mukaan big data -analytiikka ei ole terveydenhuollolle pelkästään mahdollisuus vaan välttämättömyys. Esimerkiksi Sujatha, Devi, Kiran ja Manivannan (2016) esittelevät menetelmän, joka käyttää reaaliaikaista analytiikkaa ennustamaan diabeetikkojen selviytymistodennäköisyyttä ja sairaalaokeskelun pituutta. Wang, Qiu ja Guo (2017) esittelevät puolestaan uuden telehealth-järjestelmän, joka tarjoaa reaaliaikaisia tietopäivityksiä potilaiden tilasta. Big data -analytiikkaa on käytetty terveydenhuollon alalla myös esimerkiksi terveystietojen tarjoamiseen verkossa, väärinkäytösten havaitsemiseen, lääkelöydöksiin, hoidon päätöstukijärjestelmissä, tietokoneavusteisissa diagnooseissa sekä biolääketieteellisissä sovelluksissa. Sen avulla on pystytty ennakoimaan epidemioita, hoitamaan sairauksia, parantamaan elämänlaatua ja estämään estettävissä olevia kuolemia. (Ta ym., 2016.) Terveydenhuollon data-analytiikkajärjestelmien toteutuksessa on voi olla teknisiä haasteita liittyen esimerkiksi hajautetuissa järjestelmissä esiintyviin viiveisiin (Wang ym., 2016) ja datan lähteiden monimutkaisuuteen (Ta ym., 2016). Datan lähteitä alalla ovat muun muassa biolääketieteelliset kuvat ja signaalit, genomidata, hoitotekstit ja sosiaalinen media, josta löytyviä malleja voidaan hyödyntää ennustamaan maailmanlaajuisia terveystilanteita tai tartuntatauteja. Tässä haasteena kuitenkin on, että sosiaalisista verkoista peräisin oleva data ei ole aina luotettavaa. (Ta ym., 2016.)

Koulutuksen alalla Baigin ja Jabeenin (2016) mukaan oleellisimpia ovat pedagogiikkaa tukevat analytiikkajärjestelmät. He mainitsevat muun muassa opiskelijoiden oppimiskyvyn ennakoinnin, älykkäät kurssisuositukset, dataohjautuvan oppimisanalytiikan ja henkilökohtaiset oppimissuunnitelmat. Analytiikan avulla voidaan selvittää esimerkiksi opiskelijoiden valmiutta korkeakouluun tai työuraan. (Baig & Jabeen, 2016.)

3.3 Sovelluksia urheilun päätöksenteossa

Big data on lähivuosina tullut myös urheilun pariin ja avannut urheiluorganisaatioille valtavasti mahdollisuuksia. Urheilun alalla analytiikkaa sovelletaan moninaiisiin tarkoituksiin liittyen muun muassa urheilijoiden suorituskykyyn, pelaajavalintoihin, asiakassuhteisiin, liiketoiminnan hallintaan ja loukkaantumisten ehkäisemiseen. Urheiluanalytiikkaa hyödynnetään niin joukkueiden sisäisissä analyyseissa kuin fanien toimesta vedonlyönneissä, nettisivustoilla, videopeleissä ja televisiolähetysten analyyseissa ja kommentteissa. (Davenport, 2014.)

Marr (2015) toteaa, että teknologia on tehnyt urheilusta monella tapaa älykkäämpää lähes lajissa kuin lajissa. Esimerkiksi tenniksessä hyödynnetään järjestelmää nimeltä SlamTracker, joka nauhoittaa pelaajan suorituskykyä ja tuottaa reaaliaikaisia tilastoja ja otteluanalytiikkaa. Koripalloon voi olla kiinnitetty jopa 200 sensoria, jotka tuottavat pelaajalle ja valmentajalle palautetta suorituskyvystä. Rugbyssa on käytössä GPS-järjestelmä, joka tuottaa reaaliaikaista tietoa valmentajille pelaajan keskinopeudesta, sykkeestä ja hetkellisestä suoritusasteesta verrattuna normaalitasoon, jotta mahdolliset ongelmat havaitaan ennen kuin ne ilmenevät. Järjestelmä auttaa valmentajia muun muassa päättämään pelaajavaihdoksista sekä tekemään päätöksiä, joilla ehkäistään mahdollisia loukkaantumisia. Samankaltaisia järjestelmiä käytetään myös jalkapallossa sekä monissa olympialajeissa, kuten pyöräilyssä. (Marr, 2015.)

Davenport (2014) jaottelee urheiluanalytiikan kolmeen osa-alueeseen: pelaajan tai pelin suorituskykyanalytiikkaan, urheiluliiketoiminnan analytiikkaan sekä pelaajien terveyden edistämisen ja loukkaantumisten ehkäisemisen analytiikkaan. Jokainen näistä pitää sisällään reaaliaikaisen urheiluanalytiikan sovelluksia. Urheilun päätöksenteon tueksi kerättävä data on tyypillisesti videodataa, biometristä dataa tai sijaintidataa (Davenport, 2014).

Urheiluorganisaatioiden päätöksenteon tukena erityistä mielenkiintoa on herättänyt suorituskykyanalytiikka. Stensland ym. (2014) esittelemä jalkapallovalmentajien käyttöön kehitetty Bagadus-järjestelmä on esimerkki pelaajien ja joukkueen suorituskykyä parantamaan pyrkivästä analytiikkajärjestelmästä. Sen tarkoitus on integroida jo olemassa olevat järjestelmät sekä tarjota reaaliaikainen esitys urheilutapahtumasta. Se koostuu kolmesta alijärjestelmästä: video-, sijainti- sekä analytiikka-alijärjestelmistä. Videojärjestelmä tallentaa kentältä panoramakuvaa sekä mahdollistaa tietyn yksittäisen pelaajan seuraamisen. Sijaintijärjestelmä perustuu kentälle sijoitettuihin sensoreihin, jotka jäljittävät pelaajien liikkeitä. Analytiikkajärjestelmä analysoi kahden muun järjestelmän tuottamaa dataa ja tarjoaa valmentajalle reaaliaikaista tietoa tablettiin tai puhelimeen. Tiedon avulla valmentaja voi ymmärtää paremmin vastustajia ja hioa pelitaktiikkaa, mikä parantaa joukkueen suorituskykyä. Järjestelmän haasteena on sen videoalijärjestelmän kameroiden linssien aiheuttamat vääristymät ja ongelmallinen valotus. Myös kentälle sijoitetut sensorit ovat epätarkkoja, jos pe-

laajia on paljon samassa paikassa. (Stensland ym., 2014.) Vastaavanlaiset tekniset haasteet ovat yleisiä monille suorituskykyanalytiikkajärjestelmille.

Gruetzemacher, Gupta & Wilkerson (2016) esittelevät pelaajien yksilöllistä loukkaantumisriskiä arvioivan SIPS-järjestelmän (Sports Injury Prevention Screening), joka edustaa terveyden edistämisen ja loukkaantumisten ehkäisemisen analytiikkaa. Se luo henkilökohtaisia suunnitelmia loukkaantumisten ehkäisemiseksi. Järjestelmä toimii mobiililaitteilla ja kerää niistä sensoridataa, jonka perusteella voidaan suorittaa niin fyysisiä, kognitiivisia kuin näitä yhdistäviä testejä. Testeillä mitataan neuromuskulaarista koordinaatiota ja neurokognitiota. Ne tehdään jokaiselle pelaajalle kauden alussa, jolloin kullekin määritellään perustaso. Jos kauden edetessä ilmenee tarvetta tai epäilyksiä mahdollisesta loukkaantumisesta, voidaan suorittaa uusi testi, jonka tuloksista voidaan havaita esimerkiksi mahdollinen aivotärähdys ja reagoida siihen ajoissa. Kehitteillä on ennakoivaa analytiikkaa hyödyntävä ominaisuus, joka tarjoaa urheilijalle, lääkäreille ja valmentajalle reaaliaikaista palautetta pelaajan terveydentilasta.

Urheiluliiketoiminnan osa-alueen sovellukset ovat hyvin samankaltaisia kuin liiketoiminnan alalla yleensä. Mondellon ja Kamken (2014) mukaan urheiluorganisaatiot hyödyntävät analytiikkaa esimerkiksi markkinoinnissaan, asiakashallinnassa sekä lippujen hinnoittelussa.

Davenportin (2014) mukaan haasteita urheiluanalytiikan käytölle joka osa-alueella aiheuttaa muun muassa urheilujoukkueiden perinteinen kulttuuri. Harva omistaja, manageri, valmentaja tai pelaaja tavoittelee uraa urheilun parissa siksi, että olisi kiinnostunut analytiikasta. Vaikka oleellista dataa olisi saatavilla, he eivät välttämättä käytä sitä hyväkseen, vaan luottavat omaan intuitioonsa ja kokemukseensa. Myös Lee (2016) toteaa, että analytiikkaan perustuva päätöksenteko saa urheiluorganisaatioissa osakseen suurta vastarintaa. Poikkeuksena tästä voidaan pitää moottoriurheilua, jossa teknologia on todella suuressa roolissa ja päätöksenteko on ollut dataohjautuvaa jo pitkään. Esimerkiksi Formula 1:ssä livedataa on käytetty jo 1980-luvulta lähtien (Marr, 2016).

Autourheilun kuninkuusluokka Formula 1 onkin ideaalinen ympäristö tutkia reaaliaikaisen big datan hyödyntämistä päätöksenteossa (George, Haas & Pentland, 2014), koska tiimien big dataa hyödyntävät päätöstukijärjestelmät ovat alan kehittyneimpiä edelläkävijöitä ja laji on tunnettu siitä, että jokainen päätös analysoidaan tarkasti, kisaolosuhteissa päätökset tulee tehdä nopeasti paineen alla ja suorituskyky sekä kilpailuetu ovat selkeästi mitattavissa (Aversa ym., 2018).

Livedatan hyödyntämisen alkuaikoina data tuli paketteina auton ajaessa varikkosuoralle, mutta nykyään varikkomuurilla istuvat insinöörit saavat sen reaaliajassa käyttöönsä. Data välitetään myös tallien tehtaille vain sekunnin murto-osia myöhemmin. Kaikissa F1-autoissa on noin 200 sensoria, joilla tarkkaillaan muun muassa bensamäärää, rengaspaineita, g-voimia ja kuljettajan toimintaa. Jokainen auto tuottaa pelkästään yhden kisan aikana noin 50 gigatavua dataa, jonka ansiosta tiimit säästävät aikaa ja rahaa, kun ei tarvitse kokeilla ja oppia virheiden kautta. (Marr, 2016.)

Datan avulla analysoidaan auton komponenttien teknistä suorituskykyä, kuljettajan reaktioita, varikkopysähdysten viiveitä sekä tallin ja kuljettajan välistä kommunikaatiota, joka vaikuttaa kokonaisvaltaiseen suorituskykyyn (George ym., 2014). Sitä käytetään auton säätöjen reaaliaikaiseen parantamiseen kisaviikonlopun ajosessioiden aikana sekä suorituskyvyn sekä luotettavuuden parantamiseen pidemmän aikavälin tähtäimellä (Marr, 2016). Noin 60 prosenttia kisan aikana kerätystä datasta myös hyödynnetään kisan aikana, 40 prosenttia varastoidaan myöhempää käyttöä varten (Aversa ym., 2018). Lisäksi dataa hyödynnetään simulaattoreissa, sillä tänä päivänä autojen testausta on rajoitettu säännöissä huomattavasti. Tästä syystä Formula 1-tiimit turvautuvat simulaatioihin ja data-analyysiin, jotka tarjoavat selvän kuvan auton suorituskyvystä radalla. (Marr, 2016.)

Erityisen suuri merkitys reaaliaikaisella datalla on kuitenkin kilpailun aikana. Formula 1 -tiimien ilmeisin strateginen tavoite on voittaa maailmanmestaruussarjan osakilpailuja, jotta menestyksen tuomat edut, kuten palkintorahat, sponsorisopimukset ja maailmanlaajuinen näkyvyys, pystytään maksimoimaan. Autosta reaaliaikaisesti välittyvän datan perusteella tiimit voivat tarkkailla sen osien kuntoa ja suorituskykyä, ja kuljettaja voi sitten autosta käsin muuttaa tiettyjä asetuksia, kuten moottorikarttoja, jousitusta tai jarrubalanssia. Tiimit laativat kisastrategiansa perustuen big dataa hyödyntävään päätöstukijärjestelmään. Onnistuneen strategian kannalta kriittisin päätös on varikkopysähdysten määrä ja ajankohta. (Aversa ym., 2018.)

Formula 1 ympäristönä ilmentää hyvin myös reaaliaikaiseen päätöksenteon haasteita. Päätöksentekijöillä on kovat paineet ja heillä voi olla liian paljon useita eri tehtäviä hoidettavana. Työympäristö on täynnä häiriötekijöitä, kuten ohi ajavat autot, valokuvaajat, suora televisiolähetys, sääolosuhteet ja varikkomuurin sijainti niin, ettei siitä ole suoraa näkymää kisatapahtumiin. Kommunikointi tapahtuu radion välityksellä. Myös päätöstukijärjestelmä aiheuttaa joskus ongelmia. Esimerkiksi kauden 2010 päätösosakilpailussa Ferrari hävisi maailmanmestaruuden strategisen virheen takia. Päätöstukijärjestelmä tarjosi vain kaksi vaihtoehtoa, joista kumpikaan ei ollut oikea. Ongelman aiheutti se, että järjestelmään ei voitu kesken kisan päivittää parametreja, vaikka havaittiin, että ohittaminen oli vaikeampaa kuin oli odotettu. Järjestelmä oli myös yrityskohmainen eli se keskittyi Ferrarin omista autoistaan keräämään dataa eikä kilpailijoihin. Ferrarin tiukka organisaatorakenne ja hierarkia kuitenkin edellytti, että strategiasta vastannut insinööri valitsee vaihtoehdoista toisen, mikä johti maailmanmestaruuden häviämiseen. (Aversa ym., 2018.)

4 YHTEENVETO JA JOHTOPÄÄTÖKSET

Dataa syntyy suuria ja kompleksisia määriä samaan aikaan kun perinteiset datan prosessointisovellukset eivät pysty niitä käsittelemään. Toisaalta on olemassa kasvava tarve saada operaationaalista datasta irti tietoa reaaliaikaisesti, mikä on ratkaisevaa nopeasti muuttuvissa tilanteissa. Mitä nopeammin data saadaan valjastettua oivalluksiksi, sitä suurempi hyöty siitä on arvon saamisessa, kustannuksien pienentämisessä ja tehokkuuden kasvattamisessa. Valtava datamäärä tarjoaa loistavan mahdollisuuden hallita ja käyttää sitä hyödyllisillä sovellusaloilla. Organisaatiot voivat oppia tilanteestaan ja suorituskyvystään enemmän käyttämällä big data -analytiikkaa, ja ne voivat käyttää tietoa päivittääkseen päätöksentekoprosessejaan ja parantaakseen suorituskyykyään. Jos suurta määrää dataa analysoidaan tehokkaasti, se voi johtaa ongelmien löytymiseen välittömästi ja älykkäiden ideoiden keksimiseen. (Yadranjiaghdam ym., 2016.)

Tämän tutkielman tavoitteena oli systemaattisen kirjallisuuskatsauksen keinoin selvittää big datan hyödyntämisen vaikutuksia reaaliaikaiseen päätöksentekoon ja tuoda esille, millaisia mahdollisuuksia se tarjoaa eri sovellusaloilla sekä toisaalta millaisia haasteita siihen liittyy. Tutkielman haasteena oli suoraan big datan hyödyntämistä reaaliaikaisessa päätöksenteossa käsittelevien lähteiden vähäinen määrä, joten tulokset on johdettu yleisesti big dataa päätöksenteossa käsittelevistä lähteistä sekä reaaliaikaisen big datan esimerkkisovelluksista.

Big data vaikuttaa reaaliaikaiseen päätöksentekoon monin tavoin. Sen avulla pystytään tekemään parempia päätöksiä niin taloudellisesti kuin toiminnallisestikin. Se edesauttaa organisaatioiden kasvua, ketteryyttä, tehokkuutta ja kilpailijoista erottumista. Toisaalta organisaatioiden käytössä oleva valtava big datan määrä aiheuttaa päätöksentekijöille kognitiivista ylikuormitusta. Lisäksi kognitio on reaaliaikaisessa päätöksentekotilanteessa ajallisesti, sosiaalisesti ja aineellisesti jakautunut, kun tukeudutaan sekä dataan että muihin ihmisiin. Reaaliaikaisessa päätöksenteossa myös paineet ja aikarajoitteet korostuvat.

Big datan myötä organisaatiokulttuuri muuttuu. Myös tämä voi huolestuttaa ja aiheuttaa ahdistusta päätöksentekijöissä, sillä he ei eivät ole varmoja omasta tai organisaation valmiudesta muutokseen. Kun päätöksenteosta tulee

dataohjautuvaa johtajien intuitioon perustuvan päätöksenteon sijasta, päätöksentekoprosessiin osallistuu entistä useampia tahoja. Dataa luodaan organisaation joka tasolla sekä ulkoisissa sidosryhmissä. Data-analytiikan osaajien rooli päätöksentekoprosessissa kasvaa. Big dataa hyödyntävät päätöstukijärjestelmät sisältävät monimutkaisia malleja ja sääntöjä, joiden perusteella ne ehdottavat vaihtoehtoja. Ihmisen on kuitenkin määriteltävä nämä säännöt, joten data-analyttikoiden ja päätöksentekijöiden yhteistyön on toimittava hyvin. Organisaatiokulttuurin muutos onkin yksi suurimmista big datan mukanaan tuomista haasteista.

Reaaliaikainen big data kuitenkin ennen kaikkea helpottaa päätöksentekijöiden työtä. Toimintaympäristössä tapahtuvat muutokset havaitaan automaattisesti tapahtumahetkellä, eikä ihmisen siis tarvitse olla jatkuvasti varuillaan. Jotkut päätökset voidaan jopa automatisoida kokonaan, mutta tämä ei kuitenkaan poista päätöksentekijöiden vastuuta. Monet luottamukselliset, lainopilliset ja eettiset asiat vaativat edelleen ihmisen vastaamaan päätöksentekoprosessista.

Reaaliaikainen big data tarjoaa organisaatioille mahdollisuuksia alalla kuin alalla. Taulukkoon 1 on poimittu kirjallisuudesta esiin nousevia reaaliaikaisen big datan sovellusalueita liiketoiminnan, valtion ja yhteiskunnan sekä urheilun päätöksenteon näkökulmista. Jaottelu pohjautuu pääosin Ydranjiaghdamin ym. (2016) sekä Mohamedin ja Al-jaroodin (2014) kirjallisuuskatsauksiin reaaliaikaisen big datan sovellusalueista. Kirjallisuuskatsauksissa käsitellyt sovellusalat luokiteltiin tässä tutkielmassa liiketoiminnan sekä valtion ja yhteiskunnan alle. Urheilu otettiin mukaan henkilökohtaisen kiinnostuksen vuoksi.

Taulukko 1 Reaaliaikaisen big datan sovellusalueita päätöksenteossa

Liiketoiminta	Valtio ja yhteiskunta	Urheilu
Markkinointi ja mainonta	Älykäs infrastruktuuri: esim. sähköverkko ja liikenne	Suorituskyky: esim. pelaajavalinnat ja kisastrategiat
Liiketoimintatiedon hallinta (BI)	Hallinto	Urheiluliiketoiminta
Sosiaalinen media	Vaaratilanteiden ehkäisy	Terveystiedon edistäminen ja loukkaantumisten ehkäisy
Osake- ja optiomarkkinat (petosten havaitseminen)	Hätätilat ja luonnonkatastrofit	Tulospalvelut
Tuotanto	Ympäristö	Vedonlyönti
Toimitusketjun hallinta ja logistiikka	Maanviljelys	Tv-lähetykset
Tietoturva	Tietoturva	
	Terveystiedon huolto	
	Koulutus	
	Sotilaallinen päätöksenteko	

Mahdollisuuksien kirjo on miltei rajaton. Watsonin (2014) mukaan big datalla on potentiaalia saada aikaan maailma, jossa väkiluku kasvaa, mutta palveluiden ja infrastruktuurin raskaus vähenee, terveydenhuolto paranee dramaattisesti paremman tehokkuuden, mutta pienempien investointien ansiosta, uhka yleiselle turvallisuudelle ja valtioiden rajoille lisääntyy, mutta turvallisuuden taso on silti suurempi, ja sääolosuhteet ovat jatkuvasti kovemmat, mutta niitä voidaan ennustaa ja hallita paremmin. Maailmassa on enemmän autoja, mutta vähemmän ruuhkia, enemmän vakuutuskorvaushakemuksia, mutta vähemmän petoksia, ja vähemmän luonnonvaroja, mutta runsaammin halvempaa energiaa.

Koska reaaliaikainen big data mahdollistaa historian tarkkailun sijaan asioiden seuraamisen tapahtumahetkellä, organisaatiot voivat reagoida toimintaympäristön muutoksiin välittömästi ja jopa ennakoida niitä. Tämän ansiosta voidaan muun muassa pelastaa henkiä, parantaa elämänlaatua, vähentää riskejä ja parantaa toiminnan kannattavuutta. Liiketoiminnan alalla reaaliaikainen big data tarjoaa yrityksille mahdollisuuksia muun muassa kehittää liiketoimintaprosesseja, parantaa markkinointia, kerätä sosiaalisesta mediasta arvokasta tietoa asiakkaista ja maailmasta, valvoa rahoitusmarkkinoita ja ehkäistä petoksia, parantaa tietoturvaa ja työturvallisuutta sekä löytää kokonaan uusia tapoja harjoittaa liiketoimintaa. Valtion ja yhteiskunnan näkökulmasta reaaliaikaista big data tarjoaa mahdollisuuden muun muassa kansalaisten hyvinvoinnin, elämänlaadun ja turvallisuuden parantamiseen. Infrastruktuurista voidaan rakentaa älykkäämpää ja ympäristöanalyysin avulla voidaan varautua luonnonkatastrofeihin sekä monin tavoin muun muassa torjua ilmastonmuutosta. Urheilussa reaaliaikainen big data mahdollistaa muun muassa suorituskyvyn parantamisen sekä loukkaantumisten ehkäisemisen ja terveyden edistämisen. Se myös tarjoaa urheiluliiketoiminnalle samankaltaisia mahdollisuuksia kuin liiketoiminnalle yleensä.

Reaaliaikaiseen big dataan liittyy kuitenkin myös haasteita, jotka täytyy ottaa päätöksenteossa huomioon. Erityisesti tekniset haasteet ovat hyvin yleisiä. Big datan analysointiin ja päätöksentekoprosessiin vaadittava aika on yleensä hyvin pitkä, mutta reaaliaikaisten sovellusten on aloitettava toimet nopeasti tietyn aikaikkunan sisällä (Mohamed & Al-jaroodi, 2014). Tämä tekee reaaliaikaisten järjestelmien suunnittelusta, toteutuksesta ja toiminnasta on hyvin haastavaa. Data on usein hyvin levällään, sillä sitä kerätään eri järjestelmistä ja varastoidaan eri tietovarastoihin. Haasteena on datan heterogeenisuus. Koska data voi olla eri tyyppistä, on haasteena saada se analytiikkajärjestelmän ymmärtämään muotoon. Datan suuri määrä aiheuttaa haasteen siitä, että sen joukosta osataan poimia juuri oikea ja arvokas tieto, sillä kaikki data ei ole hyödyllistä. Big data ei myöskään välttämättä kerro aina totuutta. Se on sekavaa ja joskus myös virheellistä. Varsinkin sosiaalisesta mediasta kerätty data voi olla epäluotettavaa, sillä siinä on kyse ihmisten tekemistä arvioista. Siksi data-analyysiin täytyy suhtautua tietynlaisella varauksella.

Big data luo haasteita myös yksityisyyteen ja tietoturvaan sekä eettisyyteen liittyen. Yrityksillä on hallussaan paljon tietoa niiden asiakkaista. Yksittäi-

sen ihmisen toimintaa ja vuorovaikutusta on nykyään helppo jäljittää hyvin yksityiskohtaisesti, mikä on kuitenkin eettisesti kyseenalaista. Henkilökohtaisen datan kerääminen ja hallussapito ilman perusteltua syytä on myöskin laitonta. Onkin syytä valvoa, ettei dataa käytetä väärin, sillä tietoturva on jatkuvasti enemmän uhattuna. Kun kaikki toiminta perustuu yhä enemmän verkkopohjaisille järjestelmille, on se myös altista hyökkäyksille ja haavoittuvuuksille. Eettisyyteen liittyvän haasteen luo myös se, että kun päätökset perustuvat dataan, voivat päätöksentekijät niin sanotusti piiloutua sen taakse eettisesti arveluttavissa päätöksissä ja pakoilla näin omaa vastuutaan.

Reaaliaikaiseen päätöksentekoon liittyvät haasteet nousivat parhaiten esiin Formula 1-ympäristöstä, missä big dataa hyödyntäviin päätöstukijärjestelmiin perustuvan päätöksenteon reaaliaikaisuus on kisan aikana menestyksen kannalta ratkaiseva tekijä. Reaaliaikaisen päätöksenteon haasteena on aikarajoitteiden luomat paineet, kognition ylikuormitus ja hajanaisuus sekä tilannesidonaisuus. Formula 1:ssä insinööreillä on kovat paineet ja heillä voi olla liian paljon useita eri tehtäviä hoidettavana. Työympäristö on täynnä häiriötekijöitä. Päätöstukijärjestelmä ehdottaa strategiavaihtoehtoja siihen etukäteen tallennettujen mallien ja sääntöjen perusteella. Haasteena on, että jos mallit ja säännöt on etukäteen määriteltä väärin tai puutteellisesti, niihin ei voi reaaliajassa enää vaikuttaa, mikä johtaa siihen, että järjestelmä ehdottaa huonoja ratkaisuja. Tämä korostaa päätöksentekijän lopullista vastuuta päätöksentekoprosessissa ja nostaa esiin myös organisaatiokulttuuriin liittyvän haasteen. Päätöksentekijällä on oltava valtuudet ja valmiudet tulkita järjestelmän tarjoamaa analyysia.

Reaaliaikaisen päätöksenteon haasteet ovat lopulta samat myös muilla aloilla, joten mielenkiintoinen jatkotutkimusaihe olisikin tarkastella reaaliaikaista päätöksentekoa Formula 1-ympäristössä tarkemmin ja selvittää, miten näihin haasteisiin siellä vastataan ja mitä muiden alojen organisaatiot voisivat oppia siitä.

Tässä tutkielmassa tarkasteltiin big datan hyödyntämistä reaaliaikaisessa päätöksenteossa organisaatioiden näkökulmasta ja yksilön näkökulma jätettiin tarkastelun ulkopuolelle. Big data vaikuttaa kuitenkin monin tavoin myös yksilön arkipäiväiseen elämään, joten toinen mahdollinen jatkotutkimusaihe olisikin perehtyä enemmän big datan vaikutuksiin yksittäisen ihmisen arkipäiväisissä päätöksentekotilanteissa.

LÄHTEET

- Aversa, P., Cabantous L. & Haefliger S. (2018). When decision support systems fail: Insights for strategic information systems from Formula 1. *The Journal of Strategic Information Systems*, 27(3), 221-236.
- Azvine, B., Cui, Z. & Nauck, D.D. (2005) Towards real-time business intelligence. *BT Technology Journal*, 23(3), 214-225.
- Baig, A. R. and Jabeen, H. (2016). Big Data analytics for behavior monitoring of students. *Procedia Computer Science*, 82(2016), 43-48.
- Biem A., Bouillet, E., Feng, H., Ranganathan, A., Riabov, A., Verscheure, O., Koutsopoulos, H. & Moran, C. (2010). IBM infosphere streams for scalable, real-time, intelligent transportation services. *In Proceedings of the 2010 ACM SIGMOD International Conference on Management of data*, (s. 1093-1104), Indianapolis, Indiana, USA, 6-11 June, 2010.
- Bifet, A. (2013). Mining Big Data in Real Time. *Informatica*, 37 (1), 15-20.
- Borthakur D., Gray, J., Sarma, J. S., Muthukkaruppan, K., Spiegelberg, N., Kuang, H., Ranganathan, K., Molkov, D., Menon, A., Rash, S. (2011). Apache Hadoop goes realtime at Facebook. *Teoksessa Proceedings of the 2011 ACM SIGMOD International Conference on Management of data*, (s. 1071-1080), Athens, Greece, 12-16 June, 2011.
- 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.
- Davenport, T. H. & Harris J. G. (2005). Automated Decision Making Comes of Age. *MIT Sloan Management Review*, 46(4), 83-89.
- Davenport, T. H. (2014). Analytics in sports: The new science of winning. *International Institute for Analytics*, 2, 1-28.
- Deng L., Gao J. & Vupplapati C. (2015). Building a Big Data Analytics Service Framework for Mobile Advertising and Marketing. *Teoksessa Proceedings of the 2015 IEEE First International Conference on Big Data Computing Service and Applications*, (s. 256 - 266), Washington DC, USA, 30 March - 2 April, 2015.

- Deng, L. & Gao, J. (2015). An Advertising Analytics Framework Using Social Network Big Data. *Teoksessa Proceedings of 2015 5th International Conference on Information Science and Technology (ICIST)*, (s. 470-475), Changsha, Hunan, China, 24-26 April, 2015.
- Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, 35(2), 137-144.
- George, G., Haas, M. R. & Pentland, A. (2014). Big Data and Management. *Academy of Management Journal*, 57(2), 321-326.
- Gruetzemacher, R., Gupta, A., & Wilkerson, G. B. (2016). Sports Injury Prevention Screen (SIPS): Design and Architecture of an Internet of Things (IoT) Based Analytics Health App. *CONF-IRM 2016 Proceedings*, 18.
- Hua, Y., Jiang, H. & Feng, D. (2014). Fast: Near real-time searchable data analytics for the cloud. *Teoksessa Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis* (s. 754-765), New Orleans, LA, USA, 16 - 21 November, 2014.
- Kitchin, R. (2014). The real-time city? Big Data and smart urbanism. *GeoJournal*, 79(1), 1-14.
- Labrinidis, A., & Jagadish, H. V. (2012). Challenges and opportunities with big data. *Proceedings of the VLDB Endowment*, 5(12), 2032-2033.
- 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), 20-31.
- Lee, C. (2016). A content analysis of sports analytics resistance. *Teoksessa AMCIS 2016 Proceedings: Decision Support and analytics (SIGDSA)*, 1, San Diego, CA, USA, 10 - 14 August, 2016.
- MacKenzie, D.A., Muniesa, F. & Siu, L. (2007). *Do Economists Make Markets? On The Performativity of Economics*. Princeton: Princeton University Press.
- Mahmood, T. & Afzal, U. (2013). Security Analytics: Big Data analytics for cybersecurity: A review of trends, techniques and tools. *Teoksessa 2013 2nd National Conference on Information Assurance (NCIA)* (s. 129-134). Rawalpindi, Pakistan, 11 - 12 December, 2013.
- Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C. & Byers, A. H. (2011). Big data: The next frontier for innovation, competition, and productivity. *McKinsey Global Institute*.

- Marr, B. (2015). *Big Data: Using smart big data, analytics and metrics to make better decisions and improve performance*. Chichester: Wiley.
- Marr, B. (2016). *Big Data in Practice: How 45 Successful Companies Used Big Data Analytics to Deliver Extraordinary Results*. Hoboken: Wiley.
- McAfee, A. & Brynjolfsson, E. (2012). Big data: The management revolution. *Harvard Business Review*, 90(10), 60-68.
- Merendino, A., Dibb, S., Meadows, M., Quinn, L., Wilson, D., Simkin, L. & Canhoto, A. (2018). Big data, big decisions: The impact of big data on board level decision-making. *Journal of Business Research*, 93 (2018), 67-78.
- Mishne G., Dalton, J., Li, Z., Sharma, A. & Lin, J. (2013). Fast data in the era of Big Data: Twitter's real-time related query suggestion architecture. Teoksessa *Proceedings of the 2013 ACM SIGMOD International Conference on Management of Data* (s. 1147-1158). New York, NY, USA, 22 - 27 June, 2013.
- Mohamed, N. & Al-jaroodi, J. (2014). Real-Time Big Data Analytics: Applications and Challenges. Teoksessa *International Conference on High Performance Computing & Simulation (HPCS)*, Bologna, Italy, 21 - 25 July, 2014.
- Mondello, M., & Kamke, C. (2014). The Introduction and Application of Sports Analytics in Professional Sport Organizations. *Journal of Applied Sport Management*, 6(2), 1-12.
- Mukherjee, A., Diwan, P., Bhattacharjee, P., Mukherjee, D. & Misra, P. (2010). Capital market surveillance using stream processing. Teoksessa *2nd International Conference on Computer Technology and Development (ICCTD)* (s. 577-582). Cairo, Egypt, 2 - 4 November, 2010.
- Nguyen, D. T. & Jung, J. E. (2017). Real-time event detection for online behavioral analysis of big social data. *Future Generation Computer Systems*, 66 (2017), 137- 145.
- Panian, Z. (2007). Just-in-Time Business Intelligence and Real-Time Decisioning. *International Journal of Applied Mathematics and Informatics*, 1(1), 28-35.
- Power, D. J. (2011). Challenges of Real Time Decision Support. Teoksessa F. Burstein, P. Brézillon & A. Zaslavsky (toim.) *Supporting Real Time Decision-Making. The Role of Context in Decision Support on the Move* (s. 3-12). New York: Springer.
- Power, D. J. (2014). Using 'Big data' for analytics and decision support. *Journal of Decision Systems*, 23(2), 222-228.

- Provost, F. & Fawcett, T. (2013). Data science and its relationship to big data and data-driven decision making. *Big Data*, 1(1), 51-59.
- Russom, P. (2011). Big data analytics. *TDWI Best Practices Report, Fourth Quarter*, 19, 40.
- Shi, Q. & Abdel-Aty, M. (2015). Big Data applications in real-time traffic operation and safety monitoring and improvement on urban expressways. *Transportation Research Part C: Emerging Technologies*, 58 (2015) , 380-394.
- Stensland, H. K., Gaddam, V. R., Tennøe, M., Helgedagsrud, E., Næss, M., Alstad, H. K., Mortensen, A. , Langseth, R., Ljødal, S., Landsverk, O., Griwodz, C., Halvorsen, P., Stenhaug, M. & Johansen, D. (2014). Bagadus: An integrated real-time system for soccer analytics. *ACM Transactions on Multimedia Computing, Communications and Applications*, 10, 1s, 14.
- Sujatha, V., Devi, S. P., Kiran, S. V. & Manivannan S. (2016). Big data analytics on diabetic retinopathy study (drs) on real-time data set identifying survival time and length of stay. *Procedia Computer Science*, 87, 227-232.
- Templier, M. & Paré, G. (2015). A Framework for Guiding and Evaluating Literature Reviews. *Communications of the Association for Information Systems*, 37(1), 112-137.
- Thierauf, R. J. (1982). *Decision Support Systems for Effective Planning and Control*. Englewood Cliffs: Prentice-Hall, Inc.
- Waller, M. A. & Fawcett, S. E. (2013). Data Science, Predictive Analytics and Big Data: A Revolution That Will Transform Supply Chain Design and Management. *Journal of Business Logistics*, 34(2), 77-84.
- van der Schaar, M. (2014). Real-time discovery and decision making from big data. Teoksessa *2014 IEEE International Conference on Consumer Electronics* (s. 1-3), Taipei, Taiwan, May 26-28, 2014.
- Ta, V-D., Liu C-M. & Nkabinde G. W. (2016). Big data stream computing in healthcare real-time analytics. Teoksessa *Proceedings of 2016 IEEE International Conference on Cloud Computing and Big Data Analysis (ICCCBDA)* (s. 37-42), Chengdu, China, 5 - 7 July, 2016.
- Wang, J., Qiu, M. & Guo, B. (2017). Enabling real-time information service on telehealth system over cloud-based Big Data platform. *Journal of Systems Architecture*, 72 (2017), 69-79.
- Wang, G., Gunasekaran, A., Ngai, E. W. & Papadopoulos, T. (2016). Big data analytics in logistics and supply chain management: certain investigations for research and applications. *International Journal of Production Economics*, 176, 98-110.

- Watson, H. (2014). Tutorial: Big Data Analytics: Concepts, Technologies and Applications. *Communications of the Association for Information Systems*, 34(5), 1247-1268.
- Vera-Baquero, A., Colomo-Palacios, R. & Molloy, O. (2016). Real-time business activity monitoring and analysis of process performance on big-data domains. *Telematics and Informatics*, 33(3), 793-807.
- Yadranjiaghdam, B., Pool, N., & Tabrizi, N. (2016). A survey on real-time big data analytics: Applications and tools. Teoksessa *The 2016 International Conference on Computational Science and Computational Intelligence (CSCI)* (s. 404-409), Las Vegas, Nevada, December 15-17, 2016.