

LAMK

Lahden ammattikorkeakoulu
Lahti University of Applied Sciences

BIG DATAN HYÖDYNTÄMINEN SUOMALAISSA PK-YRITYK- SISSÄ

LAHDEN AMMATTIKORKEAKOULU
Liiketalouden ala
Tietojenkäsittelyn koulutusohjelma
Opinnäytetyö
Syksy 2018
Henna Vartiainen

Tiivistelmä

Tekijä(t) Vartiainen, Henna	Julkaisun laji Opinnäytetyö, AMK Sivumäärä 33	Valmistumisaika Syksy 2018
Työn nimi Big datan hyödyntäminen suomalaisissa pk-yrityksissä		
Tutkinto Tradenomi, tietojenkäsittelyn koulutusohjelma		
Tiivistelmä <p>Opinnäytetyön toimeksiantaja oli #DigiLAHTI-hanke. Työn tarkoituksena oli tutkia suurten datamassojen hyödyntämistä suomalaisissa pk-yrityksissä ja saada selville, missä määrin data-analytiikka on otettu osaksi liiketoimintaa pienemmissäkin yrityksissä.</p> <p>Teoreettista pohjaa työlle saatiin aiemmin tehdyistä tutkimuksista, kirjallisuudesta sekä big datan käyttötapauksista eri yrityksissä. Opinnäytetyön empiirinen tutkimus toteutettiin kvalitatiivisella tutkimusmenetelmällä ja tutkimusaineistona toimi haastattelut. Haastatteluun otettiin mukaan eri kokoisia ja eri alojen pk-yrityksiä, jotta saataisiin mahdollisimman monipuolisia näkökulmia. Tutkimukseen saatiin haastattelu yhteensä viideltä eri pk-yritykseltä.</p> <p>Tutkimustuloksista selvisi, että suurin osa haastatelluista pk-yrityksissä ei hyödynnä suuria datamassoja liiketoiminnassaan. Kaksi yritystä viidestä hyödynsi big dataa, mutta kohtuullisen vähäisesti vielä. Haasteiksi koettiin muun muassa yksityisyyteen liittyvät asiat, tiedonkeruuongelmat sekä datan analysointimenetelmät. Useimmat haastateltavat olivat kuitenkin lähes yksimielisiä siitä, että big data -analytiikka tulee olemaan tulevaisuudessa yhä tärkeämpi myös pk-yritysten kannalta.</p>		
Asiasanat big data, digitalisaatio, data-analytiikka, suuret datamassat, pk-yritykset		

Abstract

Author(s) Vartiainen Henna	Type of publication Bachelor's thesis	Published Autumn 2018
	Number of pages 33	
Title of publication Big data usage in small and medium sized Finnish businesses		
Name of Degree Bachelor's Degree Programme in Business Information Technology		
Abstract <p>The thesis explores the use of big data in small and medium sized Finnish companies. The aim of the thesis was to find out to what extend data analytics is utilized in small and medium sized businesses. The thesis was commissioned by the #Digi-LAHTI project.</p> <p>The sources for the theoretical part of this thesis include literature, publications related to the subject and actual cases in which companies used big data to improve their business. The empirical part of the study was performed by applying a qualitative research method and a deductive research approach. The data was obtained by interviewing the representatives of different small and medium sized companies. The companies were selected from different industries in order to get versatile results. Five companies participated in the interview study.</p> <p>The study results show that most of the interviewed companies do not utilize big data in their business. Two out of five of the interviewed companies use big data in their business, but only in small quantities. Many of the respondents felt that gathering and analyzing the data was too difficult and expensive to do. Most of the interviewed companies were nevertheless interested in big data and thought that it will be a major asset in the future for small and medium sized businesses.</p>		
Keywords big data, digitalization, data analytics, large masses of data, small and medium sized businesses		

SISÄLLYS

1	JOHDANTO	1
2	TUTKIMUKSEN TAUSTA	2
2.1	Toimeksiantaja	2
2.2	Tutkimuskysymys ja tutkimusongelma	2
2.3	Tutkimusmenetelmät	2
3	TEOREETTINEN TIETO	4
3.1	Mitä on big data?	4
3.2	Datan muuttaminen tietämykseksi	4
3.3	Data-analytiikan tuomat hyödyt liiketoiminnalle	5
3.4	Big datan käyttötapauksia maailmalta	7
3.4.1	Airbnb	7
3.4.2	Ralph Lauren	8
3.4.3	Pieni lontoolainen teurastamo	8
3.4.4	Netflix	9
3.5	Big data ja Suomi	10
3.5.1	Big datan tilanne Suomessa ennen ja nyt	10
3.5.2	Big datan käyttötapauksia Suomessa	12
3.6	Pk-yritykset ja big datan mahdollisuudet	14
4	TUTKIMUSAINEISTO JA ANALYYSI	17
4.1	Haastattelurunko	17
4.2	Haastatteluaineisto	17
4.2.1	Yritys A	18
4.2.2	Yritys B	18
4.2.3	Yritys C	19
4.2.4	Yritys D	20
4.2.5	Yritys E	20
4.2.6	Lisähaastattelu	21
4.3	Aineiston analyysi	22
5	JOHTOPÄÄTÖKSET	25
6	YHTEENVETO	26
6.1	Yhteenveto tutkimuksesta	26
6.2	Ongelmat tutkimusprosessissa	27
6.3	Tutkimuksen validiteetti ja reliabiliteetti	27

6.4	Yleistettävyys.....	28
6.5	Suositukset jatkotutkimukselle	29
LÄHTEET		30
LIITTEET		33

1 JOHDANTO

Tietomassat kasvavat nykyään uskomatonta vauhtia ja iso osa niistä on luokittelematonta dataa. Big data on yhteisnimitys näille suurille ja jatkuvasti kasvaville tietomassoille, joita on mahdotonta analysoida ja tulkita perinteisin menetelmin. Vain pieni osa tästä kaikesta kerääntyneestä tiedosta osataan hyödyntää ja jalostaa yritysten käyttöön.

Big datalla ja sen oikeaoppisella analysoinnilla tulee tulevaisuudessa olemaan entistä tärkeämpi vaikutus yritysten kilpailukykyyn ja tehokkuuteen. Mitä paremmin yritys osaa poimia oleellisen tiedon epäoleellisesta, sitä tehokkaammin se pystyy hyödyntämään big datan sisältävää tietoa.

Tämän opinnäytetyön tarkoituksena on tutkia big datan merkitystä ja käyttöä suomalaisissa pienissä ja keskisuurissa yrityksissä. Suuret yritykset ovat jo vuosia hyödyntäneet suuria datamassoja liiketoiminnassaan ja näistä tapauksista löytyy myös paljon tutkimustietoa. Sen sijaan pienistä suomalaisista yrityksistä ei ole ainakaan julkisesti saatavilla paljon tietoa big datan käyttötapauksista.

Tutkimuksen teoreettisessa osuudessa käsitellään suurten datamassojen analytiikkaa aiheena sekä yleisesti että Suomen kannalta. Tarkoituksena on tuoda esiin erilaisia käyttötapauksia maailmalta ja Suomesta ja antaa esimerkkejä siitä, miten suuria datamassoja on mahdollista hyödyntää yritysten liiketoiminnassa. Teoriatieto saadaan aiemmista aiheeseen liittyvistä tutkimuksista, asiakirjoista, kirjallisuudesta sekä artikkeleista.

Teorian lisäksi opinnäytetyö pyrkii kvalitatiivisen tutkimuksen avulla selvittämään, miten suomalaisissa pk-yrityksissä hyödynnetään big dataa ja mitä mahdollisuuksia big datalle nähdään Suomessa. Myös mahdolliset suurten datamassojen tuomat haasteet pk-yritykselle tuodaan esiin. Tutkimusmateriaalina käytetään haastatteluista saatuja vastauksia.

2 TUTKIMUKSEN TAUSTA

2.1 Toimeksiantaja

Opinnäytetyön toimeksiantajana toimii #DigiLAHTI-hanke. Hankkeen tavoitteena on yritysten prosessien tehostus, joka tapahtuu yritysten toimintaa digitalisoimalla, kehittämällä uusia asiakas- ja käyttäjälähtöisiä tuotteita, palveluita ja toimintamalleja. Hankkeen tuotoksena syntyy digitalisaation hyödyntämiseen pk-yrityksissä tarvittavaa tietoa, toimintamalleja sekä työvälineitä. Onnistuneen digitaalisen osaamisen lisääntymisen myötä yritysten on mahdollista menestyä kansainvälisessä kilpailussa sekä luoda uutta liiketoimintaa. Hankkeen toteuttavat yhteistyössä Lahden ammattikorkeakoulu, Lappeenrannan teknillinen yliopisto ja Lahden Seudun Kehitys LADEC Oy.

2.2 Tutkimuskysymys ja tutkimusongelma

Tutkimuksen tutkimuskysymys on seuravanlainen: miten ja missä määrin suomalaisissa pk-yrityksissä hyödynnetään big dataa? Tutkimuskysymys on sekä kuvaileva, että selittävä. Tutkimusongelmana on selvittää big datan nykytilanne suomalaisissa pk-yrityksissä ja saada selvyttä sen hyödyntämistavoista sekä hyödyistä yrityksille. Myös mahdolliset haasteet, joita pk-yritykset kokevat liittyen suurten datamassojen analytiikkaan, tuodaan esille.

Tutkimus suoritetaan kvalitatiivisesti haastattelemalla viiden eri yrityksen IT-johtajia tai asiantuntijoita sekä keräämällä tietoa aiheesta eri internet-lähteistä ja aiemmista tutkimuksista. Koska tutkittavasta aiheesta ei löydy paljon aineistoa ainakaan julkisesti, on lopputuloksena tarkoitus syntyä tietoa, joka voisi antaa jonkinlaisen kuvan siitä, että hyödynnetäänkö big dataa suomalaisten pk-yritysten liiketoiminnassa jo aktiivisesti, vai onko suurten datamassojen tuomat hyödyt vielä rantautumatta pieniin yrityksiin.

2.3 Tutkimusmenetelmät

Tutkimuksen empiirisessä osuudessa käytetään kvalitatiivista eli laadullista tutkimusmenetelmää ja deduktiivista eli teorialähtöistä lähestymistapaa. Tutkimus suoritetaan haastatteluiden muodossa. Koska tarkoituksena on saada tietoa yritysten omakohtaisista kokemuksista, on laadullinen menetelmä oikea valinta tähän tutkimukseen. Haastatteluilla on mahdollista saada yksityiskohtaista tietoa eri yritysten näkökulmasta aiheeseen.

Tutkimuksen analysointimenetelmänä käytetään sisällönanalyysia. Tutkimuksen tuloksista muodostetaan tiivistetty kuvaus ja aineistoa vertaillaan keskenään etsien yhtäläisyyksiä tai

eroavaisuuksia. Tutkimuksen tuloksia analysoimalla pyritään vastaamaan tutkimuskysymykseen ja luomaan suuntaa antava kuva siitä, missä määrin suomalaisissa pk-yrityksissä hyödynnetään suuria datamassoja liiketoiminnassa. Tarkkaa kokonaiskuvaa koko Suomen big data -tilanteesta pk-yrityksissä tutkimus ei pysty luomaan, eikä se ole tarkoituskaan. Tämän vuoksi tutkimustulokset ovat lähinnä suuntaa antavia ja yksittäisten pk-yritysten näkemyksiä big datan heille tuomista hyödyistä sekä haasteista. Haastateltavien yritysten toimialat vaihtelevat paljon toisistaan, mikä mahdollistaa mahdollisimman monipuoliset tutkimustulokset.

3 TEOREETTINEN TIETO

3.1 Mitä on big data?

Big data ilmiönä lähti nousuun vuonna 2011, mutta se on ollut käsitteenä tuttu jo vuosia aiemmin. Big data käsitteellä voidaan tarkoittaa kahta asiaa. Se viittaa monipuolisen ja nopeasti kasvavan datan suureen määrään, joka aiheuttaa haasteen sekä organisaatioille että yhteiskunnalle. Toisaalta sillä viitataan myös ratkaisuihin, jotka pyrkivät vastaamaan tähän haasteeseen. Viime vuosina datamäärä on kasvanut enemmän kuin koskaan aiemmin luoden painetta yrityksille. Dataa syntyy kovalla vauhdilla muun muassa sosiaalisesta mediasta, terveydenhuollon järjestelmistä, valvontakameroista, älypuhelimista, mittalaitteista, liikkeentunnistimista, lokitiedoista ja erilaisista rekistereistä. (Salo 2013, 10.)

Data voidaan jakaa strukturoituun ja strukturoimattomaan dataan. Strukturoitu data on hyvin organisoitua ja siitä on helppo hakea etsimänsä tieto yksinkertaisillakin hakumenetelmillä. Se on helppo säilöä ja analysoida tietokannoissa ja on jaoteltu selkeästi. Strukturoimaton data on sen sijaan organisoimatonta ja vaikealukuista, perinteisillä tietokannoilla on vaikeuksia tulkita sitä. Muutamia esimerkkejä strukturoimattomasta datasta ovat pdf-tiedostot, kuvatiedostot, metatiedot, sosiaalisen median kirjoitukset, sähköpostit ja nettisivustot. Suurin osa big datasta on strukturoimatonta, eli sillä ei ole selkeää rakennetta. (Reynolds 2016, 43–44.)

Tästä strukturoimattomasta datasta tallennetaan ja analysoidaan vain murto-osa, ja tämä on ongelma, johon big dataan liittyvät ratkaisut pyrkivät tuomaan helpotusta (Salo 2013, 22–25). Pelkästä datasta ei itsessään ole hyötyä, vaan se pitää osata tallentaa, analysoida ja tulkita oikein. Kaikki kerätty data ei myöskään ole hyödyllistä, vaan siitä on osattava löytää oleellimmat asiat liiketoiminnan kannalta. On siis tiedettävä mitä etsii suuresta datamassasta ja millä tekniikoilla sen saavuttaa. Tähän tarvitaan asianmukaisia työkaluja, suunnittelua sekä osaavaa henkilöstöä.

3.2 Datan muuttaminen tietämykseksi

Jotta data saataisiin muutettua tietämykseksi ja oivallukseksi, tarvitaan matkan varrelle useita eri vaiheita. Big dataan liittyvät ratkaisut auttavat sekä datan säilömisessä että analysoinnissa. Yksi suosittu esimerkki big data -ratkaisusta on Hadoop, joka yksinkertaistettuna on kokoelma avoimen lähdekoodin ohjelmistoja ja menettelyjä, jotka auttavat suurten datamassojen kanssa työskentelyssä. Hadoop koostuu moduuleista, eli itsenäisistä osista, joilla kaikilla on oma tehtävänsä. Distributed File-System -moduuli on hajautettu

tiedostojärjestelmä, jonka tehtävänä on säilyttää dataa helposti saatavilla olevassa muodossa, jotta siihen pääsisi käsiksi eri laitteilla ja työkaluilla. MapReduce -nimisen moduulin tehtävänä on taas lukea ja suodattaa dataa tietokannasta ja muuttaa se sellaiseen muotoon, että sen voi analysoida. YARN -moduuli puolestaan hallitsee dataa säilyttävän ja analysoivan systeemin resursseja. (Marr 2018.)

Ohjelmistojen ja työkalujen lisäksi tarvitaan myös osaavia ihmisiä datan analysointia varten. Data-analyytikolla on keskeinen rooli datan analysoinnissa ja tulkitsemisessa, ja hyvä data-analyytikko osaa niin tilastotiedettä, liiketaloutta kuin koodaustaitojakin (Fenix Solutions 2018). Kuitenkin myös yrityksen henkilöstön ja johdon on hyvä olla mukana prosessissa niin suunnitteluvaiheessa kuin päätöstenkin teossa. Tietämystä ja erilaisia näkökulmia tarvitaan niin liiketaloudesta, markkinoinnista kuin teknillisestä puolesta, jotta voidaan ymmärtää, millainen data on hyödyllistä juuri oman liiketoiminnan kannalta. (Loskin 2013, 25–28.)

Ennen kuin suuria datamassoja lähdetään analysoimaan, on yrityksellä hyvä olla myös strategia. On otettava huomioon, onko yrityksellä paljon omaa dataa käytössään jo ennestään ja millä lailla se on tallennettu. Digitaalisesti tallennettu data on helpommin analysoitavissa kuin esimerkiksi paperiset arkistot, joiden saattaminen digitaaliseen muotoon vie aikaa. Jos yrityksellä ei ole ennestään suuria määriä dataa, on pohdittava, että mistä sitä saa kerättyä ja miten. Isoilla yrityksillä kuten Amazonilla, Googlella ja Walmartilla on tarpeeksi resursseja, aikaa ja rahaa säilöä ja analysoida kaikki potentiaalinen data, oli se sitten enemmän tai vähemmän hyödyllistä liiketoiminnalle. Pienemmille yrityksille kaiken mahdollisen datan kerääminen ja talletus ei taas ole realistinen tavoite, mutta se ei tarkoita, etteikö pienikin yritys voisi hyötyä suurista datamassoista. Nykyään on esimerkiksi saatavilla suuret määrät avointa dataa, jota pienetkin yritykset voivat käyttää hyväkseen. Tämä on vain yksi esimerkki siitä, mistä pienikin yritys voi saada itselleen haltuun dataa. (Marr 2015, 23–26.)

3.3 Data-analytiikan tuomat hyödyt liiketoiminnalle

Big datan on todettu tuovan merkittävää hyötyä organisaatioille oikein analysoituna, ja mahdollisuuksia on monia (KUVIO 1). Big dataa hyödyntämällä on mahdollista optimoida toimintoja, saada aikaan säästöjä, tehdä riskianalyysejä, parantaa asiakaspalvelua ja tuotekehitystä, kehittää liiketoiminnan prosesseja, saada apua päätöksentekoon, tunnistaa tietoturva-uhkia ja analysoida kuluttajakäyttäytymistä. Kun datan jalostaminen tiedoksi tehdään oikein, voidaan saada korvaamatonta etua liiketoimintaan. (Smith 2016, 8–11.)



KUVIO 1. Big datan hyödyntämismahdollisuudet (SAS Institute Oy, 2018)

Haasteita big datalle luo datan suuren määrän lisäksi muun muassa tietoturvakysymykset ja yksityisyydensuoja. Datan väärinkäyttö on mahdollista, etenkin kun kyseessä on yksilölliset tiedot, kuten henkilö- ja tilitiedot. Myös liiketoiminnan kannalta tärkeät tiedot ja algoritmit voivat herättää kiinnostusta väärinkäyttöksiin. Salassa pidettävä tiedon hallinta tai sen puute voi myös aiheuttaa haasteita datan turvalliseen säilytykseen niin fyysisesti kuin pilvipalveluissa. (Salo 2014, 50–56.)

Tietoturvan lisäksi osaamisen puute on ongelma nyt ja tulevaisuudessa. Pelkkä data ja sen jalostaminen työkaluilla informaatioksi ei riitä, vaan sitä on ymmärrettävä ja osattava hyödyntää. Tähän tarvitaan osaavia data-analytikoita, koulusta ja perehtymistä aiheeseen. (Salo 2013, 92–93.)

Vuonna 2013 Big Insights niminen yritys teki kyselyn Australiassa ja sitä ympäröivissä maissa eri liiketoiminnan johtajille. Kyselystä selvisi, että yritykset kokivat suurimmaksi haasteeksi big datan ymmärtämisen ja hyötykäytön liiketoiminnassaan. Suurten datamassojen tuomia hyötyjä ei ymmärretä täysin ja big datan konsepti on epäselvä. Myös muissa tutkimuksissa on selvinnyt samankaltaisia asioita. Datamassojen jatkuva kasvu aiheuttaa myös pelkoa siitä, että data-analytiikan käyttö liiketoiminnassa on asia, joka vaatii liikaa aikaa ja alituista huomiota. Haasteita aiheuttaa myös big data -työkalujen monimutkaisuus ja jatkuva kehittyminen. Työkalujen ja uusien teknologioiden käyttö on jatkuva oppimisprosessi, johon kaikilla yrityksillä ei ole omia resursseja. (Reynolds 2016, 106–107.)

3.4 Big datan käyttötapauksia maailmalta

Big datan käytöstä löytyy kattavia esimerkkejä ja menestystarinoita eri puolilta maailmaa. Alla käydään läpi esimerkkejä big datan hyödyntämistavoista niin suuressa kuin pienessäkin yrityksessä.

3.4.1 Airbnb

AirBnb on vuonna 2008 avattu verkkopalvelu, joka pyrkii yhdistämään matkustajat ja majoituksen vuokraajat ympäri maailmaa. Palvelu on kerännyt avautumisensa jälkeen valtavat määrät dataa käyttäjiensä mieltymyksistä ja käyttäytymisestä. Keräämäänsä dataa Airbnb käyttää palvelun laadun ylläpitämiseen. Majoittujia ja majoituksen tarjoajia on niin paljon, että ilman laajaa big datan käyttöä olisi todella vaikea tarjota majoittujalle hänen haluamanlaista majoitusta. Sama pätee myös majoituksen vuokraajan suunnalta; he eivät halua vuokrata majoitusta kenelle tahansa. (Marr 2016b, 163–167.)

Käytännössä AirBnb seuraa asiakkaansa käyttäytymistä ja hänen tekemiään toimia ja pyrkii näin määrittämään, mikä tapahtumaketju johtaa varaukseen. Tällä tavoin he voivat tehokkaasti tarjota asiakasta eniten miellyttäviä kohteita esimerkiksi vuokran hinnan, sijainnin ja varustelun mukaan. Tämä prosessi auttaa majoittujaa saamaan aina esille mieltymyksensä mukaisia kohteita, mutta se auttaa myös AirBnb:ta välittämään tietoja vuokraajille. Vuokraajat saavat tietoa muun muassa kysynnän ennakkoinnista, sesonkiajoista, parhaasta vuokrausajankohdasta sekä siitä mikä ylipäättään on sopiva vuokran hinta. Tässä onkin nähtävissä big datan voima, sillä AirBnb:llä on palvelussaan 1,5 miljoonaa kohdetta 34000 kaupungissa ja 50 miljoonaa majoituksen etsijää. (Marr 2016b, 163–167.)

AirBnb kerää dataa valtavia määriä ja käyttää niitä myös tehokkaasti palvelujen parantamiseen sekä ulkoisesti että sisäisesti. He joutuivat palkkamaan 2011 useita dataspesialisteja, sillä palvelun suuri suosion nousu aiheutti ongelmia pienelle datan käsittelytiimille. Yrityksessä tätä pidettiin elintärkeänä, sillä ilman sitä jatkuvasti laajeneva yritys ei olisi voinut toimia ja pitää palvelun laatua yllä. Ilman big datan käyttöä AirBnb:n tarjoamat palvelut sekä majoittujille että vuokraajille olisi mahdoton järjestää pelkän volyymin takia. Asiakaspalvelun parantamiseksi AirBnb on jopa ottanut käyttöön datatyökalun, jolla kaikki työntekijät voivat nopeasti tulkita dataa, tehdä päätöksiä sen perusteella ja auttaa näin asiakasta. Datan käsittely ei siis ole palvelussa pelkästään dataspesialistien käsissä. (Marr 2016b, 166.)

AirBnb:n big datan käyttöä voidaan pitää menestystarinana, he ovat käyttäneet big dataa eri muodoissa käytännössä perustamisestaan lähtien. Big datan turvin yritys on voinut uudistua ja kehittyä. Tämä osoittaa, että yrityksen data specialistit on hyvin integroitu yrityksen muihin palveluihin ja he ymmärtävät yrityksen tavoitteet. Lisäksi johto on ymmärtänyt big datan arvon alusta asti. (Marr 2016b, 167.)

3.4.2 Ralph Lauren

Big data avaa markkinoita myös uudennlaisille urheiluvaatteille ja muille puettaville asusteille. Ralph Lauren on suunnitellut älypaidan, jonka sisältämät sensorit keräävät dataa muun muassa käyttäjän liikkeistä, askeleista, kulutetuista kaloreista, sydämensykkeistä ja hengityksen tiheydestä. Ilmaiseksi ladattava sovellus seuraa kerättyä dataa ja luo sen perusteella käyttäjälleen kustomoituja kuntoharjoituksia. (Marr 2016b, 195–198.)

Seuraamalla omaa aktiivisuuttaan ja saamalla sovelluksesta ehdotuksia itselle sopiviin harjoituksiin, on käyttäjän mahdollista kohentaa kuntoaan tehokkaammin ja mahdollisesti myös ennaltaehkäistä liian kovasta kuntoilusta johtuvia vammoja. (Marr 2016b, 195–198.)

3.4.3 Pieni lontoolainen teurastamo

Big data on auttanut mm. pientä lontoolaista teurastamoa. Alkuperäisessä tekstissä, jossa kerrottiin yrityksen tarina, yrityksen ja sen työntekijöiden nimet oli muutettu yksityisyyden suojaamisen vuoksi.

Kyseinen yritys aloitti toimintansa 90-luvun puolen välin jälkeen ja toimi pitkään tasaisesti, hyvän maineen ja uskollisten asiakkaiden ansiosta. Yritys ajautui kuitenkin vaikeuksiin, kun samalle kadulle avautui iso marketti, joka alkoi kilpailemaan hinnalla, vieden näin suuren määrän yrityksen asiakkaista. Teurastamon perustaja tiesi, että hänen tuotteensa ovat huomattavasti laadukkaampia kuin marketin vastaavat. Ongelma oli kuitenkin saada tämä viestittyä asiakkaalle ja tuoda asiakkaat liikkeeseen. Hintakilpailu ei pienelle yritykselle ollut vaihtoehto. Ongelmaa alettiin yrityksessä ratkomaan big datan turvin ja tämä oli jopa yrityksen selviämiseksi ehto. Teurastamolle ei ollut mahdollista käsitellä dataa itse vaan he alkoivat työskennellä big data -konsultin kanssa vuonna 2016. Yrityksellä ei ollut aluksi valmista dataa, jonka kanssa olisi voitu työskennellä, joten sitä alettiin kerätä. Ensimmäiseksi asennettiin halvat sensorit näyteikkunaan, joiden avulla tutkittiin näyteikkunan vaikuttavuutta. Sensoreiden avulla teurastamo kykeni selvittämään, kuinka monta ihmistä kulki näyteikkunan ohi, ja moniko pysähtyi tutkimaan näyteikkunaa. Tämän lisäksi pystyttiin seuraamaan kuinka moni tuli sisälle itse liikkeeseen näyteikkunan seurauksena. Tämän tiedon avulla yritys kykeni muokkaamaan näyteikkunaa ja sen välittämää viestiä ja

esillä olevia tuotteita niin, että ne kiinnostivat ja kiinnittivät eniten huomiota asiakkaissa. (Marr 2016b, 51–55.)

Sensoreilla saadulla datalla havaittiin myös uusi, yllättävä mahdollisuus liikevaihdon lisäämiseen. Kadulla oli kaksi suosittua pubia ja iltayhdeksästä puoleen yöhön teurastamon ohi kulki lähes yhtä paljon ihmisiä kuin lounasaikaan. Tämän pohjalta yritys päättikin kokeilla ilta aukioloa. Tarkoituksena oli palvella publiin jälkeen kotiin suuntaavaa nälkäistä joukkoa. Ilta-aikaan tarjottavan ruokalistan sisältö ratkaistiin myös dataa analysoimalla Google Trends:n kautta; suosituimmat ruuat päätyivät ruokalistalle. Teurastamo toivoo jatkossa voivansa hyödyntää vielä lisää datan analysointia esimerkiksi sääennusteiden avulla tehtävien menekkiennusteiden muodossa sekä suunnitteilla olevalla kanta-asiakasjärjestelmällä, joka mahdollistaisi kohdennetun mainonnan suoraan asiakkaille. (Marr 2016b, 52.)

Alkuperäisellä sensoridatalla saadut tulokset näyttivät yritykselle, että asiakkaita ei loppupeleissä juurikaan kiinnostanut yrityksen viesteistä välittyvä tieto hinnoista. Eniten kiinnostusta saivat muut asiat, kuten ruokavinkit ja kehotus käydä liikkeessä kysymässä reseptejä ja lisää vinkkejä. Tällä tiedolla yritys tehosti viestiään ja sai asiakkaita enemmän liikkeeseen. Lisäksi ilta-aukiolo osoittautua todella suosituksi ja toi yritykselle huomattavan määrän tuloja sekä uusia asiakkaita. (Marr 2016b, 53.)

Kaikki tämän saavuttamiseen tarvittu data saatiin kerättyä verrattain halvasti ja tehokkaasti ilman suurta alkupanostusta. Suurin ongelma yrityksen datan käytössä oli vähäiset resurssit, mistä johtuen pienen yrityksen piti palkata joku muu käsittelemään ja tulkitsemaan kerättyä dataa heidän puolestaan. Tämän pienen teurastamon tapaus on hyvä esimerkki siitä, miten pieni yritys voi hyödyntää dataa liiketoiminnassaan. Se myös osoittaa, että datan keräämismetodi tai datan määrä ei ratkaise onnistumista, vaan se miten sitä osataan hyödyntää. (Marr 2016b, 54–55.)

3.4.4 Netflix

Netflix tarjoaa elokuvia ja televisiosarjoja suoratoistona miljoonille jäsenilleen. Palvelun käyttäjistä kerätään isoja määriä dataa, jonka analysoinnilla pyritään ymmärtämään ja ennakoimaan entistä paremmin heidän katselutottumuksiaan. Big datan hyödyntämisessä Netflix onkin ammattilainen, ja tästä on todisteena palvelun menestys sekä useat kymmenet miljoonat palvelun käyttäjät. Netflix alkoi käyttämään data-analytiikkaa työkalunaan jo ennen kuin yritys toimi suoratoistopalveluna. Ensimmäinen selkeä big data projekti Netflixissä nähtiin vuonna 2006, jolloin yritys järjesti kilpailun, jossa palkintona oli miljoona dollaria. Haasteena oli luoda algoritmi, jonka avulla pystyttiin ennustamaan kuinka Netflixistä

elokuvan vuokrannut asiakas tulee arvostelemaan elokuvan, asiakkaan aiempien arvostelujen perusteella. Vaikka Netflixin liiketoiminta on muuttunut täysin, he käyttävät yhä samaa algoritmia datan analysoinnin pohjana, vaikkakin algoritmia onkin muokattu ja siihen on lisätty parametrejä matkan varrella. (Marr 2016b, 17–23.)

Big dataa hyödynnetään yrityksen monissa eri osa-alueissa, mutta tärkein tavoite on juurikin ennakointi; millaisista elokuvista tai sarjoista käyttäjät pitävät? Tietoa kerätään muun muassa katsotuista elokuvista, katseluajankohdista, elokuva valintaan käytetystä ajasta, arvosteluista sekä siitä, kuinka monta kertaa elokuva pysäytetään katselun aikana. (Marr 2016b, 17–23.)

Aivan datan käytön alussa Netflix havaitsi suuren ongelman; vaikka dataa kerättiin suuret määrät asiakkaan katselutiedoista, dataa jäi myös paljon keräämättä. Kyseinen data oli itse elokuvissa oleva kuva- ja äänidata. Ongelman ratkaisemiseksi ja datan keräämiseksi Netflix maksoi katsojille, jotka korvausta vastaan katsoivat palvelussa olevia sarjoja ja elokuvia ja merkitsivät elementtejä näistä medioista. Nämä koekatsojat saivat jokainen tarkat ohjeet, mitä tulee merkitä ja kuinka, jotta saatu tieto olisi yhtenäistä. Tavoitteena oli kerätä erinäisiä piirteitä (esim. moraalisia valintoja, vahva naispääosa ja muita teemoja), joita videot sisälsivät. Tämän datan turvin Netflix on tunnistanut lähes 80000 "mikro-genreä". Palvelu voikin nyt tunnistaa käyttäjän mieltymykset paljon tarkemmin kuin kukaan muu aiemmin; palvelu ei enää tunnista pelkästään, että käyttäjä pitää tietystä genrestä, vaan se tunnistaa, minkälaista alagenreä käyttäjä suosii ja voi suositella seuraavaa katseltavaa ennen kuin käyttäjä edes itsekkään tietää. (Marr 2016b, 17–23.)

3.5 Big data ja Suomi

3.5.1 Big datan tilanne Suomessa ennen ja nyt

Suurten datamassojen hyödyntämiseen liiketoiminnassa on Suomessa herätty suhteellisen myöhään. Keskustelu ja tietoisuus big data -aiheesta alkoi kiihtyä vasta vuoden 2012 aikana ja big data olikin tuolloin esillä paljon eri tietotekniikka- ja talousaikakausilehdissä (Salo 2014, 34). Big data -ratkaisujen käyttö oli kuitenkin vielä harvassa ja siihen liittyvät koulutukset korkeakouluissa olivat vähäisiä. On toisaalta haastavaa saada täydellistä kokonaiskuvaa big datan -markkinoista Suomessa, koska sen tarkka rajaaminen on vaikeaa, ja läheskään kaikki big data -ratkaisujen käyttäjät eivät välttämättä tuo tietoa julkisuuteen. (Salo 2013, 88–89.)

Ivorio tuotti selvityksen big data -markkinoista Suomessa vuonna 2013. Selvitys toteutettiin lomakekyselyillä ja haastatteluilla, joihin osallistuivat julkishallinnon ja yksityisen sektorin toimijat, keskeisimmät palveluntarjoajat, konsultit sekä koulutus- ja tutkimussektorin edustajat. Selvityksestä kävi ilmi, että big data on vielä varsin uusi ja vaikeasti lähestyttävä asia monelle organisaatiolle. Big datan ilmiöön oli kuitenkin havahduttu; se kiinnosti ja sitä pidettiin tärkeänä asiana kilpailukyvyn edistämiseksi tulevaisuudessa. (Liikenne- ja viestintäministeriö 2013.)

Ivorion tuottamasta selvityksessä kävi myös ilmi, että big data -osaajista on pulaa. Oppilaitoksissa tähän oli jo lähdetty hakemaan ratkaisuja, mutta tulokset näkyisivät kuitenkin viiveellä vasta vuosien päästä. Helpotusta osaajapulaan voisi tuoda nykyisten vahvojen tiedonhallinnan ammattilaisten täydennyskoulutus. Joka tapauksessa, selvityksessä oltiin sitä mieltä, että koordinoituja toimenpiteitä tarvitaan nopeasti, jotta voitaisiin luoda pohja tulevaisuuden big data -ammattilaisten kehittymiselle ja kilpailukyvyn tehostamiselle. Julkisella hallinnolla todettiin myös olevan tärkeä rooli kehityksen tukemisessa, ja sen toimijat ovatkin lähteneet edistyksellisesti toteuttamaan muun muassa avoimen datan suunnitelmia. Kansallisten tietovarojen avoin saatavuus ja selkeät oikeudelliset rajat datan käsitteilyssä takaavat sen, että dataliiketoimintaa ja uutta yrittäjyyttä voi syntyä. (Liikenne- ja viestintäministeriö 2013.)

SAS Instituten toteuttaman tutkimuksen (2015) mukaan Suomi kuitenkin kulkee muita Pohjoismaita edellä suurten datamassojen hyötykäytössä. Tutkimukseen osallistuivat suurimpien yritysten IT-johtajat. Tutkimuksen mukaan Suomessa big dataa hyödynnetään etenkin asiakkuuksissa, tuotekehityksessä ja operatiivisessa toiminnassa (Virkki 2015). Kehitystä on siis tapahtunut varmasti paljonkin jo vuosien varrella, kun otetaan huomioon, että Suomi alkoi havahtua big dataan kunnolla vasta vuoden 2012 jälkeen.

Nykytilanne on jo huomattavasti parempi kuin vuosia sitten. Koulutusten saatavuus on parantunut ja big data -ratkaisuja myyviä yrityksiä on syntynyt Suomeen enemmän. Telia on lähtenyt tukemaan datan suurta ja nopeaa kasvua rakennuttamalla Helsinkiin Suomen suurinta avointa datakeskusta. Ottaen huomioon, että Suomen julkiset tietovarannot ovat myös yleisesti saatavilla, on Suomella jo pelkästään avoimen datan puolesta hyvät mahdollisuudet suurten datamassojen hyötykäyttöön. (Rihti 2016.)

Datan hyödyntämistä liiketoiminnassa pyritään lisäämään myös hallituksen periaatepäätöksellä. Tarkoituksena on mahdollistaa massadatan luoma kasvupotentiaali Suomessa vuoteen 2020 mennessä, poistaa turhat kasvun rajoitteet ja laajentaa datalähtöistä toimintaa yritysten liiketoiminnassa. Datan käyttöä mahdollistetaan muun muassa kannustavalla sääntelyllä, osaamisen lisäämisellä ja datan hyödyntämiseen perustuvien toimintamallien

kehityksen tukemisella (Valtioneuvosto 2016). Visiona on myös suomalaisella osaamisella toteutettuja kansainvälisiä datapalveluja ja liiketoimintaa, jotka houkuttelevat myös kansainvälisiä investointeja ja osaamista. Näiden toteutumiseksi Suomessa olisi edistettävä sovellus-, palvelu- ja analytiikkakehitystä, kehitettävä datan käyttöön liittyvää palveluinfrastruktuuria sekä vahvistettava osaamista kasvattavaa datapolitiikkaa. (Liikenne- ja viestintäministeriö 2016.)

Kaiken kaikkiaan näyttää siltä, että Suomessa ollaan menossa oikeaan suuntaan suurten datamassojen jalostamisessa hyötykäyttöön. Haasteita on varmasti vielä paljon, mutta jos hallituksen periaatepäätös toteutuu suunnitelmien mukaan, on Suomi varmasti erittäin hyvällä mallilla. Suomella on hyvät edellytykset big data -osaamiseen ja täältä löytyy paljon IT-ammattiosaamista. Tärkeää olisi kouluttaa lisää osaajia, jotta suurta big datan määrää pystyttäisiin hyödyntämään tehokkaammin.

Haasteita yrityksille voi aiheuttaa osaamisen ja asiantuntijoiden vähäisyyden lisäksi ymmärryksen puute aiheeseen liittyen; aloitteet big data -projekteihin saattavat useammin tulla yritysten IT-osastoilta, ei ylimmältä johdolta, mikä taas vaikuttaa osaltaan budjettiin ja strategioihin. Liika varoivaisuus ja vanhoissa ajattelutavoissa pysyminen ovat myös omiaan hidastamaan kehitystä. (Salo 2014, 34–35.)

3.5.2 Big datan käyttötapauksia Suomessa

Suomesta löytyy viime vuosien varrelta lukuisia käyttötapauksia big datan hyödyntämiseen liittyen. Nämä käyttötapaukset koskevat enimmäkseen kuitenkin suurempia yrityksiä; big datan hyödyntämisestä pk-yrityksissä ei vaikuttaisi olevan kovin paljon julkista tietoa.

Yksi suomalaisista big datan käyttötapauksista on Metsäteho Oy:n koordinoima Forest Big Data -hanke. Sen yhteydessä toteutettiin kaksi pilottia. Visiona oli puun hankintaketjun digitalisaatio. Ensimmäisen pilotin aikana syntyi toimintamalliehdotus puukuljetusreittien haajaantuneen tiestödatan yhdistämisestä ja jakamisesta eri käyttäjäryhmille. Toisessa pilotissa kokeiltiin sovellusta, joka mahdollistaa automaattisen anturi- ja kuvadatan keräämisen ajon aikana, tuulilasiin kiinnitetyllä älypuhelimella. Tarkoituksena oli testata tiestöjen olosuhdetietojen keräämistä ratkaisulla, joka soveltuu joukkoistutettuun datankeräykseen. Kerättyä tietoa ja aineistoa olisi mahdollista hyödyntää esimerkiksi ongelmatilanteiden havaitsemisessa ja teiden kunnon arvioinnissa. Tiestön ja puukuljetusten täsmällisesti suunnitellun kunnossapidon myötä hyötyinä tulisi mm. resurssien tehokkaampi käyttö ja tiestövaurioiden vähentyminen. Kustannustehokkuuden ja jatkuvuuden vuoksi tavoitteena olisi, että mahdollisimman moni teillä liikkuva pystyisi automaattisesti tuottamaan ja jakamaan dataa tiestöstä ja sen olosuhteista. (Venäläinen 2016.)

Suomalaisessa suunterveydenhuollossa on myös nähty tapauksia big datan hyödyntämisestä viime vuosina. Lohjan kaupunki ja IT-alan yritys Kuukaste Oy suunnittelivat yhdessä ratkaisun, jonka tarkoituksena oli suunterveydenhuollon ajanvarauksen tehostaminen. Hyödyksi käytettiin potilastietojärjestelmissä olevaa dataa, jota Lohjan kaupungille oli kerääntynyt 20 vuoden ajalta. Tämä data koostui muun muassa hammastarkastuksista ja muista asiakkaiden käynneistä vastaanotolla. (TIEKE ry 2016.)

Pohjolan Liikenne hyödyntää suuria datamassoja liiketoiminnassaan parantaakseen ajoturvallisuutta. Satoihin linja-autoihin on liitetty tallentimia, kameroita ja sensoreita, jotka tuottavat tietoa reaaliajassa ja auttavat tapaturmien määrän seurannassa sekä nopeuttavat tilanteisiin reagointia. Videotallenteet helpottavat myös häiriötilanteiden raportoinneissa sekä tapaturmien tarkistamisessa, koska tilanteet eivät jää enää tulkinnan varaan. Sensoridatasta on löytynyt apua myös autojen lämmityksen ajoittamisessa. Auton lämmitys käynnistyy automaattisesti sen mukaan, monelta auto on lähdössä liikkeelle ja millainen sää on. Auton lähtöajankohdasta tulee tieto järjestelmän palvelimelta ja tieto säästä saadaan Ilmatieteenlaitoksen rajapinnasta. Ajoittamalla auton lämmitys juuri oikeaan ajankohtaan saadaan aikaiseksi suuria säästöjä kustannuksiin. (Enfo 2018.)

Myös S-ryhmä hyödyntää toiminnassaan suuria datamassoja. S-ryhmä on yhteistyössä SAS Institute Oy:n kanssa kehittänyt data-analytiikkaansa ja päivittäistavarakaupan analyttistä kokonaisratkaisua, johon sisältyy muun muassa tilanoptimointiratkaisuja, kysynnän ennustamista sekä hintojen optimointia, kyetäkseen vastaamaan nykymarkkinoiden kasvaviin haasteisiin. Kaikki edellä mainitut asiat ovat kytköksissä toisiinsa; kysynnän ennustaminen vaikuttaa tuotteiden ja hyllytilan tarvittavaan määrään ja hinnat taas vaikuttavat kysynnän ennustamiseen. S-ryhmä kerää asiakasomistajarekisteritietoja, joihin sisältyy muun muassa dataa asiakkaiden ostamista tuotteista. Tämän datan avulla S-ryhmä profiloii jokaisen myymälänsä erilaisiksi tuotteiltaan ja tuoteryhmiltään asiakaskunnan mukaisesti. Jokaiselle myymälälle suunnitellaan oma tuotevalikoima, joka vastaa parhaiten kysyntää. (SAS Institute Oy 2018.)

Yle keräsi vuonna 2018 dataa noin 300 gigatavua päivässä. Tieto kerättiin suurilta osin Ylen verkkopalveluiden käytöstä, sisältäen internetsivut ja mobiilipalvelut. Yle kerää käytännössä asiakkaiden käyttötietoja, milloin on katsottu mitäkin, kuinka kauan palvelua käytetään kerralla, ja mitkä osa-alueet ovat eniten käytössä. Yle käyttää datan varastointiin ja analysointiin pilvipalvelua, koska heidän mukaansa monen pilvipalvelutarjoajan valikoimaan kuuluu datan käsittelyyn tarvittavia työkaluja. Suureksi syyksi Ylessä nostetaan myös pilvipalvelun niin kutsuttu skaalautuminen; dataa voi tiettyinä aikoina tulla suurina

piikkeinä ja pilvipalvelu tarjoaa mahdollisuuden reagoida tähän. Yle kykenee keräämälään datalla ennakoimaan suuria kävijä määriä ja osaa näin tehostaa omaa verkkopalveluaan, jotta ne eivät tukkeudu. Datan keräyksellä pyritään myös parantamaan käyttäjäkokeumusta ja tarjoamaan käyttäjälle paremmin kohdennettuja ehdotuksia. (Aho 2018.)

Yle on todennut big datan käyttönsä yhteydessä, että valmistelu on avainasemassa, big datan analysoinnista saadut tulokset ovat arvaamattomia ja valittujen datan käsittelyratkaisujen tulee olla joustavia, jotta yllättäviin tilanteisiin voidaan reagoida. Big data-analytiikan käyttöönotossa tulee myös ottaa huomioon kustannukset; aluksi on hyvä lähteä kevyellä kokeilulla liikkeelle ja katsoa, mikä tiedon keruu muoto ja analysointi tekniikka sopii yrityksen tarkoituksiin. (Aho 2018.)

3.6 Pk-yritykset ja big datan mahdollisuudet

SAP:n rahoittaman, pieniin yrityksiin keskittyneen kyselyn mukaan noin 2/3 pienistä yrityksistä ovat vasta alkutekijöissään, mitä tulee digitalisaation täydelliseen hyödyntämiseen (International Data Corporation 2017). Suurimpien yrityksiä kuten Amazonin, Facebookin ja Googlen liiketoimintamallit perustuvat pitkälti suurten datamassojen hyödyntämiseen, ja useat muutkin isot yritykset ovat saaneet lukuisia hyötyjä datan jalostamisesta liiketoiminnan tueksi. Moni pienempi yritys saattaa kuvitella, että big data kuuluu vain isommille yrityksille tai että sen käyttö on kallista. Oli yritys kuitenkin miten pieni tahansa, tiedon keräämisestä ja sen strategisesta hyödyntämisestä on apua yrityksen kasvattamiseen ja uusien asiakkaiden löytämiseen. Nykyään on myös saatavilla monia edullisia tai ilmaisia työkaluja datan analysointiin, joten budjetin ei pitäisi olla enää myöskään ongelma. Big datan käyttö liiketoiminnassa tulee yleistymään entisestään tulevaisuudessa, ja sitä hyödyntämättömät yritykset saattavat lopulta pudota pois kilpailusta tai olla kasvamatta suurimpaan potentiaaliinsa. (Marr 2016a, 8–11.)

Moni pk-yritys saattaa kuvitella, että ei voi hyödyntää big dataa, koska pieneltä yritykseltä ei kerääntäni niin suuria datamassoja. Sen sijaan, että dataa olisi ääretön määrä, tärkeämpää on saada laadukasta, omaa liiketoimintaa hyödyttävää dataa nopeasti. Tulevien trendien tunnistaminen ja uusien asiakasryhmien löytäminen ovat tärkeitä asioita myös pienen yrityksen kilpailukyvyille. (Gordon 2017.)

Pienillä yrityksillä on myös etulyöntiasema big datan hyödyntämisen suhteen tietyillä osa-alueilla. Data-analytiikasta saadut oivallukset ovat hyödyttömiä, jos yritys ei pysty laittamaan niitä käytäntöön riittävän nopeasti ja tehokkaasti. Pienten yritysten ketteruus ja nopeus voikin olla iso etu kyseisessä tilanteessa. (Marr 2016a, 11.)

Millä tavoin pienet ja keskisuuret yritykset voivat sitten hyödyntää suuria datamassoja? Yksi tärkeistä datan lähteistä on yrityksen oma verkkosivu, olettaen että sellainen on olemassa. Pieni yritys ei välttämättä tarvitse datan analysointiin suurta erikoistunutta tiimiä taikka kalliita ohjelmistoja. Omilta verkkosivuilta saatu tieto asiakkaista voi jo itsessään olla riittävää liiketoiminnan kehittämiseen ja päätösten tekoon. Google Analytics on hyvä, ilmainen ja helposti saatavilla oleva työkalu. Sen avulla voi saada monipuolista tilastotietoa ja ymmärrystä sivuston kävijöistä ja kävijämäärästä, joka taas voi auttaa esimerkiksi potentiaalisten asiakkaiden ja kohderyhmän löytämisessä. (Glidden 2017.)

Myös muuta yrityksen sisäistä dataa on mahdollista jalostaa hyötykäyttöön. Monella pk-yritykselläkin on olemassa tietokanta asiakkaista. Kyselylomakkeet, jotka lähetetään asiakkaille sähköpostitse, on hyvä tapa kerätä asiakkailta mielipiteitä ja kokemuksia yrityksen tuotteista tai palveluista. Myös asiakaspalvelutilanteissa käydyistä keskusteluista kerätty data tai myyntitilastot antavat hyödyllistä informaatiota oikeanlaisella louhinnalla ja analysoinnilla. Pelkkä yrityksen sisäinen data ei aina ole kuitenkaan riittävää. Pk-yrityksille oman sisäisen datan vähäinen määrä voi luoda haasteen. Jos yritys ei halua itse kerätä ja analysoida dataa, yksi vaihtoehto on myös data-analytiikka palveluiden ostaminen ulkopuolelta. (Marr 2016a, 62–63.)

Ratkaisuna sisäisen datan vähäiseen määrään pk-yritys voi myös hyödyntää lukuisia ulkoisia datamassoja, joista monet ovat ilmaisia. Sosiaalinen media on yksi tärkeimmistä ulkoisista datan lähteistä. Facebook tallentaa ihmisistä lukemattomat määrät tietoa. Facebook Graph rajapinnan avulla yrityksen on mahdollista kerätä tietoa käyttäjistä ja heidän kirjoittamistaan kommentteista ja julkaisuista. Facebook, Instagram ja Twitter tarjoavat kaikki myös kohdistettua mainontaa, jolla yrityksen on helpompi tuoda tuotteitaan tai palveluitaan esiin juuri oikealle kohderyhmälle. Erilaiset sosiaalisessa mediassa pidetyt kampanjat saattavat tuoda lukuisan määrän jakoja, tykkäyksiä ja kommentteja, ja nämäkin toimivat hyödyllisenä datan lähteenä yritykselle. Sosiaalisen median kautta saadun datan analysoinnilla voi myös tunnistaa tulevat trendit ja osata varautua niihin ajoissa. Twitterissä käytyjen julkisten keskustelujen louhintaa varten on myös olemassa omat työkalunsa. Näiden työkalujen avulla on mahdollista tarkkailla mitä mieltä ihmiset ovat yrityksestä tai tuotteesta. (Marr 2016a, 40–42.)

Myös säästä kerätty data on arvokasta pienemmille yrityksille. Se on ilmaista ja kaikkien saatavilla verkosta. Säädataa on mahdollista hyödyntää asiakas- tai kävijämäärän ennustamisessa sekä esimerkiksi arvioidessa, kuinka paljon jotain tuotetta täytyy olla valmiina varastossa päivän varalle. Kun pystytään ennustamaan kävijämäärä tarkasti, on myös mahdollista varautua oikealla määrällä henkilökuntaa. Data-analytiikan hyödyntäminen

tässä tapauksessa vähentää arvailun tarvetta sekä lisää kustannustehokkuutta. (Marr 2016a, 51.)

Google Trends on myös yksi datan lähde, josta pienempikin yritys voi saada hyötyä liiketoimintansa tueksi. Se on ilmainen, mutta tehokas työkalu suosituimpien hakutulosten löytämiseen. Työkalun avulla on mahdollista nähdä mikä on suosittua juuri nyt ja lähitulevaisuudessa. Oleelliset tiedot oman yrityksen kannalta on mahdollista saada esiin suodattimien avulla. (Marr 2016a, 221.)

Vaihtoehtoisesti pk-yritykset voivat myös kehittää omia datan keräystapojaan. Erilaisilla liikesensoreilla on mahdollista havaita esimerkiksi myymälässä olevien asiakkaiden liikkeitä. Tästä on mahdollista saada selville kuinka paljon asiakkaita käy kaupassa ja mitkä osastot ovat suosituimpia. Tieto asiakkaiden päivittäisestä määrästä ja siitä mitä he suosivat tuo apua muun muassa henkilöstötarpeen suunnitteluun. Myös asiakkaille suunnattu palkinto-ohjelma on hyvä tapa lisätä myyntiä, sekä samalla saada tietoa asiakkaiden ostotottumuksista, etenkin jos sen liittää osaksi verkkokauppaa. (Kh 2018.)

4 TUTKIMUSAINIESTO JA ANALYYSI

4.1 Haastattelurunko

Opinnäytetyön tutkimusosuus suoritettiin haastatteluilla. Haastattelussa oli kaksitoista kysymystä (LIITE 1), ja osaan kysymyksistä yritys pystyi vastaamaan, vaikka ei hyödyntäisikään big dataa liiketoiminnassaan tai päätöstenteeissa. Tällä pyrittiin saamaan tutkimusaineistoa myös siitä, onko big dataa hyödyntämättömillä pk-yrityksillä suunnitelmia data-analytiikan suhteen tulevaisuudessa.

Haastattelu pyrki selvittämään hyödyntävätkö pienet ja keskisuuret yritykset big dataa liiketoiminnassaan, ja jos hyödyntää, niin millä tavoin ja missä määrin. Myös yleistä näkemystä suurista datamassoista ja niiden merkityksestä pk-yritysten liiketoiminnalle tulevaisuudessa oli tarkoitus kartoittaa.

Haastattelut suoritettiin ottamalla yhteyttä yrityksiin sähköpostin kautta. Saaduista haastattelutuloksista oli tarkoituksena rakentaa suuntaa antava kuva big datan nykytilanteesta suomalaisissa pk-yrityksissä.

4.2 Haastatteluaineisto

Haastatteluihin valittiin yrityksiä eri aloilta, jotta tuloksiin saataisiin vaihtelevuutta ja mahdollisimman monipuolisia näkökulmia. Vaikka Suomen big data -markkinoista on tehty selvityksiä ja tutkimuksia, tutkimustietoa big datan hyödyntämisestä suomalaisissa pk-yrityksissä ei ole vielä juuri saatavilla ainakaan julkisesti. Alkuoletuksena kuitenkin oli, että osa pk-yrityksistäkin nykyään hyödyntää big dataa jossain määrin, tai on ainakin tutustunut aiheeseen.

Haastattelupyynnö lähetettiin sadalle pk-yritykselle. Kymmenen yritystä vastasi haastattelupyynnöön, joista viisi yritystä pystyi antamaan haastattelun. Loput viisi yritystä eivät joko hyödyntäneet big dataa liiketoiminnassaan tai tienneet aiheesta tarpeeksi voidakseen vastata kysymyksiin. Lisäksi mukaan otettiin myös erillinen haastattelu isommalta yritykseltä, jotta saataisiin erilaista näkökulmaa ja vertailua siitä, miten isommissa yrityksissä hyödynnetään big dataa. Haastatteluaineisto analysointiin etsien yhtäläisyyksiä tai eroavaisuuksia eri yritysten välillä ja tekemällä tutkittavasta ilmiöstä päätelmiä. Seuraavaksi käydään läpi kaikkien haastatteluiden tulokset yritys kerrallaan.

4.2.1 Yritys A

Ensimmäinen haastateltava yritys (yritys A) oli sähkö- ja tietoliikennepalveluiden tuottaja. Haastateltavana oli yrityksen toimitusjohtaja. Mielipide big datasta oli lyhyt ja ytimekäs:

Big data avaa yrityksille uusia mahdollisuuksia kehittää omaa liiketoimintaa sekä mahdollistaa myös uutta liiketoimintaa.

Vastauksesta ilmenee, että haastateltavan näkemys big datasta on positiivinen. Haastateltava kuitenkin koki, että tällä hetkellä big datan merkitys suomalaisten pk-yritysten liiketoiminnalle on vielä varsin vähäinen. Hänen mukaan asia voisi kuitenkin muuttua tulevaisuudessa, kunhan työkalut kehittyisivät enemmän myös pk-yritysten tarpeisiin.

Yritys A ei hyödynnä liiketoiminnassaan big dataa vielä tällä hetkellä. Big datan haasteina haastateltava piti suurta tietomäärää sekä lähteiden moninaisuutta; kun tulkintaa tehdään isosta tietomäärästä, vaaditaan myös enemmän ymmärrystä tiedon lähteestä.

4.2.2 Yritys B

Yritys B vuokraa työ- ja kokoustiloja yrityksille sekä tarjoaa asiantuntijapalveluita. Haastateltavana oli yrityksen kehitysjohtaja. Haastateltavan näkemyksenä oli, että big data kasvaa jatkuvasti esineiden internetin ja muiden dataa tuottavien toimintojen yleistyessä. Hän mainitsi myös, että big dataa hyödynnetään entistä enemmän muun muassa analytiikan ja koneoppimisen kautta.

Haastateltava oli sitä mieltä, että big datan merkitys pk-yritysten liiketoiminnassa tulee jatkossa kasvamaan huomattavasti, ja että tällä hetkellä vain pieni osa hyödyntää sitä riittävästi. Myös palveluliiketoiminta laajentuu, ja tekoäly sekä koneoppiminen tulevat vahvasti mukaan. Kysyttäessä miten suuri merkitys big datalla on haastateltavan mielestäsi pk-yritysten kilpailukykyyn nykypäivänä, asteikolla 1-5 (1 vähäinen merkitys, 5 suuri merkitys), vastaus oli 2.

Yritys B käyttää omassa liiketoiminnassaan big dataa vielä erittäin vähäisesti. Kerättävä data on lähinnä yrityksen sisäistä talousdataa ja käytössä on joitain raportoinnin pohjana olevia ratkaisuja. Konkreettista hyötyä tai yrityksen toimintatapojen muutosta big data ei ole vielä tuonut yritykselle; lähinnä vain raportointi on tehostunut. Kun kysyttiin missä määrin big data vaikuttaa päätösten tekoon yrityksessä, vastaus oli seuraava; ”*paremman talousdatan ansioista voi sanoa, että hieman*”.

Kysyttäessä big datan haasteista ja haittapuolista, esiin nousi etenkin yritysten haluattomuus muutokseen.

Haasteena yleensäkin digitalisaation vähäinen hyödyntäminen. Ei olla halukkaita muutokseen. Kilpailu tulee kovenemaan jatkossa entistä enemmän ja jos yritykset eivät nyt ole mukana hyödyntämisessä, niin tulevaisuus voi olla ainakin joillain toimialoilla haastavaa.

Kuten vastauksesta ilmenee, digitalisaation omaksunta voi tuottaa haasteita monelle yritykselle. Tämä koskee todennäköisesti etenkin pienempiä, muun kuin IT-alan yrityksiä. Halukkuus muutokseen olisi tärkeä ensimmäinen askel varmasti monelle, jos haluaa pysyä mukana kilpailussa myös tulevaisuudessa.

4.2.3 Yritys C

Yritys C on palveluita ja ohjelmistoja rakennusteollisuudelle kehittävä IT-alan yritys. Haastateltavana oli yrityksen teknologiajohtaja. Ensimmäisenä kartoitettiin haastateltavan näkemystä big data -käsitteestä.

Big data –termiä käytetään kovin vaihtelevasti kontekstista riippuen. Itse käsittän big datan pääasiallisesti erilaisen telemetriatiedon keräämistä analytiikka varten. Tieto voi sisältää joko ohjelmiston, laitteen tai vaikkapa käyttäjän käyttäytymistä kuvaavaa dataa, jota usein käytetään asiakkaiden käyttökokemuksen ja/tai tuotteen kehittämiseen.

Kysyttäessä mielipidettä big datan merkityksestä suomalaisten pk-yritysten liiketoiminnan kannalta, haastateltava ei halunnut yleistää mielipidettään koskemaan kaikkia eri toimialoja, mutta oli kuitenkin sitä mieltä, että data-analytiikka auttaa yhä enemmän ja enemmän yrityksiä ohjaamaan tuotekehitystään ja palvelutoimintaansa oikeampaan suuntaan, sekä antamaan palautetta, jota aiemmin oli ehkä vaikeaa tai mahdotonta saada. Asteikolla 1-5, big datan merkitys pk-yritysten kilpailukyvyllä nykypäivänä on haastateltavan mukaan ehkä vain tasolla 3. Hän mainitsi kuitenkin, että oikeanlaisen datan keräys ja analysointi vie oman aikansa, ja ne, jotka osaavat tämän parhaiten, hyötyvät jo nyt big datasta merkittävästi.

Yritys C hyödyntää omassa liiketoiminnassaan big dataa kohtuullisesti, tai ehkä jopa vähäisesti vielä, mutta kuitenkin laajenevissa määrin. Data kerätään tyypillisesti ohjelmiston käyttöä ja käyttäytymistä seuraavista lähteistä, joiden perusteella on mahdollista arvioida virhetilanteiden esiintymisen tiheys sekä tunnistamaan tyypillisiä ohjelmiston käyttötapoja. Muun muassa Elastic Search -ohjelmisto sekä New Relic -palvelu ovat käytössä erilaisten telemetriatietojen tallentamiseen.

Kysyttäessä onko big datan hyödyntäminen muuttanut yrityksen toimintatapoja millään tavoin, vastaus oli myönteinen. Haastateltava kertoi, että he pyrkivät jatkuvasti kehittämään toimivampia mittareita tuotekehityspäätöksiensä tueksi. Konkreettista hyötyä yritykselle on tuonut kerätyn tiedon perusteella tehty analytiikka, joka toimii tukena tuotekehitykseen liittyvien päätösten tekemisessä. Kysyttäessä missä määrin big data vaikuttaa päätösten tekkoon yrityksessä, vastaus oli ”jonkin verran” (vaihtoehdot olivat erittäin paljon/paljon/jonkin verran/hieman/ei ollenkaan).

Haastateltavan näkemys big datan haittapuolista ja haasteista liittyi etenkin big datan tuomiin yksityisyyden ongelmiin.

Yksityisyyteen liittyvät haasteet on otettava huomioon. GDPR tuo osaltaan selkeyttä tähän. Erilaisten telemetriatietojen sivutuloksena voidaan pahimmillaan pystyä luomaan hyvinkin tarkkoja profiileja käyttäjistä, vaikka niin ei tietoa kerättäessä ole edes ollut tarkoitus tietoa käyttää.

Vastauksesta esiin tulevat huolenaiheet yksityisyydestä ovat tärkeitä ottaa huomioon.

4.2.4 Yritys D

Yritys D on ICT-yritys, joka on erikoistunut muun muassa talous- ja asianhallinnan ohjelmistoihin. Yritys ei tällä hetkellä hyödynnä omassa liiketoiminnassaan big dataa, mutta voisi kuvitella sen mahdollisesti tuovan tietopääomaa rikastamaan yrityksen olemassa olevia palveluita.

Haastateltava arvioi, että big datan merkitys pk-yritysten kilpailukykyyn nykypäivänä on 4 (asteikolla 1-5). Hän pitää big dataa isona asiana; vuonna 2017 luotiin enemmän dataa kuin edellisenä 5000 vuotena yhteensä.

Jotkut tahot vertaavat tietoa energiaan – mahdollistaa liiketoiminnan suunnittamisen oikeisiin asioihin ja asiakkaisiin, sekä toki uudentyypin toiminnan > kasvun mahdollistaja.

Haittapuolina ja haasteina haastateltava mainitsi olennaisen tiedon sirpaloitumisen, tavoitettavuuden, oikeellisuuden sekä yhteismitallisuuden. Haasteena on myös käytön oivaltaminen.

4.2.5 Yritys E

Yritys E on fysioterapia-alan yritys ja haastateltavana oli yrityksen toimitusjohtaja. Yritys ei hyödynnä toiminnassaan big dataa tällä hetkellä, mutta mahdollisuuksia tulevaisuudessa

varmasti voisi olla. Haastateltavan mukaan fysioterapian ja lääketieteen puolella kaikki ihmisten tuottama aktiivisuus- ja terveysdata voisi olla hyödyllistä oikealla analysoinnilla, ja sitä kautta voitaisiin saavuttaa jopa kansanterveydellisiä vaikutuksia.

Erittäin suurten, järjestelemättömien tietomassojen, kerääminen, tallentaminen sekä hyödyntäminen tulee lisääntymään jatkuvasti tulevaisuudessa.

Kysyttäessä big datan merkityksestä suomalaisten pk-yritysten liiketoiminnassa, haastateltava oli sitä mieltä, että näiden tietojen oikeanlainen analysointi antaa yrityksille mahdollisuuden laadukkaampaan päätöksentekoon. Näin tehokkuus kasvaa ja yrityksen menestys voi parantua. Avointen tietovarastojen tai ulkoisten lähteiden hyödyntäminen voi luoda mahdollisuuden pk-yritykselle tuottaa uusia palveluita ja sitä kautta liiketoimintaa.

Haastateltava kokee, että big datan merkitys pk-yritysten kilpailukykyyn nykypäivänä on keskimäärin 2 (asteikolla 1-5). Yrityksen toimialalla tosin on merkitystä, kuinka merkittävää big data yritykselle on. Toisaalta pk-yrityksillä on vain rajalliset mahdollisuudet tuottaa ja hyödyntää big dataa, ja joillakin tätä mahdollisuutta taas ei ole ollenkaan.

Big datan haittapuolina ja haasteina haastateltava mainitsi muun muassa osaamisen; oikeanlainen big datan analysointi ja hyödyntäminen vaatii todella paljon osaamista analysoijalta sekä monimutkaisia algoritmeja. Myös tiedon nopeaa syntymisnopeutta sekä valtavien tietomäärien vaatimaa tilantarvetta voidaan pitää haittapuolena.

4.2.6 Lisähaastattelu

Lisänä tutkimukseen on sisällytetty myös haastattelu suurelta suomalaiselta teleoperaattorilta. Tämä antaa hyvää vertailupohjaa siitä, miten isot yritykset hyödyntävät big dataa tehokkaasti.

Seuraavat asiat kävivät ilmi haastattelusta. Haastateltu yrityksen edustaja on sitä mieltä, että ilman big data -ratkaisuja ei enää pärjätä kilpailussa. Niin sanotusta big datan hypevaiheesta on siirrytty todellisia liiketoiminta-arvoa tuottaviin ratkaisuihin ja seuraavan viiden vuoden aikana big datan hyödyntäminen on välttämättömyys kilpailukyvyyn kannalta. Viimeisen kolmen vuoden aikana datamassojen mahdollisimman reaaliaikainen analysointi ja analysoidun datan syöttäminen takaisin tuotantoprosessiin on tullut todella tärkeäksi, sekä teknisellä että markkinoinnillisella puolella.

Haastateltava oli sitä mieltä, että heidän tämän hetkinen big datan käyttö on kohtuullisella tasolla, mutta muuttuu parin vuoden sisällä laajaksi. Ratkaisut ovat jo nyt olemassa ja enää on kyse vain niiden levittämisestä organisaation prosesseihin ja niitä tukeviin IT-järjestelmäratkaisuihin.

Yrityksellä on käytössä monia eri big data -ratkaisuja, muun muassa muutamia Hadoop-pohjaisia ratkaisuja (avoimen lähdekoodin ohjelmisto suurien tietomäärien käsittelyyn), Apache Kafka, Apache Spark, Amazon Web Services ja Amazon Elastic MapReduce. Kaikkia näitä ratkaisuja yhdistää helppo skaalautuvuus ylös ja alaspäin tarpeen mukaan. Se tuo tehoa laskentaan sekä myös kustannussäästöjä.

Big data on tuonut haastatellulle yritykselle konkreettista hyötyä asiakasanalytiikan kautta. Merkittävää tehostumista on tullut markkinointiin ja reaaliaikaiseen palvelukokemukseen teknisesti, verkossa, kaupassa ja asiakaspalvelussa. Tämä näkyy suoraan asiakaskohtaamisten konversioissa eli kyvyssä saada asiakkaita kiinnostumaan tuotteista ja palveluista ja lopulta tekemään yrityksen kanssa kauppaa.

Yritys kerää dataa, eli lähinnä mobiiliverkon käyttötietoja, verkkonsa teknisistä komponenteista. Myös yrityksen kaikista sähköisistä palvelukanavista tulee laajasti käyttödataa. Sosiaalista mediaa ja yrityksen palveluista käytävää keskustelua hyödynnetään myös datan lähteenä. Päätöksen tekoon kerätty data vaikuttaa yrityksessä tällä hetkellä paljon, ja tulee vaikuttamaan yhä enemmän tulevaisuudessa.

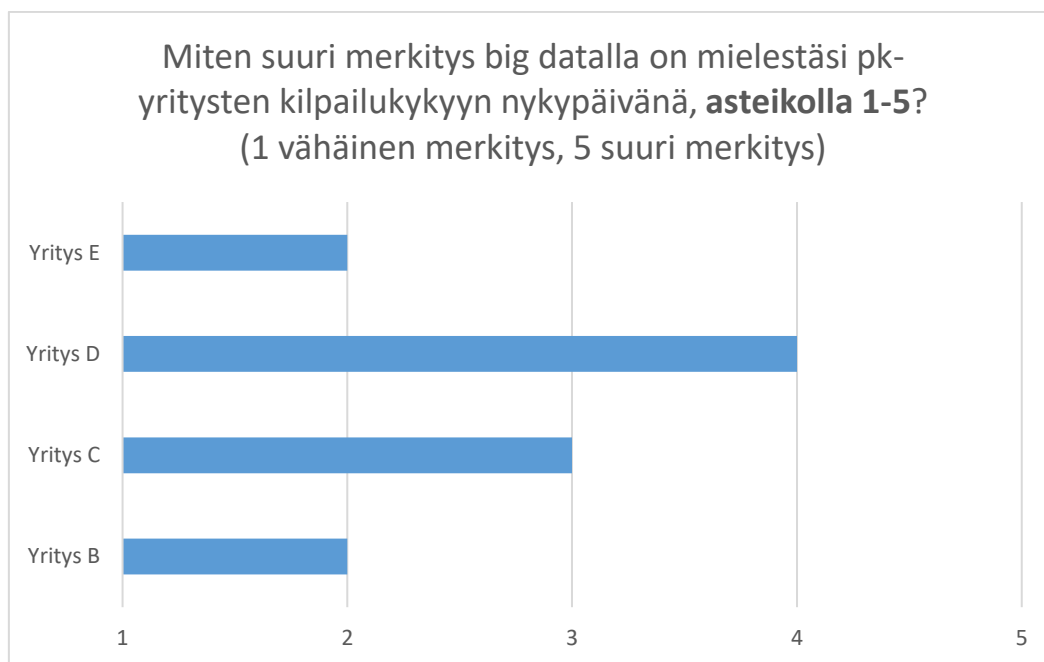
Yritys näkee, että suurimpia big datan haasteita on osaamisen puute. Tehokkaimmat ratkaisut vaativat paljon kustomointia ja laajaa osaamista koodauksesta, sekä mielellään myös tilastotieteestä. Perinteiset IT- tai tilastotiedetaustaiset henkilöt eivät suoraan sovellu big datan analysoinnin pariin. Myös liiketoiminnalta vaaditaan uutta asennoitumista data-ohjautuvaan liiketoiminnan ohjausmalliin. On mahdollista, että liiketoimintaprosessit ja niitä tukevat järjestelmät on suunniteltava kokonaan uudelleen big datan kokonaisvaltaisessa käyttöönotossa. Tämä taas toisi lisää kustannuksia ja olisi aikaa vievää.

4.3 Aineiston analyysi

Haastateltavista pk-yrityksistä kaksi hyödyntää suuria datamassoja liiketoiminnassaan, mutta kuitenkin vielä aika vähäisesti. Päätösten tekoon big data vaikutti näissä yrityksissä vain hieman tai jonkin verran. Kaikki haastatteluun osallistuneet pk-yritykset tiesivät kuitenkin aiheesta ja antoivat kattavia vastauksia big dataan liittyen. Lisähaastattelun suuri yritys puolestaan hyödyntää big dataa liiketoiminnassaan omien sanojensa mukaan kohtuullisesti, mutta yhä laajemmin tulevaisuudessa. Päätösten tekoon big data vaikuttaa kyseisessä yrityksessä paljon.

Suurin osa haastateltavista pk-yrityksistä koki, että big datan merkitys pk-yritysten kilpailukykyyn nykypäivänä on vielä suhteellisen vähäinen tai kohtuullinen (KUVIO 2). Yritys A ei osannut antaa vastausta kysymykseen. Kaikki yritykset olivat kuitenkin sitä mieltä, että big

datan merkitys tulee kasvamaan tulevaisuudessa. Yleismaallisena haasteena pidettiin datan keräämistä, sen aikaa vievyyttä ja lähteen tuntemusta; jollei lähdettä tunneta, saadun datan analysointi voi mennä pahasti pieleen. Myös riittävän osaamisen puute koetaan haasteeksi.



KUVIO 2. Vastausjakauma big datan merkityksestä pk-yritysten kilpailukykyyn

Vastauksista käy ilmi, että ulkoisen datan merkitys tulee lisääntymään kaikille yrityksille, kun sitä saadaan jatkuvasti enemmän käyttöön. Hyvänä esimerkkinä annettiin terveysalalle hyödyllisiä henkilön aktiivisuustietoja, joita eri älylaitteiden avulla kerätään jatkuvasti kasvavissa määrin ja jota terveysalan yritykset voivat hyödyntää koosta riippumatta. Samanlaista dataa saadaan kerättyä esimerkiksi ihmisten verkkokäyttäytymisestä ja ostopäätöksistä enenevissä määrin.

Haastateltavat yritykset pitivät Suomen big datan analysoinnin osaamispuutaa suurena ongelmana tällä hetkellä, mutta myös tulevaisuudessa. Ilman datan analysointia hallitsevia henkilöitä, ei datasta ole hyötyä, vaan siitä voi olla jopa haittaa. Pelkkä datan oikea analysointi ei välttämättä edes riitä, datasta tulee tunnistaa erinäiset mallit ja kaavat, joita yritys voi sitten hyödyntää. Haasteiksi koettiin myös yksityisyyteen liittyvät kysymykset sekä yritysten haluttomuus muutokseen. Edellä mainitut haasteet nousivat esiin myös kirjallisuudessa ja muissa big dataan liittyvissä kirjoituksissa, kuten opinnäytetyön teoriaosuudessa havaittiin.

Haastatellut pk-yritykset, jotka big dataa hyödynsivät, käyttivät big dataa lähinnä sisäisten tilanteiden ratkaisussa, raportoinnin tukena tai tuotekehityksessä. Yritykset pyrkivät tunnistamaan mm. virhetilanteita ja estämään niitä ennen niiden syntyä. Toinen yrityksistä ei kokenut vielä saavansa big datasta konkreettista hyötyä, kun taas toinen yritys koki, että siitä on hyötyä etenkin tuotekehitykseen liittyvissä päätösten teoissa. Molemmat yritykset olivat kuitenkin sitä mieltä, että data-analytiikalla tulee olemaan entistä suurempi rooli yritysten liiketoiminnassa tulevaisuudessa. Lisähaastattelun suuri teleoperaattori puolestaan saa merkittävää hyötyä big datasta muun muassa markkinointiin ja asiakaspalveluun.

5 JOHTOPÄÄTÖKSET

Kerätystä tutkimusaineistosta voidaan havaita, että tutkimukseen osallistuneet suomalaiset pk-yritykset eivät vielä juuri hyödynnä big dataa. Vastausten perusteella voidaan kuitenkin päätellä, että monet pitävät sitä hyödyllisenä työkaluna ja monella olisi kiinnostusta hyödyntää sitä. Kuitenkaan iso osa ei ole siihen vielä ryhtynyt resurssipulan tai muun vastaavan takia. Big data aiheena on myös selkeästi tuttu monelle pk-yritykselle, vaikka yritys ei itse sitä hyödyntäisikään liiketoiminnassaan. Moni haastateltavista oli sitä mieltä, että big datan merkitys pk-yritysten liiketoiminnassa kasvaa tulevaisuudessa entisestään, vaikka ei tällä hetkellä ehkä olekaan vielä niin iso osa suomalaisten pk-yritysten liiketoimintaa.

Kirjallisuudesta ja maailmalta saaduista esimerkeistä voidaan kuitenkin selkeästi nähdä, että big dataa voi hyödyntää yritys kuin yritys, kokoon tai toimialaan katsomatta. Ainoa kysymys onkin yrityksille, miten dataa hyödynnetään ja millaisilla työkaluilla, sillä monelle pienelle yritykselle varsinkin datan kerääminen ja käsittely voi olla todella työläs prosessi.

Lisähaastattelun ja teoreettisen tiedon perusteella on selvää, että big datalla on tärkeä merkitys etenkin isoissa yrityksissä, jotka käsittelevät suuria datamassoja. Datasta saatavia hyötyjä ja kilpailuetuja on paljon. Aiempien tutkimusten mukaan 50% big dataan käytettävistä investoinneista menee markkinointiin, kuten kuluttajakokemuksen parantamiseen ja viestinnän kohdentamiseen (Liikenne- ja viestintäministeriö, 2013). Tämä näkyi myös lisähaastattelun yrityksessä, jolle big data toi merkittävää hyötyä etenkin markkinointiin ja asiakaspalveluun. Oli myös yllättävää, kuinka monta eri big data ratkaisua lisähaastattelun yritys käytti toiminnassaan. Tästä voi päätellä, että big datan tehokas käyttö vaatii useampien eri ratkaisujen yhdistelemistä.

Big data on jo tämän päivän työkalu monella yrityksellä, mutta tulevaisuudessa sen arvo tulee jatkuvasti kasvamaan. Internet ja sosiaalinen media luovat jatkuvasti tietovirtoja ihmisille, joista yritysten on vaikea erottautua. Dataa keräämällä ja analysoimalla yritys voi tehokkaasti kohdentaa palvelujaan ja selvittää mikä todellisuudessa kuluttajia kiinnostaa ja mikä saa heidät käyttämään palvelujaan, ja laajentamaan yleisesti toimenkuvaansa.

6 YHTEENVETO

6.1 Yhteenveto tutkimuksesta

Tutkimuksen tarkoituksena oli kartoittaa big datan käyttöä suomalaisissa pk-yrityksissä. Tutkimuskysymys aseteltiin sen mukaan, että saataisiin mahdollisimman laaja-alainen käsitys yrityksen big datan käytöstä toimialasta riippumatta. Tutkimus kysymys oli, että miten ja missä määrin suomalaisissa pk-yrityksissä hyödynnetään big dataa.

Kirja-aineistolla pyrittiin luomaan käsitys big datan keräämistavoista, käyttömahdollisuuksista sekä miten sitä on jo nyt käytetty maailmalla. Lähdemateriaalin kerääminen lähti siitä olettamasta, että suuret yritykset hyödyntävät suuria datamassoja liiketoiminnassaan jo laajasti, mutta pienet yritykset vain vähän. Teoreettista tietoa saatiin kerättyä myös aiemmista tutkimuksista ja selvityksistä, jotka liittyivät suurten datamassojen hyödyntämiseen Suomessa.

Opinnäytetyön tutkimusosuuden aineisto kerättiin kvalitatiivisin menetelmin haastattelemalla eri alojen pk-yrityksiä. Haastattelukysymykset suunniteltiin mahdollisimman yleisluontoisiksi, jotta yritykset koosta ja toimialasta riippumatta voisivat osallistua haastatteluihin. Haastattelu saatiin yhteensä viideltä eri pk-yrityksen edustajalta, ja se suoritettiin sähköpostin välityksellä. Mukana oli kaksi ICT-yritystä, fysioterapialan yritys, sähkö- ja tietoliikennepalveluiden tuottaja sekä yritys, joka vuokraa tiloja yrityksille ja tarjoaa asiantuntijapalveluita. Lisäksi mukaan otettiin haastattelu yhdeltä suurelta yritykseltä, suomalaiselta teleoperaattorilta, jotta aiheeseen saataisiin erilaista näkökulmaa.

Tutkimuksessa selvisi, että haastatteluun osallistuneet suomalaiset pk-yritykset eivät vielä hyödynnä big dataa kovin suurella mittakaavalla. Kaksi viidestä haastatteluun osallistuneesta hyödynsi suuria datamassoja, mutta kohtuullisen vähäisesti vielä. Data-analytiikkaa käytettiin muun muassa raportoinnin ja päätöksenteon tukena. Toinen yrityksistä koki saavansa big datasta konkreettista hyötyä ja suunnitelmana oli hyödyntää data-analytiikkaa yhä laajemmin tulevaisuudessa. Toinen yrityksistä taas koki, että big data on tehostanut lähinnä vain raportointia. Big datan rantautuminen pienempiin yrityksiin saattaa olla vielä alkutekijöissä, mutta tietoisuutta aiheesta on kuitenkin. Suuret yritykset ovat sen sijaan hyödyntäneet big dataa tehokkaasti jo vuosia. Pienet yritykset kokevat haasteiksi muun muassa kerätyn datan määrän ja sen oikean analysoinnin. Kirjallisuudesta voidaan kuitenkin todeta, että nämä pelot ovat osittain turhia; pelkkä datan määrä ei ratkaise vaan sen laatu ja käyttötarkoitus. Suppeahkolla datalla voidaan saada todella hyviä tuloksia ai-

kaan, mikäli sen ydinasia on sopiva ja sitä tulkitaan oikein. Yrityksillä on myös mahdollisuus käyttää ulkopuolisia palveluita auttamaan datan analysoinnissa ja toimenpiteiden päättämisessä.

6.2 Ongelmat tutkimusprosessissa

Haastatteluaineiston kerääminen osoittautui tutkimuksessa ajoittain ongelmaksi. Haastatteluita saatiin tutkimukseen toivottua vähemmän. Suurin osa, joille haastattelu lähetettiin, ei vastannut. Voi olla, että tämä johtuu siitä, ettei aihe ole vielä kovin tuttu tai ajankohtainen monelle suomalaiselle pk-yritykselle. Toisaalta, haastattelukysymykset oli rakennettu niin, että yritykset, jotka eivät käytä big dataa liiketoiminnassaan, voivat myös vastata niihin. Tämän tarkoituksena oli kannustaa useampaa yritystä osallistumaan tutkimukseen ja antaa oman näkemyksensä aiheesta. Voi olla, että kaikki yritykset eivät myöskään halua kommentoida datan käyttöään, vaikka hyödyntäisivätkin big dataa liiketoiminnassaan. Jo big dataa käyttävät yritykset voivat pelätä menettävänsä saavutetun hyödyn, tämä siitä huolimatta, että haastattelumateriaalia käsitellään opinnäytetyössä vain anonymisti. On myös mahdollista, että haastateltavien lähestyminen sähköpostitse ei ollut paras tapa, koska saapunut sähköposti voi helposti jäädä huomaamatta tai päätyä roskapostiin.

Koska haastatteluita saatiin kasaan hyvin vähän, tuli tutkimuksesta aika suppea. Suuremmalla haastatteluiden määrällä olisi voitu saada enemmän tutkimustietoa, joka auttaisi vastaamaan paremmin tutkimuskysymykseen ja -ongelmaan. Big datan käytöstä pk-yrityksissä olisi tällöin saatu laajempi kuva, joka olisi ollut paremmin yleistettävissä.

Myös teoriaosuuteen liittyi jonkin verran tiedonkeruu ongelmia. Vaikka tutkimustietoa big datan hyödyntämisestä isoissa yrityksissä onkin paljon saataville, ei pk-yrityksistä löydy aiheen osalta kovin paljon luotettavaa tietoa. Myös suomenkielistä tieteellistä kirjallisuutta aiheesta ei ole vielä niin paljon saatavilla. Tämä osoittaa suuressa määrin sen, että big datan käytöstä nimenomaan pk-yritysten työkaluksi on huomattava määrä tutkimatonta potentiaalia.

6.3 Tutkimuksen validiteetti ja reliabiliteetti

Validiteetti eli pätevyys tarkoittaa tutkimusmenetelmän kykyä mitata sitä mitä sen on määrä mitata. Reliabiliteetti eli luotettavuus tarkoittaa taas mittauksen toistettavuutta ja sitä, onko aineiston keruu tehty oikein. (Tilastokeskus 2018.)

Opinnäytetyön reliabiliteettia olisi voitu parantaa, jos empiiristä tutkimusaineistoa olisi saatu enemmän. Kuitenkin, koska tutkimus on laadullinen ja perustuu yksittäisiin haastat-

teluihin, on reliabiliteettia hankala arvioida. Tutkimuksesta saatu materiaali on liian suppeaa vastatakseen tutkimusongelmaan täysin luotettavasti. Jotta tutkimus olisi voinut vastata paremmin tutkimusongelmaan, olisi haastateltavia yrityksiä täytynyt olla enemmän. Näin olisi voitu saada laajempia ja monipuolisempia näkökulmia aiheeseen.

Tutkimuksen suorittamista haastattelun muodossa sekä tutkimustuloksia voidaan kuitenkin pitää pätevinä, koska haastateltavat henkilöt olivat alansa asiantuntijoita ja vastasivat kysymyksiin asianmukaisesti. Haastatteluissa esille tulleet asiat ovat yritysten omia näkemyksiä ja kokemuksia aiheesta eikä tulkinnallisia ongelmia ei ilmennyt. Big dataa liiketoiminnassaan hyödyntävät yritykset myös antoivat konkreettisia ja tarkkoja esimerkkejä siitä, mitä hyötyjä yritys on saanut data-analytiikasta.

6.4 Yleistettävyyys

Tutkimustulokset antoivat hyviä näkökulmia siitä, miten eri suomalaiset pk-yritykset kokevat big datan ja sen hyödyntämismahdollisuudet. Koska haastateltavia yrityksiä oli kuitenkin rajallisesti, ei tuloksia voida luonnollisesti yleistää koskemaan koko Suomen tilannetta aiheen osalta. Tulokset antavat lähinnä suuntaa siitä, että big data on vielä kohtuullisen uusi alue pk-yritysten liiketoiminnassa.

Suurempi määrä haastatteluista olisi antanut laajempaa näkemystä aiheeseen ja big datan erilaisiin hyödyntämistapoihin. Muutaman yrityksen haastattelun perusteella ei voi tehdä yleistyksiä muista yrityksistä, mutta se antoi kuitenkin selkeän kuvan siitä, millä tavoin big dataa on mahdollista hyödyntää myös pk-yritysten liiketoiminnassa. Laajemmalla haastattelu otannalla pk-yritysten big datan käyttötapoja ja ratkaisuja olisi voinut myös vertailla keskenään enemmän ja saada tarkempaa kuvaa suomalaisten pk-yritysten tilanteesta. Voidaan kuitenkin todeta, että jokaisen yrityksen ollessa yksilö ja omatessa omat toimintatapansa ja asiakaskuntansa, ei tällaisia yleistyksiä ole välttämättä kannattavaa aiheesta edes pyrkiä tekemään. Big datan voidaan katsoa olevan pk-yrityksille tulevaisuuden työkalu. Suurille yrityksille se on sen sijaan jo tärkeä kilpailuvaltti. Koko ajan kiristyvässä kilpailussa tulee erottua kilpailijoista aina vain selkeämmin. Suurten datamassojen jalostamisella tiedoksi ja ymmärrykseksi yritys voi tulevaisuudessa aina vain paremmin tehostaa toimintaansa ja vaikuttaa jo olemassa olevien asiakkaiden ja potentiaalisten asiakkaiden päätöksen tekoon.

6.5 Suositukset jatkotutkimukselle

Sopiva jatkotutkimuksen aihe tälle opinnäytetyölle olisi, että miten pk-yrityksissä voitaisiin kehittää valmiutta ja keinoja data-analytiikan tehokkaalle hyödyntämiselle liiketoiminnassa. Jatkotutkimuksessa sähköisten haastatteluiden sijaan voitaisiin keskittyä muuttamaan pk-yritykseen lähemmällä tasolla ja tutkia miten juuri heidän yrityksessään voitaisiin hyötyä data-analytiikasta, ja miten se voitaisiin saada osaksi yrityksen liiketoimintaa käytännössä.

LÄHTEET

- Aho, T. 2018. Miltä näyttää Ylen big data? [Viitattu 16.11.2018] Saatavissa: <https://yle.fi/aihe/artikkeli/2018/03/07/timo-aho-milta-nayttaa-ylen-big-data>
- Enfo. 2018. Case: Digitaalisuus osana liiketoimintaa. [Viitattu 18.11.2018] Saatavissa: <https://www.enfo.fi/Asiakkaita/Pohjolan-Liikenne>
- Fenix Solutions. 2018. Mitä data-analytiikka tekee? [Viitattu 18.11.2018] Saatavissa: <https://www.fenixsolutions.fi/mita-data-analytiikka-tekee/>
- Glidden, C. 2017. Small But Mighty – How Big Data Benefits Small Business. [Viitattu 15.9.2018] Saatavissa: <https://insidebigdata.com/2017/12/17/small-mighty-big-data-benefits-small-business/>
- Gordon, A. 2017. How Your Small Business Can Make Use Of Big Data Now. [Viitattu 15.9.2018] Saatavissa: <https://www.forbes.com/sites/forbestechcouncil/2017/05/26/how-your-small-business-can-make-use-of-big-data-now/#e605d782dc6e>
- International Data Corporation. 2017. The Next Steps in Digital Transformation. [Viitattu 30.9.2018] Saatavissa: http://news.sap.com/wp-content/blogs.dir/1/files/SAP_IDC_infobrief_SMB_DX_102016.pdf
- Kh, R. 2018. 5 Innovative Ways Small Companies Can Collect Big Data. [Viitattu 18.11.2018] Saatavissa: <https://www.smartdatacollective.com/5-innovative-ways-small-companies-can-collect-big-data/>
- KUVIO 1. SAS Institute Oy. 2018. Big Data Analytics. [Viitattu 18.04.2018] Saatavissa: https://www.sas.com/en_us/insights/analytics/big-data-analytics.html
- Liikenne- ja viestintäministeriö. 2013. Big data- strategia. [Viitattu 30.11.2017] Saatavissa: <https://www.lvm.fi/lvm-mahti-portlet/download?did=139030>
- Liikenne- ja viestintäministeriö. 2016. Datan hyödyntämistä lisätään hallituksen periaatepäätöksellä. [Viitattu 20.11.2017] Saatavissa: <https://www.lvm.fi/-/datan-hyodyntamista-lisataan-hallituksen-periaatepaatoksella>
- Loskin, D. 2013. Big Data Analytics: From Strategic Planning to Enterprise Integration with Tools, Techniques, NoSQL, and Graph. USA: MK Publications.
- Marr, B. 2015. Big Data - Using SMART Big Data, Analytics and Metrics To Make Better Decisions and Improve Performance. UK: John Wiley & Sons Ltd.

Marr, B. 2016a. Big data for small business for dummies. Chichester, West Sussex: John Wiley & Sons Ltd.

Marr, B. 2016b. Big Data in Practice: How 45 Successful Companies Used Big Data Analytics to Deliver Extraordinary Results. Chichester, West Sussex: John Wiley & Sons Ltd.

Marr, B. 2018. Big Data: What Is Hadoop - An Easy Explanation For Absolutely Anyone. [Viitattu 18.11.2018] Saatavissa:

<https://www.bernardmarr.com/default.asp?contentID=1080>

Reynolds, V. 2016. Big Data For Beginners. USA: CreateSpace Independent Publishing Platform.

Rihti, M. 2016. Big data – miten voit hyödyntää sitä liiketoiminnassasi. Telia Company. [Viitattu 19.04.2018] Saatavissa: <https://www.telia.fi/yrityksille/artikkelit/artikkeli/big-data-liiketoiminnassasi>

Salo, I. 2013. Big Data, tiedon vallankumous. Jyväskylä: Docendo.

Salo, I. 2014. Big data & pilvipalvelut. Jyväskylä: Docendo.

SAS Institute Oy. 2018. Lisää tuottavuutta S-ryhmän arvoketjuun. [Viitattu 16.11.2018] Saatavissa: https://www.sas.com/fi_fi/customers/customer-interviews/s-ryhma.html

Smith, J. 2016. Data Analytics: What Every Business Must Know About Big Data and Data Science. USA: Pinnacle Publishers, LLC.

TIEKE ry. 2016. Tehoa pienistä ratkaisuista – Big datan ketterää hyödyntämistä suun terveydenhuollossa. [Viitattu 19.04.2018] Saatavissa:

<https://www.tieke.fi/pages/viewpage.action?pageId=46926040#FBD>

Tilastokeskus. 2018. Käsitteet. [Viitattu 12.10.2018] Saatavissa:

<https://www.stat.fi/meta/kas/index.html?R>

Valtioneuvosto. 2016. Valtioneuvoston periaatepäätös. [Viitattu 20.11.2017] Saatavissa:

<https://valtioneuvosto.fi/paatokset/paatos?decisionId=0900908f804c23c7>

Venäläinen, P. 2016. Puun kuljetus- ja tiestödatan digitalisaatio: Digiroad- ja kuvadatapilotit. TIEKE Tietoyhteiskunnan kehittämiskeskus ry. [Viitattu 19.04.2018]

Saatavissa: <https://www.tieke.fi/pages/viewpage.action?pageId=46926040#FBD>

Virkki, J. 2015. Suomi kulkee muita Pohjoismaita edellä Big Datassa. [Viitattu 25.11.2017]
Saatavissa: <http://www.talouselama.fi/kumppaniblogit/sas/suomi-kulkee-muita-pohjoismaita-edella-big-datassa-3472613>

LIITTEET

LIITE 1. Haastattelukysymykset

1. Mikä on näkemyksesi big datasta?
2. Mikä merkitys sillä on mielestäsi nyt ja tulevaisuudessa suomalaisten pk-yritysten liiketoiminnan kannalta?
3. Miten suuri merkitys sillä on mielestäsi pk-yritysten kilpailukykyyn nykypäivänä, asteikolla 1-5? **(1 vähäinen merkitys, 5 suuri merkitys)**
4. Mitkä ovat mielestäsi big datan haittapuolet ja haasteet?
5. Hyödyntääkö yrityksesi big dataa liiketoiminnassa? **(KYLLÄ/EI)**

Jos vastaus kysymykseen 5 oli EI:

6. Voisitko kuvitella yrityksesi hyödyntävän suuria datamassoja tulevaisuudessa? Mitä hyötyä voisit kuvitella sen tuovan yrityksellesi?

Jos vastaus kysymykseen 5 oli KYLLÄ:

7. Miten laajaa big datan käyttö on yrityksessäsi? **(laajaa/kohtuullista/vähäistä)**
8. Mitä big data -ratkaisuja/ohjelmistoja yritykselläsi on käytössä?
9. Miten ja mistä data kerätään? Mitä lähteitä käytetään? (esim. ulkoisia vai sisäisiä)
10. Onko big datan hyödyntäminen muuttanut yrityksenne toimintatapoja millään tavalla?
11. Millaista konkreettista hyötyä big data on tuonut yrityksellesi? (esim. tuotekehitys, kulojen minimointi, kustannustehokkuus etc.)
12. Missä määrin big data vaikuttaa päätösten tekoon yrityksessäsi? **(Erittäin paljon/paljon/jonkin verran/hieman/ei ollenkaan)**