

Valtteri Tuominen

**DATAN HYÖDYNTÄMINEN VERKKOKAUPAN LII-
KETOIMINTAJOHTAMISEN PÄÄTÖKSENTEOSSA**



JYVÄSKYLÄN YLIOPISTO
INFORMAATIOTEKNOLOGIAN TIEDEKUNTA
2023

TIIVISTELMÄ

Tuominen, Valtteri

Datan hyödyntäminen verkkokaupan liiketoimintajohtamisen päätöksenteossa

Jyväskylä: Jyväskylän yliopisto, 2023, 29 s.

Tietojärjestelmätiede, kandidaatintutkielma

Ohjaaja(t): Kuusio, Ari

Verkkokauppojen liiketoiminnasta kertyvän datan määrä kasvaa jatkuvasti. Tämän tutkimuksen tarkoituksena oli selvittää, miten verkkokauppojen dataa saadaan hyödynnettyä liiketoiminnan päätöksenteossa. Tutkielma suoritettiin kirjallisuuskatsauksena. Tutkielman tuloksista selvisi, että liiketoimintatiedon hallinnan avulla kertyvästä datasta voidaan saada arvokasta tietoa liiketoiminnan päätöksentekoon ja verkkokaupan liiketoiminnassa erityisesti markkinointia voidaan tehostaa massadata-analytiikan avulla. Verkkokaupan datan kaltaisen massadatan saaminen päätöksenteossa hyödynnettäväksi tiedoksi vaatii kuitenkin yritykseltä erilaisten massadatateknologioiden ja menetelmien hallitsemista.

Asiasanat: verkkokauppa, liiketoimintatiedon hallinta, massadata, massadata-analytiikka

ABSTRACT

Tuominen, Valtteri

Utilization of data in decision making of e-commerce business management

Jyväskylä: University of Jyväskylä, 2023, 29 p.

Information Systems science, Bachelor's thesis

Supervisor(s): Kuusio, Ari

The amount of data that can be collected from ecommerce businesses is increasing rapidly. The purpose of this study was to find out how this data can be utilized in decision-making. The thesis was conducted as a literature review using peer-reviewed articles gathered from Google Scholar and Scopus. The results of the thesis revealed that valuable information for business decision-making can be obtained from the data collected with the usage of business intelligence, and in the e-commerce business, marketing in particular can be enhanced with the help of big data analytics. However, transforming big data of an e-commerce business into information that can be utilized in decision-making requires the company to master various technologies and methods.

Keywords: e-commerce, business intelligence, big data, big data analytics

KUVIOT

KUVIO 1	Liiketoimintatiedon hallintajärjestelmän arkkitehtuuri	12
---------	--	----

SISÄLLYS

TIIVISTELMÄ

ABSTRACT

KUVIOT JA TAULUKOT

1	JOHDANTO.....	6
1.1	Motivointi	6
1.2	Tutkimusongelma ja tutkimuksen rajaus.....	7
1.3	Tutkimusmenetelmä	7
1.4	Keskeiset käsitteet.....	8
2	LIIKETOIMINTATIEDON HALLINTA PÄÄTÖKSENTEON TUKENA ..	10
2.1	Liiketoimintatiedon hallinta prosessi ja järjestelmät	10
2.2	Liiketoimintatiedon hallinnan hyödyt päätöksenteossa.....	13
3	VERKKOKAUPAN DATA JA LIIKETOIMINTATIEDON HALLINTA....	15
3.1	Massadata ja sen ominaispiirteet.....	15
3.2	Massadata ja liiketoimintatiedon hallinta	16
3.2.1	Massadatan kerääminen.....	18
3.2.2	Massadatan varastointi ja käsittely.....	18
4	VERKKOKAUPAN DATA PÄÄTÖKSENTEOSSA.....	20
4.1	Massadata-analytiikka verkkokaupan päätöksenteossa.....	20
4.2	Massadata-analytiikan sovelluskohteet verkkokaupassa.....	21
4.2.1	Markkinointi	21
4.2.2	Tuotantoketjun hallinta	23
4.2.3	Turvallisuus ja petoksien havaitseminen	23
5	YHTEENVETO	24
	LÄHTEET	26

1 JOHDANTO

1.1 Motivointi

Yrityksillä on mahdollisuus kerätä jatkuvasti enemmän dataa liiketoiminnastaan. Nykypäivänä onkin jopa vaikea löytää yritystä, joka ei hyödyntäisi datan keruuta liiketoiminnassaan (Chaudhuri, Dayal & Narasayya, 2011). Yritykset pyrkivät hyödyntämään kerättyä dataa saavuttaakseen kilpailuetua markkinoilla. Useimmat yritykset kuitenkin kohtaavat ongelmia liiketoimintatiedon hyödyntämisessä päätöksentekoprosesseissaan (Davenport, 2006).

Verkkokaupoista on tullut erityisesti datan louhinnalle merkittävä toimiala, sillä verkkokaupoista kerättävän käyttäjätiedon ja liiketoiminnallisen datan määrä kasvaa nopeasti hyvin suureksi (Kohavi, Mason, Parekh, & Zheng, 2004). Monet yritykset löytävät kilpailuetua hyödyntämällä suurta datamäärää muun muassa segmentoimalla asiakkaita entistä tarkemmin ja pystyvät näin markkinoimaan kohdennetummin (Du, Liu & Zhang, 2019).

Gartner (2022a) korostaa dataan pohjautuvan päätöksenteon merkitystä ja ennustaakin, että vuonna 2023 kolmasosalla suurista yrityksistä on organisatiossa analyttikkoja, jotka keskittyvät ainoastaan datapohjaiseen päätöksentekoon ja päätöksenteon mallintamiseen. Gartnerin (2022b) ennusteen mukaan myös liiketoimintatiedonhallinta- ja analytiikkaohjelmistojen markkinat ovatkin jatkuvassa kasvussa, ja ennusteen mukaan näiden ohjelmistojen markkina kasvaisivat 13,8 miljardiin dollariin vuoden 2025 loppuun mennessä.

1.2 Tutkimusongelma ja tutkimuksen rajaus

Dataan pohjautuvan päätöksenteon merkitys korostuu kilpaillussa markkinassa entisestään tulevaisuudessa ja erityisesti verkkokaupan liiketoiminnasta kerääntyä suuria määriä dataa. Datan hyödyntämisestä päätöksenteossa ja sen hyödyistä on jo tehty jonkin verran tutkimusta. Myös verkkokaupan liiketoiminnasta syntyvän massadatan hyödyntämistä lähinnä massadata-analytiikan avulla on tutkittu jonkin verran, mutta tutkimukset eivät anna kokonaisvaltaista kuvaa datan hyödyntämisestä verkkokauppojen liiketoiminnassa. Tämän tutkielman tarkoituksena onkin muodostaa kokonaisvaltainen kuva prosessista, jossa verkkokaupan liiketoiminnasta saatavilla olevaa dataa käytetään päätöksenteon tukena kehittämään liiketoimintaa. Tämän pohjalta tutkimusongelmaksi valikoitui:

Miten verkkokaupan dataa voidaan hyödyntää päätöksenteossa?

Tutkimusongelmaan lähdettiin etsimään vastausta seuraavien tutkimuskysymysten pohjalta:

- Miten liiketoimintatiedon hallinta voi parantaa päätöksentekoa yrityksissä?
- Millaista dataa verkkokaupan liiketoiminnasta kertyy ja miten sitä saadaan hyödynnettyä päätöksenteossa?
- Miten massadata-analytiikka voi kehittää verkkokaupan liiketoimintaa?

Täten tutkielman aihetta käsitellään siis lähinnä päätöksenteon näkökulmasta. Verkkokaupan dataa voitaisiin käsitellä laajemminkin, mutta tähän tutkielmaan aihetta on täytynyt rajata pääosin päätöksenteon kontekstiin. Esimerkiksi datan keräämiseen ja käsittelyyn liittyviä tietoturva-asteita tai uusimman EU:n tietosuojalain vaikutuksia verkkokaupan datan keräämiseen ei juurikaan käsitellä, jottei tutkielman aihe laajene liikaa.

1.3 Tutkimusmenetelmä

Tutkielma suoritettiin kuvailevana kirjallisuuskatsauksena. Kuvaileva kirjallisuuskatsaus antaa mahdollisuuden muodostaa laaja yleiskuva tutkimusaiheesta, sillä kuvailevassa kirjallisuuskatsauksessa käytetyt aineistot ovat laajoja ja

tutkimuskysymykset ovat väljempää kuin esimerkiksi systemaattisessa kirjallisuuskatsauksessa (Salminen, 2017). Tutkielman tavoitteena oli muodostaa laaja yleiskuva aiheesta ja tästä syystä kuvaileva kirjallisuuskatsaus valittiin kirjallisuuskatsauksen muodoksi (Salminen, 2017). Kirjallisuuskatsauksen lähteitä on haettu JYKDOK, Scopus ja Google Scholar tietokannoista hakusanoilla ”e-commerce”, ”data analytics”, ”decision making” ja ”business intelligence”, sekä erityisesti näiden erilaisilla yhdistelmillä. Aiheesta löytyy runsaasti tieteellisiä lähteitä ja tutkielmassa on pyritty pääosin käyttämään vertaisarvioituja sekä vähintään JUFO 1 luokiteltuja tieteellisiä lähteitä. Tutkielmassa on käytetty myös muuta aiheen kirjallisuutta, kuten Gartnerin uusimpia artikkeleita lähteenä.

1.4 Keskeiset käsitteet

Liiketoimintatietojen hallinnalle (engl. business intelligence) ei löydy kirjallisuudesta yksiselitteistä yleisesti käytössä olevaa määritelmää. Yleisesti liiketoimintatietojen hallinta ymmärretään laajempaan käsitteeseen, joka kattaa keinot ja menetelmät päätöksenteon kehittämiseksi (Lim ym. 2013). Nykäsen, Järvenpään ja Teittisen (2016) mukaan liiketoimintatietojen hallinnan käsitteen määrittelylle on kuitenkin yleensä kaksi näkökulmaa: teknologinen näkökulma ja prosessinäkökulma. Teknologinen näkökulma käsittelee ainoastaan datan käsittelyyn käytettäviä teknologioita eikä liiketoiminnallisia tavoitteita. Prosessinäkökulmassa pääpaino on prosessissa, jossa data muutetaan liiketoiminnassa hyödynnettäväksi tiedoksi ja teknologia on vain tukemassa tätä prosessia (Nykänen ym., 2016).

Wixomin ja Watsonin (2010, s. 13) mukaan liiketoimintatietojen hallinnalla tarkoitetaan usein teknologioita, sovelluksia ja prosesseja, joiden avulla kerätään, varastoidaan, analysoidaan liiketoimintatietoja, jonka avulla voidaan tehdä parempia päätöksiä. Chen ym. (2012) määrittelee tutkimuksessaan käsitteen BI&A eli liiketoimintatietojen hallinta ja analytiikka, joka sisältää teknologiat, käytännöt, menetelmät, järjestelmät ja sovellukset, joiden avulla analysoidaan liiketoiminnan dataa, jotta voidaan ymmärtää liiketoimintaa paremmin sekä täten tehdä parempia päätöksiä. Tämä määritelmä tukee hyvin Wixomin ja Watsonin (2010) määritelmää. Chen ym. (2012) määritelmä sisältää myös eri toimialoille soveltuvat liiketoimintakeskeiset käytännöt ja menetelmät. Pirrtimäki ja Lönnqvist (2006) sen sijaan määrittelevät liiketoimintatiedon hallinnan johtamisfilosofiaksi ja työkaluksi, jolla liiketoiminnasta kertyvä data jalostetaan tiedoksi, jonka avulla voidaan tehdä parempia päätöksiä.

Liiketoimintatietojen hallinnan päätavoitteina on mahdollistaa helppo ja interaktiivinen pääsy monimuotoiseen dataan sekä tämän datan käsittely ja muokkaaminen ja käsittely muotoon, joka mahdollistaa täsmällisen analyysin liiketoiminnasta sekä toimenpiteet analyysin pohjalta (Lim, Chen & Chen, 2013). Tässä tutkielmassa liiketoimintatietojen hallinnan määritelmä kattaa teknologiat,

menetelmät, järjestelmät ja käytännöt, joiden avulla voidaan tehdä parempia päätöksiä liiketoimintaa johdettaessa.

Verkkokauppa käsitetään yleisesti tuotteiden ja palveluiden ostamisena internetin välityksellä (Niranjanamurthy, Kavyashree, Jagannath & Chahar, 2013). Verkkokauppa voidaan jakaa kolmeen eri alalajiin, yritysten välinen kauppa (B2B), yritysten ja kuluttajien välinen kauppa (B2C) sekä kuluttajien välinen kauppa (C2C) (Bigdoli, 2002). Verkkokauppaan sisältyy kuitenkin paljon enemmän kuin vain yritysten ja asiakkaiden väliset maksutapahtumat ja verkkokaupalle voidaankin terminä löytää useita eri määritelmiä (Chaffey, 2007). Kalakota ja Whinston (1997) jakavat verkkokaupan määrittelyn neljään eri näkökulmaan: Kommunikaatio-, liiketoimintaprosessi-, palvelu- ja onlinenäkökulmaan. Kommunikaation näkökulmasta verkkokaupalla tarkoitetaan informaation, maksujen, tuotteiden tai palveluiden siirtämistä tietoverkkojen avulla. Liiketoimintaprosessin näkökulmasta verkkokauppa on työnkulun ja transaktioiden automatisoimista teknologian avulla. Palvelun näkökulmasta verkkokaupan nähdään olevan työkalu, jolla saadaan niin leikattua kuluja, kuin parannettua tuotteiden tai palveluiden laatua ja nopeutettua palveluiden tai tuotteiden toimitusta. Online-näkökulman mukaan verkkokauppa tarjoaa mahdollisuuden tuotteiden ja palveluiden ostamiseen ja myymiseen tietoverkon kautta (Kalakota & Whinston, 1997.).

Uudemmassa kirjallisuudessa verkkokaupan määritelmässä on painotettu enemmän asiakaslähtöisyyteen liittyviä asioita, kuten digitaalista arvonluontia sekä sähköistä markkinointia. Verkkokauppaan keskittyvien organisaatioiden korostetaan olevan erittäin asiakaslähtöisiä (Maity & Dass, 2014). Akter ja Wamba (2016) käyttävät verkkokaupan massadatan hyödyntämiseen liittyvässä tutkimuksessaan vieläkin laajempaa määritelmää verkkokaupasta. Tämä määritelmä kattaa kaikki verkkokaupan liiketoiminnan osa-alueet, joille datan analysoimisesta voidaan saada hyötyjä. Nämä osa-alueet kattavat toiminnot, joiden tavoitteena on saavuttaa transaktioarvoa eli kustannussäästöjä, tuottavuuden ja tehokkuuden parantamista, sekä strategista arvoa, kuten kilpailuetuja ja yrityksen suorituskyvyn parantamista. (Akter, Wamba, 2016). Koska tutkielman tavoitteena on selvittää, miten dataa voidaan hyödyntää verkkokaupan liiketoiminnassa, käytetään tässä tutkielmassa tätä Akterin ja Wamban (2016) käyttämää verkkokaupan laajennettu määritelmää.

2 LIIKETOIMINTATIEDON HALLINTA PÄÄTÖKSENTEON TUKENA

Liiketoimintatietojen hallinta käsite voi sisältää kattavasti teknologiat, menetelmät, järjestelmät ja käytännöt, joiden avulla voidaan tehdä parempia päätöksiä liiketoimintaa johdettaessa. Tämän takia liiketoimintatietojen hallinta (engl. business intelligence) käsitteenä nousikin kirjallisuutta etsiessä esille lähes kaikissa lähteissä, joissa käsitellään datan hyödyntämistä liiketoiminnan päätöksenteossa. Liiketoimintatietojen hallinta voidaan myös nähdä prosessina, jossa raakadata kerätään, käsitellään ja muunnetaan arvokkaaksi tietämykseksi liiketoiminnalle (Pirttimäki, 2007).

Kohdassa 2.1 käydään läpi, millainen liiketoimintatiedon hallintaprosessi tyypillisesti on ja minkälaisia teknologioita prosessissa yleensä käytetään sekä käsitellään myös sitä, miten prosessi voidaan soveltaa verkkokaupan datan hyödyntämiseen, jonka jälkeen käsitellään liiketoimintatiedon hallinnan hyötyjä päätöksenteon kannalta.

2.1 Liiketoimintatiedon hallinta prosessi ja järjestelmät

Liiketoimintatietojen hallintaprosessille löytyy kirjallisuudesta useita eri kuvauksia, kuten Zengin, Xun, Shin, Wangin ja Wun (2006) sekä Watsonin (2009) ja Pirttimäen (2007) kuvaukset, mutta pääpiirteittäin kaikista kuvauksista löytyy samat prosessin päävaiheet. Esimerkiksi Zengin ym. (2006) kuvauksessa liiketoimintatiedon hallintaprosessi on jaettu kuuteen vaiheeseen; liiketoimintaogelman määrittäminen, datamallin muodostaminen, datan esikäsittely, oikeiden analyysimenetelmien valinta, tulosten tulkinta ja tulosten pohjalta ongelmaan reagoiminen (Zeng ym., 2006). Tässä tutkielmassa käsitellään Pirttimäen (2007) esittämää prosessin kuvausta, sillä Pirttimäki (2007) keskittyy tutkimuksessaan

erityisesti liiketoimintatiedon hallintaprosessin tuomiin hyötyihin. Pirttimäen (2007) jakaa liiketoimintatiedon hallintaprosessin viiteen osaan:

1. Tietotarpeiden määrittäminen
2. Datan kerääminen
3. Datan prosessointi
4. Tiedon jakaminen
5. Tiedon hyödyntäminen

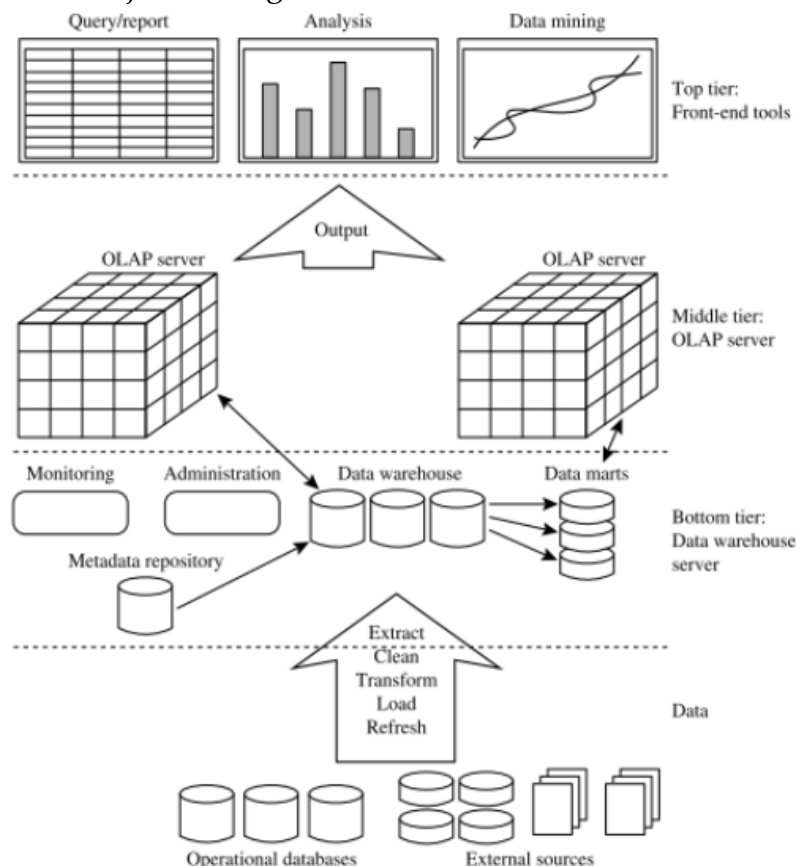
Prosessin ensimmäinen vaihe tietotarpeiden määrittäminen on kriittisen tärkeä, sillä se luo pohjan koko prosessille. On tärkeää varmistaa, että oikeanlainen informaatio päätyy päätöksentekijöille. Tietotarpeet tulee määrittää liiketoimintaan pohjautuen (Watson, 2014). Tietotarpeiden määrittämistä kuitenkin vaikeuttaa päätöksentekijöiden erilaiset tietotarpeet ja roolit organisaation sisällä (Pirttimäki, 2007)

Toinen vaihe on datan kerääminen. Perinteisissä liiketoimintatietojen hallintajärjestelmissä datan tärkeimpänä lähteenä toimii organisaation sisäiset järjestelmät (Chen ym., 2012), mutta moderneissa järjestelmissä dataa pystytään keräämään myös organisaation ulkopuolisista järjestelmistä (Pirttimäki, 2007). Sisäisistä järjestelmistä kerättävä data on usein rakenteeltaan strukturoitua, kun taas ulkopuolisista järjestelmistä kerättävä data on usein osittain strukturoitua tai täysin strukturoimatonta (Nykänen, Järvenpää & Teittinen, 2016).

Strukturoitu data voidaan määritellä niin, että sen muoto on ennalta määritettyä ja sitä pystytään sellaisenaan käsittelemään tietokoneohjelmilla, esimerkiksi tallentamaan relaatiotietokantaan, kuten SQL-tietokantaan (Baars & Kemper, 2008; Hu, Wen, Xhuan & Li, 2014). Osittain strukturoitu data taas voidaan määritellä niin, että sillä ei ole tarkkaa ennalta määritettyä muotoa, mutta data sisältää käyttäjien määrittämiä tunnisteita, mikä saa datan koneluettavaan muotoon. (Gandomi & Haider, 2015) Strukturoimatonta data sen sijaan ei noudata mitään ennalta määritettyä muotoa. Strukturoimatonta dataa voidaan kerätä monissa eri muodoissa, kuten videoina, tekstinä, aika- ja paikkatietoina (Mohamed, Najafabadi, Wah, Zaman & Maskat, 2020).

Kuviossa 1 on kuvattu tyypillinen liiketoimintatietojen hallintajärjestelmän arkkitehtuuri Ferreiran, Pedrosan ja Bernardinon (2017) mukaan. Kuten kuvio osoittaa datan kerääminen tapahtuu yhdistämällä organisaation sisäiset tietolähteet ja ulkoiset tietolähteet, muuntamalla ne sopivaan muotoon ja tallentamalla data tietovarastoihin erilaisten ETL (Extraction-Transformation-Load) menetelmien avulla (Ferreira, Pedrosa & Bernardino, 2017 ; Watson, 2009). ETL-menetelmät koostuvat kolmesta vaiheesta; datan kerääminen, muuntaminen ja lataaminen. Keräämisvaiheessa muodostetaan yhteys datan lähteisiin sekä valitaan ja haetaan tarvittava data (Hu ym., 2014). Muuntamisvaiheessa haettu data muunnetaan standardoituun muotoon ja lataamisvaiheessa muunnettu data ladataan liiketoimintatiedon hallintajärjestelmään (Ferreira ym., 2017). Tietovarastointia pidetään Liiketoimintatietojen hallinnan prosessin ytimenä (Lim ym., 2013). Tutkielman kolmannessa luvussa käsitellään tarkemmin verkkokaupan

liiketoiminnasta kertyvän datan keräämiseen, säilömiseen ja käsittelemiseen käytettäviä menetelmiä ja teknologioita.



Kuvio 1. Liiketoimintatiedon hallintajärjestelmän arkkitehtuuri (Ferreira, Pedrosa & Bernardino, 2017)

Pirttimäen (2007) kolmannessa vaiheessa kerättyä dataa käsitellään ja analysoidaan erilaisten työkalujen ja menetelmien avulla tietämykseksi (Pirttimäki, 2007). Liiketoimintatietojen hallintajärjestelmissä tällaisia menetelmiä ovat OLAP (online analytical processing) ja datanlouhinta (Watson, 2009). OLAP on yleinen liiketoimintatietojen hallintajärjestelmien analyysimenetelmä, jonka avulla tuotetaan raportteja tietystä liiketoiminnan mittarista käyttäjän ennalta määrittämien kriteerien mukaan sekä optimoidaan tiedonhakua generoimalla automaattisesti esimerkiksi SQL-kyselyjä (Olszak & Ziemia, 2006). Käyttäjät voivat vertailla esimerkiksi liiketoiminnan mittareiden tuloksia eri aikaväleillä (Watson, 2009). Datanlouhinta on laskennallinen prosessi toistuvien kaavojen löytämiseksi suurista tietojoukoista (Hu ym., 2014). Tutkielman kolmannessa tarkastellaan tarkemmin verkkokaupan datan analysointiin käytettäviä menetelmiä ja teknologioita.

Neljännessä vaiheessa tietoa jaetaan päätöksentekijöille (Pirttimäki, 2007). Tieto voidaan jakaa monissa eri muodoissa. Pirttimäen (2007) mukaan sen lisäksi, että tieto täytyy jakaa oikea aikaisesti, on myös päättäjien omat preferenssit otettava huomioon tiedon jakamisen muodossa.

Viimeisen vaiheen tavoitteena on mahdollistaa päättäjille pääsy kaikkeen tarvitsemaansa tietoon mahdollisimman nopeasti (Pirttimäki, 2007). Tieto pitää

olla helposti saatavilla, mutta kuitenkin tietoturva huomioon ottaen (Pirttimäki, 2007). Päättäjiltä saatu palaute auttaa taas uusien tietotarpeiden määrittelyssä ja liiketoimintatietojen hallinta onkin tämän takia iteratiivinen prosessi (Pirttimäki, 2007).

2.2 Liiketoimintatiedon hallinnan hyödyt päätöksenteossa

Päätöksentekoa yrityksissä lähestytään usein rationaalisen valinnan teorian pohjalta, missä eri vaihtoehtoja arvioidaan saatavilla olevan informaation perusteella ja arvioinnin pohjalta valitaan paras vaihtoehto (Nykänen ym., 2016). Olszakin ja Ziemban (2006) mukaan liiketoiminnan päätöksenteossa on aina ollut kyse erilaisten tietolähteiden hyödyntämisestä, tiedon hajaantuminen organisaatioissa on kuitenkin voinut aiheuttaa ongelmia viimeisimmän tiedon saatavuudessa päätöksentekijöille liiketoiminnan kaikilla eri alueilla. Tämän takia organisaatioiden tulisi kehittää sellaista järjestelmä infrastruktuuria, jonka avulla saadaan kokonaisvaltainen kuva liiketoiminnasta (Olszak & Ziemba, 2006). Useat teoriat ja käytännön kokemukset liiketoimintatietojen hallintajärjestelmien vastaavan tähän tarpeeseen (Olszak & Ziemba, 2006). Liiketoimintatietojen hallintajärjestelmät tukevat myös vahvasti rationaalisen valinnan kaltaista päätöksentekoa, sillä niiden avulla saadaan päättäjille tarkempaa, oleellisempaa ja ajankohtaisempaa tietoa. (Nykänen, ym 2016).

Liiketoimintatietojen hallintajärjestelmässä tietoa ovat erilaiset liiketoiminnan tilasta kertovat raportit (Chen ym., 2012). Näiden raporttien lisäksi tilastollisia- ja datanlouhintamenetelmien avulla voidaan luoda ennusteita, assosiaatio- ja regressioanalyysjä sekä havaita poikkeuksia eri liiketoiminnan osa-alueilta (Lim ym., 2013). Olszak ja Ziemba (2006) esittävät konkreettisempia esimerkkejä liiketoimintatietojen hallinnan prosessin tuottamien tietojen hyödyistä eri toimialoilla. Esimerkiksi vähittäiskaupan alalla asiakkaista saatuja tietoja voidaan hyödyntää markkinoinnissa, myyntiennusteita voidaan käyttää varastojenhallinnassa, jakelun ja logistiikan aikatauluttamisessa sekä tuotesuunnittelussa markkinasta saatujen tietojen pohjalta (Olszak & Ziemba, 2006). Verkkokaupalle ominaisia asiakasdatan hyödyntämisen keinoja käydään läpi tarkemmin tutkielman neljännessä luvussa.

Yritykset hyödyntävät liiketoimintatietojen hallintaa pääosin juuri päätöksenteon parantamiseen. Nykäsen ym. (2016) tutkimuksessa tarkasteltiin, miten ja miksi suomalaiset yritykset hyödyntävät liiketoimintatietojen hallintaa liiketoiminnassaan. Erityisesti tutkittiin liiketoimintatiedon hallintaprosessin toteutumista sekä myös sitä, miten syntyvää tietoa hyödynnetään päätöksenteossa. Tutkimuksen tuloksissa havaittiin, että tärkeimmät syyt, miksi yritykset hyödyntävät liiketoimintatietojen hallintaa ovat parempi liiketoiminnan ymmärrys, kehittynyt operatiivinen tehokkuus sekä parempi päätöksenteko. 76% prosenttia heidän tutkimukseensa vastanneista yrityksistä nostivat päätöksenteon parantamisen yhdeksi tärkeimmäksi syyksi liiketoimintatietojen hallinnan käyttöön (Nykänen ym., 2016).

Nykänen ym. (2016) tutkimuksessa yleisin tapa, jolla liiketoimintatietojen hallintaa hyödynnettiin päätöksenteossa, oli erilaisten liiketoiminnan tilasta kertovien raporttien tarkastelu, jota lähes kaikki kyselyyn vastanneet kertoivat hyödyntävänsä. Näiden raporttien avulla voidaan esimerkiksi arvioida yksittäisten työntekijöiden suoriutumista (Lim ym., 2013). Toinen yleinen hyödyntämistapa oli jatkuva liiketoimintaprossien valvonta ja kolmantena oli liiketoimintatietojen hallinnan avulla luotujen ennustavien analyysien hyödyntäminen toiminnan suunnittelussa. Tutkimuksessa 78% kyselyyn vastanneista kertoi päätöksenteon nopeutuneen ja 93% kertoi päätöksenteon laadun parantuneen liiketoimintatiedon hallinnan avulla (Nykänen ym., 2016). Voidaan siis päätellä, että liiketoimintatiedon hallinnan avulla yritykset voivat tehdä parempia päätöksiä nopeammin.

3 VERKKOKAUPAN DATA JA LIIKETOIMINTATIEDON HALLINTA

Verkkokaupan liiketoiminnassa käsitellään suuria määriä monen muotoista dataa, joka on peräisin useasta eri lähteestä (Akter & Wamba, 2016). Tämä data voidaan jakaa sen rakenteen perusteella kahteen osaan: strukturoitu ja strukturoimaton data (Akter & Wamba, 2016).

Strukturoitua dataa verkkokaupoissa on esimerkiksi yrityksen sisäisistä järjestelmistä kerääntyvä data, asiakkaiden tiedot, kuten nimi, ikä ja osoite sekä kaikki liiketoiminnallinen data kuten maksutapahtumat (Akter & Wamba, 2016). Strukturoimatonta dataa verkkokaupassa taas on esimerkiksi sivuston klikkaukset, käyttäjien ja hakujen lokitiedostot, käyttäjien luoma sisältö, kuvat, äänitiedostot ja tykkäykset. Klikkauksia kerätään asiakkaiden verkkokaupan käyttäytymisestä. Strukturoimattoman datan merkitys verkkokaupan liiketoiminnalla on hyvin suuri, sillä se pitää sisällään verkkokaupoista kerättävän käyttäjien seuraamisesta syntyvän datan lisäksi myös sosiaalisen median kautta kerääntyvän tykkäysten, linkkien, klikkausten ja sosiaalisen media sisällön datan, jonka merkitys ja määrä on kasvanut viime vuosien aikana. (Akter & Wamba, 2016).

Voidaan siis todeta, että verkkokaupan liiketoiminnan data on hyvin monimuotoista. Lisäksi kaikki nämä datan eri lähteet yhdistettynä, verkkokaupan liiketoiminnasta kerääntyvän datan määrä kasvaa myös nopeasti hyvin suureksi (Kohavi, Mason, Parekh, & Zheng, 2004). Näistä syystä massadatan (engl. big data) ja sen analysoinnin merkitys on kasvanut erityisesti verkkokaupan alalla viime vuosina (Akter & Wamba, 2016).

3.1 Massadata ja sen ominaispiirteet

Massadatan (engl. big data) suosio on myös yleisesti ollut nousussa useita vuosia ja erityisesti suuret verkkokaupan alan toimijat kuten Amazon ovat olleet pääroolissa suosion kasvattamisessa (Chen, Chiang & Storey, 2012). Massadatalle ei vielä kuitenkaan kirjallisuudessa ole yhtä yksiselitteistä tai vakiintunutta

määritelmää (Hu ym., 2014). Massadata pohjimmiltaan tarkoittaa kuitenkin enemmän kuin vain suurta määrää dataa ja massadatan määrittelemisiin vaaditaan muitakin ominaisuuksia kuin vain datan suuri määrä (Hu ym., 2014). Yleinen käytetty määritelmä on Laneyn 2001 määrittelemä kolmen V:n malli. Kolme V:tä tulee englannin kielen sanoista volyymi (engl. volume), monimuotoisuus (engl. variety) ja nopeus (engl. velocity) (Gandomi & Haider, 2015).

Volyyymillä tarkoitetaan datan suurta määrää, joka on massadatan yksi ominaisimpia tunnusmerkkejä ja yleensä ensimmäinen asia, joka nostetaan esille massadataa määriteltäessä (Gandomi & Haider, 2015). Esimerkiksi Walmart kerää joka tunti noin 2.5 petatavua (miljoona gigatavua) dataa (Gandomi & Haider, 2015).

Monimuotoisuus viittaa kerääntyvän datan useisiin eri lähteisiin ja eri muotoihin (Gandomi & Haider, 2015). Kerääntyvää dataa on lähtökohtaisesti kaikissa aiemmin mainituissa datan eri muodoissa, strukturoidussa, strukturoimattomassa ja osittain strukturoidussa (Gandomi & Haider, 2015). Kuitenkin vain noin 5 prosenttia kaikesta datasta on täysin strukturoidussa muodossa ja vaikka datan monimuotoisuus massadatan ominaisuutena on noussut vasta viime vuosina esille, ei datan monimuotoisuus itsessään ole uusi ilmiö. Sen sijaan teknologiat, jotka mahdollistavat monimuotoisen datan hyödyntämisen liiketoiminnassa, ovat uudempia innovaatioita (Gandomi & Haider, 2015).

Nopeus viittaa siihen nopeuteen, jolla dataa kerääntyy ja siihen, miten nopeasti dataa pitäisi pystyä analysoimaan ja tekemään toimenpiteitä analyysin pohjalta (Gandomi & Haider, 2015). Kuten aiemmin jo mainittu, useiden eri datan lähteitten takia, erityisesti verkkokaupan liiketoiminnasta kerääntyvä datan määrä kasvaa lyhyessäkin ajassa erittäin suureksi (Kohavi, Mason, Parekh, & Zheng, 2004).

Massadatan kolmen V:n mallia on laajennettu nykytutkimuksissa muun muassa viiden V:n malliksi, joka tuo massadatan määritelmään aikasempien ulottuvuuksien lisäksi arvon ja todenmukaisuuden (Akter & Wamba, 2016). Arvo viittaa massadatan avulla luodun tiedon arvoon ja sen mukana tuomiin informaationaaliin etuihin (Akter & Wamba, 2016). Todenmukaisuus viittaa massadatan eri lähteiden vaihtelevaan luotettavuuteen (Gandomi & Haider, 2015). Massadatan eri ulottuvuuksille ei ole yhteisiä tarkkoja rajoja, siispä massadatan käsite on suhteellinen ja riippuu organisaation koosta, toimialasta ja sijainnista. Massadatan määritelmän rajat myös kehittyvät jatkuvasti ajan myötä (Gandomi & Haider, 2015).

3.2 Massadata ja liiketoimintatiedon hallinta

Vaikka tässä tutkielmassa käytetty liiketoimintatiedon hallinnan määrittely pitää sisällään kaikki teknologiat ja metodit, joilla parannetaan päätöksentekoa, on hyvä huomioda, että perinteisten liiketoimintatietojen hallintajärjestelmien kyky varastoida ja analysoida dataa reaaliaikaisesti ei usein ole riittävä verkkokaupan

massadatan kaltaisten useiden monimutkaisten strukturoimattoman datan lähteiden soveltamiseen (Du ym., 2019). Aikaisemmin mainitut massadatan kolme keskeisintä ominaisuutta volyyymi, monimuotoisuus ja vauhti kukin vaikeuttavat massadatan hyödyntämistä omalla tavallaan (Hu ym., 2014). Monet tutkijat ovatkin väittäneet, että vaikka massadatalla on potentiaali yritysten suorituskykyä, on yritysten kohdattava monia erilaisia liiketoiminnan haasteita saadakseen massadatatista selkeää hyötyä (Akter & Wamba, 2016). Kuten aiemmin mainittiin, perinteisten liiketoimintatietojen hallintajärjestelmän datan lähteitä ovat sisäiset operatiiviset järjestelmät, kun taas massadatan lähteet ovat lähtökohtaisesti peräisin internetistä, esimerkiksi verkkosivuilta ja sosiaalisesta mediasta (Du ym., 2019; Baars & Kemper, 2008). Massadatan on myös sanottu olevan murroksellinen teknologia, joka tulee muokkaamaan liiketoimintatietojen hallintaa yleisesti (Fan, Lau & Zhao, 2015) ja Chen ym. (2012) käyttääkin massadataa hyödyntävästä liiketoimintatietojen hallinnasta käsitettä liiketoimintatietojen hallinta 2.0. Jotta perinteisten liiketoimintatietojen hallintajärjestelmien ja massadata teknologioiden hyödyt saadaan yrityksen käyttöön, täytyy yritysten siis suunnitella perinteiset liiketoimintatietojen hallintajärjestelmät ja prosessit uudelleen (Du ym., 2019).

Selkeällä liiketoiminta- ja IT-strategioiden linjauksella (engl. alignment), tarkoitetaan sitä, että massadata-analytiikan hankkeiden täytyy palvella liiketoiminnan strategiaa (Watson, 2014). Tämän takia näiden hankkeiden pitäisi lähtökohtaisesti olla liiketoiminnan tarpeisiin perustuvia (Watson, 2014).

Selkeä liiketoiminta- ja IT-strategioiden linjaus, faktoihin perustuva päätöksentekokulttuuri, vahva datainfrastruktuurin, oikeat analyttiset työkalut ja analytiikan käyttöön taitavat ihmiset voivat kuitenkin auttaa organisaatioita hyödyntämään massadataa liiketoimintatietojen hallintaprosessissaan (Watson, 2014).

Faktoihin perustuvalla päätöksentekokulttuurilla tarkoitetaan sitä, että kun massadata-analytiikan avulla saadaan tietoa liiketoiminnasta, täytyy päättäjien perustaa päätöksenteko näihin faktoihin (Watson, 2014). Päättäjät voivat edistää dataan pohjautuvaa päätöksentekokulttuuria kannustamalla tiedon jakamiseen, data-analytiikan koulutuksen saatavuuden lisääminen, tietopohjaisen päätöksenteon eduista tiedottaminen (Watson, 2014). Päättäjien täytyy myös ymmärtää, että kaikki työntekijät eivät tule omaksumaan tätä kulttuuria (Watson, 2014).

Datan saaminen päätöksenteon perustaksi vaatii myös oikeanlaista osaamista (Watson, 2014). Myös Akter, Wamba, Gunasekaran, Dubey ja Childe (2016) korostavat tarvittavan osaamisen merkitystä liiketoiminnallisen hyödyn saavuttamisessa dataan perustuvassa päätöksenteossa. Watson (2014) jakaa tarvittavat osaajat kolmeen eri rooliin; liiketoiminta-analyttikko, data-analyttikko ja data-tieteilijä. Liiketoiminta-analyttikoilla tulee olla laaja liiketoiminta-alue-tuntemus ja heidän tulee ymmärtää massadata-analytiikan mahdollisuudet arvonn luomisessa liiketoiminnalle (Watson, 2014). Datatieteilijöillä sen sijaan on oltava vahva tuntemus massadatatista ja sen eri tyypeistä sekä siitä, kuinka massadata voidaan varastoida (Watson, 2014). Heidän tehtävänä on löytää malleja ja suhteita, joita kukaan muu ei ole huomannut, ja muuttaa nämä löydöt käyttökelpoiseksi

tiedoksi, joka luo arvoa organisaatiolle (Watson, 2014). Data-analyytikoilla täytyy olla perusymmärrys sekä liiketoiminnasta, että massadatan käsittelystä (Watson, 2014).

Seuraavissa alakohdissa käsitellään vahvaan datainfrastruktuuriin liittyviä asioita kuten, millaisia teknologioita ja metodeja massadatan kerääminen ja käsittely vaatii verrattuna aikaisemmin käsitellyyn perinteiseen liiketoimintatiedon hallintaprosessiin.

3.2.1 Massadatan kerääminen

Koska massadataa kerätään useista eri lähteistä, keräämisprosessi alkaa, aivan kuten aikaisemmin käsitellyssä liiketoimintatiedon hallintaprosessissa, kaiken eri lähteiden datan keräämisestä yhteen paikkaan (Hu ym., 2014). Massadatan keräämisessä voidaan hyödyntää organisaation sisäisten järjestelmien lisäksi hakurobotteja, lokitiedostoja ja sensoreita (Hu ym., 2014). Hakurobotti käy läpi URL-osoitteita, lataa osoitteen sivulta dataa, tunnistaa kaikki osoitteen sivun URL-osoitteet ja lisää ne jonoon läpikäytäväksi (Hu ym., 2014). Hakurobotit ovat yleisiä tiedonkeruusovelluksia verkkosivustopohjaisille sovelluksille, kuten verkkohakukoneille ja verkkovälimuistille (Hu ym., 2014). Lokitiedostot ovat yksi yleisimmin käytetyistä tiedonkeruumenetelmistä, jolla tallennetaan tapahtumia määrättyyn tiedostomuotoon myöhempää analysointia varten (Hu ym., 2014). Esimerkiksi verkkokaupan verkkopalvelin tallentaa normaalisti kaikki verkkosivuston käyttäjän tekemät napsautukset, osumat, käyttöoikeudet ja muut attribuutit lokitiedostoon (Kohavi ym., 2004).

Koska massadatan määrä kasvaa jatkuvasti, massadataa hyödyntävien järjestelmien täytyy säilöä ja käsitellä massiivinen määrä dataa, kuitenkin toiminnallisuudet ja suorituskyky säilyttäen (Watson, 2014). Esimerkiksi Facebookin täytyy säilöä, käsitellä ja analysoida yli 30 petatavua pelkästään käyttäjien luomaa dataa (Hu ym., 2014). Massadatan suuren määrän takia, kerättävässä datassa on myös hyödyntämiskelvotonta dataa, jonka takia dataa pitää ennen tallentamista käsitellä sekä datan siirtämisen täytyy olla tehokasta, jotta järjestelmän suorituskyky säilyy mahdollisimman hyvänä (Hu ym., 2014). Massadatan säilömistä edeltävään käsittelyyn kuuluu aiemmin mainitun ETL-prosessin lisäksi datan puhdistaminen, joka koostuu viidestä eri vaiheesta; virhetyyppien määrittäminen, virheiden tunnistaminen datasta, virheiden korjaaminen, virheiden dokumentointi ja datan keräämismenetelmien muokkaaminen samojen virheiden välttämiseksi tulevaisuudessa (Hu ym., 2014).

3.2.2 Massadatan varastointi ja käsittely

Jotta kerätystä informaatiosta voidaan saada arvokasta tietoa, täytyy varastoinnin infrastruktuurin pystyä tallentamaan data pysyvästi ja luotettavasti, kyetä

skaalautumaan datan määrän kasvaessa, sekä tarjota soveltuva liittymä datan hakemiseen ja analysoimiseen (Hu ym., 2014). Tästä syystä Watson (2014) määrittelee vahvan datainfrastruktuurin yhdeksi kriittisimmistä tekijöistä massadatan hyödyntämiseen.

Viime vuosina suosiotaan kasvattaneet Hadoop ja Redis järjestelmät ovat mahdollisia ratkaisuja edellä mainittujen vaatimusten toteuttamiseksi (Niu, Ying, Yang, Bao, Sivaparthipan, 2021; Chen ym., 2012). Hadoop esimerkiksi on avoimen lähdekoodin ohjelmisto, joka tukee suurten tiedon määrien tallennusta ja käsittelyä. Kalliiseen omaan laitteistoon tietojen tallentamiseen sijaan Hadoop mahdollistaa suurten tietomäärien hajautetun käsittelyn suurissa palvelinklustereissa (Watson, 2014). Hadoopin skaalautuvuus, kustannustehokkuus ja joustavuus ovat sen suurimpia etuja massadatan hallintaan ja analysoimiseen. (Hu ym., 2014).

Watson (2014) ehdottaa myös massadatan tallentamiseen perinteisessä liiketoimintatietojen hallinnan prosessissa käytettyjen SQL-tietokantojen sijasta hieman uudempia noSQL-tietokantoja. NoSQL-tietokannat sopivat paremmin massadatan tallentamiseen, koska kuten todettu massadata on monimuotoista ja noSQL-tietokantoihin tallennetun data ei tarvitse olla aina yhdessä standardoidussa muodossa (Hu ym., 2014).

Hu ym. (2014) mukaan myös pilvipohjaisten palveluiden mahdollistama datan virtualisointi on yksi mahdollinen teknologia massadatan varastoinnin vaatimusten täyttämiseen. Tallennuksen virtualisointi tarkoittaa käytännössä useiden tallennuslaitteiden yhdistämistä yhdeksi laitteeksi (Hu ym. 2014). Pilvipohjaisten palveluiden voidaankin nykypäivänä katsoa olevan lähes välttämätön ratkaisu tehokkaan massadatan käsittelyn mahdollistamiseksi (Mohamed ym., 2020).

4 VERKKOKAUPAN DATA PÄÄTÖKSENTEOSSA

Aikaisemmissa kappaleissa verkkokaupan dataa käsiteltäessä massadatan merkityksen todettiin olevan suuri erityisesti nykyaikaisessa verkkokaupan liiketoiminnassa. Liiketoimintatiedon hallintaa sovellettaessa verkkokaupan liiketoimintaan, voidaankin prosessin tiedon analysointi vaiheessa verkkokaupan liiketoiminnassa ajatella olevan kyse massadata-analytiikasta. Ja kuten aiemmin mainittu vahvan datainfrastruktuurin, selkeän liiketoiminta- ja IT-strategioiden linjauksen, faktoihin perustuvan päätöksentekokulttuurin ja analytiikan käyttöön kykenevien osaajien lisäksi oikeat analyttiset työkalut auttavat luomaan arvokasta tietoa tästä massadatatista (Watson, 2014).

Tässä kappaleessa käsitellään, millaista päätöksenteon kannalta oleellista tietoa massadata-analytiikan eri menetelmien avulla voidaan verkkokaupan liiketoiminnasta saada.

4.1 Massadata-analytiikka verkkokaupan päätöksenteossa

Massadata-analytiikka voidaan jakaa kolmen erityyppiseen analytiikkaan sen tuottaman tiedon pohjalta; kuvaileva analytiikka, ennustava analytiikka ja ohjailtava analytiikka (Watson, 2014; Hu ym., 2014).

Kuvaileva analytiikka katsoo ajassa taaksepäin ja kertoo mitä on tapahtunut. Analyysi voidaan esittää esimerkiksi erilaisina raportteina (Hu ym., 2014). Watson (2014) vertaa kuvailevaa analytiikkaa perinteiseen liiketoimintatietojen hallintaan ja kuten Nykäsen ym. (2016) tutkimus osoitti, erilaisten raporttien tarkastelu oli yrityksissä yleisin tapa hyödyntää liiketoimintatietojen hallintaa.

Ennustava analytiikka keskittyy siihen mitä tulee tapahtumaan ja pyrkii sanan mukaisesti ennustamaan tulevaa (Hu ym., 2014). Ennustavan analyysin metodit ja algoritmit, kuten regressioanalyysi, koneoppiminen ja neuroverkot ovat olleet olemassa jo hetken aikaa, vasta viime vuosina erilaiset ohjelmistot ovat

helpottaneet ennustavan analyysin toteuttamista (Watson, 2014). Joissakin tutkimuksissa puhutaan myös tutkivasta analytiikasta, jossa suuresta määrästä dataa löydetään yhteyksiä eri asioiden väliltä, mutta Watsonin (2014) mukaan tutkiva analytiikka kuuluu ennustavan analytiikan termin alle.

Ohjaileva analytiikka sen sijaan pyrkii kertomaan mitä tulisi tehdä. Sen avulla pyritään etsimään optimaalisia ratkaisuja, usein rajallisten resurssien optimaaliseksi hyödyntämiseksi (Hu ym., 2014). Myös ohjailevaa analytiikkaa on tutkittu jo jonkin aikaa, mutta vasta viime vuosina sille on löydetty enemmän käytännön sovelluskohteita (Watson, 2014).

4.2 Massadata-analytiikan sovelluskohteet verkkokaupassa

Kaikkien edellä mainittujen massadata-analytiikan eri tyyppien avulla voidaan tuottaa arvokasta tietoa parantamaan päätöksentekoa ja erityisesti päätöksenteon laatua (Ghasemaghaei, Ebrahimi & Hassanein, 2018). Ghasemaghaei ym. (2018) tekemässä tutkimuksessa data-analytiikan vaikutuksista organisaatioiden päätöksentekoon selvisi, että datan määrän kasvaessa päätöksenteon laatu parani, sillä tuloksista saatiin luotettavampia, mutta päätöksenteon tehokkuus ei kuitenkaan enää noussut datan määrän kasvaessa. Tämä voi johtua siitä, että kuten aiemmin mainittu, massadatan kerääminen, käsittely ja analysointi vaatii paljon resursseja (Ghasemaghai ym., 2018). Seuraavissa kappaleissa käydään läpi konkreettisia esimerkkejä eri verkkokaupan liiketoiminnan osa-alueilta, miten massadata-analytiikan eri menetelmillä saatu tieto voi parantaa verkkokaupan liiketoiminnan päätöksentekoa.

4.2.1 Markkinointi

Markkinointi on yksi massadata-analytiikan keskeisimpiä sovelluskohteita verkkokaupan liiketoiminnassa (Akter & Wamba, 2016) ja massadata-analytiikan hyödyntämisen onkin todettu parantavan verkkokauppojen markkinointistrategiaa (Alrumiah & Hadwan, 2021). Verkkokaupan dataa voidaan hyödyntää viidellä eri markkinoinnin osa-alueilla: ihmiset, tuote, viestintä, myyntipaikka ja hinnoittelu (Fan, Lau & Zhao, 2015). Seuraavaksi tarkastellaan, millaista päätöksenteon kannalta oleellista tietoa miten markkinoinnista syntyvä dataa voidaan hyödyntää näiden eri osa-alueiden päätöksissä.

Ihmisiin eli asiakkaisiin liittyvää dataa voidaan hyödyntää muun muassa asiakkaiden segmentoinnilla ja profiloinnilla (Fan ym., 2015). Segmentoinnilla voidaan löytää asiakkaista erilaisia ryhmiä, jotka esimerkiksi jakavat samankaltaisia mielenkiinnon kohteita (Fan ym., 2015) Asiakkaita voidaan jaotella myös ostokäyttäytymisen mukaan (Du, Liu & Zhang, 2019). Du ym. (2019) esittelee tutkimuksessaan algoritmin, jolla asiakkaat jaetaan kolmeen ryhmään ostosten kokonaissumman, ostotiheyden ja viimeisimmästä ostoksesta kuluneen ajan mukaan. Tutkimuksessa algoritmia sovellettiin kiinalaiseen verkkokauppayritykseen, joka syntyneen tiedon pohjalta löysi lojaaleimmat asiakkaansa ja pystyi

keskittämään markkinoinnin resurssejaan heidän tärkeimmille asiakkailleen (Du ym., 2019). Asiakkaiden segmentointia tarkempaa tietoa voidaan saada asiakkaiden profiloinnilla (Fan ym., 2015). Profiloinnilla voidaan tunnistaa yksittäisten asiakkaiden tarpeita ja tämän pohjalta voidaan tarjota asiakkaille, juuri heidän tarpeisiinsa sopivaa tuotetta (Fan ym., 2015).

Asiakkaiden dataa käytetään myös erilaisten suosittelujärjestelmissä (Schafer, Konstan & Riedl, 1999; Alrumiah & Hadwan, 2021). Suosittelujärjestelmä eivät suoranaisesti liity organisaation päätöksentekoon, mutta niiden merkitys nykypäivän verkkokaupan markkinoinnissa on niin suuri, että niitä ei voi jättää mainitsematta. Suosittelujärjestelmät auttavat kasvattamaan verkkokaupan myyntiä konvertoimalla kaupan selaajia ostajiksi, tarjoamalla lisämyyntiä ostajille ja paremman ostokokemuksen avulla nostamalla asiakkaiden uskollisuutta (Schafer ym., 1999).

Tuotteesta ja erityisesti tuotteen suosiosta ja maineesta voidaan saada asiakkailta kerättävistä tuotearvosteluista (Alrumiah & Hadwan, 2021) erilaisilla menetelmillä louhimalla tietoa internetin sisällöistä (Fan ym., 2015). Esimerkiksi tuotteeseen liittyvien sosiaalisen median tekstejä keräämällä voidaan saada melko kattava yleiskuvaus tuotteen tämänhetkisestä suosiosta (Fan ym., 2015). Sosiaalisen median analytiikan merkitys on kasvussa, sillä kuluttajat luottavat ostopäätöksissään yhä enemmän muiden kuluttajien suositteluihin (Akter & Wamba, 2016). Tätä tietoa voidaan käyttää tuotteiden hallinnassa sekä muilla markkinoinnin osa-alueilla hyödyksi (Fan ym., 2015).

Markkinointiviestinnästä syntyvää dataa, kuten mainokset, alennuskuponkien käyttö ja alennusmyynnit, voidaan hyödyntää massadata-analytiikassa usein muiden osa-alueiden dataa yhdistämällä (Fan ym., 2015). Analysoimalla asiakkaiden dataa- ja mainosten dataa voidaan ymmärtää millä tavalla tietyn tyyppiset mainokset vaikuttavat tietynlaisiin asiakkaisiin ja täten saadaan paremmin kohdennettua mainoksia oikeille ihmisille (Fan ym., 2015). Kohdentamista voidaan kehittää myös paikkatietojen avulla (Fan ym., 2015). Asiakasdataa, markkinointiviestinnän ja myyntidataa yhdessä analysoimalla voidaan tunnistaa parhaat maantieteelliset alueet tietynlaiselle markkinoinnille. Sijaintiin perustuvien markkinointistrategioiden onkin todettu olevan erityisen tehokkaita (Fan ym., 2015).

Hinnoittelussa dataa voidaan hyödyntää dynaamisen hinnoittelustrategian perustana (Wu, Li & Xu, 2014). Dynaamisella hinnoittelulla tarkoitetaan hinnoittelustrategioita, joissa tuotteiden hinnat vaihtelevat joko ajan tai asiakkaiden mukaan (Wu ym., 2014). Tällaisen hinnoittelustrategian on todettu olevan jopa 33% voitollisempi kuin perinteisen staattisen hinnoittelustrategian (Wu ym., 2014). Esimerkiksi Ghosen ja Sundarajan (2006) tutkimuksessa Amazonin verkkokaupan julkisesti saatavilla olevaa dataa analysoimalla voitiin arvioida esimerkkinä käytetyn tuotteen hintajoustavuutta, muuttuvia kustannuksia ja hinnoittelupäätösten optimaalisuutta.

4.2.2 Tuotantoketjun hallinta

Toinen merkittävä massadata-analytiikan sovelluskohde on tuotantoketjun hallinta (Alrumiah & Hadwan, 2016). Tuotantoketjun hallinnalla tarkoitetaan valmistajien, myyjien, jakelijoiden ja asiakkaiden hallintaa ja koordinointia (Li, Chi, Hao & Yu, 2018). Analytiikan avulla voidaan arvioida tarvittavaa inventaarion määrää tulevaisuuden kysyntää ennustamalla (Wang & Huang, 2022; Li ym., 2018) ja parantaa toimitusketjun ja logistiikan läpinäkyvyyttä (Davenport, 2006; Li ym., 2018).

Ennustava analytiikka voi antaa tarkkoja tulevaisuuden kysynnän arvioita (Wang & Huang, 2022). Wang ja Huang (2022) esimerkiksi esittelivät tutkimuksessaan järjestelmän, jolla päätöksentekijät pystyvät tarkasti ennustamaan tulevaisuuden kysyntää. Tutkimuksessa käytettiin verkkokaupan dataa ennustamaan tietyn T-paitamallin myyntiä. Järjestelmän tuottamien myyntiennusteiden absoluuttinen prosenttipoikkeama oli vain 5,6 prosenttiyksikköä. Myyntiennusteista saatavaa tietoa voidaan hyödyntää dynaamisessa tuotantoketjun hallinnassa (Li ym., 2018). Dynaamisen tuotantoketjun hallinnan on osoitettu parantavan tuotantoketjun tehokkuutta ja näin parantavan koko yrityksen kilpailukykyä (Lim ym., 2018)

Datan avulla voidaan saada myös läpinäkyvyyttä yrityksen logistiikkaan ja toimitusketjuun. Täten saadaan selvitettyä mahdollisia ongelmakohtia sekä myös ratkottua niitä (Davenport, 2007).

4.2.3 Turvallisuus ja petoksien havaitseminen

Verkkokauppojen maksamisessa pankkikortit ovat avainasemassa. Pankkikorttien käyttö onkin nykyään verkkokaupan liiketoiminnassa ydinasemassa, sillä ne mahdollistavat tavaroiden ja palveluiden ostamisen reaaliajassa kaikkialta maailmasta ja käyttämällä mitä tahansa päätelaitetta (Carta, Fenu, Reforgiato & Saia, 2019). Pankkikorttien käyttöön liittyy kuitenkin riskejä, jotka voivat aiheuttaa esimerkiksi asiakkaiden luottokorttien arkaluontoisten tietojen varastamisen (Carta ym., 2019). Euroopan maksuneuvoston tuore raportti osoittaa, että tietty prosentiosuus Internetin sähköisistä maksuista liittyy petokseen. Massadata-analytiikan avulla transaktioihin liittyvän data voidaan reaaliajassa yhdistää asiakkaan ostohistoriaan, lokitietoihin ja mobiililaitteesta saatavaan maantieteelliseen sijaintiin. Esimerkiksi Visa on kehittänyt suuren massadata-analytiikkaan perustuvan petostentunnistusjärjestelmän, joka säästää arviolta jopa 2 miljardia dollaria petoksiin liittyviä tappiota vuosittain (Akter & Wamba, 2016).

5 YHTEENVETO

Verkkokaupan liiketoiminnalle oleellisen datan määrä kasvaa jatkuvasti ja verkkokaupan datasta puhuttaessa voidaankin siis puhua massadatasta. Massadatalle ei kuitenkaan ole yksiselitteistä määritelmää, mutta sen ominaisuuksina voidaan nähdä datan suuri määrä, datan suuri vaihtelevuus eli data voi olla monissa eri muodoissa sekä datan kerääntymisen suuri nopeus. Tutkimuksen tavoitteena oli kirjallisuuskatsauksen avulla löytää vastaus kysymykseen: Miten verkkokaupan dataa voidaan hyödyntää päätöksenteossa? Vastausta haettiin kolmen tutkimuskysymyksen avulla:

- Miten liiketoimintatiedon hallinta voi parantaa päätöksentekoa yrityksissä?
- Millaista dataa verkkokaupan liiketoiminnasta kertyy ja miten sitä saadaan hyödynnettyä päätöksenteossa?
- Miten massadata-analytiikka voi kehittää verkkokaupan liiketoimintaa?

Ensimmäisen tutkimuskysymyksen tarkastelussa tultiin aineiston pohjalta tulokseen, että liiketoimintatiedon hallinnan tuottamat raportit voivat parantaa yrityksen päätöksenteon laatua sekä nopeuttaa päätöksentekoprosessia.

Toisen tutkimuskysymyksen kohdalla tultiin tulokseen, että verkkokaupan liiketoiminnasta voidaan kerätä suuria määriä monimuotoista massadataa. Massadataa kerätään useista eri lähteistä, kuten asiakkaiden käyttäytymisestä verkkosivustolla ja sosiaalisesta mediasta. Tästä datasta voidaan saada verkkokaupan liiketoiminnan päätöksenteon kannalta arvokasta tietoa liiketoimintatiedonhallintaa hyödyntämällä. Todettiin kuitenkin, että massadatan hyödyntämiseen vaaditaan tiettyjä menetelmiä ja teknologioita, joita ei ole perinteisissä liiketoimintatiedon hallintajärjestelmissä. Jotta massadatasta saadaan tuotettua arvokasta tietoa, täytyy varastoinnin infrastruktuurin pystyä tallentamaan suuri määrä dataa pysyvästi ja luotettavasti, ja datan analysoimiseen käytettävien järjestelmien tulee pystyä käsittelemään tämä data hyvin nopeasti.

Kolmanteen tutkimuskysymykseen löydettiin muun muassa seuraavia tuloksia. Aineiston pohjalta todettiin, että erityisesti markkinoinnin päätöksentekoon saadaan massadata-analytiikan avulla tärkeää tietoa. Myös verkkokaupan tuotantoketjun hallintaa sekä petosten havaitsemista pystytään tehostamaan massadata-analytiikan avulla.

Tutkielman luotettavuutta pyrittiin nostamaan käyttämällä korkealaatuisia, vertaisarvioituja lähteitä. Tutkimuksen kulkua on pyritty kuvaamaan mahdollisimman tarkasti ja tutkimusmenetelmänä käytettiin yleisesti tunnettua tutkimusmenetelmää, kuvailevaa kirjallisuuskatsausta.

Jotta tutkielman aihe ei kasvaisi liian laajaksi, aihetta rajattiin johdannossa kuvatulla tavalla. Esimerkiksi tietoturva ja tietosuojalaki jätettiin tarkastelusta pääosin pois. Molemmista näkökulmista voisikin saada hyviä jatkotutkimuksia.

LÄHTEET

- Akter, S., & Wamba, S. F. (2016). Big data analytics in E-commerce: A systematic review and agenda for future research. *Electronic Markets*, 26(2), 173–194. <https://doi.org/10.1007/s12525-016-0219-0>. Viitattu 1.11.2022.
- Akter, S., Wamba, S. F., Gunasekaran, A., Dubey, R., & Childe, S. J. (2016). How to improve firm performance using big data analytics capability and business strategy alignment? *International Journal of Production Economics*, 182, 113–131. <https://doi.org/10.1016/j.ijpe.2016.08.018>. Viitattu 1.11.2022.
- Alrumiah, S. S., & Hadwan, M. (2021). Implementing Big Data Analytics in E-Commerce: Vendor and Customer View. *IEEE Access*, 9, 37281–37286. <https://doi.org/10.1109/ACCESS.2021.3063615>. Viitattu 1.11.2022.
- Baars, H., & Kemper, H.-G. (2008). Management Support with Structured and Unstructured Data – An Integrated Business Intelligence Framework. *Information Systems Management*, 25(2), 132–148. <https://doi.org/10.1080/10580530801941058>. Viitattu 1.11.2022.
- Carta, S., Fenu, G., Reforgiato Recupero, D., & Saia, R. (2019). Fraud detection for E-commerce transactions by employing a prudential Multiple Consensus model. *Journal of Information Security and Applications*, 46, 13–22. <https://doi.org/10.1016/j.jisa.2019.02.007>. Viitattu 1.11.2022.
- Chaffey, D. (2007). E-business and E-commerce Management: Strategy, Implementation and Practice. *Pearson Education*.
- Chaudhuri, S., Dayal, U., & Narasayya, V. (2011). An overview of business intelligence technology. *Communications of the ACM*, 54(8), 88–98. <https://doi.org/10.1145/1978542.1978562>. Viitattu 1.11.2022.
- Chen, Chiang, & Storey. (2012). Business Intelligence and Analytics: From Big Data to Big Impact. *MIS Quarterly*, 36(4), 1165. <https://doi.org/10.2307/41703503>. Viitattu 1.11.2022.
- Davenport, T. H. (2006). Competing on Analytics. *Harvard Business Review*, 84, 98–107.
- Du, X., Liu, B., & Zhang, J. (2019). Application of Business Intelligence Based on Big Data in E-commerce Data Analysis. *Journal of Physics: Conference Series*, 1395(1). <https://doi.org/10.1088/1742-6596/1395/1/012011>. Viitattu 1.11.2022.
- Fan, S., Lau, R. Y. K., & Zhao, J. L. (2015). Demystifying Big Data Analytics for Business Intelligence Through the Lens of Marketing Mix. *Big Data Research*, 2(1), 28–32. <https://doi.org/10.1016/j.bdr.2015.02.006>. Viitattu 1.11.2022.

- Ferreira, T., Pedrosa, I., & Bernardino, J. (2017). Business Intelligence for E-commerce: Survey and Research Directions. *Advances in Intelligent Systems and Computing*. https://doi.org/10.1007/978-3-319-56535-4_22. Viitattu 1.11.2022.
- Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, 35(2), 137-144.
- Gartner, 2022a "Incorporating the Skills and Capabilities Your D&A Organization Needs" haettu osoitteesta: <https://emtemp.gcom.cloud/ngw/globalassets/en/publications/documents/incorporating-the-skills-and-capabilities-your-danda-organization-needs-ebook.pdf> . Viitattu 2.10.2022
- Gartner, 2022b "Software Market Insights: Business Intelligence (BI) and Data Analytics" haettu osoitteesta: <https://emtemp.gcom.cloud/ngw/globalassets/en/digital-markets/documents/software-market-guide-bi-data-analytics.pdf>. Viitattu 19.10.2022
- Ghasemaghahi, M., Ebrahimi, S., & Hassanein, K. (2018). Data analytics competency for improving firm decision making performance. *The Journal of Strategic Information Systems*, 27(1), 101-113. <https://doi.org/10.1016/j.jsis.2017.10.001>. Viitattu 1.11.2022.
- Ghose, A., & Sundararajan, A. (2006). Evaluating Pricing Strategy Using e-Commerce Data: Evidence and Estimation Challenges. *Statistical Science*, 21(2), 131-142. <https://doi.org/10.1214/088342306000000187>. Viitattu 1.11.2022.
- Hu, H., Wen, Y., Chua, T.-S., & Li, X. (2014). Toward Scalable Systems for Big Data Analytics: A Technology Tutorial. *IEEE Access*, 2, 652-687. <https://doi.org/10.1109/ACCESS.2014.2332453>. Viitattu 1.11.2022.
- Kohavi, R., Mason, L., Parekh, R., & Zheng, Z. (2004). Lessons and Challenges from Mining Retail E-Commerce Data. *Machine Learning*, 57(1-2), 83-113. <https://doi.org/10.1023/B:MACH.0000035473.11134.83>. Viitattu 1.11.2022.
- Kalakota, R., & Whinston, A. B. (1997). *Electronic Commerce: A Manager's Guide*. Addison-Wesley Professional.
- Li, L., Chi, T., Hao, T., & Yu, T. (2018). Customer demand analysis of the electronic commerce supply chain using Big Data. *Annals of Operations Research*, 268(1), 113-128. <https://doi.org/10.1007/s10479-016-2342-x>. Viitattu 1.11.2022.
- Li Zeng, Xu, L., Shi, Z., Wang, M., & Wu, W. (2006). Techniques, Process, and Enterprise Solutions of Business Intelligence. *2006 IEEE International Conference on Systems, Man and Cybernetics*, 4722-4726. <https://doi.org/10.1109/ICSMC.2006.385050>. Viitattu 1.11.2022.

- Lim, E.-P., Chen, H., & Chen, G. (2013). Business Intelligence and Analytics: Research Directions. *ACM Transactions on Management Information Systems*, 3(4), 1–10. <https://doi.org/10.1145/2407740.2407741>. Viitattu 1.11.2022.
- Maity, M., & Dass, M. (2014). Consumer decision-making across modern and traditional channels: E-commerce, m-commerce, in-store. *Decision Support Systems*, 61, 34–46. <https://doi.org/10.1016/j.dss.2014.01.008>. Viitattu 1.11.2022.
- Mohamed, A., Najafabadi, M. K., Wah, Y. B., Zaman, E. A. K., & Maskat, R. (2020). The state of the art and taxonomy of big data analytics: View from new big data framework. *Artificial Intelligence Review*, 53(2), 989–1037. <https://doi.org/10.1007/s10462-019-09685-9>. Viitattu 1.11.2022.
- Niranjnamurthy, M., Kavyashree, N., Jagannath, S., & Chahar, D. (2013). Analysis of e-commerce and m-commerce: advantages, limitations and security issues. *International Journal of Advanced Research in Computer and Communication Engineering*, 2(6), 2360-2370.
- Niu, Y., Ying, L., Yang, J., Bao, M., & Sivaparthipan, C. B. (2021). Organizational business intelligence and decision making using big data analytics. *Information Processing & Management*, 58(6), 102725. <https://doi.org/10.1016/j.ipm.2021.102725>. Viitattu 1.11.2022.
- Nykänen, E., Järvenpää, M., & Teittinen, H. (2016). Business intelligence in decision making in Finnish enterprises. *Nordic Journal of Business*, 65(2), 24-44. http://njb.fi/wp-content/uploads/2016/09/Nykanen_Jarvenpaa_Teittinen.pdf . Viitattu 1.11.2022.
- Olszak C. M., Ziemba E. (2006). Business Intelligence Systems in the Holistic Infrastructure Development Supporting Decision-Making in Organisations. *Interdisciplinary Journal of Information, Knowledge, and Management*, 1, 47-58.
- Pirttimäki, V. (2007). Business intelligence as a managerial tool in large finnish companies. Väitöskirja, Tampere University of Technology.
- Salminen, A. 2011. Mikä kirjallisuuskatsaus? Johdatus kirjallisuuskatsauksen tyypeihin ja hallintotieteellisiin
- Schafer, J. B., Konstan, J., & Riedl, J. (1999). *Recommender Systems in E-Commerce*. 9.
- Watson, H. J. (2014). Tutorial: Big Data Analytics: Concepts, Technologies, and Applications. *Communications of the Association for Information Systems*, 34. <https://doi.org/10.17705/1CAIS.03465>. Viitattu 1.11.2022.
- Watson, H. J. (2009). Tutorial: Business Intelligence – Past, Present, and Future. *Communications of the Association for Information Systems*, 25. <https://doi.org/10.17705/1CAIS.02539> . Viitattu 1.11.2022.

Wu, J., Li, L., & Xu, L. D. (2014). A randomized pricing decision support system in electronic commerce. *Decision Support Systems*, 58, 43–52.
<https://doi.org/10.1016/j.dss.2013.01.015>. Viitattu 1.11.2022.