



Koneoppiminen sosiaalisen median datan hyödyntämisessä johdon laskentatoimessa

Lappeenrannan–Lahden teknillinen yliopisto LUT

Tuotantotalouden kandidaatintutkielma

2023

Valtteri Ahola

Tarkastaja: Yliopisto-opettaja Leena Tynninen

TIIVISTELMÄ

Lappeenrannan–Lahden teknillinen yliopisto LUT

LUT Teknis-luonnontieteellinen

Tuotantotalous

Valtteri Ahola

Koneoppiminen sosiaalisen median datan hyödyntämisessä johdon laskentatoimessa

Machine learning in utilizing social media data in management accounting

Tuotantotalouden kandidaatintyö

2023

46 sivua, 5 kuvaa ja 1 taulukko

Tarkastaja: Yliopisto-opettaja Leena Tynninen

Avainsanat: johdon laskentatoimi, sosiaalinen media, data-analytiikka, massadata, koneoppiminen, tekstilouhinta, päätöksenteko

Keywords: management accounting, social media, data analytics, big data, machine learning, text mining, decision-making

Tämä kirjallisuuskatsauksena toteutettu kandidaatintyö tarkastelee, kuinka sosiaalisen median dataa voitaisiin hyödyntää johdon laskentatoimen tehtävissä koneoppimisen avulla. Tavoitteena on selvittää, miten sosiaalisen median palveluista kerättyä tekstidataa voisi hyödyntää johdon laskentatoimen eri osa-alueilla ja voiko koneoppimista hyödyntää tämän datan prosessoinnissa. Työn tutkimuskysymyksiin vastattiin vertailemalla teoreettisia artikkeleita ja yritys esimerkkejä aiheeseen liittyen.

Tutkimuksessa havaittiin, että sosiaalisen median dataa on mahdollista hyödyntää teorian perusteella kaikilla johdon laskentatoimen osa-alueilla, erityisesti budjetoinnissa ja strategiaseurannassa. Käytännön tasolle sosiaalisen median datan hyödyntäminen johdon laskentatoimen tehtävissä ei ole vielä edennyt laajasti lukuun ottamatta budjetointia ja siinä myynnin ennustamista, jossa sitä on jo jonkin verran hyödynnetty. Koneoppimista oli käytetty teoriaa tukevissa kokeellisissa tutkimuksissa sekä yhdessä yritysmerkissä, joissa datan prosessoinnissa oli hyödynnetty tekstilouhintaprosessia ja koneoppimisalgoritmeja. Johtopäätöksenä voidaan todeta, että sosiaalisen median datan merkitys johdon laskentatoimen tehtävien kannalta on merkittävä mutta hyödyntämätöntä potentiaalia on vielä runsaasti. Koneoppimisen rooli sosiaalisen median datan prosessoinnissa on suuri ja mahdollistaa datan luokittelun sekä analysoinnin tukemaan johdon laskentatoimen tehtäviä.

Sisällysluettelo

Tiivistelmä

1	Johdanto.....	3
1.1	Työn tausta ja tavoitteet	3
1.2	Tutkimusmenetelmä ja aiheen rajaus	4
1.3	Työn rakenne.....	5
2	Sosiaalinen media ja koneoppiminen	6
2.1	Data-analyysi.....	6
2.2	Koneoppiminen analyysin tukena	12
3	Johdon laskentatoimi organisaatioissa	17
3.1	Lähestymistapoja johdon laskentatoimeen	17
3.2	Ohjaus	20
3.3	Päätöksenteko.....	22
4	Sosiaalisen median data ja johdon laskentatoimi	24
4.1	Teoreettinen viitekehys	24
4.2	Case-esimerkkejä	30
4.3	Havaintoja teoriasta ja caseista	33
5	Johtopäätökset	37
	Lähteet	40

1 Johdanto

Yritykset ja organisaatiot hakevat lähtökohtaisesti kasvua liiketoimintoihinsa, tehokkuutta ja parempaa kannattavuutta. Johdon laskentatoimella on tässä merkittävä rooli. Johdon laskentatoimen päärooleihin kuuluu kerätyn datan prosessoimista ja analysointia yrityksen sisäisille sidosryhmille, erityisesti johdolle. Tuotetun tiedon pääasiallinen tarkoitus on päätöksenteossa avustaminen. (Ikäheimo, Malmi & Walde 2019, 127–128) Digitalisaation kehitys viime vuosikymmeninä on vaikuttanut merkittävästi johdon laskentatoimen työhön (Holtkemper 2020, 11).

1.1 Työn tausta ja tavoitteet

Yksi suurimmista muutoksista johdon laskentatoimen työssä on ollut saatavilla olevan datan määrän eksponentiaalinen kasvu. Tämä data on avannut uusia tapoja organisaatioille analysoida omaa toimintaansa sekä kehittää liiketoimintaa. Enää yritykset eivät keskity talousjohtamisessa pelkkään menneisyyteen pohjautuvaan taloudelliseen dataan, kuten edellisten vuosien tilinpäätöstietoihin, vaan nykyään tämän rinnalle on tullut yhä enemmän niin sanottua ei-taloudellista dataa. (Holtkemper 2020, 11) Yksi esimerkki saatavilla olevan datan määrän kasvusta on niin kutsuttu massadata (*big data*). Tässä työssä käytämme tästä englanninkielistä ilmausta eli big data. Big datalla viitataan suuriin digitaalisiin tietovarantoihin, jotka kasvavat nopealla vauhdilla (Schintler & McNeely 2022, 75). Merkittäviä big datan lähteitä ovat muun muassa maailmanlaajuisten tietoverkkojen kehityksen seurauksena syntyneet sosiaalisen median palvelut (Oracle, 2023). Saito ja Gupta (2022) toteavat sosiaalisen median potentiaalinen yritysten kannalta olevan merkittävä, sillä se tarjoaa ainutlaatuista dataa asiakkaiden mielipiteistä ja käytöksestä, jota yritysten itse tuottama data ei sisällä. Tästä syystä sosiaalisesta mediasta kerättyä dataa on mahdollista käyttää myös johdon laskentatoimen työn apuna esimerkiksi budjetoinnin ja optimaalisen hinnoittelun tukena (Appelbaum, Kogan, Vasarhelyi & Yan 2017, 32; Saito & Gupta 2022, 1).

Sosiaalisen median datalle ominaista on sen jäsentämättömyys ja suuri määrä, jotka aiheuttavat sen, että merkittävä osa datasta on hyödytöntä. Suurimpana haasteena datan hyödyntämisessä onkin prosessoida se muotoon, jossa sillä on mahdollista tuottaa informatiivista

lisäarvoa yrityksille. (Goncalves & Cornelius Smith 2018, 75–76) Tähän apuna voidaan käyttää tekoälyä ja sen koneoppimisen osa-aluetta (Hassan, Ahamed & Ahmad 2022, 240).

Työn tavoitteena on selvittää, kuinka sosiaalisesta mediasta saatavaa dataa olisi mahdollista hyödyntää johdon laskentatoimen tehtävissä ja voiko koneoppimista hyödyntää datan prosessoinnissa. Koska johdon laskentatoimi on käsitteenä laaja, määritellään erikseen myös johdon laskentatoimi ja sen osa-alueet. Datan hyödyntämisprosessiin kuuluu olennaisena osana myös datan keräys ja prosessointi analysoitavaan muotoon (Khan, Jan & Farman 2019, 16). Tästä syystä työssä käsitellään myös datan keräys- ja käsittelyprosessia sekä määritellään datan lähde eli sosiaalinen media käsitteenä. Työn tavoitteeseen pyritään vastaamaan seuraavilla tutkimuskysymyksillä:

1. *Minkälaista dataa sosiaalinen media tarjoaa johdon laskentatoimeen?*
2. *Kuinka koneoppimista voidaan hyödyntää datan prosessoinnissa?*
3. *Millä tavoin yritykset hyödyntävät sosiaalisen median dataa johdon laskentatoimessaan?*

Digitalisaation ja analysoitavan datan määrän kasvun vaikutusta johdon laskentatoimen tehtävissä on jo tutkittu melko kattavasti. Aiempi aiheen käsittely oli kuitenkin melko laajaa ja tutkimusta liittyen nimenomaan sosiaalisen median rooliin johdon laskentatoimessa ei löytynyt kattavasti, joka oli merkittävä motiivi tämän aiheen käsittelyyn. Myös sosiaalisen median jatkuva kasvu luo siitä mielenkiintoisen datan lähteen, jota on syytä tarkastella lähemmin johdon laskentatoimen kannalta (Zachlod, Samuel, Ochsner & Werthmüller 2022).

1.2 Tutkimusmenetelmä ja aiheen raja

Työ toteutetaan kirjallisuuskatsauksena. Kirjallisuuskatsaus tarkoittaa metodia, jossa olemassa olevia tutkimuksia kootaan yhteen ja luodaan näin perustaa uudelle tutkimustulokselle (Salminen 2011, 3). Työ pohjautuu aiheeseen liittyvien tieteellisten artikkeleiden ja teoriakirjallisuuden ympärille. Lisäksi työssä viitataan myös muihin aiheeseen liittyviin nettiaartikkeleihin sekä neljää case-yritystä, Beyond The Arc, Zeus, Art ja Utility, koskeviin yritys-esimerkkeihin. Case-yrityksistä Artin, Zeuksen ja Utilityn nimet oli lähdeaineistoissa muutettu luottamuksellisuuteen vedoten. Teoreettisten lähteiden etsimisessä on hyödynnetty lähinnä LUT Primo sekä Google Scholar hakupalveluita. Lähteiden etsinnässä hyödynnetyt

hakusanat liittyvät pääasiassa johdon laskentatoimen, big datan, data-analytiikan ja koneoppimisen ympärille ja niitä on pyritty keräämään mahdollisimman laajasti ympäri maailmaa. Johdon laskentatoimeen liittyvässä teoriakirjallisuudessa on käytetty sekä suomenkielistä, että englanninkielistä kirjallisuutta, jotta käsite voidaan määritellä mahdollisimman kokonaisvaltaisesti. Lähteiden relevanttiuden varmistamiseksi työhön kerättyjen tieteellisten artikkeleiden on pyritty olemaan vuodelta 2017 tai uudempia. Työn neljännen luvun teoriakirjoitusten ja yritysesimerkkien vertailua varten toteutettiin artikkeleiden etsintää avainsanapareilla ”*management accounting and social media data*” sekä ”*social media data in financial management*”. Haun tuloksella löytyi yhteensä 13 tieteellistä artikkelia ja yritysesimerkkiä, joita luvussa referoidaan.

Työ on rajattu käsittelemään nimenomaan sosiaalista mediaa datan lähteenä ja kuinka dataa voisi hyödyntää johdon laskentatoimen tehtävissä. Saatavilla olevasta datasta keskitytään tekstipohjaiseen dataan. Sosiaalinen median alustoista työssä käsitellään palveluita, joissa merkittävässä osassa on käyttäjien sisällön luonti tekstipohjaisesti sekä sosiaalinen verkostoituminen. Datan keräys toteutetaan koneoppimista hyödyntämällä.

1.3 Työn rakenne

Työ koostuu viidestä pääluvusta. Ensimmäinen luku on johdanto, jossa luodaan pohjaa työhön alustamalla aiheen käsittelyä sekä kertomalla tavoitteet, tutkimusmenetelmät, aiheen rajaus ja rakenne. Toinen luku käsittelee sosiaalista mediaa datan lähteenä sekä koneoppimista. Luvussa määritellään sosiaalinen media sekä minkälaista informatiivista dataa se sisältää organisaatioiden kannalta. Luvussa määritellään myös koneoppiminen ja tuodaan esiin sen sovellusalue, joka voisi olla datan prosessoinnin kannalta relevantti. Kolmannessa luvussa määritellään johdon laskentatoimi käsitteenä ja tarkastellaan sen eri alueita organisaatioiden toiminnan kannalta. Luku on jaettu alakappaleisiin johdon laskentatoimen eri osa-alueiden mukaan. Neljäs luku käsittelee datan hyödyntämistä teorian ja yritysesimerkkien avulla sekä vertailee näistä tehtyjä havaintoja. Tulokset sosiaalisen median datan hyödyntämisestä johdon laskentatoimen tehtävissä kootaan taulukkoon, joka on jaoteltu johdon laskentatoimen eri osa-alueiden perusteella. Viimeinen kappale käsittelee johtopäätökset, jossa vastataan päätutkimuskysymyksiin ja tuodaan esiin omaa pohdintaa liittyen tehtyihin havaintoihin.

2 Sosiaalinen media ja koneoppiminen

Luvussa käsitellään sosiaalista mediaa datan lähteenä, sosiaalisen median data-analyysiprosessia sekä koneoppimista datan prosessointikeinona. Tarkoituksena on, että lukija ymmärtää luvun luettuaan, mitä dataa sosiaalisesta mediasta on mahdollista saada, kuinka sosiaalisen median data-analyysiprosessi etenee, mitä koneoppiminen tarkoittaa ja kuinka sitä voidaan käytännössä hyödyntää sosiaalisen median datan prosessoinnissa.

2.1 Data-analyysi

Sosiaalinen media tarkoittaa kokoelmaa erilaisia internetsivuja ja applikaatioita, jotka on luotu informaation jakamista ja verkostoitumisen kehittämistä varten digitaalisissa ympäristöissä (Osborne-Gowey 2014). Digitalisaation edetessä kansainvälisten tietoverkkojen (*world wide web*) toisessa murrosvaiheessa eli niin sanotun web 2.0:n seurauksena verkko-ympäristöt kehittyivät aikaisempaa asiakaskeskeisiksi ja pyrkimyksenä oli kehittää uusia tapoja kommunikoida sekä jakaa ja vaihtaa informaatiota ihmisten kesken. Tämän seurauksena sosiaalisen median voidaan ymmärtää syntyneen. Sosiaaliseen mediaan kuuluu useita kategorioita kuten keskustelualustat (Skype), jakelualustat (YouTube), mikroblogit (Twitter) ja sosiaaliset verkostot (Facebook ja LinkedIn). (Power & Phillips-Wren 2011) Kuten jo aikaisemmin mainittiin, tässä työssä keskitytään palveluihin, joissa merkittävässä osassa on käyttäjien sisällön luonti tekstipohjaisesti sekä sosiaalinen verkostoituminen. Näistä suosituimpia alustoja ovat Facebook ja Twitter (Haya 2019).

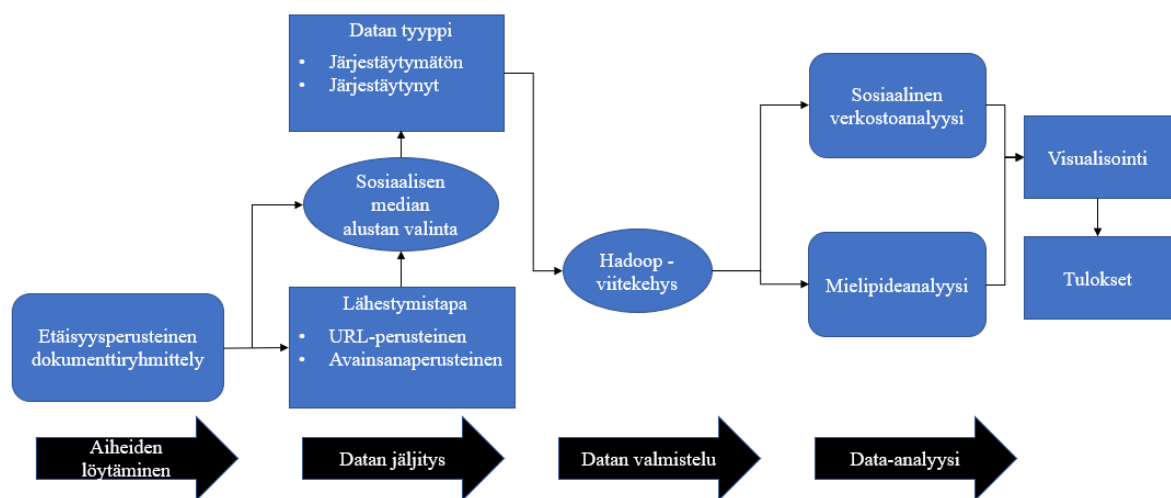
Sosiaalisen median tarjoaman datan määrä on valtava ja jatkuvasti kasvava, johtuen ole-massa olevien käyttäjien suuresta määrästä ja uusien jatkuvasta kasvusta. Vuonna 2020 sosiaalisen median palveluissa oli yhteensä arviolta 3,6 miljardia käyttäjää, joka on noin 46 % koko maailman väestöstä. Tämän luvun ennakkointiin edelleen kasvavan. (Zachlod et al. 2022; The World Bank 2023) Sosiaalisen median data käsittää alustojen käyttäjien luomaa materiaalia kuten kommentteja, tekstipostauksia ja videoita. Dataa syntyy jatkuvasti lisää alustoilla tapahtuvan kommunikoinnin seurauksena. Alustat edesauttavat datan määrän kasvua linkittämällä käyttäjäprofileita toisiin muun muassa samankaltaisten kiinnostuksenkohteiden perusteella. (Hayat, Daud, Alshdadi, Banjar, Abbasi, Bao & Dawood 2019)

Sosiaalisen median datan voidaan käsittää lukeutuvan niin kutsuttuun big dataan. Kuten aikaisemmin mainittiin, big data on kompleksista tietoaineistoa perustuen sen ennenäkemättömään ja järjestäytymättömään määrään sekä moninaisuuteen. (Schintler & McNeely 2022, 75) Datan järjestäytymättömyys tarkoittaa, että sitä ei ole jäsenneilty mitenkään, jolloin halutun informaation löytäminen vaatii datan voimakasta muokkausta. Datan kasvusta kertoo jotain se, että sen määrän on arvioitu kaksinkertaistuvan noin kahden vuoden välein (Oracle 2023). Datan ainutlaatuisuus ja laajuus tekee siitä organisaatioiden kannalta hyvin hankalaa käsitellä ja saattaa lisäarvoa tuottavaan muotoon (Nakashololo & Iyamu 2019). Rajoituksista huolimatta sen tarjoama hyöty on selkeä: sosiaalisesta mediasta kerätyn datan avulla on mahdollista kerätä kuluttajien mielipiteitä ja reaktioita, joita ei yritysten omaan toimintaan liittyvässä informaatioissa kuten asiakastilauksiin, taloudellisiin lukuihin ja tilaustietoihin perustuvassa datassa ole. Näistä syistä datan implementoinnin tuomat mahdollisuudet ovat selkeästi sen tuomia haasteita suurempia ja sosiaalisen median implementointi datan lähteeksi on selkeä potentiaalista lisäarvoa tuottava tekijä. (Saito & Gupta 2022)

Facebook on sosiaalisen median alusta, jonka tarkoituksena on helppo yhteydenpito ja tiedon jakaminen sosiaalisten verkostojen välillä. Facebook mahdollistaa kuvien, musiikin, videoiden, artikkeleiden ja omien mielipiteiden jakamisen niin monen ihmisen kanssa kuin tiedon jakaja vain haluaa (Webwise 2023). Facebookin potentiaali datan lähteenä perustuu sen yleisyyteen ja luonteeseen: lähes kaikilla yrityksillä on Facebook-sivut, jolloin palvelussa toimiminen on tuttua. Facebook perustuu yksittäisten ihmisten mielipiteiden ilmaisuihin kommenttien ja julkaisuiden muodossa, jolloin kuluttajien aidot mielipiteet tulevat esiin. (Hayat et al. 2019) Lisäksi palvelu on julkinen, jolloin data on helposti ladattavissa. Kommenttien ja julkaisujen lisäksi Facebook sisältää muun muassa sukupuoleen, koulutukseen, suhdestatukseen ja käyttäjien kykyihin liittyvää informaatiota profiilitietojen kautta. Vuonna 2023 Facebookissa oli arviolta 3 miljardia aktiivista käyttäjää kuukaudessa tehden siitä näin suosituimman sosiaalisen median alustan maailmassa. Lisäksi arviolta yli 200 miljoonaa yritystä käyttää Facebookia, joka indikoi sen potentiaalista myös yrityksille hyödyllisenä kanavana. (Luu 2023)

Tieteellisessä keskustelussa sosiaalisen median dataa ja sen hyödyntämistä lähestytään **sosiaalisen median data-analyysin** (*social media data analysis*) kautta. Termi ”sosiaalisen median data-analyysi” voidaan määritellä tarkoittavan ”nousevaa tieteidenvälistä tutkimusalaa, jonka tavoitteena on yhdistää, laajentaa ja soveltaa erilaisia metodeja sosiaalisen median

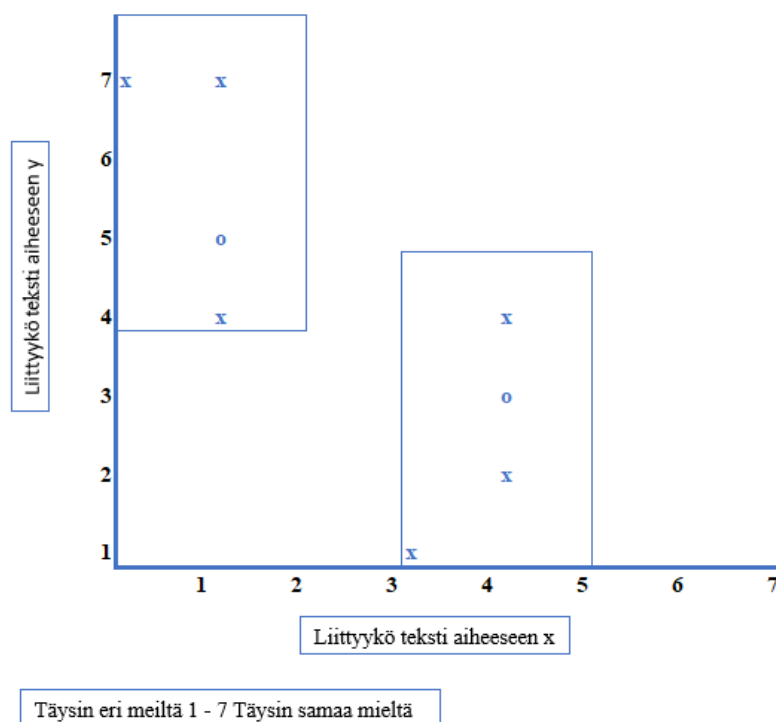
data-analyysiä varten” (Zeng, Chen, Lusch & Li 2010, 14). Stieglitz, Mirbabaie, Ross ja Neuberger (2018) kehittivät nelivaiheisen viitekehksen ohjaamaan sosiaalisen median data-analyysiä. Kehyksen neljä vaihetta ovat: aiheiden löytäminen, datan jäljitys, datan valmistelu ja data-analyysi. Analyysin jälkeen datan pitäisi teorian mukaan olla sellaisessa muodossa, että siitä on mahdollista saada hyödyllistä informaatiota (Stieglitz et al. 2018, 158). Viitekehys on kuvattu alla:



Kuva 1. Sosiaalisen median data-analytiikan viitekehys (mukaillen Stieglitz et al. 2018)

Viitekehys määrittelee analyysin ensimmäiseksi esivaiheeksi **aiheiden löytämisen**. Tarkentamalla aihetta, johon kerättävän datan halutaan liittyvän, helpottuu analyysin seuraava vaihe, jossa määritellään tarkemmat yksityiskohdat käytettävälle datalle. Chinnovin, Kerschken, Mesken, Stieglitzin ja Trautmannin (2015) mukaan yksi relevantti aiheiden etsimistekniikka on etäisyysperusteinen dokumenttiryhmittely. Prosessi alkaa määrittelemällä ensin tekstidatan aiheen sekä piirteet, joilla dataa verrataan muihin ja algoritmin, jolla ryhmittely toteutetaan. Yksi käytetyimmistä ryhmittelyalgoritmeista on k-keskiarvon klusterianalyysi (*k-means clustering*), jossa aineistoja verrataan keskenään perustuen niiden kykyyn vastata ennalta määrättyihin kysymyksiin. Algoritmi määrittelee ensin niin monta ryhmäkeskipistettä kuin ryhmiä halutaan muodostaa ja jakaa sitten havainnot ryhmiin sen mukaan, mitä ryhmäkeskipistettä lähimpänä ne ovat (Wu 2012, 7).

Sijoittelua voidaan havainnollistaa esimerkiksi kaksiulotteisessa koordinaatistossa, jossa x- ja y-akselilla on kaksi eri kysymystä ja vastaushavainnot kuvaavat pisteitä koordinaatistossa. Havainnot jakautuvat akselistolle sen mukaan, mitä mieltä yksittäiset havainnot ovat eri akselien kysymyksistä. Perustuen havaintojen etäisyyksiin valituista ryhmäkeskipisteistä, algoritmi jakaa havainnot lopulta ryhmiin. Tuloksena samaan aihepiiriin liittyvä data on ryhmitelty erilleen muusta datasta. Aiheiden löytämävaiheessa datan ryhmittely toteutetaan vielä hyvin yleisellä tasolla ja prosessin päätarkoitus onkin tässä vaiheessa auttaa datalähteen valinnassa ja löytää tietyn aihepiirin keskustelun yhteydessä käytettäviä avainsanoja, joita voidaan hyödyntää viitekehyksen myöhemmissä vaiheissa. (Pauha 2019, 7–8; Stieglitz et al. 2018, 158) K-keskiarvon klusterianalyysin perusidea on kuvattu alla:



Kuva 2. K-keskiarvon klusterianalyysin perusperiaate (mukaillen Pauha 2019)

Datan jäljitysvaihe käsittää datan lähteen, keräystavan ja tyyppin valinnan. Lähteen valinnassa tärkeintä on huomioida, minkälaista dataa halutaan ja missä muodossa. Esimerkiksi Twitteristä saatava data perustuu julkisiin twiitteihin, kun taas Facebook-data on julkista vain sellaisten kirjoitusten osalta, jotka on määritelty julkaisijan toimesta julkisiksi. Toisaalta Facebook-kirjoitukset eivät ole pituutensa osalta rajoitettuja verrattuna Twitter

twiitteihin. (Stieglitz et al. 2018) Datan keräystavan valinnassa tärkeintä on päättää, millä tavalla dataa halutaan kerätä. Dataa voidaan kerätä esimerkiksi avainsana tai -aihe perusteisesti. Tämä tarkoittaa sitä, että esimerkiksi Facebook postauksia kerätään sen perusteella, sisältävätkö ne tiettyyn aihepiiriin kuuluvia sanoja. Jotta keräys olisi mahdollisimman tehokasta, avainsanojen valinnassa ja päätettäessä, montako avainsanaa halutaan huomioida, on oltava huolellinen (Stieglitz, Dang-Xuan, Bruns & Neuberger 2014, 94).

Vaihtoehto avainsanaperusteiselle datan keräykselle voi olla esimerkiksi URL-perusteinen datan keräys. Tämä perustuu siihen, että datan keräyksessä pyritään keskittymään sellaisiin sosiaalisen median kirjoituksiin, jotka sisältävät hyperlinkkejä, jolloin data-analyysissä huomioidaan myös URL-linkkien takana olevaa dataa. Tämä voi olla erittäin hyödyllinen datan keräystapa lisäämällä potentiaalisen hyödynnettävän tiedon määrää entisestään ja mahdollisesti avaamalla ajatuksia sosiaalisen median kirjoitusten taustalla aikaisempaa yksityiskohteisemmin. (Stieglitz et al. 2014) Datatyypin valinta liittyy siihen, minkä muotoista dataa halutaan hyödyntää. Sosiaalisen median data koostuu niin sanotuista järjestäytyneestä ja järjestäytymättömästä datasta. Järjestäytynyt data on tarkempaa ja koostuu esimerkiksi käyttäjätילוista sekä tykkäysten, katselukertojen ja klikkausten määrästä. Järjestäytymätön data koostuu jo aikaisemmin mainituista käyttäjien luomista kirjoituksista. (Kirwan 2022) Dataa voidaan muuntaa tiettyyn formaattiin ja varastoida tietokantoihin. Tyypillisesti datasta on myös mahdollista jäljittää, kuka sen on julkaissut ja milloin sekä minkä tyyppistä se on (oma julkaisu tai kommentti toiseen julkaisuun) (Hayat et al. 2019, 36959–36961).

Datan valmisteluvaiheessa kerättyä dataa visualisoidaan havainnollisempaan muotoon. Datan visualisointi on tärkeä vaihe, koska päätöksenteko on usein nopeaa, joten datan on oltavan selkeää ja sisältää vain tarvittavaa informaatiota, jotta se loisi todellista hyötyä päätöksenteon tueksi (Anuncia, Gohel & Vairamuthu 2020, 7). Tässä vaiheessa suurimman haasteen luo erityisesti sosiaalisen median datan suuri määrä ja vaihtelevuus: datan määrän kasvaessa sen käsittely vaikeutuu ja asettaa entistä suurempia vaatimuksia datatyökaluille (Anuncia et al. 2020, 17). Myös sosiaalisen median datan suuri vaihtelevuus perinteisestä tekstidatasta esimerkiksi sijaintitietoihin, vaikeuttaa sen visualisoimista merkittävästi (Zahedi & Costas 2018). Yksi keino suurten datamäärien käsittelyyn on niin kutsuttu Hadoop-viitekehys. Hadoop on Apache Software Foundationin kehittämä avoimen lähdekoodin ohjelmisto, joka on kehitetty suurien hajautettujen tietomäärien käsittelyyn. Mallin etuna on, että se kykenee skaalautumaan yksittäisistä palvelimista useisiin tietokoneisiin ja hyödyntää

näin niiden laskentatehoa samaan aikaan. Näin luodulla useiden palvelimien verkostolla suurten datamäärien käsittely yhtäaikaaisesti on tehokasta. Hadoop hyödyntää MapReduce-menetelmää, jonka idea pohjautuu siihen, että kerätty data pilkotaan osiin ja tieto prosessoidaan pienissä ryhmissä samanaikaisesti. Saadut tulokset kootaan myöhemmin yhteen kerätväiksi. (Turkington, Deshpande & Karanth 2016, 11–13)

Verrattuna perinteisiin tietokantoihin, jotka edellyttävät tarkkaan määriteltyä ja järjestäytyntä dataa, Hadoop ja MapReduce toimivat parhaiten järjestäytymättömän datan käsittelyssä, sillä ne eivät vaadi, että datan pitäisi olla rakenteeltaan tietynlaista. Sosiaalisen median datan sovellusalueella tätä menetelmää hyödyntämällä suuresta määrästä järjestäytymätöntä dataa voidaan suodattaa pois merkityksettömät asiat ja syöttää merkityksellinen tieto avain-arvo-pareina analysoitavaksi. (Turkington et al. 2016, 13–15) Avain-arvo-parit tarkoittavat tapaa varastoida dataa, jossa dataa voi järjestellä ja lajitella avainseteinä ja näille avainseteille voidaan määrittää arvoja. Esimerkkinä käy tekstidata, jossa siinä esiintyvät sanat ovat avaimia ja tiettyjen sanojen esiintymismäärät toimivat näiden avaimien arvoina (Turkington et al. 2016, 53–54).

Kun data on kerätty ja prosessoitu, on vuorossa viitekehyyksen viimeinen vaihe eli **data-analyysi**, jossa muokatusta datasta pyritään saamaan merkityksellisiä havaintoja (Stieglitz et al. 2014, 94). Viitekehys tarjoaa erilaisia näkökulmia analyysiä varten. Eräs teoriapohjainen näkökulma on mielipide/tunneanalyysi (*opinion-based/sentiment analysis*). Näkökulman perustana on se, että datan keräyksen tarkoituksena on saada muodostettua kuvaa sosiaalisen median käyttäjien mielipiteistä johonkin tiettyyn aiheeseen liittyen (Pozzi, F. & Pozzi, F.A., 2017). Analyysi voidaan toteuttaa esimerkiksi keräämällä tiettyyn aihepiiriin kuuluvia tekstejä sosiaalisesta mediasta ja tilastoida, kuinka paljon negatiivisia ja positiivisia ilmauksia teksteissä esiintyy. Tällä tavalla voidaan pyrkiä luomaan tilastollisia havaintoja suurten käyttäjäjoukkojen suhtautumisesta eri aiheisiin. Käytännössä mielipideanalyysin toteuttaminen suuresta datamäärästä vaatii erilaisia sovelluksia datan keräämiseksi ja kokoamiseksi. Yksi mahdollisuus on käyttää koneoppimisperusteista algoritmia, joka opetetaan luokittamaan aineistoa negatiiviseen, positiiviseen ja neutraaliin ryhmään (Younis 2015, 45). Tietyn aihepiirin tekstien keräämisessä voidaan hyödyntää esimerkiksi koneoppimiseen perustuvaa ohjattua tekstiluokittelua (Stieglitz & Dang-Xuan 2013, 1295–1297). Koneoppimiseen palataan tarkemmin seuraavassa luvussa.

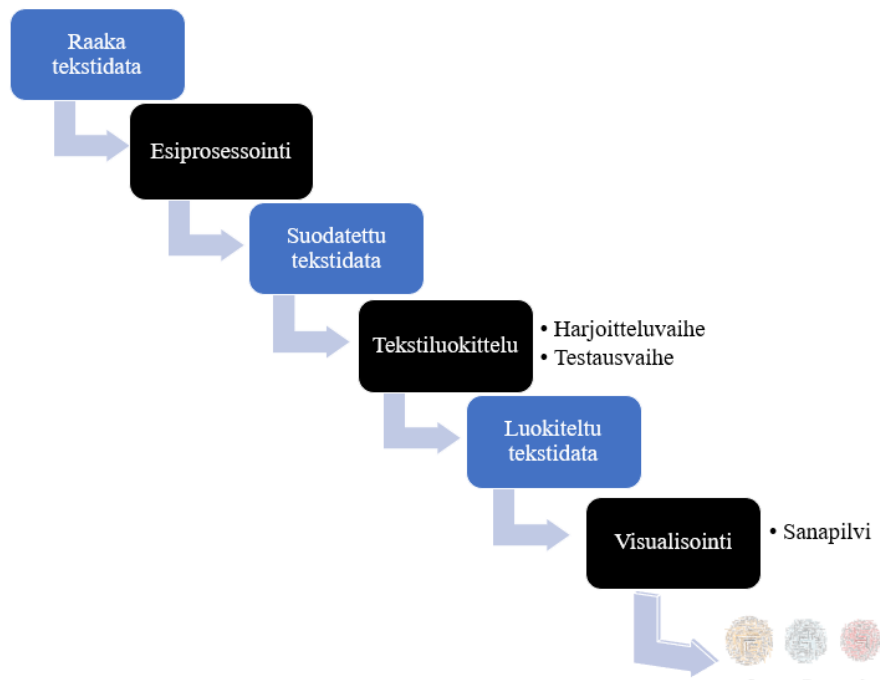
Toinen näkökulma data-analyysiin on sosiaalinen verkostanalyysi (*social network analysis*), jossa lähtökohtana on tunnistaa datan avulla erilaisia sosiaalisia ryhmiä (Stieglitz & Dang-Xuan 2013, 1298–1299). Ryhmien tunnistaminen voi olla hyödyllistä erityisesti sosiaalisen median valtavassa kentässä, jossa tiettyyn ryhmään kuuluvat yksilöt usein ajattelevat samalla tavalla, jolloin dataa voi olla järkevämpiä tutkia käyttäjäkohtaisuuden sijaan ryhmäkohtaisesti ja näin vähentää data-analyysin monimutkaisuutta (Bedi & Sharma 2016, 115). Muun muassa aikaisemmin mainittu k-keskiarvon klusterianalyysi on mahdollinen metodi eri ryhmien tunnistamiseen. Yhdistämällä sosiaalisesta mediasta kerättyä järjestäytyntä ja järjestäytymätöntä dataa, voidaan datan lähteet, eli palvelun käyttäjät, tunnistaa ja luokitella heidät ryhmiin joidenkin valittujen tekijöiden kuten iän, poliittisen mielipiteen tai sosiaalisen aseman perusteella. (Bedi & Sharma 2016, 122–123) Tunnistamalla erilaisia sosiaalisia ryhmiä, voitaisiin analyysiin yhdistää myös mielipide- ja aiheanalyysi, ja näin pyrkiä tilastoi- maan eri ryhmien mielipiteitä valittuihin aiheisiin. Yritysten näkökulmasta tämä analysoin- titapa voisi olla potentiaalinen keino eri asiakasryhmien ja heidän tarpeidensa määrittelemis- sessä.

2.2 Koneoppiminen analyysin tukena

Kun puhutaan perinteisestä tietokoneohjelmoinnista, algoritmikäsite on keskiössä. Algori- tmi on tarkka kuvaus tehtävän prosessin suorittamisesta, joka koostuu yksittäisistä järjes- tyksessä olevista tehtävistä, joilla ennalta määritetty syöte saadaan vietyä haluttuun loppu- tulokseen. Algoritmien avulla olemme pystyneet helpottamaan jokapäiväistä elämäämme hyödyntämällä niiden tarjoamia mahdollisuuksia monien eri tehtävien hoidossa. (Alpaydin 2020, 1–2) On kuitenkin tehtäviä, joita algoritmille on hankalaa, jopa mahdotonta, määrittellä suoraan. Voimme käyttää esimerkkinä kuluttajakäyttäytymistä: on mahdotonta määrittää yk- sioikoista kaavaa sille, miten ihmiset tekevät ostopäätöksiään. Tiedämme, että täysin satun- naista se ei ole, vaan siihen voivat vaikuttaa esimerkiksi vuodenaika, maantieteellinen si- jainti ja muut ostokset, joiden avulla on mahdollista tunnistaa joitain johdonmukaisuuksia ostokäyttäytymiseen liittyen. Tämä ei välttämättä riitä yksityiskohtaisen prosessin tunnistam- iseen mutta näitä johdonmukaisuuksia ja kaavoja tunnistamalla on mahdollista kuitenkin luoda jonkintasoinen approksimaatio. (Alpaydin 2020, 2–3) Tässä esiin astuu koneoppimi- nen (Rebala, Ravi & Churiwala 2019, 1–2).

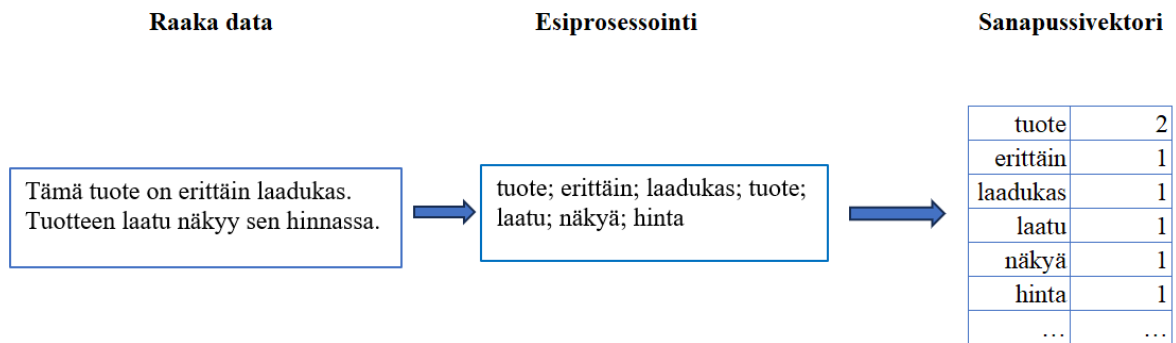
Koneoppiminen on osa tekoälyä, jonka perusideana on kehittää ohjelmiston toimintaa ennakkotiedon ja ohjelmiston käyttäjän toiminnan perusteella. Perusideana on lähteä liikkeelle hyvin yleisluontoisesta mallista, jossa ohjelmalle on määritetty paljon muuttujia ja muuttujien järjestystä muuttamalla ohjelma voi teoriassa tehdä lukuisia eri toimintoja. Oppimisen kautta ohjelma muokkaa arvoja näille muuttujille niin, että ne mukautuvat mahdollisimman hyvin oppimisprosessissa käytetylle datalle. Tämän perusteella muuttajat päätyvät tiettyihin arvioihin ja syntyy lopulta algoritmi, jossa alkuparametrit ovat saaneet tarkemman määrittelyn perustuen käytettyyn dataan. Tässä ilmenee myös se, miksi koneoppiminen liittyy niin vahvasti tekoälyyn. Ollakseen älykäs, muuttuvassa tilassa olevan systeemin tulee pystyä muuttumaan. Muutoskykyisyys vaatii sitä, että ohjelman on suunniteltu tarjoamaan ratkaisuja lukuisiin eri skenaarioihin. (Alpaydin 2020, 2–3) Koneoppimisessa algoritmit siis kehittävät itseään etsimällä tietomassoista korrelaatioita sekä kaavoja, luovat näiden pohjalta ennusteita siitä, kuinka prosessin kuuluisi edetä ja muokkaavat sitten toimintaansa näiden mukaan (IBM 2023).

Koneoppimisen sovellusten avulla sosiaalisen median datankeräyksen haasteet liittyen datan suureen määrään ja arvokkaan datan tunnistamiseen on mahdollista ratkaista (T.K., Annavarpu & Bablani 2021, 1). Keskeinen koneoppimisen sovellus tähän liittyen on **tekstilouhinta** (Hassani, Beneki, Unger, Mazinani & Yeganegi 2020). Tekstilouhinta tarkoittaa automatisoitua prosessia, jonka ideana on tunnistaa ja kerätä järjestytmätöntä tekstidataa perustuen sen sisäisten vuorovaikutussuhteiden tunnistamiseen sekä muokata tätä järjestytyneeseen muotoon hyödyllisen informaation tunnistamiseksi. Perinteisen hakusanaperusteisen tekstietsinnän sijaan, tekstilouhinnassa ideana on tunnistaa datan sisäistä, vielä tunnistamatonta, informaatiota algoritmeja hyödyntämällä. (Younis 2015, 44–45) Tekstilouhinnan ymmärtämiseksi, se voidaan esittää prosessina, joka on kuvattu alla:



Kuva 4. Tekstilouhintaprosessin eteneminen (mukaillen Younis 2015)

Prosessin alussa on suuri määrä järjestäytymätöntä tekstidataa esimerkiksi käyttäjien kommentteja Facebookista. Ensin dataa prosessoidaan poistamalla välimerkit ja turhat sanat kuten virkkeiden lopetukset sekä muut yleisesti käytetyt sanat kuten partikkelit (Younis 2015, 44–45). Kun data on putsattu, on seuraavaksi löydettävä soveltuva malli tekstin esittämiselle. Eräs yleinen malli on esittää tekstiä ominaisuusvektorien muodossa, jossa jokainen vektori esittää havaintoaineiston osaa, kuten tekstikappaletta, sanastoina, jossa jokainen sana saa numeroarvon, jolla se yksilöllistetään muista sanoista (Hassani et al. 2020, 6). Eräs yksinkertainen vektorimalli on niin kutsuttu sanapussimalli (*bag-of-words model*), jossa havainnot esitetään vektorimuodossa, joka sisältää kussakin kappaleessa esiintyvien sanojen tunnistamisarvot ja niiden esiintymismäärät. (Zhao & Mao 2018) Sanapussimalli on hyvin yksinkertainen eikä tunnista esimerkiksi eri sijamuotoja samaksi sanaksi. Tästä syystä, mallin käyttökelpoisuuden parantamiseksi, havaintoaineistolle on suositeltavaa tehdä esiprosessointia, jossa turhat sanat ja merkit poistetaan sekä sanat esitetään perusmuodoissa pienellä kirjoitettuna (Great Learning 2022). Sanapussimallin rakennetta on havainnollistettu alla olevassa kuvassa:



Kuva 3. Sanapussivektori tekstidatan esittämiseksi (mukaillen Balodi 2019).

Kun tekstidatalle on löydetty sopiva esittämismalli, on seuraavana vuorossa luokittelu. Luokittelua varten on löydettävä jokin tekstissä esiintyvä piirre ja luokittelufunktio, jolla luokittelu on mahdollista toteuttaa (Younis 2015, 45). Käytännön sovelluksena voidaan käyttää esimerkiksi koneoppimis pohjaista tekstiluokittelua. Tekstilukittelu koostuu kahdesta vaiheesta: harjoittelu- ja testausvaiheesta. Harjoitteluvaiheessa algoritmillemme määritetään manuaalisesti loogisia sääntöjä, joita hyödyntämällä se oppii luokittelemaan tekstiä. Algoritmillemme voidaan syöttää ensin dataa, joka on valmiiksi luokiteltu esimerkiksi perustuen siihen, mihin aihepiiriin se liittyy. Yritysten näkökulmasta luokittelu voisi perustua yrityksen eri tuotteisiin ja palveluihin. Manuaalisesti luokitellun aineiston perusteella algoritmi oppii datan sisäisten vuorovaikutussuhteiden kautta mallin, jonka avulla tekstiä luokitellaan. (Hassan et al. 2022) Se mihin algoritmi luokittelunsa perustaa, riippuu valitusta koneoppimisen tekniikasta.

Yksi mahdollisuus algoritmin rakentamisessa on käyttää tukivektorikonemallia (*support vector machine*). Malli luokittelee tekstivektoreita muodostamalla aineistoon päätösrajan kohtaan, jossa jakotason ja sen molemmilla puolilla olevien marginaalitasojen välimatka on mahdollisimman suuri. Yksikään näyte ei saa jäädä marginaalitasojen väliin. Toisella puolella olevat vektorit eivät kuulu haluttuun ryhmään ja toisella puolella olevat kuuluvat. (Kadhim 2019, 279–280) Kun algoritmi on määrittänyt päätösrajan aihekategorialle, jota halutaan analysoida, tarvitsee haluttu tekstidata vain muuntaa vektorimuotoon ja tarkistaa, kummalle puolelle rajaa se kuuluu (Kadhim 2019, 279).

Tekstilukittelun testausvaiheessa tätä oppimisalgoritmin muodostamaa luokittelumallia testataan luokittelemattomaan dataan. Ensin valitaan avainsanat luokittelua varten. Sitten mallia harjoitetaan syöttämällä siihen dataa, joka luokitellaan manuaalisesti. Tämän

perusteella tukivektorikoneen algoritmi oppii, miten luokittelua halutaan toteuttaa ja hyödyntää tätä uuden datan luokittelussa. Oppimisvaiheen jälkeen algoritmin pitäisi itse pystyä analysoimaan syötettyä tekstiä ja arvioimaan, mihin annetuista aiheista data eniten liittyy ja kuinka suurella todennäköisyydellä. Algoritmin toimivuus on vahvasti sidoksissa oppimisdatan määrään: mitä enemmän oppimista harjoitetaan, sitä enemmän skenaarioita algoritmi oppii liittyen tekstiluokitteluun, jolloin algoritmin kyky tunnistaa dataa entistä paremmin kasvaa. (MonkeyLearn 2023) Kun data on luokiteltu, se visualisoidaan esimerkiksi sanapilven avulla. Sanapilvi on graafinen esitys kerätyistä avainsanoista, jossa enemmän käytetyt sanat näkyvät suurempina kuin vähemmän käytetyt (Younis 2015, 46). Käytännössä jokainen määritelty sanaluokka voidaan esittää omana sanapilvenään, jolloin voidaan kätevästi visualisoida, mitkä sanat ovat esiintyneet eniten eri luokissa.

3 Johdon laskentatoimi organisaatioissa

Luvussa käsitellään johdon laskentatoimen käsitettä sekä johdon laskentatoimen roolia organisaatioiden toiminnan kannalta. Luvun luettuaan lukijan tulisi ymmärtää, mitä johdon laskentatoimi käsitteenä tarkoittaa sekä mikä rooli sillä on organisaatioiden toiminnassa.

3.1 Lähestymistapoja johdon laskentatoimeen

Yritykset ja organisaatiot toimivat usein hyvin yksilöllisesti ja päätöksentekotilanteet ovat monimutkaisia. Nämä lainalaisuudet ovat luoneet tarpeen laskentatoimen alalle. Laskentatoimen (*accounting*) ideana onkin tuottaa luotettavaa numeromääräistä informaatiota organisaatiolle toiminnan tueksi (Ikäheimo et al. 2019, 12). Ikäheimo et al. (2019) luonnehtivat laskentatoimea prosessiksi sekä järjestelmäksi, jonka tarkoituksena on kerätä, mitata, tallentaa ja kommunikoida informaatiota organisaation toiminnan ja päätöksenteon tueksi. Määrittelyssä painotetaan erityisesti kirjanpidon roolia taloudellisten tapahtumien tallentamisessa ja luokittelussa. Laskentatoimea voidaan lähestyä jakamalla se kahteen osa-alueeseen: rahoittajien laskentatoimeen ja johdon laskentatoimeen. Schuster, Heinemann ja Cleary (2021, 2) määrittelevät rahoittajien laskentatoimen (*financial accounting*) keskittyvän ulkoisten sidosryhmien, kuten sijoittajien, rahoittajien ja yritysten asiakkaiden palvelemiseen. Sen tarkoituksena on antaa oikea ja riittävä kuva yrityksen taloudellisesta asemasta ja tuloksesta lakien ja määräysten velvoittamalla tavalla. Ikäheimo et al. (2019, 13) ovat omassa määritelmässään hyvin samoilla linjoilla ja korostavat sisäisten ja ulkoisten sidosryhmien palvelemista sekä lakisääteisen taloudellisen tiedon tuottamista yrityksen toiminnasta.

Johdon laskentatoimi (*management accounting*) liittyy informaation tuottamiseen yrityksen sisäiseen käyttöön. Verrattuna rahoittajien laskentatoimeen, johdon laskentatoimea eivät säätele ulkoiset normit, vaan keskiössä on tuotettavan informaation merkityksellisyys ja hyödyllisyys yritysjohdolle tuomaan lisäarvoa päätöksentekotilanteisiin. (Ikäheimo et al. 2019, 13–14) Ikäheimo et al. (2019, 130) kuvailevat johdon laskentatoimea kokonaisuudeksi, joka sisältää tiedon kokoamista, hallitsemista ja analysointia sekä tämän prosessoidun tiedon viestimistä relevantille kohderyhmälle yrityksen sisällä, merkittävimmässä määrin juuri johdolle. Tieto voi olla rahamääräisen lisäksi myös ei-rahamääräistä sekä historiaan

perustuvaa tai ennustettua. Rahamääräiseen informaatioon kuuluu esimerkiksi valmistuksen juoksevat kulut, jakelukulut, hallintokulut ja liikevaihto. Ei-rahamääräiseen tietoon kuuluu esimerkiksi asiakasdata, tuotantoprosessin laatu sekä oikea-aikaisuus kysyntään nähden ja tieto liittyen työntekijöiden pysyvyyteen. (Ikäheimo et al. 2019, 131) Dahal (2019) määrittelee johdon laskentatoimen alaksi, jonka tehtäviin kuuluu päätöksentekoon osallistuminen sekä taloudellinen raportointi ja neuvonta. Näiden tehtävien tarkoituksena on tukea yritysjohtoa organisaation strategian suunnittelussa ja toteuttamisessa.

Ikäheimo et al. (2019, 131) kuvaavat johdon laskentatoimen roolin jakautuvan kolmeen osaluokkaan: kustannus- ja kannattavuuslaskentaan, ohjaukseen ja päätöksentekoon. Kustannus- ja kannattavuuslaskenta liittyy tuottojen ja kulujen kohdistamiseen yrityksen tuotteille, palveluille, asiakkaille ja muille relevanteille sidosryhmille. Sitä kuvataan johdon laskentatoimen perustana. Ohjaus liittyy päätöksenteon tukemiseen ja tähän kuuluu muun muassa strategiasuunnittelu ja budjetointi. Päätöksenteossa johdon laskentatoimen rooli kuvataan korostuvan erityisesti hinnoittelu- ja investointipäätöksissä (Ikäheimo et al. 2019, 131–133). Atkinson, Kaplan, Matsumura ja Young (2012, 26–27) kuvaavat johdon laskentatoimeja prosessiksi, jossa tuotettua relevanttia tietoa hyödynnetään yrityksissä niin päätöksenteossa, resurssien jakamisessa kuin myös toteutetuista päätöksistä seuranneiden tulosten valvonnassa ja arvioinneissa huomioimatta myöskään hyvistä suorituksista palkitsemista. Näiden prosessien kautta pyritään edistämään yrityksen strategian toteutumista kehittämällä operatiivista toimintaa sekä tekemällä päätöksiä liittyen esimerkiksi uusiin tuotteisiin ja asiakkaisiin sekä motivoimalla työntekijöitä (Atkinson et al. 2012, 34).

Myös Schuster et al. (2021) korostavat määrittelyssään johdon laskentatoimeja informaation lähteenä johdon päätöksenteossa. He painottavat sitä, että päätöksenteossa avustaminen on relevanttia erityisesti lyhyellä aikavälillä ja merkittävin tieto on rahamääräistä, käytännössä siis liiketoiminnan kuluihin ja menoihin liittyvää (Schuster et al. 2021, 3–4). Määritelmässään Schuster et al. (2021) jakavat johdon laskentatoimen työn kahteen päätarkoitukseen: päätöksentekoon ja päätöksiin vaikuttamiseen. Sen lisäksi, että johdon laskentatoimen tarkoitus on tukea päätöksentekoa tarjoamalla tarvittavan informaation johdolle, jotta päätöksenteko ylipäättänsä on mahdollista, on sillä myös rooli vaikuttaa muiden tekemiin päätöksiin tilanteessa, jossa päätöksentekijöitä on useita, osa päätöksistä on siirretty muille tehtäväksi esimerkiksi ajan puutteen vuoksi ja eri päätöksentekijöillä on erilaisia tavoitteita, jotka

saattavat ohjata heidän päätöksentekoaan epäedulliseen suuntaan koko organisaation näkökulmasta (Schuster et al. 2021, 3–5)

Lambert ja Sponem (2012, 571–572) lähestyvät johdon laskentatoimen määrittelyä tapaus-tutkimuksen kautta tunnistuen neljä tyyliä johdon laskentatoimen toiminnalle organisaatioissa käytännössä sekä riskejä, joita kukin vaihtoehto mahdollisesti voi tuoda. Erittelevässä tyyliässä johdon laskentatoimi on organisoitu erillisenä tietolähteenä suhteessa operatiiviseen johtoon, jolloin johto ei tukeudu pelkästään siihen vaan säilyttää myös oman ajattelun päätöksissään vahvana. Ongelmana esitetään riski sisäisen kontrollin puutteesta. Suojelevassa tyyliässä johdon laskentatoimi on suunniteltu välietapiksi tulevien operatiivisten johtajien uralla, jotta he oppisivat paremmin johdon laskentatoimen informaation merkityksestä päätöksenteossa, jota he tulevaisuudessa tulevat tekemään. Yhteistyökeskeisessä tyyliässä johdon laskentatoimi on organisoitu työskentelemään tiiviissä yhteistyössä operatiivisen johdon kanssa organisaation tavoitteiden saavuttamiseksi. Riskinä tässä esitetään operatiivisen johdon omien intressien ajamista heidän päästessä vahvemmin mukaan johdon laskentatoimen työhön. Itsevaltaisessa tyyliässä johdon laskentatoimen päärooli on keskittää valta yrityksessä ylimmälle johdolle. Riskinä tässä järjestelyssä esitetään kasvavaa kuilua käytännön toiminnan ja lopullisen päätöksenteon välillä. (Lambert & Sponem 2012, 572–584) Appelbaum et al. (2017, 32) tiivistävät johdon laskentatoimen roolin pyrkimykseen vastata kysymyksiin: mitä on tapahtunut, mitä tulee tapahtumaan ja mikä on optimaalisin ratkaisu. Tutkimuksessa tuodaan esiin, kuinka datan saatavuuden ja ajankohtaisuuden kasvaessa johdon laskentatoimen rooli tulee siirtymään enemmän tulevaisuuden ennakointiin pelkän historiatiedon tarkastelun sijaan (Appelbaum et al. 2017, 29).

Tulevissa alaluvuissa käsittelemme ohjaukseen ja päätöksentekoon liittyviä sisältöjä. Kustannus- ja kannattavuuslaskennan teknisten yksityiskohtien käsittelyn sivuutamme, koska ne eivät kuulu olennaisena osana työn näkökulmaan. Vaihtoehtoinen argumentti kustannus- ja kannattavuuslaskennan sivuuttamiselle on myös johdon laskentatoimen työn muutos: digitalisaation aiheuttama tietojärjestelmien kehittyminen 1900-luvun lopulta 2000-luvulle siirtyäessä on muokannut johdon laskentatoimen roolia huomattavasti. Kun aikaisemmin keskiössä oli tiedon kerääminen, nykyään pääroolissa on kyky analysoida ja jalostaa suuria tietomassoja raportoitavaan muotoon. Tämän takia esimerkiksi kustannuslaskennan rooli on nykyään enemmän muuta työtä tukevaa eikä enää johdon laskentatoimen ytimessä. (Dahal 2019) Merkittävässä roolissa ovat olleet yritysten sisäiset tiedonkeruujärjestelmät, kuten

ERP (*enterprise resource planning systems*), jotka ovat mahdollistaneet tarkemman, tehokkaamman, ajantasaisemman ja kokonaisvaltaisemman datankeräyksen sekä sisäisen raportoinnin (Pervan & Dropulic 2019). Johdon laskentatoimen osa-alueet on havainnollistettu alla olevaan kuvaan:



Kuva 5. Johdon laskentatoimen osa-alueet

Tässä työssä keskitytään johdon laskentatoimen osa-alueista ohjauksen ja päätöksenteon alle kuuluviin strategiasuunnitteluun ja -seurantaan, budjetointiin, hinnoitteluun sekä investointeihin. Osa-alueet, joita työssä ei käsitellä on merkattu yllä olevaan kuvaan keltaisella. Kuvassa esitetty osa-aluejako perustuu Ikäheimo et al. (2019) määritelmään johdon laskentatoimesta.

3.2 Ohjaus

Merkittävä osa yrityksen pitkántähtäimen ohjaukselta on strategiasuunnittelu. Strategia kuvastaa toimenpiteitä, joilla organisaation resurssit ja toimintaympäristö sovitetaan yhteen toimenpiteiksi parhaalla mahdollisella tavalla tavoitteiden saavuttamista varten. Strategian keskiössä ovat päätökset siitä, mihin markkinoille yrityksen tulisi keskittyä ja miten. (Atkinson et al. 2012, 29) Johdon laskentatoimen rooli on tiedon tuottamisessa strategian toteutumisen seurannassa. Kustannus- ja kannattavuuslaskennan avulla saadaan tietoa, mitkä toiminnot tuottavat parhaiten ja miten kustannukset jakautuvat. Tätä tietoa yritys voi hyödyntää strategisessa päätöksenteossa liittyen esimerkiksi joidenkin toimintojen lakkauttamiseen tai uusiin liiketoimintoihin investoimiseen. (Ikäheimo et al. 2019, 148–154) Frezatti, Aguiar, Guerreiro ja Gouvea (2011) korostavat johdon laskentatoimen tuottaman tiedon merkitystä

strategian laatimisen kannalta. Strategiasuunnittelun kuvataan liian usein olevan etäistä organisaation todelliseen toimintaan verrattuna, koska strategiasuunnittelijat nojaavat liikaa pelkkään tietojärjestelmien tuottamaan informaatioon. Tämä johtaa todennäköisemmin huonoon strategiaan, joka ei edesauta haluttuihin tavoitteisiin pääsyä. Tutkimuksen keskeinen havainto oli, että merkittävä syy huonolle strategialle oli juuri johdon laskentatoimen riittämättömän implementointi suunnitteluprosessiin. (Frezatti et al. 2011, 249) Johdon laskentatoimen tuottaman taloudellisen ja ei-taloudellisen informaation avulla, liittyen yrityksen sisäiseen toimintaan mutta myös operatiiviseen ympäristöön, strategiassa voidaan huomioida paremmin operatiiviset rajoitteet sekä mahdollisuudet ja näin parantaa sen laatua (Atkinson et al. 2012, 29).

Kun strategia on valmis, luodaan tämän pohjalta toiminnallinen suunnitelma strategian jalkauttamiseksi käytäntöön. Budjetti on tässä merkittävässä roolissa. Budjetti on johdon laskentatoimen tärkeä työkalu, jonka avulla kustannuksia ja tuottoja pyritään jakamaan organisaation eri tasoille niin tarkasti kuin mahdollista (Ikäheimo et al. 2019, 162–163). Pervan ja Dropulic (2019) korostavat budjetointia johdon laskentatoimen perinteisimpänä tehtävänä, josta on ajansaatossa laajennuttu tuottamaan informaatiota myös laajemmin organisaation sisällä. On tavallista, että organisaatioiden toiminnassa hyödynnetään useita eri budjetteja. Usein budjetoidaan esimerkiksi myyntimääriä ja -hintoja, valmistusta, kustannuksia sekä investointeja. Nämä niin kutsutut osabudjetit kootaan yhteen tulos-, tase- ja rahoitusbudjeteiksi, ja näin pyritään ennustamaan yrityksen operatiivisen toiminnan vaikutusta yrityksen taloudelliseen tulokseen sekä tekemään mahdollisia muutoksia osabudjetteihin, jos näyttää siltä, että menot ja tulot eivät kohtaa halutulla tavalla (Schuster et al. 2021, 227). Budjetteja laatiessa on tavallista asettaa jokin taloudellinen ja toiminnallinen tavoite, jotka budjettia laatiessa halutaan täyttyvän. Taloudellinen tavoite voi liittyä esimerkiksi haluttuun oman pääoman tuottoon ja toiminnallinen tavoite tiettyyn tuotantomäärään (Ikäheimo et al. 2019, 163).

Budjetin asettamiseen liittyy myös haasteita. Endenich ja Trapp (2020) kuvaavat yhdeksi suurimmista ongelmista löysän budjetoinnin. Työntekijät yliarvioivat helposti kuluja ja aliarvioivat tuottoja budjetoinnissa varmistaakseen helpomman tavoitteisiin pääsyn ja sitä kautta tästä saatavien palkkioiden maksimoinnin (Endenich & Trap 2020, 318–320). Johdon laskentatoimella on tässä merkittävä vaikutus, koska sen tuottaman kokonaisvaltaisen ja organisatorisen tiedon avulla, budjetointiprosessin vastuuhenkilöt voivat hallita prosessia

omaksi edukseen entistä paremmin. Tämä luo yritysjohdolle vastuun löysän budjetoinnin hillitsemiseen. (Namazi & Rezaei 2023)

3.3 Päätöksenteko

Hudson (2015) esittää, että rationaalisen päätöksenteon toteutumiseksi organisaatioiden liikejohtossa, päätöksentekijöiden on pakko huomioida statistinen informaatio eri lähteiden, kuten tilinpäätösten, analyttisten indikaattoreiden ja ennusteiden kautta. Tässä prosessissa johdon laskentatoimen rooli informaation lähteenä on merkittävä.

Yhtenä tärkeimmistä päätöksentekomuodoista on hinnoittelu. Hinnoittelu on usein laaja-alaista ja sisältää päätöksiä liittyen yleiseen hintatasoon eri asiakasryhmille, mahdollisiin määrälennuksiin tietyille asiakkaille ja organisaation sisäisiin siirtohintoihin (Ikäheimo et al. 2019, 195). Hinnoittelun roolia organisaatioiden toiminnassa kuvastaa hyvin siihen liittyvät tavoitteet. Sunarni ja Ambarriani (2019) tarkastelevat tapaustutkimuksessaan 52 teollisuusyrityksen hinnoitteluprosessia Yogyakartan alueella Indonesiassa. Tutkimuksen mukaan yrityksillä oli monta tavoitetta hinnoitteluun liittyen, joista tärkeimmät olivat markkina-aseman säilyttäminen pitkällä tähtäimellä ja tavoitellun tuloksen saavuttaminen lyhyemmällä tähtäimellä. Perinteisesti hinnoittelu on perustunut kustannusinformaatioon. Monissa taloustieteen malleissa liittyen optimaalisen hinnan määrittämiseen juuri kustannukset ovat kriittisin tekijä. Tätä kutsutaan kustannusperusteiseksi hinnoitteluksi, jossa hinnoittelun perustana käytetään tuotteen tai palvelun aiheuttamia tuotantokustannuksia. (Ikäheimo et al. 2019, 196–197) Tämän lisäksi kokonaisvaltaisen hinnoittelun saavuttamiseksi on päätöksissä tärkeä huomioida myös ulkoiset tekijät kuten asiakaskäyttäytyminen, kilpailijoiden hinnoittelu sekä lakien muodostamat rajoitteet (Schuster et al. 2021, 61).

Laaja-alaisen tiedon hyödyntäminen hinnoittelussa ei ole kuitenkaan helppoa. Thrane, Jarmatz, Fetahi Laursen ja Kornmaaler (2019) lähestyvät tätä haastetta kahdesta näkökulmasta: kirjanpidollisesta ja markkinalähtöisestä. Kirjanpidollisessa näkökulmassa tuodaan esiin kuinka liiallinen kustannuksiin keskittyminen hinnoittelussa voi johtaa todennäköisemmin hinnoittelun epäonnistumiseen, jos tällöin unohdetaan markkinoiden tarjoama tieto, kuten kilpailijoiden hinnoittelu, ja mahdollisuus oppia siitä. Näkökulmassa keskustellaan myös myyntiyksikön roolista hinnoittelussa. Vaikka heillä on todennäköisesti runsaasti arvokasta tietoa asiakkaista hinnoittelun tueksi, saattavat heidän omat intressinsä olla eriävät

strategisista tavoitteista ja näin heidän voimakas integraationsa hinnoitteluprosessiin pitäisi kyseenalaistaa. Markkinanäkökulma keskittyy ulkoisen informaation hyödyntämiseen ja sen tuomiin haasteisiin. Esimerkiksi asiakasperusteinen tieto on usein vaillinaista ja epävarmaa johtuen muun muassa siitä, että kuluttajat ovat harvoin valmiita kertomaan todellista maksuvalmiuttaan. (Thrane et al. 2019, 117–119) Kuluttajapohjainen tieto on kuitenkin hinnoittelussa erittäin tärkeää ja hyvin prosessoituna sillä on suurta hyötypotentiaalia. (Ingenbleek & Van Der Lans 2013).

Toinen merkittävä osa organisaatioiden päätöksentekoa ovat investointipäätökset. Investointi tarkoittaa jonkin yritystoimintaan liittyvän hyödykkeen kuten tehtaan, laitteen, kiinteistön, teknisen järjestelmän tai uuden toimintatavan hankkimista tulevaisuuden tuottojen varmistamiseksi. Investointien valinnan merkitystä kuvaavat niiden pitkä kesto, suuri oman pääoman sitoutuminen ja kannattavuuden epävarmuus. Investoinnit ovat merkittäviä pitkän ajan satsauksia, jotka voivat luoda hyvinkin tarkat tulevaisuuden raamit yrityksen liiketoiminnan suuntaviivojen kannalta. (Ikäheimo et al. 2019, 202–204) Suuri pääoman sitoutuneisuus korostaa investointipäätösten huolellista valmistelua: epäonnistuessaan, ne muuttuvat helposti pelkäksi rasitteeksi. Epävarmuus liittyy hankalaan tulevaisuuden ennakointiin: Koska investoinnit suuntautuvat pitkälle tulevaisuuteen ja tulevaisuuden tapahtumat ovat epävarmoja, on investoinnin lopullisten vaikutusten tarkka ennustaminen hyvin haasteellista. (Ikäheimo et al. 2019, 204–205) Tähän liittyy erityisesti tulevaisuuden tuottojen ennustaminen. Koska investointien vaikutus organisaatioon on suuri, on niiden valinta hyvin monivaiheinen prosessi, joka sisältää eri vaihtoehtojen etsimistä, esittämistä ja analysointia ennen lopullista valintaa (Schuster et al. 2021, 10).

Kun lopullista investointipäätöstä pohditaan, on informaatiolla keskeinen rooli. Johdon laskentatoimen kirjallisuus korostaa erityisesti investointilaskelmien merkitystä investointipäätösprosessin kannalta. Investointilaskelmat koostuvat investoinnin alkupääomasta, vuotuisista tuloista ja menoista, jäännösarvosta sekä investoinnin ajallisesta pituudesta. Tämän avulla taloudellinen näkökulma tuodaan mukaan päätöksentekoon. (Ikäheimo et al. 2019, 208) Myös ei-rahamääräisellä informaatiolla on rooli eri investointivaihtoehtoja arvioitaessa. Esimerkiksi tulevaisuuden tuottoja arvioitaessa, informaatio ulkoisesta toimintaympäristöstä kuten asiakkaiden kulutuspreferenssien muutoksista, kilpailijoiden toiminnasta ja yleisen taloustilanteen kehittymisestä ovat merkityksellisiä (Epstein & Lee 2006, 98).

4 Sosiaalisen median data ja johdon laskentatoimi

Luvussa käsitellään, kuinka sosiaalisen median datan hyödyntämisestä on kirjoitettu tieteellisissä artikkeleissa ja minkälaisia yritys esimerkkejä aiheeseen liittyen löytyy. Teorian ja yritys esimerkkien tulokset tuodaan lopuksi yhteen ja kootaan taulukkoon.

4.1 Teoreettinen viitekehys

Arnaboldi, Busco ja Cuganesan (2017b) tutkivat sosiaalisen median suhdetta rahoituksen ja johdon laskentatoimeen sekä jakoivat tutkimuksensa perusteella aiheen kolmeen näkökulmaan. Ensimmäisessä näkökulmassa sosiaalinen media on datan lähde, jonka tuomalla informaatiolla on mahdollista luoda uusia tehokkuusindikaattoreita laskentatoimen työhön. Agostinon ja Sidrovan (2016) mukaan tehokkuusindikaattorit voivat olla esimerkiksi taloudellisia (*financial*), tunteisiin perustuvia (*sentimental*) tai sisällöllisiä (*content*). Tutkimus lähtee liikkeelle asetelmasta, jossa sosiaalisen median vaikutuksia organisaatioiden toimintaan mitataan tilanteessa, jossa organisaatiot käyttävät sosiaalista mediaa liiketoiminnallisiin tarkoituksiin.

Taloudelliset indikaattorit liittyvät sosiaaliseen mediaan investoimisen tuomiin taloudellisiin hyötyihin. Esimerkkinä käytettiin sosiaalisen median markkinointiin investoidun pääoman tuottoastetta. Tunteisiin perustuvat indikaattorit määrittelevät sosiaalisen median keskustelujen sävyn (positiivinen vai negatiivinen) vaikutusta tiettyyn tahoon. Esimerkkitietona käytettiin suhdelukua, joka kertoo positiivis- ja negatiivissävyytteisten twiittien määrän keskinäisen suhteen. Sisällölliset indikaattorit puolestaan määrittelevät sosiaalisen median keskustelujen aiheita. Tutkimuksessa määritettiin yhdeksi mittariksi tietyn sanan tai lauseen ai-
nutlaatuisuus. Mittaus voidaan toteuttaa esittämällä sosiaalisen median viestejä sanavekto-
reina ja mittaamalla, mitkä sanat esiintyvät kirjoituksissa useimmiten ja näin päätellä kirjoi-
tusten aiheita. (Agostino & Sidrova 2016, 41–42) Vaikutusten mittaaminen on tärkeää,
koska sosiaalisen median data nähdään organisaatioiden kannalta arvokkaana tehokkuuden
arvioimisessa, liiketoiminnan kasvattamisen tukemisessa sekä kannattavuuden parantami-
sessa (Agostino & Sidrova 2016, 39). Toinen näkökulma käsittelee, kuinka sosiaalisen me-
dian tarjoaman datan informaatiota voidaan suojata. On tärkeää pystyä hallitsemaan

sosiaalisesta mediasta kerättyä tietoa ja varmistaa, että se on suojattu kilpailevilta toimijoilta riittävän hyvin mutta samalla mahdollistaa sen jakaminen relevanttien sidosryhmien kesken niin organisaation sisällä kuin sen ulkopuolellakin (Arnaboldi et al. 2017b, 768). Kolmas, eli viimeinen, näkökulma keskittyy siihen, kuinka sosiaalisen median data luo perinteistä poikkeavan tietolähteen ja vaikuttaa päätöksentekoprosessiin toisaalta parantamalla rationaalista päätöksentekoa uutta tietoa tarjoamalla mutta samalla aiheuttamalla myös kasvavaa epävarmuutta datan suuren määrän ja sen sisältämän ei-relevantin tiedon takia (Arnaboldi et al. 2017b, 769).

Arnaboldi, Azzone ja Sidorova (2017a) keskustelevat artikkelissaan sosiaalisen median datan tuovan potentiaalista hyötyä organisaatioiden eri toimijoille tarjoamalla informaatiota liittyen asiakkaisiin ja kilpailijoihin. Näkemystä avataan rajaobjektiteorian kautta (*boundary objects theory*), jonka ideana on yhdistää organisaation eri toimijoita yhteen luomalla molemminpuolista hyötyä rajaobjektin avulla. Esimerkkinä käytettiin markkinoinnin ja viestinnän yksiköjä, jotka ovat perinteisesti hyödyntäneet sosiaalisen median informaatiota: kun organisaation sisällä on näin jo luodut työkalut ja toimintatavat datan keräämiseen, on muiden sisäisten tahojen mahdollista hyödyntää näitä valmiita rakenteita ja toimintatapoja omassa työssään ja näin nopeuttaa uuden tietolähteen implementointia verrattuna tilanteeseen, jossa lähdettäisiin liikkeelle tyhjästä (Arnaboldi et al. 2017a, 821–823). Tutkimuksen mukaan kiinnostus muiden organisaation sisäisten toimijoiden kesken liittyen sosiaalisen median dataan on kasvanut. Johdon laskentatoimen kuvattiin olevan osa näitä toimijoita, mutta tietoa liittyen käytännötason hyödyntämiseen johdon laskentatoimessa on vielä vähän ja uutta tutkimusta tarvitaan lisää (Arnaboldi et al. 2017a, 825).

Bhimani (2020) mainitsee artikkelissaan sosiaalisen median datan kuuluvan niin sanottuun jäljitysdataan (*trace data*), joka voi liittyä esimerkiksi klikkausten määrään ja käyttäjien kirjoituksiin sosiaalisen median alustoilla. Tällä ei nähdä olevan suoraa linkitystä taloudellisiin vaikutuksiin, sen sijaan roolin nähdään olevan ennemminkin perinteistä taloudellista dataa tukeva. (Bhimani 2020, 15) Johdon laskentatoimen kannalta sosiaalisen median datan avulla pyritään keräämään tietoa alustojen käyttäjistä jäljittämällä heidän kommunikointiaan ja tapaa, jolla he palvelua käyttävät. Tämän datan avulla pyritään varmistamaan, onko operatiiviseen toimintaan liittyneen datan perusteella tehdyt johtopäätökset ja linjat päätöksenteossa oikeat vai pitäisikö suuntaa muuttaa. Tällä tavalla voidaan vahvistaa kokonaiskuvaa liittyen

siihen, ollaanko menossa oikeaan suuntaan organisaation suurien tavoitteiden näkökulmasta (Bhimani 2020, 16).

Appelbaum et al. (2017) esittävät artikkelissaan sosiaalisen median datan tehtäväksi johdon laskentatoimen kannalta ohjailta päätöksentekoa parantamalla ennusteita ja helpottamalla valintojen tekemistä. Sosiaalista mediaa voidaan käyttää esimerkiksi optimaalisen markkinoitibudjetin määrittelemisessä sekä ehkäisemään riskiä resurssien suuntaamisesta väärille markkinasegmenteille (Appelbaum et al. 2017, 32). Artikkelin kuvaava sosiaalisen median tiedon mahdollistavan esimerkiksi liiketoimintaympäristön ja taloudellisten olosuhteiden muutoksiin perustuvien mallien uudelleen arvioimisen. Artikkelissa tuodaan myös esiin näkökulma sosiaalisen median datan roolista raaka-ainetoimittajien valintaan liittyen. Monet teollisuusyritykset kohtaavat haasteita raaka-ainetoimittajan löytämiseksi hyvällä hinta-laatusuhteella. Hyödyntämällä sosiaalisen median dataa tässä prosessissa, voidaan valintaa helpottaa (Appelbaum et al. 2017, 36). Artikkelin ei suoraan avannut syitä tähän mutta näkisin sosiaalisen median datan mahdollistavan tiedon keräämisen toimittajien laadusta ja maineesta, joita vertaamalla heidän hinnoitteluunsa, on mahdollista suorittaa kokonaisvaltaisempaa vertailua toimittajien kesken ja tehdä näin luotettavampia, tietoon perustuvia, päätöksiä.

Myös Jeong, Yoon ja Lee (2019) demonstroivat sosiaalisen median datan roolin tärkeyttä organisaatioiden päätöksenteon kannalta. Tutkimuksessa kerättiin arvosteluja liittyen Samsung -merkkiseen puhelimeen sosiaalisen median keskustelualusta Redditistä. Luokittelemalla havaintoaineistoa, tutkimuksessa onnistuttiin keräämään käyttäjien palautetta liittyen puhelimen huonoon akunkestoon, kosketuksen tunnistamiseen sekä muun muassa kameraan. Lisäksi tutkimuksessa mitattiin käyttäjien yleistä tyytyväisyyttä liittyen puhelimen eri ominaisuuksiin sekä niiden tärkeyteen käyttäjien näkökulmasta hyödyntämällä mielipideanalyysiä ja mahdollisuusalgoritmia (*opportunity algorithm*). Mahdollisuusalgoritmin avulla määritettiin, missä määrin tuotteen eri ominaisuuksilla oli käyttäjien näkökulmasta, asteikolla 1–10, parantamisen varaa. (Jeong et al. 2019, 283–287) Edellä mainitulla informaatiolla on mielestäni mahdollista tukea esimerkiksi yrityksen budjetointia: kun tiedetään, mitkä ominaisuudet tuotteessa ovat asiakkaiden mielestä tärkeitä, mutta eivät tarpeeksi hyvällä tasolla, on luonnollista suunnata budjettia näiden ominaisuuksien kehittämiseen tuotekehityksessä.

Jeong et al. (2019) tutkimuksessa hyödynnettiin sosiaalisen median data-analyysin viitekehityksen vaiheita sekä tekstilouhintaprosessia. Aluksi valittiin datan lähde ja tyyppi. Tämän

jälkeen raakadataa kerättiin avainsanaperusteisesti tekstilouhintaa hyödyntämällä. Dataa luokiteltiin tunnistamalla aiheita ja niiden tärkeyttä, joista puhelimen ominaisuuksiin liittyen keskusteltiin sosiaalisessa mediassa. Tässä hyödynnettiin tunneanalyysiä. Kerättyä dataa analysoitiin lopuksi edellä mainitun mahdollisuusalgoritmin avulla ja saadut tulokset pisteytettiin. (Jeong et al. 2019, 282–283) Myös Uyar (2021) mainitsee artikkelissaan sosiaalisen median datan merkityksestä budjetoinnissa, koska sitä hyödyntämällä on mahdollista vähentää riskiä resurssien allokoinnista väärille markkinasegmenteille. Sosiaalisen median datan avulla on myös mahdollista saada tietoa kuluttajien kulutustottumusten muutoksista ja markkinoiden uusista trendeistä, joita hyödyntämällä johdon laskentatoimen on mahdollista tehdä realistisempia analyysyjä niin liiketoiminnan suunnitteluun kuin kustannusten hallintaan liittyen (Uyar 2021, 375).

Askar, Aboutabl ja Galal (2022) käsittelevät artikkelissaan sosiaalisen median datan hyödyntämistä pankkialan yritysten näkökulmasta. Artikkelissa tuodaan esiin, kuinka sosiaalisen median datasta on mahdollista saada hyödyllistä tietoa strategiasuunnitteluun. Keräämällä keskusteluita liittyen muiden kilpailevien pankkien tarjoamiin palveluihin ja tuotteisiin, on mahdollista saada yksityiskohtaista tietoa siitä, mitä mieltä potentiaaliset asiakkaat ovat kilpailijoiden tarjonnasta ja mitä näistä kannattaisi implementoida omaan toimintaan parantaakseen asiakastyytyväisyyttä ja asemaa markkinoilla. Artikkelissa painotetaan myös sitä, kuinka siirtymä kohti sosiaalisen median datan hyödyntämistä vaatii uusia analytiikkatyökaluja, jotta suuria määriä dataa eri lähteistä ja eri muodoissa on mahdollista kerätä ja analysoida (Askar et al. 2022, 409, 412).

Saito ja Gupta (2022) tutkivat artikkelissaan kolmea erilaista talousjohtamisessa hyödynnettävää teoreettista mallia ja pohtivat, tuoko sosiaalisen median datan lisääminen näihin malleihin lisäarvoa. Ensimmäinen teoreettinen malli oli tuottojohtamismalli (*revenue management model*) hotelliliiketoiminnan näkökulmasta. Mallin ideana on, että hotellien päivittäistä liikevaihtoa arvioidaan perustuen historialliseen varausdataan, jonka perusteella huoneelle pyritään asettamaan mahdollisimman optimaalinen päivän hinta tuottojen maksimoimiseksi. Tutkimuksen mukaan sosiaalisen median muuttujan arvon parantuessa, esimerkiksi positiiivisten sanojen määrän kasvaessa yritykseen liittyvissä sosiaalisen median postauksissa, todennäköisyys huonevarausten tekemiseen sekä siksi myös odotetun myynnin määrä kasvaa. Näin sosiaalisen median datan mukaan ottaminen liikevaihdon hallintamalliin nähtiin kannattavaksi ja sosiaalisen median muuttujan arvoja mittamalla, voitiin arvioida paremmin

optimaalista huonehintaa (Saito & Gupta 2022, 4–6). Toisena mallina tutkimus esitteli korkotasomallin (*interest rate model*), jonka ideana oli arvioida markkinakorkoja keräämällä talousaiheisia uutisia ja niistä saatua tietoa tulevaisuuden talousnäkymistä (pessimistisiä tai optimistisiä). Mallia kerrottiin voivan hyödyntää esimerkiksi pääomasijoitusyhtiöissä, joiden toimintaan markkinakorot vaikuttavat merkittävästi: tällaiset organisaatiot sijoittavat usein esimerkiksi valtionlainoihin, joiden hinta on vahvasti sidoksissa yleisiin markkinakorkoihin, joten niissä esiintyvät muutokset ovat tärkeä ottaa huomioon arvioitaessa sijoitusstrategioita (Saito & Gupta 2022, 8). Tutkimuksessa todettiin, että perinteisten talousuutisten lisäksi sosiaalisen median alustat, erityisesti Twitter, ovat yleisiä talouskeskustelun alustoja ja muun taloudellisen informaation julkistuspaikkoja, joten myös niiden huomioiminen on yleistä korkotasoa arvioitaessa tärkeää (Saito & Gupta 2022, 10).

Viimeisenä mallina tutkimus esitteli ultranopean kaupankäynnin mallin rahoitusmarkkinoilla (*high-frequency trading equity market model*). Ultranopea kaupankäynti tarkoittaa rahoitusinstrumenteilla käytävää kauppaa, jossa kaupankäynti toteutetaan tietokoneohjelmien avulla automaattisesti hyvin nopeissa, jopa sekunnin kymmenesosien, sykleissä (Chen 2021). Yleensä tämä on osa suuret resurssit omaavien sijoittajien, kuten sijoituspankkien ja institutionaalisten sijoittajien, sijoituskenttää ja suuressa roolissa heidän liiketoiminnassaan. Perinteinen malli pyrkii arvioimaan kaupankäynnin kohteena olevien osakkeiden hintoja perustuen lainsäädännön muutoksiin ja sen aiheuttamiin vaikutuksiin sijoituskäyttäytymisessä markkinoilla. (Saito & Gupta 2022, 11–12) Saito ja Gupta (2022) mukaan myös sosiaalisen median dataa voidaan käyttää sijoituskäyttäytymisen arvioinnissa. Esimerkiksi sijoittajien postaukset liittyen tulevaisuuden näkymiin voivat olla merkityksellisiä arvioidessa sijoituskäyttäytymistä ultranopeassa kaupankäynnissä ja sen vaikutuksia osakkeiden hintaan. Poliitikkojen ilmoitukset ja sosiaalisen median verkostovaikutukset äänestyskäyttäytymisessä voivat muuttaa paikallista poliittista linjaa ja näin vaikuttaa sijoittamisen säännöksiin. Siksi tutkimus korostaa sosiaalisen median datan implementoimisen merkitystä institutionaalisten sijoittajien talousjohtamiseen arvioitaessa osakkeiden hinnan muutoksia ultranopeassa kaupankäynnissä ja sopivien sijoitusstrategioiden löytämistä. (Saito & Gupta 2022, 11)

Duan, Vasarhelyi, Codesso ja Alzamil (2023) käsittelevät tutkimuksessaan sosiaalisen median datan hyödyntämistä valtion laskentatoimen informaatiojärjestelmissä. Tutkimuksen tarkoituksena oli kerätä dataa Facebookista ja Twitteristä liittyen New Yorkin siisteyteen. Twitteristä kerättiin paikkatietoja sisältäviä twiittejä ja Facebookista postauksia sekä

kommentteja New Yorkin eri asukasryhmien ryhmistä. Ensin dataa karsittiin poistamalla turhat täytesanat kuten pisteet ja välimerkit. Datan keräyksessä käytettiin määritteinä postauksen ajankohtaa, käyttäjien tunnistenumeroa, seuraajien määrää sekä tykkäysten määrää postauksissa. Merkityksettömän datan poistamiseksi määritettiin avainsanoja, joiden perusteella karsittiin. Datan prosessointi toteutettiin hyödyntämällä koneoppimisalgoritmeja, joiden avulla datasta saatiin karsittua vielä paremmin tutkimuksen aiheeseen liittymättömät postaukset. Varsinainen analyysi toteutettiin tunneanalyysin avulla. (Duan et al. 2023, 5–8) Tutkimuksen tulos oli, että kerätyn datan yleisin sävy oli negatiivinen. Yksityiskohtaisemman tiedon tuottamiseksi, kerätyt datasetit jaettiin neljään ryhmään perustuen kaupungin siisteyttä tarkasteleviin osa-alueisiin: kadut, metro, kodittomat ja parkkeeraus. Kodittomat viittasivat tässä New Yorkin kodittomiin asukkaisiin, jotka aiheuttavat artikkelin mukaan siisteysongelmia kaupungissa. Valtaosa datasta liittyi katujen siisteyteen. (Duan et al. 2023, 9–11)

Tutkimuksen perusteella sosiaalisen median dataa voidaan hyödyntää kuntien reaaliaikaisessa raportoinnissa ja kulujen sekä resurssien allokoinnissa budjetointia varten. Artikkelin mukaan moni kunta perustaa budjetointipäätöksensä juuri tuottamiensa palveluiden laadun ja tehokkuuden maksimointiin, jolloin on tärkeää tunnistaa kohteet, jotka vaativat panostamista (Duan et al. 2023, 13). Duan et al. (2023) tutkimuksessa hyödynnettiin tekstilouhinta-prosessia ja sosiaalisen median data-analyysin viitekehystä. Koska tutkimusasetelma oli tiedossa, oli datan lähde valmiiksi annettu. Lähestymistapana datan keräyksessä käytettiin paikkatietoinformaatiota ja kerätty data oli järjestäytymätöntä. Aluksi dataa karsittiin edellä kuvatulla tavalla, jonka jälkeen sitä luokiteltiin koneoppimisen avulla. Lopuksi dataa analysoitiin hyödyntämällä tunneanalyysiä. (Duan et al. 2023, 5–7)

Cui, Gallino, Moreno ja Zhang (2018) suorittivat teoreettisen tutkimuksen, jonka avulla demonstroitiin, kuinka hyödyntämällä Facebook-dataa päivittäisen myynnin ennustamisessa, voidaan merkittävästi vähentää ennusteen suhteellista virheprosenttia ja näin parantaa budjetointia. Tutkimuksessa perinteisen operatiivisen datan rinnalle kerättiin sosiaalisen median tekstidataa postauksien ja kommenttien muodossa yhteistyöyrityksen Facebook-sivuilta. Kerätystä datasta analysoitiin sen informatiivisuutta määrittelemällä, paljonko uniikkeja sanoja datasetit sisälsivät. Data jaettiin myös ryhmiin sen mukaan, oliko tekstin sävy positiivinen, neutraali vai negatiivinen. (Cui et al. 2018, 1753–1754) Datan keräämisen ja analyysin jälkeen tutkimuksessa toteutettiin vertailuennusteet myyntiin liittyen, joista toisessa

hyödynnettiin vain perinteistä operatiivista dataa ja toisessa myös sosiaalisen median dataa. Perinteiseen operatiiviseen dataan pohjautuva ennuste perustui matemaattiseen kaavaan, joissa tietyn päivän myyntimäärät olivat edellisviikon toteutuneiden myyntien, promootioiden ja mainonnan funktioita. Sosiaalisen median dataa hyödyntävässä ennusteessa alkupe-
räiseen malliin lisättiin sosiaalisen median muuttuja esimerkiksi kommenttien keskimääräinen sävy tai esiintyvien sanojen määrä. Lopullisten ennustustulosten muodostamiseksi tutkimuksessa sovellettiin koneoppimisalgoritmeja, joissa hyödynnettiin muun muassa tuki-
vektorimallia. Alkuperäisiä matemaattisia pohjamalleja sovellettiin näihin algoritmeihin, joiden avulla lopulliset myyntiennusteet muodostettiin.

(Cui et al. 2018, 1754–1758) Tutkimuksen perusteella lisäämällä sosiaalisen median muuttuja malliin, myyntiennusteen tarkkuus kasvoi merkittävästi. Tutkimuksen johtopäätöksissä syinä tähän esitettiin muun muassa seuraavaa: kun käyttäjä kommentoi yrityksen Facebook-sivulle, kommentti näkyy myös kyseisen käyttäjän ystävien Facebook-sivulla. Jos kommentti on positiivinen, voi kommentin näkeminen saada muut käyttäjät kiinnostumaan yrityksestä sekä sen tuotteista ja näin johtaa lisääntyneeseen myyntiin. Toisaalta negatiivinen kommentointi voi aiheuttaa päinvastaisen reaktion. (Cui et al. 2018, 1759–1760)

4.2 Case-esimerkkejä

Demonstroidakseen tekstianalyysin mahdollisuuksia asiakasdatan keräämisen kannalta, **Beyond The Arc** (2011), yritysjohtamisen konsulttiyritys, keräsi sosiaalisen median palveluista, Facebookista ja Twitteristä, kommentteja Bank of America - nimiseen yhdysvaltalaiseen pankkiin liittyen. Casessa onnistuttiin keräämään yhteensä 41 000 kommenttia, jotka karsittiin 9000 uniikkiin kommenttiin, joista 88 % olivat lyhyitä, tunteisiin perustuvia twiittejä ja loput 12 % pidempiä Facebook-kommentteja. Hyödyntämällä tekstilouhintaprosessia ja rahoitusalan tuntemusta, konsulttiyritys onnistui löytämään keskeisiä teemoja, jotka liittyivät asiakkaiden huoliin pankin palvelutasosta. Yksi merkittävimmistä teemoista oli asiakkaiden valitukset teknisiin ongelmiin liittyen kuten ongelmat uusien pankkikorttien lähettämisisä edellisten vanhennuttua, rahaliikenteen viivästyksset sekä virheelliset tuplaveloitukset. Näiden kuvattiin aiheuttavan negatiivisuutta asiakkaissa ja lisäävän todennäköisyyttä pankin vaihtoon (Beyond the Arc, 2011).

Johdon laskentatoimen kannalta tällaisen informaation kerääminen voisi olla merkityksellistä erityisesti budjetoinnin kannalta ja indikoisi, että varoja täytyisi suunnata enemmän järjestelmien kehittämiseen ja ylläpitoon, jotta ne toimisivat moitteettomasti suurellakin kuormituksella ja mahdolliset häiriöt voitaisiin korjata nopeasti ilman, että asiakkaille koituisi tästä kohtuutonta haittaa. Myös tilastoimalla saadun palautteen määrän jakautumista positiiviseen ja negatiiviseen palautteeseen esimerkiksi tunneanalyysin avulla sekä vertaamalla tätä toteutuneeseen asiakaspoistumaan, voisi olla mahdollista luoda ennustetta asiakkaiden määrän kehitymisestä sosiaalisen median kommentointia seuraamalla. Tätäkin voisi hyödyntää budjetoinnissa, kun arvioidaan, paljonko keskimääräisesti yksi asiakas tuottaa pankille rahaa tietyllä aikavälillä.

Agostino ja Sidrova (2017) toivat esiin laskentatoimen kirjallisuuden artikkelissaan yritys-esimerkkinä tietoliikennepalveluita tarjoavan **Zeus**-nimisen yrityksen. Esimerkki käsitteli sosiaalisen median datan hyödyntämistä johdon laskentatoimessa. Vuonna 2014 Zeus päätti avata uuden organisaatioyksikön, jonka tarkoituksena oli hyödyntää sosiaalisen median dataa aikaisempaa laajemmin yrityksen toiminnassa. Yksikkö oli pyhitetty pelkän sosiaalisen median datan analysointiin. Yksiköllä oli oma valvontakeskus, joka koostui näytöistä, joissa esitettiin sosiaalisen median dataa Twitteristä ja Facebookista sekä mittareita, joiden tarkoitus oli tuoda dataa helpommin seurattavaan muotoon. Sosiaalisesta mediasta kerätty data koostui yksittäisistä keskusteluista yrityksen Facebook-sivuilta, Twitter-sivun seurannasta, merkittävistä paikallisuutisista ja kilpailija-analyysistä liittyen heidän omiin postauksiinsa ja asiakkaiden kommentteihin. Seurattavia mittareita oli yhteensä neljä: verkkotunnuksen sijoitus (*domain ranking*), tunnearvosana (*sentiment score*), tavoitettavuus (*reach*) ja sitoutuvuus (*engagement*). Verkkotunnuksen sijoitus listasi aiheita, joista keskusteltiin eniten sosiaalisessa mediassa. Tunnearvosana jakoi jokaisen kommentin sävyiltään, joko neutraaliin, positiiviseen tai negatiiviseen ryhmään. Tavoitettavuusmittari indikoi ihmisten määrää, jotka olivat nähneet postauksia sosiaalisessa mediassa ja sitoutuvuus määrittäi, kuinka paljon huomiota nämä postaukset olivat saaneet (kommenttien määrä Facebook postauksissa ja Twitter twiittien uudelleenjakomäärät). (Agostino & Sidrova 2017, 784) Näiden mittareiden avulla Zeus analysoi sosiaalisen median keskusteluja ymmärtääkseen lisää asiakkaista ja heidän tarpeistaan sekä luodakseen ennusteita ja ennakoidakseen trendejä (Agostino & Sidrova 2017, 785).

Zeuksen edustajan mukaan erityisen tärkeää oli tehdä ennusteita siitä, milloin asiakkaat ovat valmiita uuteen tarjoukseen sekä analysoida kilpailijoiden vastatarjouksia. Suurimman haasteen yritykselle toi datan kerääminen ja analysointi sosiaalisesta mediasta, sillä käsiteltävä data sisälsi myös paljon merkityksetöntä dataa. (Agostino ja Sidrova 2017, 785) Asiakkaiden tarpeisiin ja mielipiteisiin perustuva data voisi Zeuksen tapauksessa parhaiten tuoda lisäarvoa strategiasuunnittelussa, kun mietitään toimintatapoja tarjousten lähettämissykleistä niin olemassa oleville kuin myös potentiaalisille asiakkaille. Keräämällä tietoa kilpailijoiden tarjouksista seuraamalla heidän postauksiaan sosiaalisessa mediassa, voidaan kerättyä informaatiota hyödyntää myös omien tarjouksien sisältöä ja hintaa arvioitaessa. Toisaalta edellä mainittuja sosiaalisen median seurantamittareita voisi mahdollisesti hyödyntää myös strategian toteutumisen arvioinnissa analysoitaessa asiakastyytyvää asiakastytyvää organisaation tarjoamista tuotteista ja palveluista.

Arnaboldi et al. (2017a) toivat artikkelissaan osana laskentatoimen tiedejulkaisuja esiin tapauksitutkimukset kahdesta organisaatiosta, jotka hyödyntävät sosiaalista mediaa sisäisesti eri osastojen kesken. Ensimmäisessä esimerkissä käsiteltiin voittoa tuottamatonta järjestöä nimeltä **Art**, joka tuottaa taidetta suuressa italialaisessa kaupungissa. Art oli perustanut sosiaalisen median tilit Facebookiin, Twitteriin ja YouTubeen jo 2010 - luvun taitteessa mutta aluksi sosiaalisen median hyödyntämiseen ei kiinnitetty suurta huomiota. Vasta myöhemmin sosiaalisen median potentiaalinen positiivinen vaikutus organisaation toiminnan kannalta huomattiin, jonka seurauksena yritys päätti investoida valvontajärjestelmään, jolla seurattiin Artin sosiaalisen median sivujen kommenttien ja tykkäysten määriä sekä keskustelun tunnetasoa (Arnaboldi et al. 2017a, 830).

Sosiaalisesta mediasta kerättyä dataa hyödynnettiin erityisesti tukemaan suunnittelua: Artilla oli käytössä digitaalinen suunnitteludokumentti, joka sisälsi kahden vuoden suunnitelman toimintaan liittyen. Tähän kuuluivat ennusteet tulevasta liikevaihdosta ja myytyjen lippujen määristä sekä näiden lopulliset toteumat, joita kaikkia oli mahdollista vertailla graafisesti keskenään. Artin julkaisemat postaukset sosiaalisessa mediassa esitettiin samalla kartalla, jossa kuvattiin myös toteutuneiden myyntien määrää maantieteellisesti tarkoituksena seurata sosiaalisen median kampanjoiden vaikutusta toteutuneeseen myyntiin ja parantaa ennustettavuutta. Tällä haluttiin tutkia, minkälaista materiaalia Artin kannattaisi julkaista sosiaaliseen mediaan edistääkseen sillä myyntiä. Uuden sosiaalisen median valvontatyökalun avulla saatiin esimerkiksi linkitettyä ostettujen lippujen määrä maksettuihin

kampanjapostauksiin perinteisen sosiaalisen median aktiivisuuden seurannan lisäksi (Arnaboldi et al. 2017a, 832). Tämän valvontatyökalun kehittämisen tavoitteena oli mahdollistaa sosiaalisen median datan hyödyntäminen laajemmin organisaation sisällä pelkän markkinoinnin lisäksi esimerkiksi budjetoinnissa: datan avulla pyrittiin arvioimaan eri lipputyypien myyntiä ja näin parantamaan tulevien tuottojen ennustettavuutta (Arnaboldi et al. 2017a, 833).

Toinen esimerkki käsitteli italialaista energia-alan yritystä nimeltä **Utility**. Yritykseen perustettiin sisäinen tiimi, jonka yhtenä tarkoituksena oli hallita yrityksen uusia sosiaalisen median tilejä eri alustoilla kuten Facebookissa ja Twitterissä (Arnaboldi et al. 2017a, 834). Toisena tavoitteena tiimillä oli hallita ja kerätä informaatiota sosiaalisesta mediasta liiketoiminnallisiin tarkoituksiin. Kuten Artin tapauksessa, myös Utilityssä hankittiin, automaatioita hyödyntävä, sosiaalisen median valvontatyökalu, jolla pystyttiin keräämään ja analysoimaan sosiaalisen median dataa, kuten tekstiä, linkittämällä se tiettyyn käyttäjään, sekä muuttamaan dataa tehokkuusmittareiksi. (Arnaboldi et al. 2017a, 835). Näistä mittareista tuotettiin raportteja, joita jaettiin yrityksen eri osastoille kuten henkilöstöhallintoon, tutkimukseen ja kehitykseen sekä taloushallintoon aina tarpeen mukaan. Raporttien arvo tuli esiin erityisesti niiden tarjoamassa konkreettisessa ja jäljitettävässä tiedossa: jokaiselle kerätylle kommentille oli mahdollista löytää julkaisija ja kerätä heistä myös lisätietoa julkisten profiilien sisältämän tiedon avulla. Tämä mahdollisti tarkemman analysoinnin, kun saatiin dataa kokonaisvaltaisempaa tietoa asiakkaiden mielipiteistä verrattuna esimerkiksi anonyymeihin kyselyihin (Arnaboldi et al. 2017a, 835). Sosiaalisen median analysointi koottujen raporttien ja seurantamittareiden avulla voisi tuoda johdon laskentatoimen tehtävissä eniten lisäarvoa juuri strategiaseurannassa, kun analysoidaan yrityksen toiminnan laatua asiakasnäkökulmasta.

4.3 Havaintoja teoriasta ja caseista

Relevantteja artikkeleita teoriaksi ja yritysesimerkeiksi löytyi yhteensä 13 kappaletta. Näistä seitsemän oli vuodelta 2018 tai vanhempia, loput tätä uudempia. Kerätyistä lähteistä viidessä tutkimusmenetelmänä oli käytetty kirjallisuuskatsausta, neljässä tapaustutkimusta, kahdessa kokeellista tutkimusta ja yhdessä haastattelua. Yksi käytetyistä lähteistä oli blogi.

Artikkeleiden haun perusteella, sosiaalisen median data johdon laskentatoimen työssä on aiheena vielä suhteellisen tuore ja tieteellisiä kirjoituksia tähän liittyen on rajattu määrä saatavilla. Löydettyjen artikkeleiden perusteella on kuitenkin todettava, että sosiaalinen media on tietolähteenä relevantti myös johdon laskentatoimen työn kannalta. Sosiaalisen median datalla nähdään selkeä tukeva rooli perinteisen organisaation toimintaan liittyvän taloudellisen datan rinnalla tuomalla uutta tietoa kilpailijoista ja asiakkaista sekä parantamalla näin toiminnan onnistumisen mitattavuutta ja päätöksenteon laadukkuutta. Mahdollisuus verrata uutta dataa sosiaalisesta mediasta perinteiseen johdon laskentatoimessa käytettyyn talousdataan avaa mahdollisuuden uusien vuorovaikutussuhteiden löytämiselle ja näin myös paremmalle liiketoiminnan ennustettavuudelle.

Koneoppimisen ja sosiaalisen median data-analyysin hyödyntäminen tuli esiin vaihtelevasti teoreettisissa artikkeleissa ja yritys esimerkeissä. Teoreettisissa artikkeleissa, joissa oli tehty kokeellista tutkimusta, datan keräys ja prosessointi seurasi kohtuullisen tarkasti sosiaalisen median data-analyysin viitekehystä. Näissä tutkimuksissa datan prosessoinnissa oli hyödynnetty melko yleisesti myös tekstilouhintaprosessia ja vain valitut koneoppimisalgoritmit vaihtelivat. Artikkeleissa, joissa puhuttiin sosiaalisen median datan hyödyntämisestä täysin teoreettisesti, ilman kokeellista tutkimusta, ei juuri näistä mainittu. Kerätyissä yritys esimerkeissä koneoppimista ja sosiaalisen median data-analyysiä ei käsitelty suoraan lukuun ottamatta yhtä esimerkkiä, jossa datan keräämistä oli toteutettu tekstilouhinnan avulla. Muissa esimerkeissä datan keräyksessä mainittiin hyödynnettäneiden automatisoituja järjestelmiä, joten voisi olettaa, että sosiaalisen median datan luonteen vuoksi, myös nämä järjestelmät hyödyntäisivät koneoppimis pohjaisia algoritmeja, vaikka niistä ei suoraan mainittukaan.

Kokonaisuutena voidaan todeta, että koneoppimis pohjaisten algoritmien hyödyntäminen on käyttökelpoinen tapa käsitellä sosiaalisen median dataa analysoitavaan muotoon. Tätä tukee erityisesti se, että teoreettisissa tutkimuksissa sosiaalisen median datan keräys ja prosessointi oli lähes varauksetta toteutettu koneoppimista hyödyntämällä. Sosiaalisen median data-analyysin viitekehys ja tekstilouhintaprosessi soveltuvat erinomaisesti teoreettiseksi tueksi datan prosessointiin. Tutkitusta aineistosta tehdyt havainnot on koottu alapuolella olevaan taulukkoon ja avattu tarkemmin sen jälkeen. Taulukossa käsitellyt johdon laskentatoimen osat alueet on järjestetty suurimmasta päätöstopasta pienimpään: strategia ja investointipäätökset voivat olla ajalliselta kestoaltaan jopa usean vuoden pituisia, kun taas hinnoittelua voidaan

parhaimmassa tapauksessa tehdä jopa päivätasolla. Budjetointi on puolestaan näiden välillä ja sitä tarkastellaan yleensä joitain kertoja vuodessa. (Ikäheimo et al. 2019, 203, 162, 150)

Taulukko 1. Sosiaalisen median datan hyödyntäminen johdon laskentatoimessa teoreettisen viitekehyyksen ja yritysmerkkin perusteella.

Johdon laskentatoimen osa-alueet (joita käsiteltä tässä työssä)	Sosiaalisen median datan hyödyntäminen teoreettisen viitekehyyksen perusteella	Sosiaalisen median datan hyödyntäminen yritysmerkkin perusteella	Koneoppimisen hyödyntäminen datan prosessoinnissa teoreettisen viitekehyyksen perusteella	Koneoppimisen hyödyntäminen datan prosessoinnissa yritysmerkkin perusteella
Strategiasuunnittelu ja -seuranta	<ul style="list-style-type: none"> Tietoa asiakaskäyttäytymisestä ja kilpailijoista oman toiminnan suunnittelemiseksi Seurantamittareiden muodostus tukemaan strategian toteutumisen seuranta 	<ul style="list-style-type: none"> Tietoa asiakaskäyttäytymisestä ja kilpailijoista oman toiminnan suunnittelemiseksi Seurantamittareiden muodostus tukemaan strategian toteutumisen seuranta 	<ul style="list-style-type: none"> Ei mainittu 	<ul style="list-style-type: none"> Ei mainittu suoraan (mainittiin yleisesti vain automatisoidut järjestelmät)
Investointipäätösten- teko	<ul style="list-style-type: none"> Sijoitusinvestointipäätösten tukeminen arvioimalla yleisiä talousnäkyviä 	<ul style="list-style-type: none"> Ei mainittu 	<ul style="list-style-type: none"> Ei mainittu 	<ul style="list-style-type: none"> Ei mainittu
Budjetointi	<ul style="list-style-type: none"> Asiakkaiden preferenssit resurssien oikeaan kohdentamiseen. Tiedon keräämistä asiakassegmenttien ja markkinointibudjetin määrittämiseen ja kohdentamiseen. Tuottojen budjetointi myyntiä ennustamalla 	<ul style="list-style-type: none"> Asiakkaiden preferenssit resurssien oikeaan kohdentamiseen Tuottojen budjetointi myyntiä sekä asiakasmäärien muutoksia ennustamalla 	<ul style="list-style-type: none"> Koneoppimis- pohjaiset algoritmit datan luokittelussa ja analysoinnissa 	<ul style="list-style-type: none"> Koneoppimis- pohjaiset algoritmit datan luokittelussa ja analysoinnissa
Hinnoittelu	<ul style="list-style-type: none"> Sosiaalisen median datan hyödyntäminen optimaalisen hinnan määrittelyssä seuraamalla keskusteluja 	<ul style="list-style-type: none"> Kilpailijoiden tarjousten arviointi ja huomointi omassa hinnoittelussa 	<ul style="list-style-type: none"> Ei mainittu 	<ul style="list-style-type: none"> Ei mainittu

Tässä työssä määritettiin aikaisemmin johdon laskentatoimen osa-alueiden jakautuvan kustannus- ja kannattavuuslaskentaan, ohjaukseen sekä päätöksentekoon. Näistä keskityttiin erityisesti johdon laskentatoimen rooliin organisaatioissa ohjauksen ja päätöksenteon alle kuuluvissa budjetoinnissa, hinnoittelussa, investointipäätösten teossa sekä strategiasuunnittelussa ja -seurannassa. Tähän jakoon peilaten, teoreettiset artikkelit painottavat datan roolia erityisesti budjetoinnin osalta. Artikkeleissa painotettiin kuluttajien mielipiteiden keräämistä liittyen olemassa oleviin tuotteisiin ja palveluihin sekä kehitysehdotuksiin mietittäessä,

mihin tuotteisiin ja niiden ominaisuuksien kehittämiseen halutaan suunnata resursseja. Sosiaalisesta mediasta kerätyillä kommentteilla kuvattiin olevan myös mahdollista ennustaa paremmin myyntiä. Asiakkaista kerätyn tiedon avulla kerrottiin voitavan määrittää tarkemmin asiakassegmentit sekä saada parempi ymmärrys ihmisten tämänhetkisestä tietämyksestä yritykseen liittyen ja näin arvioida paremmin markkinointibudjetin suuruuden tarpeellisuutta sekä parasta mahdollista kohdentamista eri kuluttajaryhmien välillä.

Teoreettisten kirjoitusten mukaan, käyttökelpoinen kohde sosiaalisen median datalle johdon laskentatoimen kannalta oli myös strategiasuunnittelu ja -seuranta. Artikkeleissa mainittiin muun muassa sosiaalisen median datan roolista organisaatioiden toimintaan liittyvien seurantamittareiden muodostamisessa asiakastyytyvääsyyttä mitattaessa. Pitkän tähtäimen strategiasuunnittelun kannalta sosiaalisen median mainittiin tarjoavan myös kulutustottumusten muutoksiin liittyvää tietoa sekä mahdollistavan liiketoimintaympäristön ja taloudellisten olosuhteiden muutoksiin liittyvien mallien parantamisen. Lyhyen tähtäimen strategiasuunnitelman eli toimintasuunnitelman kannalta artikkeleissa mainittiin kilpailijoista kerätyn tiedon merkitys sekä tieto siitä, miltä osin potentiaaliset asiakkaat pitävät kilpailijoiden tuotteita parempina. Näiden tietojen merkitys nähtiin erityisesti oman toiminnan suunnittelussa. Myös hinnoittelusta mainittiin muutamissa artikkeleissa, erityisesti hotelliliiketoimintaan liittyvässä esimerkissä, jossa optimaalisen huonehinnan määrittämistä pystyttiin tukemaan seuraamalla käyttäjien kirjoituksia sosiaalisessa mediassa. Investointipäätösten osalta artikkeleissa mainittiin suoraan sosiaalisen median datan hyöty institutionaalisten sijoittajien sijoitusinvestointipäätöksissä, sillä sosiaalisen median postausten avulla oli mahdollista arvioida yleisiä talousnäkyelmiä ja niiden vaikutuksia yleiseen korkotasoon.

Yritysesimerkeissä painotettiin, johdon laskentatoimen osa-alueisiin peilaten, asiakasperusteisen datan hyödyntämistä erityisesti budjetoinnissa arvioitaessa tulevaisuuden myyntituotot lipunmyyntiä sekä asiakasmäärien muutoksia ennustamalla ja resurssien allokoimista asiakaspalautteen perusteella. Sosiaalisen median datan arvo tuli esiin myös strategiasuunnittelussa arvioitaessa kuluttajien valmiuksia vastaanottaa uusia tarjouksia sekä seuraamalla kilpailijoiden postauksia sosiaalisessa mediassa. Esimerkeissä mainittiin myös sosiaalisen median datan hyödyntäminen tehokkuusmittareiden muodostamisessa organisaatioiden sisäiseen raportointiin sekä kilpailijoiden tarjouksien arvioinnissa, joiden perusteella voidaan todeta, että dataa on myös hyödynnetty strategian toteutumisen seurannassa sekä hinnoittelussa.

5 Johtopäätökset

Kirjallisuuskatsauksena laaditun kandidityön tarkoituksena oli selvittää, kuinka sosiaalisesta mediasta saatavaa dataa on mahdollista hyödyntää johdon laskentatoimen tehtävissä ja voiko koneoppimista hyödyntää datan prosessoinnissa. Aihetta lähestyttiin avaamalla sosiaalisen median käsitettä sekä tutkimalla sosiaalisen median datan hyödyntämistä data-analyysimallin avulla. Tämän lisäksi työssä käsiteltiin koneoppimisen käsitettä ja tuotiin esimerkkinä tekstilouhintaprosessi. Työssä määritettiin myös johdon laskentatoimi osa-alueineen. Artikkeleiden ja muiden sähköisten lähteiden perusteella tutkittiin, miten sosiaalisen median datan hyödyntämistä johdon laskentatoimen kannalta on käsitelty teoreettisen viitekehyksen ja yritysmerkien perusteella sekä vertailtiin näistä tehtyjä havaintoja. Työssä vastattiin kolmeen tutkimuskysymykseen, joista ensimmäinen oli:

1. Minkälaista dataa sosiaalinen media tarjoaa johdon laskentatoimeen?

Työn alussa määritettiin, että sosiaalisen median datasta keskitytään nimenomaan tekstipohjaiseen dataan, koska se nähtiin johdon laskentatoimen kannalta kaikista hyödyllisimmäksi. Myös aiheen rajauksen kannalta tällainen jako oli järkevää tehdä. Käsitellessä sosiaalisen median datan hyödyntämistä johdon laskentatoimessa, korostui erityisesti käyttäjien postausten ja kommenttien merkitys, jotka liittyvät käyttäjien mielipiteisiin tietyn organisaation toiminnasta. Myös käyttäjien profiilitiedot sekä kommenttien sijainti- ja aikatiedot voivat olla hyödyllisiä tarkempaa analyysiä tehdessä eri asiakasryhmistä. Sosiaalinen media sisältää runsaasti tekstipohjaista dataa, josta relevantin sisällön löytäminen ilman datan käsittelyä ei ole mahdollista.

Teoriaosiossa tarkasteltiin koneoppimisen teoriaa ja kuinka sitä voisi hyödyntää nimenomaan tekstidatan prosessointiin, joka oli työn kannalta relevanttia. Toisena tutkimuskysymyksenä oli:

2. Kuinka koneoppimista voidaan hyödyntää datan prosessoinnissa?

Koneoppimisen rooli tulee esiin erityisesti datan prosessointi- ja analysointivaiheessa. Koneoppimista hyödyntävien algoritmien avulla tekstidataa on mahdollista luokitella aiheperusteisesti ja näin karsia pois turhat sanat. Teoriaosiossa esimerkkinä tuotiin koneoppimista hyödyntävä tukivektorikonemalli, jolla luokittelua on mahdollista toteuttaa. Aiheluokittelun

jälkeen dataa voisi jatkojalostaa luokittelemalla sitä tunneperusteisesti. Sosiaalisen median data-analytiikkaa käsittelevässä kappaleessa mainittiin tunneanalyysi, jota on mahdollista toteuttaa koneoppimisen avulla. Teoreettisia artikkeleita ja yritysesimerkkejä kerätessä sosiaalisen median datan hyödyntämisestä johdon laskentatoimen työhön liittyen, datan merkittävin arvo nähtiin sen potentiaalisissa tuoda uutta lisätietoa asiakkaiden mielipiteistä liittyen omaan yritykseen sekä kilpailijoihin. Koneoppimisen avulla toteutettu tekstiluokittelu ja tunneanalyysi mahdollistavat sosiaalisen median käyttäjien kommenttien keräämisen ja luokittelun tiettyyn organisaatioon liittyen ja datan tunnetason analysoinnin, jotka nostavat datan informatiivisuutta sekä käyttökelpoisuutta ja mahdollistavat näin sen hyödyntämisen myös johdon laskentatoimessa päätöksenteon tukena.

Artikkeleiden keräysosiossa vertailtiin, kuinka teorian mukaan sosiaalisen median dataa voisi hyödyntää johdon laskentatoimen tehtävissä ja kuinka yritykset tätä ovat jo hyödyntäneet. Viimeisenä eli kolmantena tutkimuskysymyksenä oli:

3. Millä tavoin yritykset hyödyntävät sosiaalisen median dataa johdon laskentatoimessaan?

Kokonaisuudessaan artikkelit painottivat sosiaalisen median datan tukevaa roolia organisaatioiden liiketoiminnallisen datan rinnalla, jota johdon laskentatoimen työssä on perinteisesti hyödynnetty. Teoreettisten artikkeleiden perusteella sosiaalisen median datalla on mahdollista tukea kaikkia johdon laskentatoimen osa-alueita, eniten artikkeleissa mainittiin datan potentiaali budjetin määrittämisessä sekä strategiaseurannassa ja vähiten investointipäätöksissä, joista sivuttiin ainoastaan sijoitusinvestointeja. Yritysesimerkeissä mainittiin sosiaalisen median datan hyödyntäminen erityisesti budjetoinnissa mutta tämän lisäksi myös hinnoittelussa sekä strategiasuunnittelussa ja -seurannassa. Huomionarvoista johtopäätöksissä oli yritysesimerkkien vähyys verrattuna teoreettisiin kirjoituksiin: kerätyistä 13 lähteestä kolme soveltui yritysesimerkeiksi ja loput käsitelivät teoreettista viitekehystä.

Ristiriita teoreettisten artikkeleiden runsauden ja yritysesimerkkien vähyys välillä kuvastaa hyvin sitä, kuinka aihetta on käsitelty. Sosiaalisen median datan hyödyntämisestä johdon laskentatoimessa on kirjoitettu jo kohtuullisesti teoreettisella tasolla, mutta käytännön tason toimintaan se ei vielä yhtä laajasti jalkautunut. Löydetyt yritysesimerkit painottivat budjetointia ja siinä erityisesti myynnin ennustamista, joten voidaan todeta, että tältä osin käytäntöön jalkautumista on jo jonkin verran havaittavissa. Kuten teoreettiset esimerkit osoittavat,

sosiaalisen median datalla on kuitenkin paljon hyödyntämätöntä potentiaalia tukea johdon laskentatoimen työtä tarjoamalla uuden informaatiolähteen organisaatioiden toiminnan kehittämiseksi.

Sytä siihen, miksi sosiaalisen median dataa johdon laskentatoimessa ei ole vielä laajasti hyödynnetty voi johtua monesta asiasta mutta näkisin yhdeksi merkittäväksi syyksi perinteisissä toimintatavoissa pysymisen, joka nostaa kynnyistä uusien ratkaisuiden implementointiin. Arnaboldi et al. (2017a) kuvasivat hyvin tätä ongelmaa tutkimuksensa johtopäätöksissä, joissa mainittiin, kuinka taloushallinnon ammattilaiset ovat melko vastahakoisia sosiaalisen median suhteen omassa työssään ja pitävät sitä ennemminkin vain markkinoinnin ja viestinnän kannalta merkityksellisinä, keskittyen siksi yleensä vain perinteiseen taloudelliseen dataan (Arnaboldi et al. 2017a, 822, 843).

Kun mietitään, kuinka käytännössä sosiaalisen median data voisi parhaiten tukea perinteistä talousdataa johdon laskentatoimen työssä, näkisin näiden keskinäisen integraation merkityksen suureksi. Muuttamalla sosiaalisen median tekstidataa numeerisiksi mittareiksi, joilla seurataan esimerkiksi keskustelujen tunnetasoa sekä merkittävimpiä esiintyviä aiheita ja linkittämällä nämä perinteiseen talousdataan kuten toteutuneeseen myyntiin, voitaisiin paremmin analysoida, löytyykö eri datojen välillä uusia johdonmukaisuuksia kuten muuttuuko myynnin laskiessa jokin sosiaalisen median mittarin arvo tiettyyn suuntaan. Tällä tavalla voitaisiin määrittää, löytyykö sosiaalisen median aktiivisuutta seuraavista mittareista joitain indikaattoreita, joita tilastoimalla voitaisiin ennustaa tulevaisuuden myyntiä perustuen seurantamittareiden arvoihin. Toinen näkökulma voisi olla esimerkiksi optimaalisen hinnan määrittäminen dynaamisesti seuraamalla sosiaalisen median mittareiden arvoja ja hyödyntämällä kerättyä tietoa niiden vaikutuksesta myyntiin, jotta voitaisiin paremmin ymmärtää asiakkaiden reaaliaikainen maksuvalmius ja suhteuttaa hinta tähän. Tätä näkökulmaa sivuttiinkin jo teoreettisessa viitekehysessä (Saito ja Gupta 2022, 4–6).

Työn päätelmien perusteella, sosiaalisen median datan hyödyntämistä johdon laskentatoimen tehtävissä tulisi tutkia vielä lisää. Jatkotutkimuksissa voitaisiin selvittää haastatteleamalla eri alan yritysten talousosastoja, ovatko he hyödyntäneet sosiaalisen median dataa ja miten tai jos eivät niin mikä tähän on syynä. Koska työ toteutettiin kirjallisuuskatsauksena ja perustui julkisten lähteiden varaan, on mahdollista, että haastatteluiden avulla löytyisi uusia yritys esimerkkejä, joissa sosiaalisen median dataa on hyödynnetty johdon laskentatoimessa enemmän kuin työn havainnot osoittavat.

Lähteet

Agostino, D. & Sidorova, Y. (2017) How social media reshapes action on distant customers: some empirical evidence. *Accounting, auditing & accountability journal*. 30, 4, 777–794. [Viitattu 1.8.2023]. Saatavilla: <https://doi.org/10.1108/AAAJ-07-2015-2136>

Agostino, D. & Sidorova, Y. (2016) A performance measurement system to quantify the contribution of social media: new requirements for metrics and methods. *Measuring business excellence*. 20, 2, 38–51. [Viitattu 1.8.2023]. Saatavilla: <https://doi.org/10.1108/MBE-05-2015-0030>

Alpaydin, E. (2020) *Introduction to machine learning*. 4. p. Cambridge, MIT press.

Anuncia, S.M., Gohel, H.A. & Vairamuthu, S. (2020) *Data Visualization: Trends and Challenges Toward Multidisciplinary Perception*. Singapore, Springer.

Appelbaum, D., Kogan, A., Vasarhelyi, M. & Yan, Z. (2017) Impact of business analytics and enterprise systems on managerial accounting. *International journal of accounting information systems*. 25, 29–44. [Viitattu 1.8.2023]. Saatavilla: <https://doi.org/10.1016/j.accinf.2017.03.003>

Arnaboldi, M., Azzone, G. & Sidorova, Y. (2017a) Governing social media: the emergence of hybridised boundary objects. *Accounting, auditing & accountability journal*. 30, 4, 821–849. [Viitattu 1.8.2023]. Saatavilla: <https://doi.org/10.1108/AAAJ-07-2015-2132>

Arnaboldi, M., Busco, C. & Cuganesan, S. (2017b) Accounting, accountability, social media and big data: revolution or hype? *Accounting, auditing & accountability journal*. 30, 4, 762–776. [Viitattu 1.8.2023]. Saatavilla: <https://doi.org/10.1108/AAAJ-03-2017-2880>

Askar, M. A. A., Aboutabl, A. & Galal, A. (2022) USING SOCIAL MEDIA ANALYTICS IN THE BANKING INDUSTRY TO DRIVE BUSINESS ADVANTAGE. *Journal of Southwest Jiaotong University*, 57, 4, 405-415. [Viitattu 1.8.2023]. Saatavilla: <https://doi.org/10.35741/issn.0258-2724.57.4.36>

Atkinson, A.A., Kaplan, R.S., Matsumura, E.M. & Young, S.M. (2012) *Management accounting*. 6. p. Harlow, Pearson Education Limited.

Balodi, T. (2019) An Optimum Approach Towards the Bag of Words with Code Illustration in Python, AnalyticSteps. [Verkkoartikkeli]. [Viitattu 16.8.2023]. Saatavilla: <https://www.analyticsteps.com/blogs/an-optimum-approach-towards-the-bag-of-words-with-code-illustration-in-python>

Bedi, P. & Sharma, C. (2016) Community detection in social networks. Wiley interdisciplinary reviews. Data mining and knowledge discovery. 6, 3, 115–135. [Viitattu 1.6.2023]. Saatavilla: <https://doi.org/10.1002/widm.1178>

Beyond the Arc. (2011) Bank of America insights make a case for social media data analytics. [Verkkoartikkeli]. [Viitattu 1.6.2023]. Saatavilla: <https://beyondthearc.com/blog/2011/customer-experience/bank-of-america-insights-make-a-case-for-social-media-data-analytics>

Bhimani, A. (2020) Digital data and management accounting: why we need to rethink research methods. Journal of management control. 31, 1-2, 9-23. [Viitattu 1.7.2023]. Saatavilla: <https://doi.org/10.1007/s00187-020-00295-z>

Chen, J. (2021) What is High-Frequency Trading (HFT)? How it works and example. [Verkkoartikkeli]. [Viitattu 10.8.2023]. Saatavilla: <https://www.investopedia.com/terms/h/high-frequency-trading.asp>

Chinnov, A., Kerschke, P., Meske, C., Stieglitz, S. & Trautmann, H. (2015) An overview of topic discovery in Twitter communication through social media analytics. Twenty-first Americas Conference on Information Systems, Puerto Rico.

Cui, R., Gallino, S., Moreno, A. & Zhang, D. J. (2018) The Operational Value of Social Media Information. Production and operations management. 27, 10, 1749–1769. [Viitattu 1.7.2023]. Saatavilla: <https://doi.org/10.1111/poms.12707>

Dahal, R.K. (2019) Changing role of management accounting in 21st Century. Review of Public Administration and Management. 7, 3, 1-8.

Duan, H.K., Vasarhelyi, M.A., Codesso, M. & Alzamil, Z. (2023) Enhancing the government accounting information systems using social media information: An application of text mining and machine learning. International journal of accounting information systems. 48, 100600. [Viitattu 1.7.2023]. Saatavilla: <https://doi.org/10.1016/j.accinf.2022.100600>

Endenich, C. & Trapp, R. (2020) Ethical Implications of Management Accounting and Control: A Systematic Review of the Contributions from the Journal of Business Ethics. *Journal of business ethics*. 163, 2, 309–328. [Viitattu 15.6.2023]. Saatavilla: <https://doi.org/10.1007/s10551-018-4034-8>

Epstein, M. J. & Lee, J. Y. (2006) *Advances in management accounting*. Vol. 15. Oxford, JAI Press.

Frezatti, F., Aguiar, A.B., Guerreiro, R. & Gouvea, M.A. (2011) Does management accounting play role in planning process? *Journal of business research*. 64, 3, 242–249. [Viitattu 1.7.2023]. Saatavilla: <https://doi.org/10.1016/j.jbusres.2009.11.008>

Goncalves, M. & Cornelius Smith, E. (2018) Social media as a data gathering tool for international business qualitative research: opportunities and challenges. *Journal of transnational management*. 23, 2-3, 66–97. [Viitattu 1.6.2023]. Saatavilla: <https://doi.org/10.1080/15475778.2018.1475181>

Great Learning. (2022) An Introduction to Bag of Words (BoW) | What is Bag of Words? [Verkkoartikkeli]. [Viitattu 18.8.2023]. Saatavilla: <https://www.mygreatlearning.com/blog/bag-of-words/>

Hassan, S.U., Ahamed, J. & Ahmad, K. (2022) Analytics of machine learning-based algorithms for text classification. *Sustainable Operations and Computers*. 3, 238-248. [Viitattu 1.6.2023]. Saatavilla: <https://doi.org/10.1016/j.susoc.2022.03.001>

Hassani, H., Beneki, C., Unger, S., Mazinani, M.T. & Yeganegi, M.R. (2020) Text Mining in Big Data Analytics. *Big data and cognitive computing*. 4, 1, 1. [Viitattu 1.6.2023]. Saatavilla: <https://doi.org/10.3390/bdcc4010001>

Hayat, M.K., Daud, A., Alshdadi, A.A., Banjar, A., Abbasi, R.A., Bao, Y. & Dawood, H. (2019) Towards Deep Learning Prospects: Insights for Social Media Analytics. *IEEE access*. 7, 36958-36979.

Holtkemper, O. (2020) *Digitization of the management accounting function: a case study analysis on manufacturing companies*. 1. p. Wiesbaden, Springer Fachmedien Wiesbaden.

Hudson, R. (2015) *Decision-making: processes, behavioral influences and role in business management*. New York, Novinka.

IBM. (2023) What is machine learning? [Verkkoartikkeli] [Viitattu 1.6.2023]. Saatavilla: <https://www.ibm.com/topics/machine-learning>

Ikäheimo, S., Malmi, T. & Walden, R. (2019) Yrityksen laskentatoimi. 8. p. Helsinki, Alma Talent Oy.

Ingenbleek, P.T.M & van der Lans, I.A. (2013) Relating price strategies and price-setting practices. *European journal of marketing*. 47, 1/2, 27–48. [Viitattu 1.7.2023]. Saatavilla: <https://doi.org/10.1108/03090561311285448>

Jeong, B., Yoon, J. & Lee, J. (2019) Social media mining for product planning: A product opportunity mining approach based on topic modeling and sentiment analysis. *International Journal of Information Management*. 48, 280-290. [Viitattu 16.6.2023]. Saatavilla: <https://doi.org/10.1016/j.ijinfomgt.2017.09.009>

Kadhim, A.I. (2019) Survey on supervised machine learning techniques for automatic text classification. *The Artificial intelligence review*. 52, 1, 273-292. [Viitattu 1.6.2023]. Saatavilla: <https://doi.org/10.1007/s10462-018-09677-1>

Khan, M., Jan, B. & Farman, H. (2019) Deep Learning: Convergence to Big Data Analytics. 1. p. Singapore, Springer Singapore.

Kirwan, K. (2022) Structured vs. Unstructured Data: What You Need to Know. [Verkkoartikkeli]. [Viitattu 18.8.2023]. Saatavilla: <https://segment.com/blog/structured-vs-unstructured-data/>

Lambert, C. & Sponem, S. (2012) Roles, Authority and Involvement of the Management Accounting Function: A Multiple Case-study Perspective. *The European accounting review*. 21, 3, 565-589. [Viitattu 16.6.2023]. Saatavilla: <https://doi.org/10.1080/09638180.2011.629415>

Lua, A. (2023) 21 Top Social Media Sites to Consider for Your Brand in 2023, Buffer. [Verkkoartikkeli]. [Viitattu 1.6.2023]. Saatavilla: <https://buffer.com/library/social-media-sites/>

MonkeyLearn. (2023) Text Classification Using Support Vector Machines (SVM). [Verkkoartikkeli]. [Viitattu 1.6.2023]. Saatavilla: <https://monkeylearn.com/text-classification-support-vector-machines-svm/>

Nakashololo, T. & Iyamu, T. (2019) Understanding big data analytics and the Interpretive approach for analysis purposes. *Journal of Contemporary Management*. 16, 1, 272-289. [Viitattu 15.6.2023]. Saatavilla: <https://hdl.handle.net/10520/EJC-15cb158ace>

Namazi, M. & Rezaei, G. (2023) Modelling the role of strategic planning, strategic management accounting information system, and psychological factors on the budgetary slack. *Accounting forum*. ahead-of-print, ahead-of-print, 1-28. [Viitattu 11.6.2023]. Saatavilla: <https://doi.org/10.1080/01559982.2022.2163040>

Oracle. (2023) What is Big Data? [Verkkoartikkeli]. [Viitattu 11.6.2023]. Saatavilla: <https://www.oracle.com/big-data/what-is-big-data/>

Osborne-Gowey, J. (2014) What is Social Media. *Fisheries*. 39, 2, 55–55. [Viitattu 1.6.2023]. Saatavilla: <https://doi.org/10.1080/03632415.2014.876883>

Pauha, T. (2019) Ryhmittelyanalyysi uskonnontutkijan työkaluna. *Uskonnontutkija*. 8, 2. [Viitattu 1.6.2023]. Saatavilla: <https://doi.org/10.24291/uskonnontutkija.v8i2.88564>

Pervan, I. & Dropulić, I. (2019) THE IMPACT OF INTEGRATED INFORMATION SYSTEMS ON MANAGEMENT ACCOUNTING: CASE OF CROATIA. *Management*. 24, 1, 21–38. [Viitattu 15.6.2023]. Saatavilla: <https://doi.org/10.30924/mjcmi.24.1.2>

Power, D. J. & Phillips-Wren, G. (2011) Impact of Social Media and Web 2.0 on Decision-Making: Decision Making in Web 2.0 and Social Media Environments. *Journal of decision systems*. 20, 3, 249–261.

Pozzi, F. & Pozzi, F.A. (2017) Sentiment analysis in social networks. 1. p. Amsterdam, Morgan Kaufmann.

Rebala, G., Ravi, A. & Churiwala, S. (2019) An Introduction to Machine Learning. 1. p. Cham, Springer International Publishing.

Saito, T. & Gupta, S. (2022) Big data applications with theoretical models and social media in financial management. *Annals of operations research*. 1–23. [Viitattu 1.7.2023]. Saatavilla: <https://doi.org/10.1007/s10479-022-05136-x>

Salminen, A. (2011) Mikä kirjallisuuskatsaus?: johdatus kirjallisuuskatsauksen tyypeihin ja hallintotieteellisiin sovelluksiin. Vaasa, Vaasan yliopisto.

Schintler, L. A. & McNeely, C. L. (2022) *Encyclopedia of Big Data*. Cham, Springer International Publishing AG.

Schuster, P., Heinemann, M. & Cleary, P. (2021) *Management accounting*. Cham, Switzerland, Springer.

Stieglitz, S. & Dang-Xuan, L. (2013) Social media and political communication: a social media analytics framework. *Social network analysis and mining*. 3, 4, 1277-1291. [Viitattu 20.5.2023]. Saatavilla: <https://doi.org/10.1007/s13278-012-0079-3>

Stieglitz, S., Dang-Xuan, L., Bruns, A. & Neuberger, C. (2014) Social Media Analytics: An Interdisciplinary Approach and Its Implications for Information Systems. *Business & information systems engineering*. 6, 2, 89-96. [Viitattu 20.5.2023]. Saatavilla: <https://doi.org/10.1007/s12599-014-0315-7>

Stieglitz, S., Mirbabaie, M., Ross, B. & Neuberger, C. (2018) Social media analytics – Challenges in topic discovery, data collection, and data preparation. *International Journal of Information Management*. 39, 156-168. [Viitattu 20.5.2023]. Saatavilla: <https://doi.org/10.1016/j.ijinfomgt.2017.12.002>

Sunarni, C.W. & Ambarriani, A.S. (2019) The Pricing Practices: Management Accounting Perspective. *Review of Integrative Business and Economics Research*. 8, 84.

The World Bank. (2023) Population, total. [Verkkoartikkeli]. [Viitattu 11.6.2023]. Saatavilla: <https://data.worldbank.org/indicator/SP.POP.TOTL?end=2021&start=2020>

Thrane, S., Jarmatz, M., Fetahi Laursen, M. & Kornmaaler, K. (2019) A practice-based approach to collective decision-making in pricing. *Qualitative research in accounting and management*. 16, 1, 117-143. [Viitattu 1.7.2023]. Saatavilla: <https://doi.org/10.1108/QRAM-03-2018-0019>

T.K., B., Annavarapu, C.S.R. & Bablani, A. (2021) Machine learning algorithms for social media analysis: A survey. *Computer science review*. 40, 100395. [Viitattu 1.6.2023]. Saatavilla: <https://doi.org/10.1016/j.cosrev.2021.100395>

Turkington, G., Deshpande, T. & Karanth, S. (2016) *Hadoop*. Birmingham, Packt Publishing, Limited.

- Uyar, M. (2021) The Role of Business Analytics in Transforming Management Accounting Information into Cost Performance. *Ege akademik bakıs.* 21, 4, 373-389. [Viitattu 15.7.2023]. Saatavilla: <https://doi.org/10.21121/eab.1015665>
- Webwise. (2023) Explained: What is Facebook. [Verkkoartikkeli]. [Viitattu 1.6.2023]. Saatavilla: <https://www.webwise.ie/parents/explained-what-is-facebook-2/>
- Wu, J. (2012) *Advances in K-means Clustering A Data Mining Thinking.* Berlin, Springer Berlin Heidelberg. [Viitattu 1.6.2023]. Saatavilla: <https://doi.org/10.1007/978-3-642-29807-3>
- Younis, E.M. (2015) Sentiment analysis and text mining for social media microblogs using open-source tools: an empirical study. *International Journal of Computer Applications.* 112, 5, 44-48. [Viitattu 15.5.2023]. Saatavilla: <https://www.ijcaonline.org/archives/volume112/number5/19665-1366>
- Zachlod, C., Samuel, O., Ochsner, A. & Werthmüller, S. (2022) Analytics of social media data – State of characteristics and application. *Journal of business research.* 144, 1064-1076. [Viitattu 1.6.2023]. Saatavilla: <https://doi.org/10.1016/j.jbusres.2022.02.016>
- Zahedi, Z. & Costas, R. (2018) General discussion of data quality challenges in social media metrics: Extensive comparison of four major altmetric data aggregators. *PloS one; PLoS One.* 13, 5, e0197326–e0197326. [Viitattu 15.6.2023]. Saatavilla: <https://doi.org/10.1371/journal.pone.0197326>
- Zeng, D., Chen, H., Lusch, R. & Li, S. (2010) Social Media Analytics and Intelligence. *IEEE intelligent systems.* 25, 6, 13-16. [Viitattu 5.6.2023]. Saatavilla: <https://doi.org/10.1109/MIS.2010.151>
- Zhao, R. & Mao, K. (2018) Fuzzy Bag-of-Words Model for Document Representation. *IEEE Transactions on Fuzzy Systems.* 26, 2, 794-804. [Viitattu 20.5.2023]. Saatavilla: <https://doi.org/10.1109/TFUZZ.2017.2690222>