

Tekoälytehosteisen big data -analytiikan hyödyntäminen kilpailija-analyysissä

Markkinoinnin
pro gradu -tutkielma

Laatija:
Salla Nurmi

Ohjaaja:
KTT Helena Rusanen

5.5.2025
Pori

Turun yliopiston laatujärjestelmän mukaisesti tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck -järjestelmällä.

Pro gradu -tutkielma

Oppiaine: Markkinointi

Tekijä: Salla Nurmi

Otsikko: Tekoälytehosteisen big data -analytiikan hyödyntäminen kilpailija-analyysissä

Ohjaaja: KTT Helena Rusanen

Sivumäärä: 102 sivua + liitteet 8 sivua

Päivämäärä: 5.5.2025

Kilpailija-analyysi on tärkeä strateginen työkalu, jonka avulla yritys voi pyrkiä erottautumaan kilpailijoistaan ja saavuttamaan kilpailuetua. Tänä päivänä kilpailu markkinoilla on kovempaa kuin koskaan, jonka myötä myös erottautuminen on vaikeutunut. Tästä syystä kilpailija-analyysin merkitys tänä päivänä on erityisen tärkeä.

Edistyneet teknologiat, kuten tekoäly ja big data -analytiikka, ovat viime vuosikymmenen aikana muuttaneet yritysten toimintaa yhä datavetoisemmaksi, ja niiden tuomia hyötyjä kartoitetaan nyt kovaa vauhtia eri aloilla, myös markkinointitieteessä. Vaikka tekoälyn ja big data -analytiikan tuomia mahdollisuuksia on jo melko laajalti tutkittu strategisen markkinoinnin ja johtamisen yhteydessä, niiden merkitystä kilpailija-analyysin kontekstissa on toistaiseksi tarkasteltu niukasti.

Tämän pro gradu -tutkielman tavoitteena on selvittää, mitä mahdollisuuksia tekoälytehostettu big data -analytiikka tuo kilpailija-analyysiin. Tutkielmassa yhdistetään tekoäly, big data -analytiikka ja kilpailija-analyysin prosessimalli teoreettiseksi viitekehykseksi ja vastataan tarpeeseen ymmärtää syvemmin edistyneiden teknologioiden merkitystä kilpailija-analyysin prosessin eri vaiheissa.

Tutkimus toteutettiin laadullisena tutkimuksena, jonka aineisto kerättiin kahdeksan asiantuntijahaastattelun avulla. Haastattelut olivat puolistrukturoituja teemahaastatteluita, joiden teemat perustuivat tutkimuksen teoreettiseen viitekehykseen. Saatu aineisto litteroitiin ja analysoitiin temaattisen sisällönanalysin avulla.

Tutkimuksen tulokset osoittavat, että tekoälytehosteinen big data -analytiikka tuo merkittäviä hyötyjä kilpailija-analyysin prosessiin mahdollistamalla prosessin tehostumisen, automaation, skaalautuvuuden ja multimodaalisen datan hyödyntämisen. Nämä teknologiat kykenevät keräämään ja käsittelemään suuria tietomassoja kilpailijoista tehokkaasti, tunnistamaan trendejä ja tekemään ennustuksia kilpailijoiden liikkeistä. Tämän lisäksi ne tukevat päätöksentekoa rakenteistamalla analyysiä, tarjoamalla johtopäätöksiä, toimintakehotuksia ja visuaalisia esityksiä tuloksista. Lisäksi havaittiin, että teknologian kehityksen myötä kilpailija-analyysin perinteinen lineaarinen prosessi on muuttumassa kohti dynaamisempaa, limittäistä ja jatkuvasti päivittyvää prosessia, jossa analyysin vaiheet sulautuvat toisiinsa. Myös kilpailija-analyttikon työn nähtiin siirtyvän manuaalisen työn vähentyessä kohti analyysiprosessin loppupäätä eli tulkinnan ja johtopäätösten tekemistä.

Tutkimus tarjoaa uutta teoreettista ymmärrystä siitä, miten tekoäly ja big data -analytiikka muuttavat kilpailija-analyysin prosessia, mitä hyötyjä näistä edistyneistä teknologioista on eri kilpailija-analyysin vaiheissa, ja antaa suosituksia liikkeenjohdolle teknologian hyödyntämiseen. Kilpailija-analyysin prosessin viitekehyksen näkökulmasta tutkimus syventää käsitystä kilpailija-analyysin prosessin kehityksestä kohti jatkuvaa, joustavaa ja reaaliaikaista analytiikkaa.

Avainsanat: Big data -analytiikka, Tekoäly, Kilpailija-analyysi

SISÄLLYS

1	Johdanto	7
1.1	Tutkimuksen tausta	7
1.2	Tutkimusaukko	8
1.3	Tutkimuksen tavoite ja rajaukset	10
1.4	Tutkimuksen keskeisimmät käsitteet	12
1.5	Tutkimuksen rakenne	13
2	Tekoälytehosteinen big data -analytiikka kilpailija-analyysissä	14
2.1	Tekoälytehosteisen big data -analytiikan osa-alueet	14
2.1.1	Datan ja big datan määritelmä	14
2.1.2	Big data -analytiikan prosessi	16
2.1.3	Tekoälyn määritelmä ja algoritmit	19
2.1.4	Tekoälyn ja big data -analytiikan integraatio	22
2.1.5	Tekoälyn hyödyntäminen big data -analytiikan prosessin eri vaiheissa	24
2.2	Kilpailija-analyysin prosessi	26
2.2.1	Yrityksen kilpailijoiden määrittäminen	26
2.2.2	Kilpailija-analyysin määritelmä	28
2.2.3	Kilpailija-analyysin toteuttaminen prosessimallin mukaan	29
2.3	Tekoälytehosteisen big data –analytiikan hyödyt kilpailija-analyysissä	34
2.4	Teoreettinen viitekehys	36
3	Metodologia	39
3.1	Tutkimuksen tieteenfilosofiset lähtökohdat	39
3.2	Tutkimusote	40
3.3	Tutkimusmenetelmä	41
3.4	Haastateltavien valinta ja kuvaus	44
3.5	Aineiston keruu	45
3.6	Aineiston analysointi	47
3.7	Tutkimuksen luotettavuuden arviointi	47
3.8	Tutkimuksen eettisyyden arviointi	50
4	Tekoälytehosteisen big data -analytiikan tuomat mahdollisuudet kilpailija-analyysissä	52
4.1	Tekoälytehosteinen big data -analytiikka kilpailijatiedon tarpeen määrittelyssä	52
4.1.1	Tekoälytehosteinen big data -analytiikka ideoinnin tukena tiedon tarvemäärittelyssä	52
4.1.2	Tiedon tarpeiden määrittelyn tehostuminen tekoälyn avulla	53
4.1.3	Generatiivisen tekoälyn haasteet ja rajoitteet tiedon tarpeiden määrittelyssä	54
4.2	Tekoälytehosteinen big data -analytiikka kilpailijatiedon keräämisessä	56
4.2.1	Kilpailijatiedon keräämisen tehostuminen tekoälytehostetun big data -analytiikan avulla	56
4.2.2	Tekoälytehosteisen big data -analytiikan mahdollistama automaatio kilpailijatiedon keruussa	59
4.2.3	Tekoälyn ja big data -analytiikan mahdollistama multimodaalisuus kilpailijatiedon keruussa	60

4.3 Tekoälytehosteinen big data -analytiikka oleellisen kilpailijatiedon seulonnassa	63
4.3.1 Tekoälytehosteisen big data -analytiikan tuomat mahdollisuudet kilpailijatiedon seulonnassa	63
4.3.2 Haasteet kilpailijatiedon seulomisessa tekoälytehosteisesti	64
4.4 Tekoälytehosteinen big data -analytiikka kilpailijatiedon analysoinnissa	66
4.4.1 Tekoälytehosteisen big data -analytiikan tuoma tehostushyöty kilpailijoiden analysoinnissa	66
4.4.2 Esimerkki reaaliaikaisesta kilpailijatietotaulusta	69
4.4.3 Tekoälytehosteinen big data -analytiikka kilpailijoiden markkinakäyttäytymisen analysoinnissa	71
4.4.4 Tekoälytehosteinen big data -analytiikka kilpailijoiden liikkeiden ennakoinnin ja ennustamisen tukena	72
4.4.5 Tekoälytehosteinen big data -analytiikka kilpailijoiden heikkouksien ja vahvuuksien analysoinnissa	74
4.5 Tekoälytehosteisen big data -analytiikan tuomat hyödyt kilpailijatiedon tulkinnassa ja johtopäätöksissä	76
4.5.1 Tekoälytehosteisen big data -analytiikan mahdollistaman visualisoinnin hyöty tulkintoja tehdessä	77
4.5.2 Tekoälyn luomat toimintakehotukset kilpailija-analyysissä	78
4.6 Kilpailija-analyysin prosessin muutos dynaamiseksi kokonaisuudeksi	79
4.7 Tulosten tarkastelu	80
5 Johtopäätökset	83
5.1 Teoreettiset johtopäätökset	83
5.2 Liikkeenjohdolliset suositukset	87
5.3 Tutkimuksen arviointi	89
5.4 Jatkotutkimusmahdollisuudet	89
6 Yhteenveto	91
Lähteet	93
Liitteet	103
Liite 1. Haastattelukysymykset	103
Liite 2. Esitietomateriaali	105
Liite 3. Tietosuojalomake	107
Liite 4. Suostumuslomake	109

KUVIOT

Kuvio 1 Data, informaatio ja tieto	14
Kuvio 2 Big data -analytiikan prosessi	17
Kuvio 3 Tekoälyn käsitteen alle kuuluvat alakäsitteet	20
Kuvio 4 Liiketoiminta-analytiikan kehitys kohti tekoälytehostettua analytiikkaa	23
Kuvio 5 Kilpailukentän hahmottaminen	26
Kuvio 6 Kilpailija-analyysin kuusivaiheinen prosessi	30
Kuvio 7 Kilpailija-analyysi	32
Kuvio 8 Vertailuanalyysin eri osa-alueet	33
Kuvio 9 Teoreettinen viitekehys tekoälytehosteisen big data -analytiikan hyödyntämisestä kilpailija-analyysissä	37
Kuvio 10 Haastattelun 6 esittämä markkinatietotaulu	69
Kuvio 11 Tekoälytehosteisen big data -analytiikan tuomat mahdollisuudet kilpailija-analyysiin	84

TAULUKOT

Taulukko 1 Tekoälyn hyödyntäminen big data -analytiikan prosessin eri vaiheissa	24
Taulukko 2 Tekoälytehosteinen big data -analytiikka kilpailija-analyysissä	34
Taulukko 3 Tutkimuksen operationalisointitaulukko	43
Taulukko 4 Haastatteluiden tiedot	46
Taulukko 5 Tutkimuksen empiirisen osuuden tulokset	81

1 Johdanto

1.1 Tutkimuksen tausta

Yrityksen strategian luonnissa yksi tärkeimmistä kulmakivistä on yrityksen ulkoisen ympäristön analyysi, johon kilpailija-analyysi kuuluu olennaisena osana (Yaneva 2020, 26; Adom ym. 2016, 116). Kilpailijoita analysoimalla yritys voi rakentaa kilpailustrategiansa ja asemoitua markkinoilla erottuakseen ja saavuttaakseen kilpailuetua (Porter 1980). Nykyisin markkinoilla vallitsee hyperkilpailu, jossa kilpailu on poikkeuksellisen intensiivistä, muutokset tapahtuvat nopeasti ja saavutettua kilpailuetua on vaikea ylläpitää. Erottautuminen kilpailijoista on entistä haastavampaa, mikä korostaa nopean reagointikyvyn merkitystä. (Knudsen ym. 2021, 367.) Yritykset, jotka kykenevät ajoissa havaitsemaan kilpailuympäristössä tapahtuvia muutoksia ja mukauttamaan sisäisiä prosessejaan niiden pohjalta, voivat säilyttää asemansa markkinoilla (Kamkankaew ym. 2022, 14). Siksi kilpailija-analyysi ja sen kehittäminen ovat nousseet tärkeään rooliin tänä päivänä, jotta yritykset voivat jatkossakin erottautua markkinoilla ja säilyttää kilpailuasemansa (Wittman 2024).

Ennen yritykset saattoivat kerätä kilpailijatietoa vuosittain tai neljännesvuosittain, mutta tänä päivänä pysyäkseen mukana markkinoiden äkillisissä muutoksissa, yritykset tarvitsevat reaaliaikaista kilpailijatietoa (Wittman 2024, 1293). Kilpailija-analyysin prosessi koostuu kuudesta vaiheesta: tiedon tarpeiden määrittely, tiedon keräys, oleellisen tiedon seulonta, tiedon analysointi, tulkinta ja johtopäätökset ja tiedon jakaminen päätöksentekijöille (Pirttilä 2000). Kilpailija-analyysin tekemisen edellytys on siis kilpailijoista saatavilla oleva data, sen kerääminen ja analysointi. Digitalisaation edetessä datan määrä on kasvanut huomattavasti, erityisesti viime vuosikymmenten aikana. Tämän eksponentiaalisen kasvun myötä datasta on tullut yksi arvokkaimmista resursseista yrityksille. Datan määrä ja muoto ovat muuttuneet niin valtavasti, että datan sijasta nykyään käytetään yleistynyttä termiä big data tai massadata, sillä se luonnehtii paremmin datan luonnetta tänä päivänä. (Monino 2021, 256.)

Datalla ei itsessään kuitenkaan ole arvoa, vaan arvo syntyy siitä, miten yritys pystyy hyödyntämään dataa liiketoimintansa tukemiseksi (Agarwal ym. 2024, 138). Samaan aikaan, kun datan määrä on kasvanut, ovat myös sen laatu ja analysointimenetelmät parantuneet. Merkittävin tekijä analysointimenetelmien edistyksessä oli, kun tekoälyalgoritmeja integroitiin osaksi perinteistä data-analytiikkaa. Tätä kehittyntä

teknologiaa kutsutaan muun muassa tekoälytehosteiseksi big data -analytiikaksi (engl. AI empowered big data analytics) (Kumar ym. 2022) tai älykkääksi data-analytiikaksi (engl. intelligent analytics) (Zong – Guan 2025). Big data -analytiikka käsittää suurten datamäärien keräämisen, analysoinnin ja tulkinnan (Sun – Sun 2019, 162–163). Tekoäly puolestaan kattaa erilaisia tekniikoita, kuten generatiivisen tekoälyn, koneoppimisen ja syvät neuroverkot, jotka mahdollistavat koneiden suorittavan tehtäviä, jotka aiemmin vaativat ihmisen älykkyyttä (Haenlein – Kaplan, 2019, 5; Kreutzer – Sirrenberg 2020, 4). Näiden yhdistelmä, eli tekoälytehosteinen big data -analytiikka, pystyy analysoimaan suuria data-aineistoja sekä huomaamaan niistä toistuvia ja uusia kaavoja aiempaa tehokkaammin (Kumar ym. 2022, 877–878; Rahmani ym. 2021, 18–20).

Tekoälyn ja big data -analytiikan on muiden edistyneiden teknologioiden mukana tunnistettu olevan ajureina sekä neljännessä teollisessa vallankumauksessa (Ivanov ym. 2021, 2070) että niin kutsutulla markkinoinnin 5.0 aikakaudella. Tällä aikakaudella tarkoitetaan markkinoinnin uutta edistysvaihetta, jossa markkinoinnin teorian ja käytännön kehittymisen keskiössä ovat kehittyneet teknologiat, kuten tekoäly ja big data -analytiikka. Kuten muillakin aloilla, myös markkinointitieteessä on alettu keskittymään yhä enemmän näiden kehittyneiden teknologioiden tuomien mahdollisuuksien tutkimiseen. (Bakator ym. 2023, 66.) On arvioitu, että markkinoinnin kyetessä hyödyntämään tekoälytehosteista big data-analytiikkaa, se voi mahdollistaa uusia menestystekijöitä yrityksille (Agarwal ym. 2024, 137). Monet markkinoinnin tutkijat ovatkin tuoneet esille tarpeen tutkia edistyneitä teknologioita, kuten tekoälyä ja big dataa, markkinoinnin eri osa-alueilla, kuten strategisessa ja operatiivisessa markkinoinnissa (Basu ym. 2023; Petrescu – Krishen 2023; Duan ym. 2019).

Vaikka big datan ja tekoälyn hyötyjä strategisessa päätöksenteossa (esim. Trunk ym. 2020) ja markkinoinnin eri osa-alueilla (esim. Huang & Rust 2021) on tutkittu melko kattavasti, niiden merkitystä kilpailija-analyysissä ymmärretään yhä rajallisesti (Basu ym. 2023; Hatzijordanou ym. 2019, 440). Wittmann (2024, 1296) korostaakin, että aihetta on tarpeen tutkia syvällisemmin, jotta voidaan ymmärtää, miten nämä teknologiat voivat tehostaa kilpailija-analyysin prosesseja ja sitä kautta tukea yritysten kilpailuetua.

1.2 Tutkimusaukko

Basun ym. (2023) teettämässä kirjallisuuskatsauksessa tulevaisuuden tarpeelliseksi tutkimussuunniksi markkinointianalytiikan tutkimuskentällä nousivat esimerkiksi

edistyneiden teknologioiden, kuten tekoälytehosteisen big data -analytiikan, integroituminen markkinointianalytiikkaan ja integraation vaikutus strategiseen markkinointiin. He painottavat, että markkinointianalytiikan on tieteenalana tärkeää pysyä teknologisen kehityksen mukana, ja tästä syystä olisi tärkeää tutkia sitä, miten näitä teknologioita pystytään hyödyntämään markkinoinnin kehittämisessä.

Hatzijordanou ym. (2019, 440) tutkivat puolestaan kilpailija-analyysin tutkimuskenttää systemaattisen kirjallisuuskatsauksen avulla. Suurin osa tutkimuksista käsitteli kilpailijoiden tunnistamista ja analyysien viitekehyksiä (22 tutkimusta). Toiseksi eniten käsiteltiin kilpailija-analyysin tekemisen käytäntöjä (20 tutkimusta) ja kolmanneksi suosituin tutkimusaiheklusteri oli kilpailija-analyysin rooli osana organisaation toimintaa (16 tutkimusta). Tässä systemaattisessa kirjallisuuskatsauksessa ei tullut ilmi tutkimuksia, jossa tutkittaisiin edistyneiden teknologioiden hyödyntämistä kilpailija-analyysissä

Wittmann (2024, 1296) esittää, että kilpailija-analyysin merkitys korostuu tämän päivän dynaamisessa liiketoimintaympäristössä ja että reaaliaikainen kilpailijatieto on entistä tärkeämpää. Hän nostaa esiin tekoälyn mahdollisuudet kilpailijoiden tunnistamisessa ja markkinasegmenttien muutosten seuraamisessa, mutta toteaa, että tekoälyn hyödyntäminen kilpailija-analyysissä vaatii vielä lisätutkimusta.

Kilpailija-analyysi voidaan nähdä osaksi laajempaa toimintaympäristöanalyysiä tai kilpailullista analyysiä. Näistä erityisen tunnettu on Porterin viiden kilpailuvoiman malli, jossa tarkastellaan yrityksen toimintaympäristöä viidestä näkökulmasta: uusien kilpailijoiden uhka, toimialan sisäisen kilpailu, korvaavien tuotteiden tai palveluiden uhka, ostajien neuvotteluvoima ja toimittajien neuvotteluvoima. (Porter 2008, 80.) Robinson ym. (2021, 2) totesivat tutkimuksessaan, että edistyneiden teknologioiden, kuten tekoälyn, hyödyntämistä ei ole tutkittu kokoavasti toimintaympäristön analyysissä, mihin kilpailija-analyysikin siis kuuluu. Tämäkin korostaa, että kilpailija-analyysiä ei ole tutkittu kovinkaan paljoa edistyneiden teknologioiden näkökulmasta.

Chakraborti ja Dey (2019) toivat artikkelissaan ilmi, että nykyiset kilpailijatiedon järjestelmät keskittyvät pääasiassa keräämään tietoa kilpailijoista ja kilpailijatiedon raportoimiseen, mutta niissä on puutteita big datan hallinnassa, tiedon laadun parantamisessa sekä päätöksenteon tukemisessa tehokkaalla tavalla. He myös nostivat esiin tutkimusaukon: tekstin klusterointia tai automaattisen tekstin tiivistämisen (ATS)

teknologioita (molemmat ovat tekoälypohjaisia algoritmeja) ei ole hyödynnetty kilpailija-analyysin kontekstissa. Tähän vastatakseen he kehittivät näitä algoritmeja hyödyntävän järjestelmän ja käyttivät sitä analysoimaan ja tiivistämään big datasta kerättyä kilpailijatietoa. Tulokset osoittivat, että tekoäly tuotti laadukkaampia tiivistelmiä kuin perinteiset menetelmät. Vaikka artikkelissa tutkitaan klusterointia ja tekstin tiivistämistä, Chakraborti ja Dey (2019) suosittelevat, että muita edistyneempiä analyysimenetelmiä pitäisi myös tutkia kilpailija-analyysin kontekstissa.

On kuitenkin joitain tutkimuksia, joissa erilaisia edistyneempiä teknologioita, jotka yhdistävät eri suhteessa big dataa ja tekoälyn eri algoritmeja, tutkitaan osana tiettyä kilpailija-analyysin vaihetta. Tutkimuksissa on tutkittu muun muassa seuraavia aiheita: asiakkaiden kilpailijoita koskevien mielipiteiden selvittäminen (Taherdoost – Madanchian 2023), kilpailijoiden patenttihakkeiden (Suominen ym. 2017) ja omien ja kilpailevien tuotteiden vertailu (Jin ym. 2016), markkinoiden rakenteen ja kilpailija-asetelmien kartoitus (Guo ym. 2017) sekä kilpailijoiden hinnoittelustrategioiden seuraaminen (Tong-On ym. 2021). Nämä kaikki ovat kuitenkin hyvin spesifejä tutkimuksia, ja kohdistuvat tiettyihin tekoälyalgoritmeihin tai tiettyihin kapeisiin kilpailija-analyysin osa-alueisiin, eivätkä käsittele koko kilpailija-analyysin prosessia.

1.3 Tutkimuksen tavoite ja rajaukset

Tutkimuksen tarkoituksena on selvittää tekoälytehosteisen big data -analytiikan tarjoamia mahdollisuuksia kilpailija-analyysin prosessissa.

Tutkimuskysymys on:

- Millaisia mahdollisuuksia tekoälytehosteinen big data -analytiikka tuo kilpailija-analyysiin?

Tutkimuskysymys jakaantuu kahteen osaongelmaan:

- Millaisia mahdollisuuksia tekoälytehosteinen big data -analytiikka tuo kilpailijatiedon tarpeiden määrittelyyn, keräämiseen ja seulontaan?
- Millaisia mahdollisuuksia tekoälytehosteinen big data -analytiikka tuo kilpailijatiedon analysointiin ja tulkintaan?

Tutkimuskysymykseen vastataan kahden osaongelman avulla. Ensimmäinen osaongelma keskittyy selvittämään, mitä mahdollisuuksia tekoälytehosteinen big data -analytiikka tuo kilpailijatiedon tarpeen määrittelyyn, keräämiseen ja seulomiseen. Toisessa osaongelmassa keskitytään puolestaan kilpailijoista kerätyn big datan analysointiin, tulkintaan ja johtopäätösten tekemiseen.

Tutkimus tuo uutta ymmärrystä strategisen markkinoinnin, kilpailija-analyysin ja markkinointianalytiikan tutkimusalueisiin. Se laajentaa markkinoinnin teoriaa tutkimalla, miten yritykset voivat hyödyntää tekoälyä ja big data -teknologioita markkinoinnin päätöksenteossa. Koska kilpailija-analyysin tekemiseen vaaditaan valtavasti dataa kilpailijoista (Porter 1980, 71), tutkimus on rajattu koskemaan big dataa. Myös sen sijaan, että tutkimus käsittelee pelkästään big data -analytiikkaa, se on rajattu käsittelemään kehittyneempää big data -analytiikka, johon on integroitu tekoäly, sillä tänä päivänä tekoäly ja big data -analytiikka kietoutuvat yhä vahvemmin yhdeksi kokonaisuudeksi (Zaripova ym. 2023, 2).

Kilpailija-analyysin viitekehyksenä tutkimuksessa hyödynnetään Pirttilän (2000) kuusivaiheista kilpailija-analyysin prosessimallia. Tutkimus on rajattu koskemaan mallin viittä ensimmäistä vaihetta: tiedon tarvemäärittelyä, keräämistä, seulontaa, analyysiä ja johtopäätöksiä. Prosessin kuudes vaihe, eli tiedon jakaminen, on rajattu tutkimuksen ulkopuolelle, sillä se sijoittuu tutkimusasetelman kannalta enemmän organisaation sisäisen viestinnän kenttään, eikä suoraan liity siihen, miten tekoälytehostettu big data -analytiikka vaikuttaa kilpailijatiedon hyödyntämiseen organisaatiossa.

Tutkimus on rajattu käsittelemään kattavasti lähes koko kilpailija-analyysin prosessia, eikä esimerkiksi vain yksittäistä osaa prosessista, koska kilpailija-analyysin prosessi on jatkuva, eli kaikki sen prosessin eri vaiheet ovat sidoksissa toisiinsa. Edellisen vaiheen onnistuminen on edellytys seuraavan vaiheen onnistumiselle ja vastaavasti jos edellisessä vaiheessa epäonnistutaan, vaikuttaa se seuraavan vaiheen hyödyllisyyteen. (Pellissier – Nenzhelele 2013, 6.) Tämän takia edistyneiden teknologioiden vaikutukset voivat alkaa jo varhaisessa vaiheessa, esimerkiksi tiedon keruussa tai suodattamisessa, mutta niiden hyöty realisoituu usein myöhemmissä vaiheissa, kuten analyysissä ja päätöksenteossa. Tämän vuoksi laajemman prosessin tarkastelu yksittäisen prosessivaiheen sijaan tarjoaa selkeämmän kuvan siitä, miten edistyneitä teknologioita voidaan hyödyntää kilpailija-analyysissä.

Vastauksia tutkimuskysymykseen kerätään asiantuntijahaastatteluiden avulla. Tutkimusmenetelmänä on laadullinen teemahaastattelu, jossa aineisto kerätään haastatteleamalla data-analytiikan ja tekoälyn ammattilaisia, jotka ovat hyödyntäneet näitä edistyneitä teknologioita liiketoiminnan kehittämisen kontekstissa.

1.4 Tutkimuksen keskeisimmät käsitteet

Big data: Big data tarkoittaa valtavia, nopeasti kasvavia ja monimuotoisia tietoaaineistoja (De Mauro ym. 2014). Se eroaa perinteisestä datasta erityisesti 3V-mallin mukaisesti: suuren määrän, nopean käsittelytarpeen ja monimuotoisuuden perusteella (Laney 2001). Myöhemmin big datan ominaisuuksiksi on lisätty myös arvo, jota yritykset sen avulla pyrkivät saavuttamaan, totuudenmukaisuus ja rakenteettomuus (Dijcks 2012; Schroeck ym. 2012; Suthaharan 2014). Big dataa hyödynnetään eri aloilla tarjoamaan oivalluksia ja tukemaan päätöksentekoa analytiikan ja tekoälyn avulla (Monino 2021, 263).

Big data -analytiikka: Big data -analytiikka tarkoittaa menetelmiä, järjestelmiä ja teknologioita, joiden avulla big datasta voidaan jalostaa yritykselle arvokasta tietoa (Saggi – Jain 2018, 768). Siinä analysoidaan suuria datamääriä älykkäillä laskentamalleilla, jotka tunnistavat merkityksellisiä yhteyksiä ja ennustavat tulevia tapahtumia. Tämä perustuu edistyneisiin laskennallisiin algoritmeihin ja tilastollisiin menetelmiin, kuten koneoppimiseen, neuroverkkoihin ja muihin analyysitekniikoihin, joiden avulla datan rakenteet ja yhteydet voidaan havaita tehokkaasti. (Nguyen ym. 2023, 525–526.)

Tekoäly: Tekoäly voidaan määritellä järjestelmän kyvyksi tulkita oikein ulkoista dataa, oppia tästä datasta ja käyttää oppimaansa saavuttaakseen tiettyjä tavoitteita ja tehtäviä joustavan sopeutumisen avulla (Kaplan – Haenlein 2019b). Keskeisiä tekniikoita ovat koneoppiminen, neuroverkot, syväoppiminen ja generatiivinen tekoäly, joiden avulla tekoäly esimerkiksi analysoi dataa, ennustaa tapahtumia (Kreutzer – Sirrenberg 2020, 4) ja tuottaa erilaisia sisältöjä, kuten tekstiä, kuvia ja videoita (Feuerriegel ym. 2024, 111–115). Tekoälyä sovelletaan muun muassa robotiikassa, autonomisessa ajamisessa ja luonnollisen kielen käsittelyssä (Sarker 2021, 4).

Tekoälytehosteinen big data -analytiikka: Nykyään big data -analytiikka on pääosin tekoälytehosteista, eli sen eri prosesseissa hyödynnetään tekoälyalgoritmeja. Tekoäly mahdollistaa merkittävän datan käsittelykapasiteetin ja edistyneitä tilastollisia

menetelmiä, minkä takia tekoälyn ja big data -analytiikan integraatio on tuonut kokonaan uuden aikakauden data-analytiikkaan. (Davenport 2018, 74.)

Kilpailija-analyysi: Kilpailija-analyysi on olennainen osa markkinointisuunnittelua (Adom ym. 2016, 124). Kilpailija-analyysi on prosessi, jonka tarkoituksena on tuottaa yrityksen strategisen johtamisen tueksi hyödyllistä tietoa sen kilpailijoista (Porter 2004, 47). Kilpailija-analyysin tarkoituksena on kilpailijoiden nykyisen toiminnan analysointi, tulevien aikeiden ennakointi ja mahdollisten uhkien havainnointi (Bennet 2003, 336). Pirttilä (2000) esittää kilpailija-analyysin kuusivaiheisena prosessina, johon kuuluvat tiedon tarpeiden määrittely, tiedon keräys, oleellisen tiedon seulonta, tiedon analysointi, tulkinta ja johtopäätökset ja tiedon jakaminen päätöksentekijöille.

1.5 Tutkimuksen rakenne

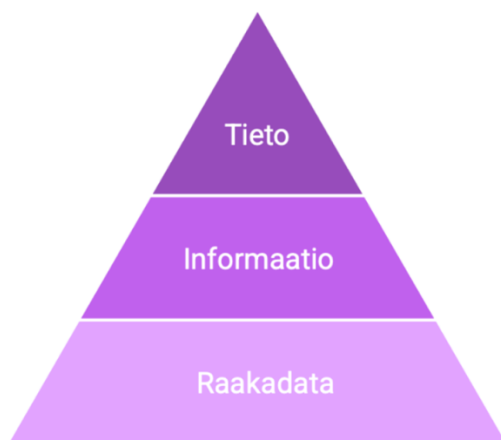
Tutkimus koostuu viidestä pääluvusta. Luku yksi esittelee tutkimuksen taustan, tutkimusaukon, tavoitteet ja rajaukset. Luku kaksi käsittelee teoreettista viitekehystä aiempaan kirjallisuuteen pohjautuen. Tutkimuksen teoreettinen viitekehys käsittelee big data -analytiikkaa, tekoälyä ja niiden integraatiota, jonka jälkeen syvennyttään kilpailija-analyysin tutkimuskirjallisuuteen. Lopuksi kirjallisuuskatsauksessa selvitetään, mitä tällä hetkellä tiedetään tekoälytehosteisen big data -analytiikan hyödyntämisestä kilpailija-analyysin kontekstissa ja esitellään tutkimuksen teoreettinen viitekehys kuviona. Luku kolme on metodologialuku, joka kuvaa tieteenfilosofiset lähtökohdat, tutkimusotteen ja tutkimusmenetelmät, aineistonkeruun, analyysin sekä tutkimuksen luotettavuuden ja eettisyyden arvioinnin. Luvussa neljä käsitellään tutkimuksen empiirisen osuuden eli haastattelujen tuloksia. Luku viisi kokoaa yhteen teoreettiset johtopäätökset, suositukset liikkeenjohdolle, tutkimuksen arvioinnin ja jatkotutkimusmahdollisuudet. Työ päättyy yhteenvetoon.

2 Tekoälytehosteinen big data -analytiikka kilpailija-analyysissä

2.1 Tekoälytehosteisen big data -analytiikan osa-alueet

2.1.1 Datan ja big datan määritelmä

Digitalisaation myötä asiakkaista saadaan valtava määrä dataa (Fitz-enz – Mattox 2014, 17). Data sekoitetaan usein käsitteisiin tieto ja informaatio. Nämä kolme hyvin toisiaan lähellä olevaa termiä pystytään kuitenkin erottamaan ja esittämään pyramidikuvaajan avulla. Pyramidin pohjana on data, eli vielä toistaiseksi käsittelemätön havainto, esimerkiksi luku, sana tai mittaustulos. Data on ikään kuin raaka-ainetta, jolla ei itsessään ole arvoa. Pyramidin keskellä on informaatio. Informaatiota syntyy, kun dataa käsitellään, järjestellään ja tulkitaan, ja näin datasta voidaan saada kokonaisuus, joka voi olla merkityksellinen. Tästä kootusta informaatiosta voidaan saada tietoa, jota kuvaa pyramidin huippu. Tämä tieto auttaa päättäjiä saamaan oivalluksia, ratkaisemaan ongelmia ja tekemään päätöksiä. (Monino 2021, 258.) Pyramidi on havainnollistettu kuviossa 1.



Kuvio 1 Data, informaatio ja tieto

Dataa voi syntyä nykypäivän digitalisoituneessa yhteiskunnassa hyvin monella tapaa. Sen perusteella, mistä lähteestä data syntyy, voidaan muodostaa kolme luokkaa: koneiden tuottama data, ihmisten tuottama data ja yritysten tuottama data. Koneiden tuottamalla datalla viitataan esimerkiksi tietokoneista, sensoreista, satelliiteista ja esineiden internetistä (engl. Internet of Things) saatavaan dataan. (Saggi – Jain 2018, 768.) Ihmisten tuottamalla datalla puolestaan viitataan nimensä mukaisesti ihmisten itsensä tuottamaan

dataan, joka on luotu joko tietoisesti tai tiedostomatta. Reaaliaikainen data syntyy esimerkiksi silloin, kun asiakas klikkaa sivustoilla, tekee ostoksia, katsoo viihdettä, hakee tietoa internetistä, valitsee evästeitä tai tekee sosiaalisen median interaktioita. Näistä liikkeistä jää jäljitettävissä oleva datajälki. Ihmisiltä itseltään voidaan myös kerätä tietoa, jonka he itse antavat eri toimenpiteiden kautta. Näitä tietoja ovat esimerkiksi ikä, sukupuoli, asuinpaikka sekä psykografiset tiedot, kuten harrastukset ja mielipiteet. (Saggi – Jain 2018, 768; Hinds ym. 2018, 17–18.) Kolmas datan lähde on yritysten tuottama data. Tällä datalla tarkoitetaan esimerkiksi transaktiotietoja, yritysten sisäisiä tietoja ja julkisten toimijoiden tietoja. Liiketoimintatiedolla (engl. business intelligence) viitataan siihen, onko data liiketoiminnalle arvokasta, tarjoaako se oivalluksia ja ymmärrystä ja tukeeko se päätöksentekoa. Mikäli data auttaa näissä kaikissa edellä mainituissa aspekteissa, se on hyödyllistä dataa. (Wu ym. 2016, 7.)

Digitaalisen vallankumouksen seurauksena dataa syntyy aiempaan verrattuna eksponentiaalisesti enemmän ja se on muodoltaan erilaista. Tämä on johtanut siihen, että datan käsitettä on muutettu, ja datan sijaan nykyään puhutaankin yleisemmin big datasta (De Mauro ym. 2014). Big datalle on ehdotettu viimeisten vuosikymmenten aikana useita eri määritelmiä. Tämä johtuu big data käsitteen nopeasta ja epäjärjestelmällisestä kehitymisestä. Yksi yleisimmistä tavoista määritellä big data on luettelemalla sen ominaisuuksia. (De Mauro ym. 2014.) Tämä tapa määritellä big data sen ominaisuuksien kautta sai alkunsa, kun Laney (2001) esitteli kehyksen 3 V:n mallista, jossa kuvaillaan big datan ominaisuuksia. Määritelmän ensimmäinen V on datan suuruus (engl. volume) ja sillä viitataan siihen, että big data on määrällisesti ennennäkemättömän suurta. Toinen V (engl. velocity) kuvaa, kuinka nopeaa big datan liike on. Tällä tarkoitetaan, että tiedon käsittelynopeus pyrkii olemaan reaaliaikaista. Kolmas V viittaa datan monimuotoisuuteen (engl. variety), eli tiedot vaihtelevat eri muodoissa, kuten esimerkiksi video, ääni, teksti, kuva, linkki tai twiitti. (Laney 2001; Monino 2021, 262–263; Madden 2012, 4.)

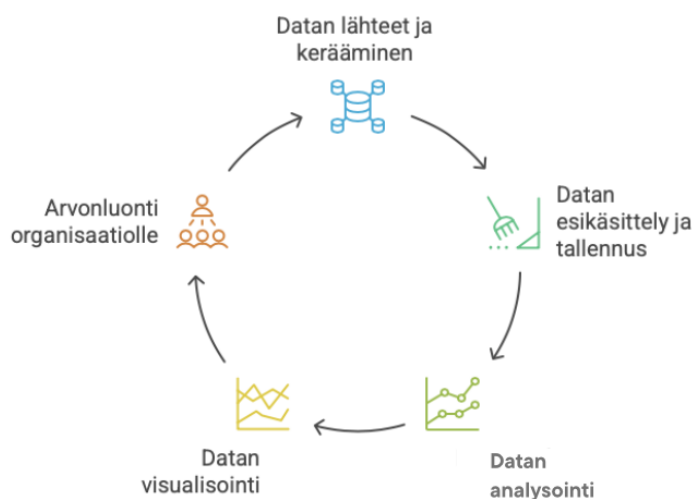
Laney'n 3 V:n kehyksen jälkeen ominaisuuksiin on myöhemmin lisätty myös esimerkiksi datan arvo (engl. value) (Dijcks 2012), totuudenmukaisuus (engl. veracity) (Schroeck ym. 2012) ja rakenteettomuus (engl. unstructuredness) (Suthaharan 2014). Datan arvolla viitataan siihen, että big datan avulla yritykset koittavat saada itsellensä arvoa ja arvokkaita oivalluksia päätöksenteon tueksi (Monino 2021, 263). De Mauro ym. (2014) määritteli big datan juuri arvon kautta. Heidän määritelmänsä mukaan big data edustaa

tietoresursseja, joille on ominaista niin suuri volyymi, nopeus ja vaihtelu, että se vaatii erityistä teknologiaa ja analyttisiä menetelmiä muuttaakseen sen arvoksi. Totuudenmukaisuus puolestaan on tärkeää puhuttaessa big datasta, sillä data voi usein olla epätäydellistä ja hyvien päätösten tekemiseksi päätösten taustalla olevan datan tulee olla mahdollisimman luotettavaa ja totuudenmukaista (Lukoianova – Rubin 2013, 12–13). Rakenteettomuus viittaa siihen, miten nykypäivänä saatava data ei välttämättä noudata tiettyä rakennetta. Tästä syystä perinteiset tietokannat eivät välttämättä pysty käsittelemään sitä (Suthaharan 2014, 70–73).

2.1.2 Big data -analytiikan prosessi

Big data -analytiikka on prosessi, jossa hyödynnetään edistyneitä analyysimenetelmiä ja mallinnusta valtavien, monimuotoisten tietomassojen käsittelyyn. Sen päätavoitteena on paljastaa piileviä yhteyksiä ja tuottaa uusia oivalluksia datasta hyödyntämällä datan louhintaa. (Nguyen ym. 2023, 525–526.) Datan louhinta keskittyy olemassa olevan datan analysointiin sekä yhteyksien ja kaavojen tunnistamiseen. Sillä tarkoitetaan prosessia, jossa pyritään löytämään datasta säännönmukaisuuksia ja tekemään niiden pohjalta ennusteita aiemmin tuntemattomista ilmiöistä. (Witten 2011, 4.) Big data voi tuoda suurta arvoa reaaliaikaiseen päätöksentekoon ja liiketoiminnalle, mikäli sitä pystytään prosessoimaan, analysoimaan ja tulkitsemaan oikein (Wilson 2019, 4). Edellisessä luvussa todettiin, kuinka datasta saadaan arvoa vasta, kun se on muutettu tiedoksi, eli oivalluksiksi ja ratkaisuuksi. Tästä prosessista on kyse, kun puhutaan big data -analytiikasta. Eli big data –analytiikalla viitataan keinoihin, järjestelmiin ja tekniikoihin, joiden avulla datasta pystytään saamaan loppujen lopuksi arvoa yritykselle. (Saggi – Jain 2018, 768.)

Big data -analytiikkaa kuvataan usein vaiheittain etenevänä prosessina. Tätä prosessia on kuvattu kuviossa 2.



Kuvio 2 Big data -analytiikan prosessi (mukaillen Saggi – Jain 2018)

Ensimmäinen prosessin vaihe on datan lähteet ja kerääminen, eli tunnistetaan, mistä big dataa syntyy. Dataa saadaan hyvin monesta eri lähteestä tänä päivänä ja nämä lähteet voidaan jakaa koneiden tuottamaan dataan, ihmisten tuottamaan dataan ja yritysten tuottamaan dataan. Dataa voidaan kerätä esimerkiksi verkkoharavoinnilla web-lokeista, sosiaalisesta mediasta, pilvipalveluista, reaalityetokannoista ja internetalustoista. Myös erilaisista koneiden sensoreista kerätty data on yleinen datan lähde tänä päivänä. (Saggi – Jain 2018, 768.)

Ennen tallentamista kerättyä dataa esikäsitellään ja valmistellaan erilaisilla teknisillä järjestelmillä, jotta siitä saadaan tallennus- ja analysointikelpoista. Tällaisia järjestelmiä ovat esimerkiksi erilaiset tietokantaratkaisut, jotka mahdollistavat datan tehokkaan käsittelyn, rakenteistamisen ja varastoinnin suurissa mittakaavoissa. (Saggi – Jain, 2018.) Tallennusvaiheessa dataa voidaan käsitellä reaaliaikaisesti ja hajauttaa useille palvelimille. Samalla dataa vielä puhdistetaan ja integroidaan analyysia varten (Devi – Siddharth 2015; Dean – Ghemawat, 2008; Saggi – Jain, 2018).

Kun data on saatu esikäsiteltyä ja tallennettua, pystytään sitä alkaa analysoimaan. Tämän analyysiprosessin tarkoituksena on löytää datasta merkityksellistä tietoa käyttämällä esimerkiksi ennustavaa analytiikkaa, semanttista analyysia, koneoppimista sekä tilastollisia analyysimenetelmiä ja -teknologioita. (Saggi – Jain 2018, 771.) *Ennustava analytiikka* on analytiikan osa-alue, joka käyttää dataa ja tunnistaa siitä suhteita, numeroita, tekstisisältöä, verkkokäyttäytymistä tai vaikka visuaalisia kuvia hyödyntämällä esimerkiksi luokittelua ja regressiota. Tämän jälkeen ennustava

analytiikka soveltaa näitä löydöksiä tulevien trendien, tapahtumien ja käyttäytymismallien ennustamiseen. Nämä ennusteet ilmaisevat todennäköisyyksiä tietyn käyttäytymisen tai tapahtuman esiintymiselle tulevaisuudessa. (Hair 2007, 304–305.) *Semanttinen analyysi* puolestaan on prosessi, jossa analysoidaan tekstidataa tunnistamalla sanojen ja lauseiden merkityksiä sekä niiden välisiä suhteita. Tässä pystytään hyödyntämään klusterointia, eli menetelmää, jossa data luokitellaan eri ryhmiin samanlaisuuksien perusteella. Semanttinen analyysi voi keskittyä yksittäisten sanojen merkityksiin tai tarkastella lauseiden ja pidempien ilmaisujen merkityksiä. (Cambria – White 2014, 54.) *Koneoppimisessa* järjestelmät käsittelevät dataa itsenäisesti ja oppivat siitä. Ne kykenevät tekemään päätelmiä ja löytämään oivalluksia joko ohjatusti tai täysin autonomisesti ilman erillistä ihmisen ohjelmointia. (Goodfellow ym. 2016, 83.)

Big data -analytiikan sovellusalueita ovat esimerkiksi sosiaalisen median analytiikka, mobiilianalytiikka, käyttäytymisanalytiikka, videoanalytiikka, visuaalinen analytiikka, äänianalytiikka ja tekstianalytiikka. Sosiaalisen median analytiikka käsittelee reaaliaikaista big dataa, joka on peräisin sosiaalisen median sivustoilta, kuten Facebook, Instagram ja LinkedIn. Dataa saadaan myös blogeista ja mikroblogeista. Myös erilaiset arviointisivustot, kuten TripAdvisor, tuottavat tärkeää sosiaalista dataa. (Saggi – Jain 2018, 772.) Mobiilianalytiikassa analysoidaan dataa, jota saadaan, kun ihmiset käyttävät puhelimillaan eri sovelluksia ja alustoja (Yazti – Krishnaswamy, 2014 1–2). Käyttäytymisanalytiikalla puolestaan viitataan yksilöiden ja yhteiskunnallisten ryhmien käyttäytymismallien analysointiin. Videoanalytiikassa hyödynnetään edistyneitä teknologioita, joiden avulla koneet pystyvät analysoimaan videoiden sisältöjä. Tekstianalytiikassa suuria tekstiaineistoja analysoidaan esimerkiksi koneoppimisen avulla. Äänianalytiikka purkaa äänitiedostoja ja tunnistaa niistä tietoa esimerkiksi puheanalyysin avulla. (Saggi – Jain 2018, 772.)

Datan analysoinnin jälkeen data yleensä pyritään esittämään mahdollisimman havainnollistavassa muodossa, jotta siitä on helpompi tehdä oivalluksia. Tähän tarvitaan datan visualisointia. (Saggi – Jain 2018, 772.) Kehittyneemmät visualisointitekniikat mahdollistavat data-arkiston reaaliaikaisen integroinnin suoraan visualisointityökaluun, jossa voidaan tehdä tulkintaa helpottavia visuaalisia analyyskejä (Gonçalves ym. 2023, 14). Perinteisempiä visuaalisia raportointityökaluja ovat kaaviot ja graafit, jotka tiivistävät ja selkeyttävät analyysin tuloksia. Näillä työkaluilla voidaan tuottaa tiivistettyjä yhteenvedoja, joissa korostetaan datan keskeisiä havaintoja selkeillä

visuaalisilla esityksillä, kuten pylväsdiagrammeilla, diagrammeilla, verkostoilla, kartoilla tai viivakaavioilla. (Saggi – Jain 2018, 777.)

Big data -analytiikan viitekehyksen viimeisenä palasena ja näin ollen myös koko prosessin tavoiteltuna saavutuksena on arvon luominen organisaatiolle. Saggi ja Jain (2018) jakavat arvon saamisen kolmeen eri osaan big data -analytiikan yhteydessä. Ensimmäinen arvo löydetään. Tässä vaiheessa yritys kerää tietoa eri datalähteistä ja käsittelee sen niin, että se saadaan käyttökelpoiseksi päätöksenteossa. Tämän jälkeen dataa hyödynnetään strategisten tavoitteiden, kuten markkinoinnin, tehostamiseen. Tavoitteena on hyödyntää dataa niin, että siitä syntyy todellista lisäarvoa yritykselle ja asiakkaille, kuten esimerkiksi mahdollisuus kohdistaa kampanjoita tarkemmin. Ja lopuksi kerätty tieto muutetaan konkreettiseksi oivalluksiksi ja toimintasuunnitelmiksi. Tämä voi tarkoittaa sitä, että yritys käyttää dataa strategisten päätösten tukena ja saa aikaan liiketoimintahyötyjä, kuten kustannusten säästöjä ja liikevaihdon kasvua. (Saggi – Jain 2018, 775–780.)

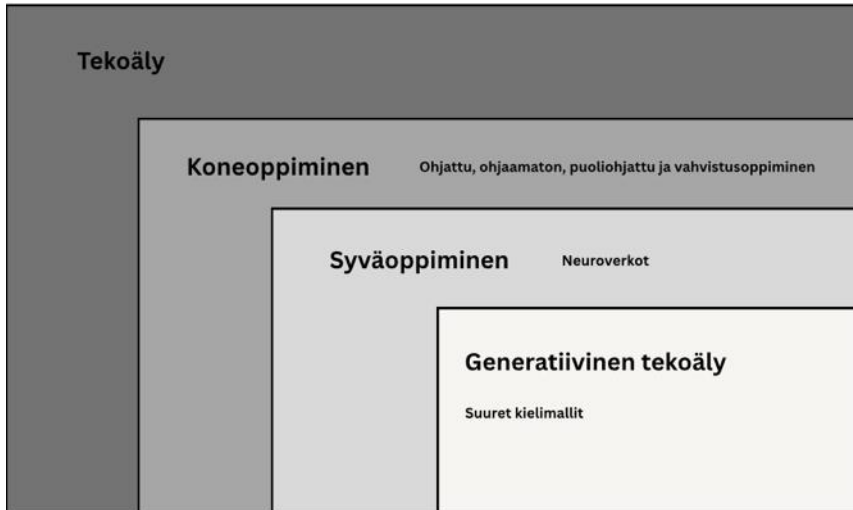
2.1.3 Tekoälyn määritelmä ja algoritmit

Tekoälylle ei ole tiettyä vakiintunutta määritelmää, vaan useat eri tutkijat määrittelevät sen eri tavoilla (Duan ym. 2019, 63). Yksi usein käytetty määritelmä tekoälylle on Haenleinin ja Kaplanin (2019b) tekoälyn määritelmä. He määrittelevät tekoälyn järjestelmän kyvyksi tulkita oikein ulkoista dataa, oppia tästä datasta ja käyttää oppia tiettyjen tavoitteiden ja tehtävien suorittamiseen joustavan mukautumisen avulla (Haenleinin ja Kaplanin 2019b, 17). Toinen usein siteerattu määritelmä tekoälylle on Russelin ja Norvigin (2016) määritelmä, jonka mukaan tekoäly on teknologisten komponenttien kokonaisuus, joka kerää, käsittelee ja toimii datan pohjalta tavoilla, jotka simuloivat ihmisen älykkyyttä. Kuten ihmiset, tekoälyratkaisut voivat soveltaa sääntöjä ja oppia ajan myötä hankkimalla uutta dataa ja tietoa.

Tässä tutkielmassa tekoäly ymmärretään Kaplanin ja Haenleinin (2019) määritelmän mukaisesti järjestelmänä, joka kykenee tulkitsemaan ulkoista dataa, oppimaan siitä ja käyttämään oppimaansa tavoitteiden saavuttamiseen. Tämä määritelmä soveltuu tutkielmaan siksi, että kilpailija-analyysin tehostamisen kannalta tekoälyn ei välttämättä tarvitse jäljitellä ihmismäistä ajattelua. Sen sijaan olennaista on järjestelmän kyky käsitellä suuria määriä ulkoista dataa, tunnistaa olennaiset ilmiöt ja tukea päätöksentekoa objektiivisesti ja tehokkaasti, joka ei aina ole inhimillisen ajattelun lopputulos. Kilpailija-analyysi on luonteeltaan datalähtöinen ja kontekstisidonnainen tehtävä, jossa parhaat

ratkaisut eivät välttämättä synny inhimillisen ajattelun jäljittelystä, vaan dataan perustuvasta rationaalisesta optimoinnista.

Tekoälystä puhuttaessa viitataan sen erilaisiin algoritmeihin, joita ovat esimerkiksi koneoppiminen, syväoppiminen ja generatiivinen tekoäly. Banh ja Strobel (2020) loivat viitekehyksen kuvaamaan tekoälyn ja sen alakäsitteiden hierarkkista järjestystä ja niiden välisiä suhteita. Tätä on havainnollistettu kuviossa 3.



Kuvio 3 Tekoälyn käsitteen alle kuuluvat alakäsitteet (mukaillen Banh – Strobel 2023, 2.)

Tekoäly on mallin yläkäsite, eli se kattaa alleen kaavion alemmat käsitteet. Alkuvaiheessa tekoäly perustui ennalta määriteltyihin sääntöihin ja päättelyketjuihin, mutta kehityksen edetessä koneet pystyivät oppimaan itsenäisesti niille syötetystä datasta ilman, että ihminen ohjelmoi niille valmiita toimintamalleja. Tätä kutsutaan *koneoppimiseksi*. Mitchell (1997) määritteli koneen oppimisen siten, että kone käyttää aiempaa dataa sille annetun tehtävän suorittamiseksi ja tätä koneen suoriutumista tehtävästä mitataan suorituskykymittareilla. (Goodfellow ym. 2016, 83.)

Koneoppiminen voidaan jakaa eri tyyppeihin sen mukaan, kuinka ihmisen toimesta ohjattua tämä koneen omatoiminen oppiminen on. Näitä eri tasoja ovat ohjattu oppiminen, ohjaamaton oppiminen, puoliOhjattu oppiminen ja vahvistusoppiminen. (Nguyen ym. 2023, 527–528; Mohammed ym. 2016, 7–11.) Ohjattu oppiminen on sitä, että kone opetetaan tiettyjen syötettyjen esimerkkien avulla ennustamaan tai tunnistamaan haluttuja asioita datan pohjalta (Sarker ym. 2020, 12). Ohjaamaton oppiminen on puolestaan sitä, että kone analysoi ja oppii datasta itsenäisesti ilman, että ihminen on ohjannut sitä antamalla sille malleja tai ohjeita, mitä sen tulisi löytää. Kone

siis etsii itsenäisesti kaavoja, rakenteita ja yhtäläisyyksiä sen pohjalta, mitä se huomaa. (Kreutzer – Sirrenberg 2020, 7.) Ohjattu oppiminen ja ohjaamaton oppiminen eroavat siis toisistaan sen perusteella, käytetäänkö ennakkotietoja vai ei. (Nguyen ym. 2023, 527–528.) Väliin jää puoli-ohjattu oppiminen, joka on näiden kahden tekniikan yhdistelmä. Siinä hyödynnetään ennalta määriteltyä dataa, eli esimerkiksi kuva kissasta ja määritelmä, joka kertoo tekoälylle, että kuvassa on kissa, mutta myös täysin määrittelemätöntä dataa, kuten kuva koirasta, ilman määritelmää, että kyseessä on koirakuva. (Sarker ym. 2020, 14.) Vahvistusoppimisessa puolestaan ei ole olemassa optimaalisia ratkaisuja, jotka olisi aluksi syötetty, vaan koneen on itse kokeiltava erilaisia ratkaisuja ja opittava virheistään ja voitoistaan (Kreutzer – Sirrenberg 2020, 7; Mohammed ym. 2016, 22), ja näin automaattisesti etsittävä optimaalisinta toimintatapaa (Sarker 2021, 4).

Syväoppiminen on koneoppimisen alaluokka, jossa hyödynnetään syviä neuroverkkoja. Neuroverkoilla tarkoitetaan toisiinsa kytkeytyneitä prosessoreita, joita nimitetään neuroneiksi. (Kreutzer – Sirrenberg 2020, 4.) *Syvät neuroverkot* ovat neuroverkkoja, joissa on hyvin monta niin kutsuttua piilokerrosta, eli todella monia kerroksia toisiinsa linkittyneitä neuroneita, jotka käsittelevät saatua dataa pala palalta (Kreutzer – Sirrenberg 2020, 5–8). Nämä syvät piilokerrokset mahdollistavat sen, että koneet voivat oppia monimutkaisia ja abstrakteja asioita ja piirteitä. Tästä syystä syvät neuroverkot ovat laajentaneet koneoppimisen mahdollisuuksia ratkaista yhä monimutkaisempia asioita. (Goodfellow ym. 2016, 83–84.)

Koneoppimisen ja erityisesti syväoppimismallien kehittyminen on mahdollistanut, että koneet pystyvät ymmärtämään ihmisten luonnollista puhetta ja tekstiä. Tätä kutsutaan koneiden *luonnollisen kielen käsittelyksi* (engl. Natural Language Processing, NLP). *Sentimenttianalyysi* on yksi luonnollisen kielen käsittelyn muoto. Se käyttää koneoppimisalgoritmeja havaitakseen ja luokitellakseen tekstin tietoa automaattisesti keskittyen tunnepitoisuuden tai asenteiden määrittämiseen. Sentimenttianalyysin ja luonnollisen kielen käsittelyn yhdistäminen on johtanut kehittyneempien algoritmien ja menetelmien kehittämiseen, ja tämän myötä ihmiskielen analysoiminen ja ymmärtäminen on helpottunut. (Srinivas ym. 2023, 468.)

Kuvion 3 sisimmässä nurkassa on *generatiivinen tekoäly*. Generatiivinen tekoäly on tekoälyn muoto, joka pystyy sille annetun kehotteen kautta luomaan uutta sisältöä, kuten kuvia, tekstiä tai ääntä. Yksi merkittävimmistä generatiivisen tekoälyn muodoista ovat

kielimallit. Kielimallit ovat tekoälymalleja, jotka on koulutettu ymmärtämään ja tuottamaan ihmiskieltä. Niiden kouluttamisessa on hyödynnetty ohjattua oppimista. *Suuret kielimallit* ovat puolestaan kielimalleja, jotka on opetettu valtavan suurella opetusaineistolla, johon voi kuulua esimerkiksi kirjoja, artikkeleita, verkkosivuja ja muuta kieltä sisältävää materiaalia. Suurin osa generatiivisista tekoälyistä perustuvat suuriin kielimalleihin. Nämä mallit tuottavat vastauksia käyttäjän luonnollisen kielen komentoihin, eli prompteihin. (Feuerriegel ym. 2024, 111–115.) Suurista kielimalleista esimerkkejä ovat GPT-mallit (Engl. Generative Pre-trained Transformer). Suuret kielimallit voidaan ohjelmoinnilla muuttaa keskusteleviksi generatiivisen tekoälyn systeemeiksi, kuten Open AI:n julkaisema ChatGPT. (Riemer – Peter, 2024, 3–4).

2.1.4 Tekoälyn ja big data -analytiikan integraatio

Aiemmissä luvuissa käsiteltiin ensin big data -analytiikkaa ja tekoälyä ja niiden eri mekanismeja. Tänä päivänä näiden kahden eri teknologian rajapinta on kuitenkin häilyvä, sillä niitä hyödynnetään yhdessä toisiaan täydentäen. (Davenport 2018, 74.) Big dataa ei enää pidetä pelkästään itsenäisenä entiteettinä, vaan se on riippuvainen tekoälyn, koneoppimisen ja syväoppimisen integraatiosta, jotka vievät big data -analytiikkaa eteenpäin (Nguyen ym. 2023, 528). Koneoppimisen ja syväoppimisen integrointi big datan analysointiin on avannut uusia mahdollisuuksia tutkimuksessa ja käytännön sovelluksissa. Tämä integraatio on muokannut tapaa käsitellä ja ymmärtää suuria tietoaaineistoja. Keskeinen muutos on tekoälyn tuoma tehokkuus ja nopeus, jolla tekoäly kykenee käsittelemään suuria datamääriä. Perinteiset data-analyysimenetelmät ovat usein hitaita, eivätkä ne aina sovellu modernin big datan moninaisen luonteen käsittelyyn. Tekoälyalgoritmit pystyvät vastaavasti käsittelemään monimuotoista dataa ja vielä murto-osassa perinteisten menetelmien käyttämästä ajasta. (Zaripova ym. 2023, 1–4.)

Toinen tekijä on se, että big data vaatii reaaliaikaista analytiikkaa, koska se on jatkuvasti muuttuvaa, sisältää ajallisia, sijaintiin liittyviä ja hierarkkisia rakenteita, ja sen tiedot virtaavat jatkuvasti. Tässäkin tekoäly on osoittautunut tehokkaaksi ratkaisuksi. (Özdemir – Hekim 2018, 66.) Big data -analytiikan ja tekoälyn synergiaa kutsutaan kirjallisuudessa tekoälytehosteiseksi big data -analytiikaksi (engl. AI Empowered Big Data Analytics) (Kumar ym. 2022) tai älykkääksi data-analytiikaksi (engl. intelligent analytics) (Zong – Guan 2025). Tekoälyn ja big datan integraatiota on kuvattu mullistavaksi, sillä big data tarjoaa ennen näkemättömän suuren tietomassan ja tekoälyn algoritmit pystyvät

käsittämään tätä datamassaa täysin uudella, tehokkaammalla tavalla (Agrawal ym. 2024, 138–139).

Davenportin (2018) tutkimuksessa näkyy data-analytiikan kehittyminen vaiheittain, ja kuinka big data ja sitten tekoälyalgoritmit mullistavat sitä. Tätä on kuvattu kuviossa 4 vaiheittaisena mallina, jossa on eri data-analytiikan aikakausia.



Kuvio 4 Liiketoiminta-analytiikan kehitys kohti tekoälytehostettua analytiikkaa (mukailen Davenport 2018)

Aiemmin analytiikka oli perinteistä analytiikkaa, jossa hyödynnettiin pääosin kuvailevaa analytiikkaan, joka perustuu menneiden havaintojen analyysiin. Nämä teknologiat käyttivät rajallisia, ”puhdistettuja” tietoaaineistoja yhdessä yksinkertaisten tilastollisten mallien kanssa, ja ne olivat ennalta jäsenneiltyjä ja toistuvia. (Nguyen ym. 2023, 525–526.) Ennen edistyneitä teknologioita suurin osa analytiikasta tehtiin käsin, joten se oli todella hidasta ja työlästä. Big data -analytiikkateknologioiden kehittyttyä analytiikasta tuli nopeaa ja monipuolista. Tänä päivänä big data -analytiikka on tekoälypohjaista. Tämä tarkoittaa, että big dataa hyödyntävään analytiikkaan integroitiin kehittyneempi tekoäly osaksi prosesseja. Tekoällyn eri muotojen täyden potentiaalin hyödyntäminen vaatii valtavia määriä dataa, merkittävää datan käsittelykapasiteettia ja edistyneitä tilastollisia menetelmiä, minkä takia tekoällyn ja big data -analytiikan integraatio on tuonut kokonaan uuden aikakauden. Tässä aikakaudessa korostuu vahvasti myös tekoällyn mahdollistama tietotyön automaatio. (Davenport 2018, 74.) Tekoällyn avulla big datasta saadaan enemmän irti, sillä ihmisen tekemänä big datan analysointi on hyvin hidasta ja työlästä, mutta tekoällyn ominaisuudet pystyvät tekemään sen paljon tehokkaammin (Gupta ym. 2018).

2.1.5 Tekoölyn hyödyntäminen big data -analytiikan prosessin eri vaiheissa

Luvun 2.1.2. kuvaajassa kuvattiin big data -analytiikan prosessia datan tuotannosta arvon luontiin asti. Tässä luvussa syvennytään tutkimuskirjallisuuteen siitä, mitä hyötyjä kehittynyt tekoöly on tuonut tämän big data -analytiikan prosessin eri vaiheisiin. Taulukkoon 1 on kerätty viimeaikaisia tutkimuksia, jotka havainnollistavat tekoölyn merkitystä tässä prosessissa.

Taulukko 1 Tekoölyn hyödyntäminen big data -analytiikan prosessin eri vaiheissa

Analytiikan vaihe	Tekoölyn funktio	Lähde
Datan kerääminen	Optimoi keräämisprosessia, tunnistaa relevantit lähteet ja suodattaa turhaa dataa	Wang ym. 2023; Yin ym. 2024
Datan käsittely	Parantaa datan luokittelua, priorisointia ja puhdistusta, mikä lisää prosessointitehokkuutta ja reaaliaikaisuutta	Wang ym. 2023; Yin ym. 2024; Sun & Sun 2024
Datan tallennus	Parantaa tietoturvaa ja yksityisyyttä, säästää tilaa ja tehostaa tallennusta	Lian & Zhang 2022
Datan analysointi	Tunnistaa piileviä kaavoja ja monimutkaisia yhteyksiä, tukee päätöksentekoa ja mahdollistaa reaaliaikaisen analyysin	Suominen ym. 2017; Rahmani ym. 2021
Datan visualisointi	Suosittelee automaattisesti sopivia visualisointeja, tukee interaktiivisuutta ja syvempää oivallusta	Yin ym. 2024

Datan keräämisen kontekstissa tekoöly voi optimoida keräämisprosessia hyödyntämällä graafisia neuroverkkoja, jotka tunnistavat sensorien välisiä yhteyksiä ja parantavat datan laatua (Wang ym. 2023, 1–4). Tekoöly voi myös löytää käyttäjän analyysitavoitteisiin tarvittavat datalähteet ja suodattaa pois turhan datan (Yin ym. 2024, 2).

Datan käsittelyn vaiheessa tekoöly voi puolestaan auttaa vähentämään kerättyyn dataan usein liittyviä epätarkkuuksia ja vääriä havaintoja (Wang ym. 2023, 1–4). Myös datan puhdistaminen saadaan optimoitua tekoölyn avulla (Yin ym. 2024, 2). Sun ja Sun (2024) tutkivat tutkimuksessaan, voidaanko datan luokittelusta ja käsittelystä saada nopeampaa ja resurssitehokkaampaa kahden eri tekoölyalgoritmin avulla. Datan käsittelyllä tutkimuksessa viitattiin sen luokitteluun ja priorisointiin. Tutkimuksen tuloksena oli, että datan luokittelutarkkuus parani ja se oli optimaalisempaa. Lisäksi datan käsittelyn

reaaliaikaisuus parani, kun tehtävien ajoitus ja resurssien käyttö optimoitiin tekoälymallin avulla. Näiden parannusten avulla prosessointitehokkuus kasvoi ja tehtävien suoritusajat lyhenivät.

Tekoälyalgoritmeja on myös tutkittu big datan tallennuksen yhteydessä. Lian ja Zhang (2022) tutkivat syvien neuroverkkojen tuomia mahdollisuuksia big datan tallentamisessa ja hallinnassa. Heidän tutkimuksensa mukaan tämä lähestymistapa paransi big datan tallentamiseen liittyvää tietoturvaa ja yksityisyyttä, tilan säästöä ja tehokkuutta.

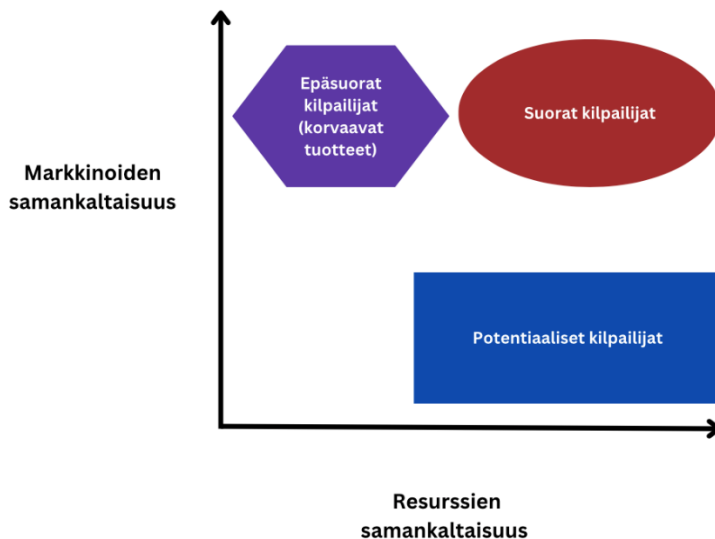
Big data -analytiikan analysointivaiheessa puolestaan esimerkiksi koneoppimisesta on hyötyä. Ohjaamattoman oppimisen avulla strukturoimattomasta datasta voidaan siis löytää piileviä kaavoja, rakenteita ja teemoja. Näistä löydöksistä pystytään järjestämään ja ryhmittelemään kokonaisuuksia, joista voidaan saada arvokasta informaatiota. (Suominen ym. 2017.) Tutkimukset ovat osoittaneet, että tekoäly ylittää perinteiset menetelmät valtavien tietomäärien tehokkaassa käsittelyssä. Koneoppimisen ja syväoppimisen avulla tekoäly voi tunnistaa datasta monimutkaisia kuvioita ja yhteyksiä. Lisäksi tekoäly mahdollistaa reaaliaikaisen datan analysoinnin, mikä on erityisen tärkeää dynaamisissa ja nopeasti muuttuvissa ympäristöissä. Tekoäly tarjoaa myös tehokkaita keinoja optimaalisten ratkaisujen löytämiseen sekä tukee monimutkaista päätöksentekoa analyysin pohjalta. Näiden ominaisuuksiensa ansiosta tekoäly auttaa ratkaisemaan big datan analysoinnin keskeisiä haasteita, kuten tietojen monimuotoisuutta, valtavaa tiedon määrää ja analyysien hitausongelmia. (Rahmani ym. 2021, 18–20.)

Tekoälystä on hyötyä myös datan visualisoinnissa. Se voi analysoida käyttäjän tarpeita, datan ominaisuuksia ja kontekstia automaattisesti ja suositella sopivaa visualisointia, kuten kaaviotyyppejä ja datan muunnoksia, ilman ihmisen manuaalista työtä. Tekoäly mahdollistaa interaktiivisen ja mukautuvan vuorovaikutuksen, jossa käyttäjä voi helposti tutkia dataa ja löytää syvällisempiä oivalluksia esimerkiksi esittämällä kysymyksiä koneelle. Tekoäly voi myös yhdistää visualisointituloksiin lisätietoa, kuten taustalla olevia syy-seuraussuhteita, mikä voi auttaa käyttäjää ymmärtämään dataa ja sen sisältöä syvällisemmin. (Yin ym. 2024, 3–5.)

2.2 Kilpailija-analyysin prosessi

2.2.1 Yrityksen kilpailijoiden määrittäminen

Kilpailijoita tutkittaessa ensiaskel on tunnistaa, ketkä ovat ne kilpailijat, joita yrityksen tulisi ottaa huomioon kilpailija-analyysissään (Porter 1980, 50). Kilpailija -käsite voidaan määritellä hyvin monella tavalla, ja se, ketkä nähdään yrityksen kilpailijoina, voi vaihdella. Peteraf ja Bergen (2001) loivat viitekehysten Chenin (1996) viitekehysten pohjalta, jonka avulla kilpailijoita voidaan tunnistaa luokittelemalla kilpailijoita sen suhteen, ovatko niiden resurssit samanlaisia ja missä määrin ne täyttävät saman asiakastarpeen. Tämä viitekehys on havainnollistettu kuviossa 5.



Kuvio 5 Kilpailukentän hahmottaminen (mukaillen Peteraf – Bergen 2001)

Kuviossa 5 pystyakselilla olevalla markkinoiden samankaltaisuudella viitataan siihen, kuinka vahvasti kilpaileva yritys on läsnä samoilla markkinoilla kuin vertailtava yritys. Tämä määritelmä on alun perin Chenin (1996, 106) luoma määritelmä markkinoiden yhteneväisyydelle kilpailijatunnistamisessa. Peteraf ja Bergen (2001) kuitenkin laajensivat käsitettä koskemaan myös sitä, missä määrin kilpailevat yritykset vastaavat samoihin asiakastarpeisiin. Tämä laajennus tehtiin siksi, koska se on paremmin linjassa markkinointikirjallisuuden kanssa ja laajentaa kilpailijoiden määritelmää. Vaaka-akseli puolestaan kuvaa sitä, miten samankaltaiset resurssit kilpailevilla yrityksillä on käytössään. Resurssien samankaltaisuus tarkoittaa, että missä määrin tietty kilpailija omaa yrityksen kanssa verrattavissa olevia strategisia voimavaroja. (Peteraf –Bergen

2002, 160.) Yritysresurssit sisältävät esimerkiksi kaikki varat, kyvykkyudet, organisaatioprosessit, yrityksen ominaisuudet, tiedot ja osaamisen. (Barney 1991, 101).

Yritykset, joiden resurssipohja on samanlainen ja jotka toimivat samanlaisilla markkinoilla, ovat yrityksen suoria alan sisäisiä kilpailijoita. Vastaavasti yritykset, joiden resurssit ovat kilpailevan yrityksen kanssa samankaltaisia, mutta jotka eivät vastaa samoihin asiakastarpeisiin, ovat ruudukon kaakkoiskulmassa. Nämä ovat potentiaalisia uusia tulokkaita markkinoille. Esimerkiksi pitopalvelu ja paikallinen ravintola ovat resurssipohjaltaan hyvin samankaltaisia: molemmilla on keittiövälineitä, ravintola-alan osaajia ja arvostettu brändi. Ne eivät kuitenkaan vastaa samanlaiseen asiakastarpeeseen, mutta voivat kuitenkin laajentaa helposti toimintaansa myös toisen kilpailijan toiminnan puolelle, koska niillä on jo suuri osa vaadittavista resursseista. Näiden lisäksi tärkeitä huomioon otettavia kilpailijoita mallin mukaan ovat myös luoteiskulmassa olevat yritykset, joilla ei ole samankaltaiset resurssit, mutta jotka palvelevat samoja asiakastarpeita. Tähän luokkaan kuuluvat epäsuorat kilpailijat, kuten korvaavat tuotteet. Esimerkiksi lentoyhtiöiden kilpailijoiksi luokituvat myös muut liikenneyhtiöt, kuten bussi- ja junayhtiöt, sillä asiakastarve on sama: vaivaton kuljetus. Viimeisenä ovat yritykset, joilla ei ole samoja resursseja eivätkä ne toimi samoilla markkinoilla, eli ne ovat kilpailukentän ulkopuolella. Ei kuitenkaan ole poissuljettua, etteivätkö nämä yritykset, jotka eivät juuri tällä hetkellä ole kilpailijoita, tulisivat tulevaisuudessa alalle. (Peteraf – Bergen 2002, 161–162.)

Myös erityisen tunnetun Porterin (1980, 50) määritelmän mukaan kilpailijoiksi voidaan määritellä nykyisten samalla toimialalla toimivien kilpailijoiden lisäksi myös yritykset, jotka eivät juuri tällä hetkellä toimi kyseisellä toimialalla, mutta jotka voisivat tulla alalle esimerkiksi alhaisten kustannusten takia. Tämän lisäksi huomioon tulisi ottaa myös yritykset, jotka saisivat synergiaetuja alalle tulosta nykyisen harjoittamansa liiketoiminnan kautta sekä yritykset, joille olisi luonnollinen jatkumo siirtyä kyseiselle alalle, kun ottaa huomioon heidän nykyisen liiketoimintansa. Porterin mukaan myös asiakkaat ja toimittajat voivat olla potentiaalisia uhkia, sillä ne voivat alkaa itse harjoittaa kilpailevaa toimintaa. Myös yritysostot ja fuusiot on hyvä ottaa huomioon, sillä ne voivat hetkessä muuttaa heikon kilpailijan vahvaksi toimijaksi markkinoilla tai vahvistaa entisestään olemassa olevaa vahvaa asemaa. (Porter 1980, 50.)

2.2.2 Kilpailija-analyysin määritelmä

Porterin (1980; 1985) merkittävien kilpailija-analyysiä koskevien julkaisujen jälkeen kilpailija-analyysi on ollut keskeinen osa yritysten strategista päätöksentekoprosessia (Hatzijordanou ym. 2019, 415). Kilpailija-analyysin määritelmät ovatkin suurimmaksi osaksi peräisin 80- ja 90-luvuilta, jolloin Porter ensimmäistä kertaa loi viitekehyksen kilpailija-analyysille. Kilpailija-analyysin yhteydessä käytetään myös usein termiä kilpailijaseuranta. Porterin (2004, 47) mukaan kilpailija-analyysi on prosessi, jonka tarkoituksena on tuottaa yrityksen strategisen johtamisen tueksi hyödyllistä tietoa sen kilpailijoista. Bennet (2003, 336) määritteli kilpailija-analyysin sen keskeisten osien kautta. Hänen määritelmänsä mukaan kilpailija-analyysin tarkoituksena on kilpailijoiden nykyisen toiminnan analysointi, tulevien aikeiden ennakointi ja mahdollisten uhkien havainnointi. Rouach ja Santi (2001, 553) puolestaan korostivat määritelmässään tiedon keräämistä, käsittelyä, tallentamista ja jakamista organisaation kaikille tasoille, jotta jokainen organisaatiossa ymmärtää yrityksen position ja yritys pystyy suojautumaan kilpailun uhilta. Bernhardt (1994, 13) puolestaan korosti omassa määritelmässään kilpailija-analyysin analyttistä luonnetta, jossa dataa kerätään esimerkiksi kilpailijoiden suorituskyvystä, aikeista, osaamisesta ja asemasta ja tämä data muunnetaan hyödynnettäväksi oman yrityksen strategisessa johtamisessa.

Kilpailijaseurantaa on myös lähestytty taloudellisen kilpailijaseurannan (engl. competitor accounting) näkökulmasta, jossa painotetaan kilpailijoiden taloudellisten kyvykkyyksien analysoimista (Heinen – Hoffjan 2005, 18). Taloudellinen kilpailija-analyysi sisältää esimerkiksi eri taloudellisten tunnuslukujen, kuten sijoitetun pääoman tuottoprosentin, käyttökateen ja liikevoittoprosentin analysoinnin sekä yrityksen taserakenteen arvioinnin keskeisillä mittareilla. Lisäksi kilpailijoiden strategisten ja operatiivisten suunnitelmien, kuten investointien ja markkinointikampanjoiden arviointi, kuuluu olennaisena osana kilpailija-analyysiin. (Järvenpää ym. 2017, 443–448.)

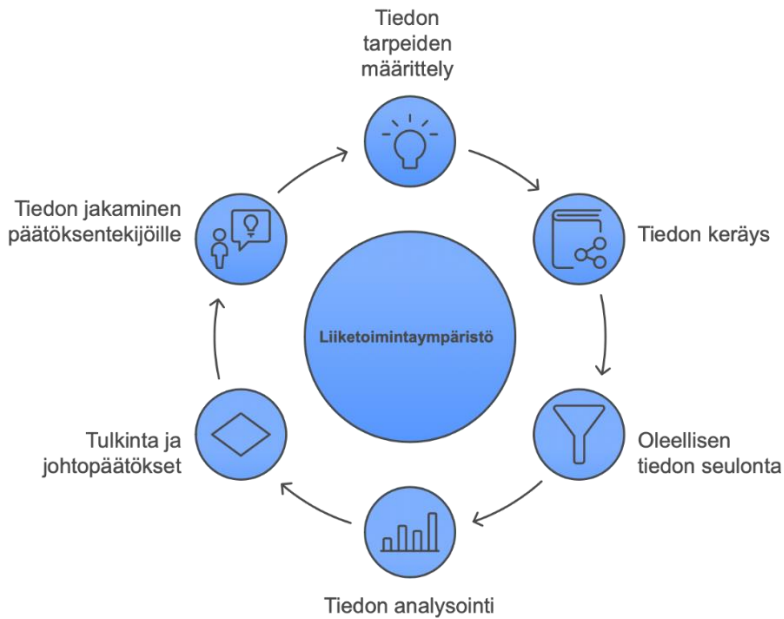
Kilpailija-analyysi voidaan nähdä olevan olennainen osa sekä strategisen johtamisen että markkinoinnin johtamisen kirjallisuutta. Nag ym. (2007) kehittivät sittemmin yleisesti hyväksytyyn määritelmän strategisesta johtamisesta. Tässä määritelmässä ulkoisen ympäristön huomioon ottaminen on olennainen osa strategista johtamista. Täten kilpailija-analyysin voidaan nähdä olevan olennainen osa strategista johtamista. Myös markkinointisuunnitelman kehittämiseen kuuluu olennaisesti ympäröivän maailman,

yrittäjien tilanteen ja markkinoiden analysointi, sisältäen myös kilpailijoiden analysoinnin (Kotler – Keller 2015, 50).

Tässä tutkielmassa kilpailija-analyysi nähdään Porterin (1980, 73) määritelmän mukaisesti prosessina, jonka tarkoituksena on tuottaa yrityksen strategisen johtamisen tueksi hyödyllistä tietoa sen kilpailijoista. Kilpailija-analyysi nähdään myös Bernhardtin tapaan (1994, 13) luonteeltaan hyvin analyttisenä. Nämä määritelmät valikoituivat tutkimukseen, sillä tänä päivänä kilpailija-analyysi perustuu hyvin paljolti dataan, ja se muistuttaa yleistä data-analytiikan prosessia. Tästä syystä kilpailija-analyysin käsitteleminen prosessinomaisesti on johdonmukaista. Tutkimuksessa on tarkoitus tarkastella kilpailija-analyysiä nimenomaan big data -analytiikan ja tekoälyn näkökulmasta, jolloin analyysin luonne on myös hyvin analyttistä.

2.2.3 Kilpailija-analyysin toteuttaminen prosessimallin mukaan

Useat tutkijat ovat kuvailleet kilpailija-analyysin tekemistä prosessiksi, joka sisältää eri vaiheita. Porter (1980, 73) kuvasi työssään ensimmäisenä tätä kilpailija-analyysin prosessia, joka sisältää tiedon keräämisen, kokoamisen, luokittelun, analysoinnin ja viestinnän. Myös Freeman (1999) kuvasi kilpailija-analyysin tekemistä jatkuvaksi ja moniulotteiseksi prosessiksi. Prosessi alkaa tiedon systemaattisella keräämisellä useista eri lähteistä, sekä yrityksen sisältä että sen ulkopuolelta. Tämän jälkeen tieto jäsenellään ja analysoidaan, ja lopuksi sitä yhdistellään ja viestitään eteenpäin. Myös Fletcher ja Donaghyn (1994) painottavat määritelmässään kilpailija-analyysin prosessoimaista luonnetta. Heidän mukaansa kilpailija-analyysi voidaan kuvata prosessina, jossa kattavasti eri lähteistä kerätty tieto kilpailijoista kootaan tarkasti ja suunnitellusti järjestelmään, jotta sitä voidaan myöhemmin hyödyntää päätöksenteon tukena. Prosessissa myös tietojen pitäminen ajantasaisena ja käyttötarkoitukseensa sopivana on hyvin tärkeää ja keskeistä. Samoin Pirttilä (2000) käsitteli kilpailija-analyysiä monivaiheisena prosessina. Hän loi kuusivaiheisen prosessikuvauksen kilpailija-analyysin toteuttamisesta, joka kuvataan kuviossa 6.



Kuvio 6 Kilpailija-analyysin kuusivaiheinen prosessi (mukaillen Pirttilä 2000)

Prosessin ensimmäisessä vaiheessa lähdetään liikkeelle määrittämällä, mitä tietoa yritys tarvitsee kilpailijoistaan, eli määritetään, mikä tieto on arvokasta. Myös se, kenestä tieto kerätään, eli ketkä ovat yrityksen kilpailijoita, on osa tätä ensimmäistä vaihetta. (Pirttilä 2000.) Lähtökohta on, että kattava kilpailija-analyysi vaatii suuren määrää dataa (Porter 1980, 71). Yleisellä tasolla tärkeiksi asioiksi, joita kilpailijoista tulisi kartoittaa, on määritelty esimerkiksi kilpailijoiden heikkoudet ja vahvuudet, tavoitteet, strategiat ja niiden toteutus sekä tulevaisuuden toiminnan ennustaminen, erityisesti markkinoilla tapahtuvien kilpailijoiden strategisten muutosten yhteydessä (Kotler – Keller 2016; Gilligan – Wilson 2003, 178). Myös kilpailijoiden markkinointitaktiikoista on tärkeää kerätä tietoa. (Kotler – Keller 2016). Järvenpään ym. (2017, 447) mukaan kilpailijaseurannassa tulisi kerätä tietoa neljästä osa-alueesta: taloudellisesta menestyksestä (esim. kannattavuus, kustannusrakenne ja kasvu), strategiasta, rakenteesta ja resursseista (esim. henkilöstöresurssit, brändi, asiakassuhteiden luonne, laatukilpailumenestys ja tuotantokapasiteetti), taloudellisesta asemasta (rahoitus- ja taserakenteet) sekä toimintasuunnitelmista ja niihin vaikuttavista tekijöistä (esim. investoinnit, yritysostot, tutkimus ja tuotekehitys, hinta- ja laatukilpailukyky eri osa-alueilla ja kilpailukykykäyttämisen muutokset, kuten tuoteportfolion muutokset, markkinointikanavat ja henkilöstön koulutus).

Seuraavassa prosessin vaiheessa tietoa kerätään. Tässä kohtaa on tarpeen miettiä, mistä lähteistä tämä aiemmassa vaiheessa määritelty tarpeellinen tieto voidaan saada. (Pirttilä 2000.) Tarvittavaa tietoa voidaan kerätä julkisista lähteistä, kuten vuosikertomuksista, internetistä ja mediasta, mutta myös epävirallisemmista lähteistä, kuten esimerkiksi tavarantoimittajilta, entisiltä ja nykyisiltä työntekijöiltä sekä yhteisiltä asiakkailta (Porter 1980, 73; Ward 1992, 110). Calof ym. (2017, 666) tutkivat eurooppalaisten yritysten tiedonhankintaa kilpailijatiedustelussa, ja kyselyissä erityisen tärkeänä kanavana vastaajat pitivät: internetiä (67 %), julkaisuja (60 %), kaupallisia tietokantoja (52 %), asiakkaita (40 %), yrityksen omia työntekijöitä (39 %) ja sisäisiä tietokantoja (35 %). Myös alan asiantuntijoita, konferensseja, sosiaalista mediaa, toimittajia, järjestöjä ja yhteiskunnallisia toimijoita pidettiin osittain tärkeinä kilpailijatiedon keräämisen lähteinä.

Seuraavana prosessin vaiheena on oleellisen tiedon seulonta, eli turhasta datasta seulotaan tärkeä tieto, jotta data on mahdollisimman hyödyllistä. Kuten missä tahansa datan hallinnassa, tässäkin erilaiset informaatiojärjestelmät ovat ratkaisevassa asemassa onnistuneeseen tiedon seulontaan. (Saggi – Jain 2018, 770.) Datan hallintaa ja seulontaa käsiteltiin tarkemmin luvussa 2.1.2.

Kun oleellinen data on saatu seulottua, prosessissa siirrytään varsinaiseen analysointivaiheeseen, jossa voidaan hyödyntää erilaisia kilpailija-analyysimalleja. Kirjallisuudessa on luotu useita analysointimalleja. Se, mitä kilpailija-analyysi menetelmää sovelletaan, riippuu esimerkiksi johdon tarpeista, toimialasta sekä resursseista, joita yrityksellä on käytössään (Simmonds, 1981). Tutkimukset jopa suosittelevat, että hyödynnettäisiin useampaa analyysimallia (Guilding 1999).

Yksi analyysimenetelmistä on tunnettu Porterin (1980) kilpailija-analyysi, joka pyrkii analysoimaan kilpailijoita luomalla niille reaktioprofiilit. Reaktioprofiilien avulla pyritään arvioimaan ja ennakoimaan kilpailijoiden mahdollisia strategisia liikkeitä, reaktioita ja vastatoimia muiden toimijoiden strategisiin toimiin sekä toimintaympäristön muutoksiin. Viitekehyksessä tarkastellaan kilpailijan tulevaisuuden tavoitteita, nykyistä strategiaa, kyvykkyyksiä sekä yrityksen ja toimialan oletuksia. Viitekehys on avattu kuviossa 7.

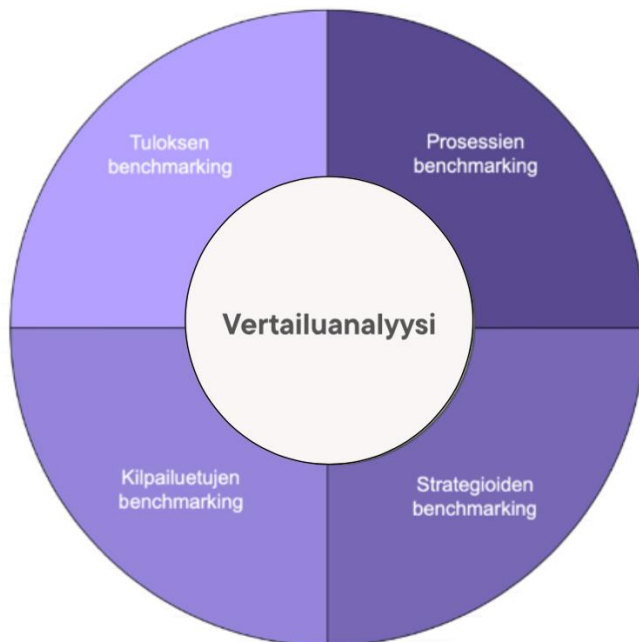


Kuvio 7 Kilpailija-analyysi (mukaillen Porter 1980, 63–67)

Mallissa lähdetään liikkeelle kilpailijoiden tulevaisuuden tavoitteiden hahmottamisella. Tavoitteet voivat olla taloudellisia tai ei-taloudellisia, mutta niiden ymmärtäminen voi auttaa hahmottamaan kilpailijan lyhyen ja pitkän tähtäimen suunnitelmia (Pirttilä 2000, 76), mikä puolestaan mahdollistaa kilpailijan strategisten liikkeiden ennakoinnin (Porter 2004, 50–51). Nykyisen strategian analyysissä keskitytään erityisesti kilpailijan markkina-asemaan, taloudelliseen tilanteeseen, tuotteisiin ja palveluihin sekä markkinakäyttäytymiseen (Pirttilä 2000, 78–79). Kyvykkyyksillä Porter viittaa kilpailijan resurssien ja osaamisen kokonaisuuteen. Kyvykkyyksien arviointiin sisältyy kilpailijan vahvuuksien ja heikkouksien tunnistaminen. (Porter 1980, 63–67.) Resurssit voivat olla sekä aineellisia, kuten koneet ja laitteet, että aineettomia, kuten henkilöstön osaaminen ja tuotekehitysosaaminen (Pirttilä 2000, 75). Resurssien ja osaamisen kokonaisuus vaikuttaa siihen, miten tehokkaasti kilpailija pystyy toteuttamaan omia strategisia liikkeitään, puolustautumaan kilpailijoiden toimenpiteitä vastaan ja mukautumaan

markkinoiden muutokseen (Porter 1980, 63–67). Oletukset kuvaavat sitä, miten kilpailija näkee itsensä ja asemansa. Porterin mukaan kilpailijan oletukset omasta asemastaan ja toimialasta vaikuttavat sen strategiaan valintoihin ja reaktioihin. Jos nämä oletukset ovat epätarkkoja, ne voivat muodostaa ns. sokeita pisteitä, joita kilpailija voi hyödyntää. Analysoimalla näitä oletuksia voidaan ennakoida kilpailijan käyttäytymistä ja löytää strategisia etuja. (Porter 1980, 57–63.)

Toinen analysointimalli on vertailuanalyysi, jossa yritystä verrataan sen suoriin kilpailijoihin (Wright ym. 2019, 52). Vertailuanalyysi eriytetään kirjallisuudessa usein eri osa-alueisiin: prosessien vertailu, strategioiden vertailu, kilpailuetujen vertailu ja tuloksien vertailu (Prasnikar ym. 2005, 262; Anand ym. 2008). Nämä osa-alueet ovat havainnollistettu kuviossa 8.



Kuvio 8 Vertailuanalyysin eri osa-alueet (mukaillen Prasnikar ym. 2005)

Prosessianalyysin tarkoituksena on pystyä analysoimaan, miten kilpailijat ja muut yritykset suunnittelevat, toteuttavat ja valvovat liiketoimintaprosessejaan. Kilpailuetujen vertailun päätavoitteena taas on kerätä tietoa tekijöistä, joihin kilpailijoiden ja muiden yritysten kilpailuedut perustuvat. Kilpailijoiden suorituskykyä arvioidaan ja verrataan omaan suorituskykyyn esimerkiksi tasapainotetun tuloskortin (Engl. balanced score card) avulla. Strategioiden vertailun tehtävä on puolestaan tunnistaa tehokkaita strategioita,

joita muut yritykset ovat käyttäneet saavuttaakseen samankaltaisia tavoitteita. (Prasnikar ym. 2005, 262–269.)

Viidennessä kilpailija-analyysin prosessin vaiheessa aiemman analyysien pohjalta tehdään johtopäätöksiä ja oivalluksia. Tässä vaiheessa on hyvin tärkeää, että analyysivaihe on luotettava, jotta johtopäätökset pohjautuvat oikeaan tietoon. Lopuksi tehdyt päätökset jaetaan organisaatiossa niille tahoille, joita kyseinen päätös koskee, jotta mahdolliset muutokset saadaan sisällytettyä jokaiselle tarvittavalle yrityksen tasolle. (Pirttilä 2000, 19.)

2.3 Tekoälytehosteisen big data –analytiikan hyödyt kilpailija-analysissä

Erilaisia tekoälyn algoritmeja ja big data -analytiikan teknologioita on jonkun verran tutkittu kilpailija-analyysin kontekstissa. Big datan aikakaudella dataa on saatavilla enemmän kuin koskaan, ja tekoälyn tuomat mahdollisuudet sen analysoimiseksi ovat mullistaneet yritysten strategista työtä, mukaan lukien yrityksen ulkoisen ympäristön analysoimisen. Tekoälytehostettua big data -analytiikkaa on kilpailija-analyysin kontekstissa tutkittu ainakin seuraavilla osa-alueilla: asiakkaiden mielipiteiden selvittäminen kilpailijoista (Taherdoost & Madanchian, 2023), kilpailijoiden patenttihakkeet (Suominen ym. 2017), innovaatioiden ja teknologisten painopisteiden ymmärtäminen (Wang ym. 2020), omien ja kilpailevien tuotteiden vertailu (Jin ym. 2016), markkinoiden rakenteen ja kilpailija-asetelmien kartoitus (Guo ym. 2017) sekä kilpailijoiden hinnoittelustrategioiden seuraaminen. (Tong-On ym. 2021). Taulukko 2 koostaa nämä tutkimukset, niiden aiheet, keskeiset menetelmät ja niiden löydökset.

Taulukko 2 Tekoälytehosteinen big data -analytiikka kilpailija-analysissä

Tutkijat	Tutkimuksen aihe	Keskeiset menetelmät	Keskeiset löydökset
Taherdoost & Madanchian (2023)	Tekoälypohjainen sentimenttianalyysi kilpailija-analysissä	Sentimenttianalyysi, tekoäly	Sentimenttianalyysi auttaa ymmärtämään kilpailijoiden asiakkaiden näkemyksiä eri osa-alueilla
Suominen ym. (2017)	Big data ja ohjaamaton koneoppiminen kilpailijoiden patenttianalysissä	Ohjaamaton koneoppiminen, big data, patenttianalyysi	Algoritmit löytävät piileviä teemoja patenttidatasta, jotka eivät ole ilmeisiä perinteisillä menetelmillä
Jin ym. (2016)	Tekoälypohjainen big data -analytiikka tuotteiden internetarvosteluista	Big data -analytiikka, tekoäly	Tekoäly mahdollistaa tuotteiden ominaisuuksien vertailun asiakkaiden mielipiteiden perusteella

Guo ym. (2017)	Automaattinen ja reaaliaikainen kilpailija-analyysi tekoälyn ja big datan avulla	Luonnollisen kielen käsittely, verkkorobotit, ohjattu ja ohjaamaton koneoppiminen, visualisointi	Automaattinen järjestelmä auttaa markkinan ja kilpailija-asetelmien kartoituksessa sekä tunnistamaan asiakassegmenttien päällekkäisyyksiä
Tong-On ym. (2021)	Tekoälypohjaiset järjestelmät kilpailijoiden hinnoittelustrategioiden seuraamisessa	Tekoälypohjainen hinnoittelustrategian seuranta	Tekoälypohjaiset teknologiat ovat tehokkaampia hintaseurannassa kuin perinteiset menetelmät

Taherdoost ja Madanchian (2023) tutkivat tutkimuksessaan tekoälypohjaisen sentimenttianalyysin hyödyntämistä kilpailija-analyysin tekemisessä. Sentimenttianalyysi voi tutkimuksen mukaan tarjota tietoa kilpailijoiden asiakkaiden ajatuksista, kuten esimerkiksi siitä, mitä he ajattelevat kilpailijan hinnoista, asiakaspalvelusta, kilpailijoilta saamasta arvosta, asiakaskokemuksen ominaisuuksista ja tuotteista. Sentimenttianalyysin avulla voidaan siis vertailla, miten asiakkaat kokevat brändin tuotteet tai palvelut verrattuna kilpailijoihin. (Taherdoost ja Madanchian 2023, 2.)

Suominen ym. (2017) puolestaan tutkivat voiko big dataa analysoivan ohjaamattoman koneoppimisen avulla saada oivalluksia kilpailijoiden patenteista. Patenttidataa voidaan hyödyntää kilpailija-analyysissä, sillä se kertoo, millä alueella kilpailijat ovat tällä hetkellä tekemässä tutkimusta ja mitä he pitävät tärkeinä oman liiketoimintansa kannalta. (Wang ym. 2020, 10). Heidän tutkimuksensa tuloksista kävi ilmi, että algoritmi löysi patenttidatasta piileviä teemoja, jotka eivät olleet ilmeisiä perinteisillä menetelmillä, mikä mahdollisti tarkemman ja monipuolisemman kuvan yritysten innovaatioista ja teknologisista painopisteistä.

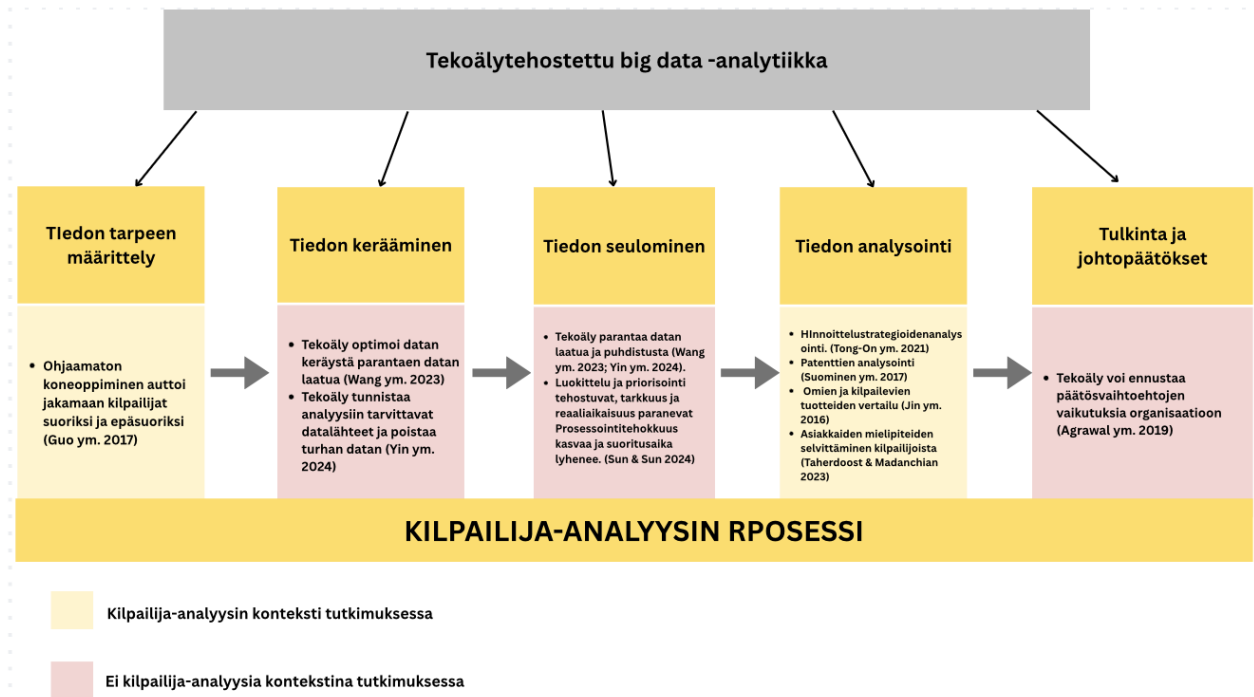
Yksi tapa ymmärtää kilpailijoiden heikkouksia ja vahvuuksia on analysoida asiakkaiden jättämiä tuotteiden internetarvosteluja. Jinin ym. (2016) tutkimuksessa tekoälypohjaista big data -analytiikkaa käytettiin tunnistamaan ja vertailemaan asiakkaiden mielipiteitä, jotka keskittyvät tiettyihin tuoteominaisuuksiin. Tämä prosessi mahdollisti tuotesuunnittelijoiden saavan syvällistä tietoa siitä, miten heidän tuotteensa ja kilpailijoiden tuotteet vertautuvat asiakkaiden näkökulmasta, ja tunnistamaan mahdolliset vahvuudet ja heikkoudet tuotesuunnittelussa.

Guon ym. (2017) tutkimuksessa pyrittiin kehittämään kustannustehokas, automatisoitu ja reaaliaikainen järjestelmä, joka hyödynsi tekoälyä ja big dataa kilpailijoiden analysoimiseen. Heidän kehittämänsä järjestelmä käytti datan keräämiseen verkkorobotteja, jotka keräsivät big dataa yhteensä 1 381 sovelluksesta, 100 892 käyttäjäkommentista ja 95 705 käyttäjäprofiilista. Järjestelmä hyödynsi luonnollista kielen käsittelyä analysoimaan tekstimuotoisia tietoja, kuten sovellusten kuvauksia ja käyttäjäarvioita, jotta se oli käyttökelpoista koneoppimisen algoritmeille. Ohjaamatonta koneoppimista hyödynnettiin sovellusten kuvausten analysointiin ja niiden yhteisten ominaisuuksien tunnistamiseen, jotta kilpailijat voitiin luokitella suoriin, epäsuoriin ja potentiaalsiin. Tämän avulla pystyttiin kartoittamaan markkinoiden rakennetta sekä kilpailijoiden välisiä eroja ja samankaltaisuuksia. Visualisointiin käytettiin multidimensionaalista skaalausta, joka visualisoi kilpailuasetelmat toimintojen samankaltaisuuden perusteella. Guon ym. (2017) kehittämä järjestelmä mahdollisti markkinan ja kilpailija-asetelmien kartoituksen, ja se auttoi tunnistamaan sovellusten vahvuuksia ja heikkouksia. Tämän lisäksi se auttoi oivaltamaan eroja ja samankaltaisuuksia kilpailijoiden välillä sekä tunnistamaan asiakassegmenttien päällekkäisyyksiä.

Tong-On ym. (2021) tutkivat puolestaan tekoälypohjaisten järjestelmien kykyä seurata kilpailijoiden hinnoittelustrategioita. Tutkimuksesta selvisi, että nämä teknologiat olivat tehokkaampia hintaseurannassa kuin perinteiset menetelmät, mikä mahdollisti hotellien mukauttaa omia hintojaan ja säilyttää kilpailukykyä markkinoilla.

2.4 Teoreettinen viitekehys

Tutkielman teoreettinen viitekehys rakentuu kirjallisuuskatsauksessa esitellyistä keskeisistä käsitteistä ja teorioista, jotka luovat pohjan tutkimuksen tarkastelulle. Viitekehys yhdistää aiempaa tutkimustietoa ja teoreettisia näkökulmia, jotka auttavat hahmottamaan tutkimusaiheen keskeisiä ulottuvuuksia. Viitekehys on esitetty kuviossa 9.



Kuvio 9 Teoreettinen viitekehys tekoälytehosteisen big data -analytiikan hyödyntämisestä kilpailija-analyysissä

Teoreettinen viitekehys kuvaa kilpailija-analyysiä etenevänä prosessimallina. Taulukko kuvaa, miten tekoälytehostettu big data -analytiikka yhdistetään jokaiseen kilpailija-analyysin viiteen eri vaiheeseen. Tätä on kuvattu nuolilla. Jokaisen vaiheen alle on kerätty, mitä näistä viidestä vaiheesta on tutkittu tekoälytehosteisen big data -analytiikan kontekstissa, eli miten tekoälytehosteista big data -analytiikkaa voidaan hyödyntää kussakin prosessin vaiheessa. Keltaisella pohjalla näkyvät tutkimukset ovat nimenomaan kilpailijatietoon liittyviä tutkimuksia ja vaaleanpunaisella taustalla olevat ovat yleisesti esimerkiksi data-analytiikan kontekstin tutkimuksia, joissa ei keskitytä erityisesti juuri kilpailijadatan hallintaan ja analysointiin, vaan jonkun muun datan hallintaan. Vaikka yleiset artikkelit datan keräämisestä ja siitä, miten tekoälytehosteinen big data -analytiikka voi tässä olla hyödyksi, pätevät myös mitä todennäköisimmin kilpailijadatan kontekstissa, voi kilpailijadatan keräämisessä ja seulonnassa olla myös jotain erityisiä piirteitä, joita näissä muuta dataa käsittelevissä tutkimuksissa, ei tule ilmi.

Aiempi tutkimus osoittaa, että tekoälytehosteinen big data -analytiikka voi tehostaa tiedonkeruuta ja seulontaa esimerkiksi optimoimalla datan laadun ja poistamalla tarpeettoman tiedon (Yin ym. 2024; Wang ym. 2023; Sun – Sun 2024). Kuitenkin, kun tarkastellaan tekoälyn hyödyntämistä kilpailijadatan analysoinnissa, tutkimus on ollut

hajanaista ja keskittynyt pääasiassa yksittäisiin osa-alueisiin, kuten patenttidatan analysointiin, tuotteiden ja hintojen vertailuun sekä asiakasnäkemyksen kartoittamiseen kilpailijoista (Tong-On ym. 2021; Suominen ym. 2021; Jin ym. 2016; Taherdoost – Madanchian 2023).

Kilpailija-analyysi on kuitenkin huomattavasti laajempi prosessi kuin pelkkä yksittäisten tietojen vertailu. Tällä hetkellä ei ole kattavaa tutkimusta siitä, kuinka tekoälytehosteista big data -analytiikkaa voisi hyödyntää koko kilpailija-analyysin prosessissa. Lisäksi tulkinta, johtopäätösten tekeminen ja tiedon jakaminen ovat keskeisiä vaiheita, joista ei löytynyt tutkimusta kilpailijatiedon kontekstissa. Vaikka tekoälyn on todettu kykenevän tekemään ennusteita datan perusteella (Agrawal ym. 2019), sen merkitystä kilpailijatiedon tulkinnaissa ja kilpailijatietoa koskevien strategisten päätösten tukemisessa ei ole vielä systemaattisesti tutkittu.

3 Metodologia

3.1 Tutkimuksen tieteenfilosofiset lähtökohdat

Se, miten todellisuus tutkimuksessa hahmotetaan, vaikuttaa olennaisesti tiedon rakentumiseen. Siksi myös tutkijan oma näkemys todellisuudesta on tärkeä avata ja käsitellä. Tutkimuksen taustalla olevat tieteenfilosofiset oletukset ohjaavat tutkimuksen toteuttamista. Tärkeitä tutkimuksessa käsiteltäviä tieteenfilosofisia lähtökohtia ovat tutkimuksen ontologiset, epistemologiset ja metodologiset lähestymistavat. (Eriksson – Kovalainen 2008, 11–12.)

Ontologia on filosofinen käsite, jossa pohditaan todellisuuden rakennetta ja luonnetta. Ontologiassa tutkija ilmaisee käsityksensä, mitä hän uskoo jonkun asian tai ilmiön olemassaolosta. Esimerkiksi miten tutkija hahmottaa tutkittavan aiheen ja miten hän ajattelee pystyvänsä osoittamaan tutkittavan asian todeksi. (Puusa – Juuti, 2020.) Ontologia voi olla subjektivistista tai objektivistista. Määrällinen tutkimus on usein ontologisesti objektivistista, eli siinä nähdään, että todellisuus on olemassa ihmisistä riippumatta. Tämä tarkoittaa, että tämän näkökulman mukaan maailmassa on ilmiöitä, rakenteita ja asioita, jotka ovat olemassa, vaikka kukaan ei havainnoisi tai ajattelisi niitä. Vastaavasti subjektivistisessa, toisella nimellä konstruktivistisessa ontologiassa, jota käytetään yleensä laadullisessa tutkimuksessa, todellisuuden nähdään muodostuvan jokaisen ihmisen subjektiivisten kokemuksen kautta sosiaalisessa vuorovaikutuksessa, joka voi muuttua kontekstin ja ajan myötä. (Eriksson – Kovalainen 2008, 13–14.) Kun tutkijat tekevät subjektivistista laadullista tutkimusta, he omaksuvat ajatuksen moninaisista todellisuuksista. Eri tutkijat näkevät todellisuuden eri tavoin, samoin kuin tutkittavat yksilöt. (Creswell – Poth 2016, 20–21.) Tässä tutkimuksessa oletuksena on subjektivistinen ontologia, eli todellisuus nähdään muodostuvan yksilöiden subjektiivisten kokemusten ja käsitysten kautta. Haastateltavat asiantuntijat muodostavat näkemyksensä aiemman kokemuksensa, tiedon ja organisaationsa käytäntöjen pohjalta. Haastatteluiden kautta saatavat näkökulmat edustavat asiantuntijoiden henkilökohtaisia ja kontekstisidonnaisia käsityksiä, jotka voivat muuttua ajan myötä esimerkiksi teknologian ja liiketoimintaympäristön kehittyessä. (Eriksson – Kovalainen 2008, 13–15.)

Epistemologialla puolestaan viitataan tieto-oppiin. Epistemologia käsittelee kysymyksiä, kuten mitä tieto on, mitkä ovat tiedon lähteet ja rajat (Eriksson – Kovalainen 2008, 15),

ja mitä ja miten voimme tietää jostakin asiasta. Grix 2002, 180). Epistemologialla viitataan siis oletuksiimme koskien tiedon luonnetta ja sitä, miten voimme tietoa hankkia. (Puusa – Juuti, 2020). Myös epistemologia voidaan jakaa objektivistiseen ja subjektivistiseen näkemykseen. Objektivistisen näkemyksen mukaan on mahdollista, että maailma on olemassa itsenäisesti ja riippumattomana teorioista. Subjektivistisen epistemologisen näkökulman mukaan ulkoinen maailma ei ole tavoitettavissa muuten kuin omien havaintojemme ja tulkintojemme kautta. Tässä tutkimuksessa myös epistemologinen näkökulma on subjektivistinen, eli tieto muodostuu ihmisten sosiaalisissa ja kulttuurisissa konteksteissa. Tiedon nähdään tässä tutkimuksessa muodostuvan haastateltavien henkilöiden kokemuksista ja tulkinnoista tekoälytehosteisen big data -analytiikan hyödyntämisestä kilpailija-analyysissä. (Eriksson – Kovalainen 2008, 15–16.) Tämä tutkimus on siis luonteeltaan tulkitseva, sillä tutkimuksessa sekä ontologinen että epistemologinen lähestymistapa ovat subjektivistisia (Grix 2002, 184).

3.2 Tutkimusote

Metodologia ja epistemologia liittyvät käsitteellisesti läheisesti toisiinsa. Siinä missä epistemologia keskittyy enemmän tieteen luonteen filosofiseen näkökulmaan, eli millaiseksi tiedon luonne ymmärretään (Eriksson – Kovalainen 2008, 15), metodologia puolestaan soveltaa näitä oletuksia käytäntöön. Metodologia kertoo, miten kutakin tutkimusongelmaa tutkitaan ja millaisia menetelmiä tiedon hankintaan ja analysointiin voidaan hyödyntää. Metodologia perustuu siis tutkijan ontologisiin ja epistemologisiin lähtökohtiin, eli siihen, millainen käsitys hänellä on todellisuudesta ja tiedon muodostumisesta. Nämä lähtökohdat sitten määrittävät, millaisia lähestymistapoja ja tutkimusmenetelmiä tutkimuksessa käytetään ja mitä rajoituksia niiden käytössä on. Metodologiset ratkaisut vaikuttavat siihen, mitä käytännön tutkimusmenetelmiä valitaan käytettävän tutkimuksessa, eli miten tutkimuksessa hankitaan ja analysoidaan tietoa todellisuudesta. (Puusa – Juuti, 2020; Eriksson – Kovalainen 2008, 15–16.)

Tutkimukset jaetaan karkeasti laadullisiin ja määrällisiin tutkimuksiin. Laadullisen tutkimuksen lähtökohtana on, että todellisuus on sosiaalisesti konstruoitunut eli laadullisessa tutkimuksessa pyritään ymmärtämään ihmisten kokemuksia ja tulkintoja tietyssä kontekstissa. (Eriksson – Kovalainen 2008, 4–5.) Tämä tarkoittaa, että tutkimuksen kohteena ovat ihmisten kokemukset, ajatukset, tunteet ja niiden merkitykset

(Puusa – Juuti, 2020). Laadullisissa tutkimuksissa kerätty tieto, eli data, on ensisijaisesti ei-määrällistä (Saldana 2011, 3).

Tämä tutkimus on laadullinen tutkimus, sillä tutkimusaiheesta ei ollut vielä juurikaan olemassa tieteellistä tutkimusta tekoälytehosteella, ja tällaisessa tilanteessa laadullinen lähestymistapa on erityisen hyödyllinen, koska aiheesta on vain vähän aiempaa tietoa tai vakiintuneita käsityksiä (Eriksson – Kovalainen 2008, 4-5). Laadulliset tutkimukset voivat olla joko tutkivia, kuvaavia, selittäviä, arvioivia tai näiden yhdistelmiä. Kuvailevat tutkimuskysymykset alkavat usein sanoilla ”kuka”, ”mitä”, ”missä”, ”milloin” tai ”miten”. Myös tiedonkeruun aikana esitettävät kysymykset noudattavat tätä rakennetta. Tämä tutkimus on ensisijaisesti kuvaileva, sillä tavoitteena on kuvata, mitä mahdollisuuksia tekoälytehosteinen big data -analytiikka voi tuoda kilpailija-analyysiin. Tutkimus on samalla selittävä, koska sen tarkoituksena ei ole vain kuvailla tekoälytehosteisen big data -analytiikan merkitystä, vaan myös selittää ja analysoida sen vaikutuksia kilpailija-analyysiin ja asiakaskokemuksen hallintaan. (Saunders ym. 2019, 180, 186.)

Tämä tutkimus hyödyntää abduktiivista lähestymistapaa. Abduktiivinen päättely tarkoittaa sitä, että tutkimus ei etene suoraviivaisesti teoriasta aineistoon tai aineistosta teoriaan, vaan tutkimuksessa liikutaan edestakaisin teorian ja aineiston välillä. Tässä lähestymistavassa vuorotellaan empirian ja teorian välillä, muodostaen vuoropuhelun niiden välille. Tätä kautta tutkittavaa ilmiötä on mahdollista tarkastella jo aiemmin tutkittujen teorioiden pohjalta, mutta lisäten siihen mahdollisesti uutta tietoa tai laajentaa vanhaa. Abduktiivinen teorian kehittäminen on siis avoin ja aineistoon herkästi reagoiva, mutta se hyödyntää myös olemassa olevia teorioita inspiraationa sekä apuna kaavojen ja merkitysten tunnistamisessa. (Saunders ym. 2019, 158.)

3.3 Tutkimusmenetelmä

Tässä tutkimuksessa haastattelut toteutettiin puolistrukturoituina teemahaastatteluina. Haastatteluiden avulla tiedonkeruusta voi saada syvällistä ja vuorovaikutteista. Haastatteluissa voidaan tarkentaa kysymyksiä tilanteen mukaan, korjata mahdollisia väärinkäsityksiä, selkeyttää ilmaisuja ja käydä vastavuoroista keskustelua haastateltavan kanssa. (Tuomi – Sarajärvi 2018.)

Haastattelututkimukset voidaan jakaa niiden struktuurin perusteella strukturoituun ja avoimiin strukturoimattomiin haastatteluihin. Tämä tarkoittaa sitä, kuinka ennalta määritelty haastattelun kulku on. Näiden kahden väliin jää puolistrukturoitu haastattelumuoto, jota tämä tutkimus noudattaa. Tämä tarkoittaa, että haastattelussa osa kysymyksistä on ennalta määrätty, mutta haastattelussa on tilaa myös vastaajan omalle pohdinnalle. (Fisher 2010, 175.) Tämän tutkimuksen puolistrukturoidut haastattelut ovat tarkemmalta määritelmältään teemahaastatteluja, eli haastatteluissa seurataan tiettyjä teemoja, joiden ympärille kysymykset on muodostettu. Muista puolistrukturoiduista haastatteluista poiketen, joissa kysymysten muoto ja järjestys saattavat olla samat kaikille osallistujille, teemahaastattelussa tätä ei edellytetä. (Hirsjärvi – Hurme, 2022.)

Puolistrukturoitu teemahaastattelu mahdollistaa haastattelujen joustavuuden, sillä tutkija voi suunnitella haastattelukysymykset etukäteen, mutta haastateltavien vastauksia ei ohjata tiukasti tiettyihin kysymyksiin, vaan keskustelu rakentuu enemmän teemojen ympärille (Hirsjärvi – Hurme, 2022). Tämä on tärkeää, kun tutkitaan monimutkaisia aiheita, jotka saattavat vaatia syvällistä pohdintaa, kuten tekoälypohjaisen big data -analytiikan hyödyntämistä kilpailija-analyysissä. Puolistrukturoitu haastattelu valikoitui menetelmäksi myös siitä syystä, että sen avulla voi löytää vastauksen ”miten” ja ”mitä” -alkuisiin kysymyksiin, joita tutkimuksessani käsitellään. (Eriksson – Kovalainen 2008, 82–83). Puolistrukturoidun teemahaastattelun mukaisesti tässä tutkimuksessa laaditaan teemoihin liittyviä kysymyksiä, jotka esitetään pääosin samassa muodossa ja järjestyksessä (Tuomi – Sarajärvi 2018). Tutkimuksessa tarkasteltavat teemat on muodostettu tutkittavan ilmiön sekä siihen liittyvien käsitteiden perusteella, joita on käsitelty tutkimuksen teoreettisessa osuudessa.

Teemahaastattelussa tärkeimpiä tehtäviä on haastatteluteemojen ja haastattelurungon rakentaminen. Jotta tutkimuksessa saataisiin kerättyä olennaista aineistoa tutkittavasta aiheesta, haastatteluteemat on johdettava teoriasta ja tutkimuksen teoreettisesta viitekehyksestä. Toisin sanoen ne pohjautuvat jo aiemmin saatavilla olevaan tietoon tutkittavasta ilmiöstä. (Hirsjärvi – Hurme, 2022.) Taulukossa 3 on esitetty tutkimuksen operationalisointitaulukko, jossa kysymykset on jaettu eri teemoihin ja joka havainnollistaa teoreettisten käsitteiden yhteyden tutkimuksen empiirisessä osiossa teemoittain kerättävään tietoon. Nämä teemat perustuvat tutkimuksen teoreettiseen viitekehykseen.

Taulukko 3 Tutkimuksen operationalisointitaulukko

Tutkimuskysymys	Osaongelmat	Teoria ja käsitteitä	Teemat ja kysymykset
Millaisia uusia mahdollisuuksia tekoälytehosteinen big data -analytiikka tuo kilpailija-analyysiin?	Millaisia mahdollisuuksia tekoälytehosteinen big data -analytiikka tuo kilpailija-analyysin tiedon tarpeiden määrittelyyn, keräämiseen ja seulontaan?	Big datan hallinta (käsite sisältää: big datan lähteet, kerääminen ja tallennus) Kilpailijoita koskevan big datan kerääminen ja seulonta tekoälytehosteisesti Kilpailija-analyysin prosessi: tiedon tarpeen määrittely, kerääminen ja seulominen	1. Tekoälytehosteisen big data -analytiikan tuomat mahdollisuudet yrityksen tiedon tarpeen määrittelyssä 2. Tekoälytehosteisen big data -analytiikan tuomat mahdollisuudet kilpailijatiedon keräämisessä 3. Tekoälytehosteisen big data -analytiikan tuomat mahdollisuudet oleellisen kilpailijatiedon seulomisessa Kysymykset: 1-5
	Millaisia mahdollisuuksia tekoälytehosteinen big data -analytiikka tuo kilpailijatiedon analysointiin, johtopäätösten tekemiseen ja tulkintaan?	Kilpailijoita koskevan big datan analysointi, tulkinta ja johtopäätösten tekeminen tekoälytehosteisesti Kilpailija-analyysin prosessi: analysointi, tulkinta ja johtopäätökset	4. Tekoälytehosteisen big data -analytiikan tuomat mahdollisuudet kilpailijatiedon analysoimisessa 5. Tekoälytehosteisen big data -analytiikan tuomat mahdollisuudet kilpailija-analyysin johtopäätösten ja tulkintojen tekemisessä Kysymykset: 6-14

Operationalisointitaulukossa vasemmalla on esitetty tutkimuskysymys, josta osaongelmat on johdettu. Osaongelmat on jaettu teoreettisen viitekehyksen mukaan kahteen osaan. Ensimmäinen osaongelma tutkii tekoälytehosteisen big data -analytiikan tuomia mahdollisuuksia kilpailijatiedon tarvemäärittelyssä, keräämisessä ja seulomisessa. Toisessa osaongelmassa keskitytään puolestaan kilpailijoista kerätyn big datan analysointiin, tulkintaan ja johtopäätösten tekemiseen tekoälytehosteisesti. Teoreettisessa viitekehysessä kuvataan, mitä mahdollisuuksia tekoälytehosteinen big data -analytiikka tuo viiteen eri kilpailija-analyysin prosessin vaiheeseen, ja nämä viisi eri prosessin vaihetta toimivat myös haastatteluteemojen taustalla: (1) Tekoälytehosteisen big data -analytiikan hyödyt ja mahdollisuudet yrityksen alustavassa tiedon tarpeen määrittelyssä kilpailija-analyysiä tehtäessä (2) Tekoälytehosteisen big data -analytiikan hyödyt ja mahdollisuudet kilpailijatiedon keräämisessä (3) Tekoälytehosteisen big data -analytiikan hyödyt ja mahdollisuudet oleellisen kilpailijatiedon seulomisessa (4) Tekoälytehosteisen big data -analytiikan hyödyt ja mahdollisuudet kilpailijatiedon

analysoimisessa ja (5) Tekoälytehosteisen big data -analytiikan hyödyt ja mahdollisuudet kilpailija-analyysin johtopäätösten ja tulkintojen tekemisessä.

Haastattelukysymykset on laadittu mahdollisimman suoriksi, jotta ne kannustavat haastateltavia kertomaan näkemyksiään vapaasti, sillä useimmiten suorat kysymykset tuottavat enemmän keskustelua kuin epäsuorat. Kysymykset on pyritty pitämään myös mahdollisimman neutraaleina. Neutraalit kysymykset pyrkivät välttämään haastattelijan ennako-oletuksia. (Eriksson – Kovalainen 2008, 84–86.) Tarkentavat kysymykset ja esimerkit on lisätty tukemaan haastateltavan ajattelua ja helpottamaan vastaamista. Niiden tarkoitus on siivittää ajattelua ja auttaa haastateltavaa hahmottamaan kysymysten laajuus, jotta haastateltava pystyy syventämään vastaustaan omista kokemuksistaan käsin (ks. liite 1). Myös joitain haastattelurungosta poikkeavia kysymyksiä kysyttiin joissain tilanteissa puolistrukturoidun haastattelumenetelmän raamien mukaisesti, mikäli haastateltava esimerkiksi poikkesi rajattujen kysymysten aiheista.

3.4 Haastateltavien valinta ja kuvaus

Laadullisessa tutkimuksessa osallistujat valitaan usein tarkoituksenmukaisesti tutkimuksen tavoitteiden mukaisesti. Tässä tutkimuksessa haastateltavien valinta perustui harkinnanvaraiseen näytteeseen (Creswell 2007, 125). Tällä tarkoitetaan, että tutkija valitsee haastateltavat sen perusteella, keiden arvioidaan tuottavan merkityksellistä tietoa tutkimukseen (Saunders ym. 2007, 226). Tarkoituksenmukainen näyte valikoitui haastateltavien valintamenetelmäksi, sillä tutkimuksessa haastateltavilta vaaditaan asiantuntemusta tutkittavasta aiheesta.

Potentiaalisia haastateltavia ja organisaatioita etsittiin hakukoneiden avulla. Näissä hauissa käytettiin relevantteja hakusanoja, kuten ”tekoäly ja big data -analytiikka yrityksen toimintaympäristön analysoimisessa” ja ”tekoäly ja big data kilpailija-analyysissä”. Myös LinkedInistä etsittiin asiantuntijoita hakusanoilla ”AI analyst” ja ”big data & AI analyst”. Alustava yhteydenotto organisaatioihin tapahtui sähköpostitse tai LinkedInin kautta. Yhteydenpidon tavoitteena oli esitellä tutkimuksen teemat ja aihealueet tiiviisti ja kysyä lupaa haastatteluun. Haastateltaville lähetettiin lyhyt yhteenveto tutkimuksessa käsiteltävistä aiheista ja teemoista (liite 2), suostumuslomake (liite 4) ja tietosuojailmoitus (liite 3).

Haastatteluun valittiin haastateltavia, joilla oli mahdollisimman laaja ammatillinen tai akateeminen kokemus data-analytiikasta, big datasta ja tekoälystä liiketoiminnan kontekstissa. Haastateltavien valinnassa painotettiin myös heidän kokemustaan yrityksen kilpailijoiden tai laajemman toimintaympäristön analysoimisesta näiden edellä mainittujen edistyneiden teknologioiden avulla. Tähän potentiaalisten haastateltavien osaamisen selvitystyöhön käytettiin LinkedIniä, jossa käyttäjät parhaimmillaan avaavat työuraansa hyvin yksityiskohtaisesti.

Esiselvityksen jälkeen potentiaalisimmille haastateltaville lähetettiin esitietomateriaali tutkimuksessa käsiteltävistä aiheista (liite 2), ja painotettiin, että tutkimusta varten etsitään asiantuntijoita, joilla on osaamista kyseisistä aiheista. Tämän esitietomateriaalin avulla tutkimukseen kutsutut pystyivät peilaamaan omaa osaamistaan ja mahdollista kontribuutiotaan tutkimukselle. Tällä toimintatavalla pyrittiin varmistamaan, että tutkimukseen löydetään henkilöitä, joilla on kokemusta ja osaamista tutkittavasta aiheesta.

Haastatteluihin osallistui yhteensä kahdeksan suomalaisissa yrityksissä toimivaa asiantuntijaa kolmesta eri taustasta: tekoäly- ja data-analytiikkapalveluita tarjoavia konsultteja, joiden yritykset kehittävät ja myyvät tekoäly- ja dataratkaisuja, sekä organisaatioiden edustajia, jotka hyödyntävät tekoälyä ja data-analytiikkaa omassa työssään, mutta eivät tarjoa sitä palveluna. Lisäksi mukana oli yksi tutkija, joka keskittyy muun muassa tekoälyn ja big data -analytiikan tutkimukseen liiketoimintakontekstissa. Kaikki haastateltavat (pl. tutkija) olivat johtavassa asemassa organisaatiossaan. Haastatteluajat vaihtelivat puolesta tunnista hieman yli tuntiin.

3.5 Aineiston keruu

Haastatteluiden sujuvan toteuttamisen kannalta on tärkeää suunnitella etukäteen käytännön järjestelyt, kuten ajankohta, paikka, arvioitu kesto sekä tarvittavat välineet (Hirsjärvi – Hurme 2010). Haastattelut toteutettiin yksilöhaastatteluina etäyhteyttä hyödyntäen joko Zoom- tai Teams-ohjelmiston avulla. Haastattelut äänitettiin haastateltavien luvalla kahdella eri tavalla: tietokoneeseen asennetulla nauhoitusohjelmistolla sekä varmuuden vuoksi myös puhelimen sovelluksella.

Haastatteluissa pyrittiin luomaan keskusteleva ilmapiiri, jossa haastateltavat saivat vapaasti jakaa näkemyksiään ja kokemuksiaan tekoälytehosteisen big data -analytiikan

hyödyntämisestä kilpailija-analyysissä. Tutkija ohjasi keskustelua ennalta määriteltyjen teemojen pohjalta.

Haastateltaville ei annettu tarkempia ennakkotietoja kuin tutkimuksessa käsiteltävät teemat ja aiheet (liite 2), eikä heiltä edellytetty valmistautumista etukäteen. Haastattelut pidettiin 5.2.2025–3.3.2025 välillä. Jokaista haastateltavaa haastateltiin kerran. Taulukossa 4 on esitetty tarkemmat tiedot haastatteluista.

Taulukko 4 Haastatteluiden tiedot

Tunnus	Asema organisaatiossa	Toimiala/tutkimusala	Haastattelun kesto
H1	Operatiivinen johtaja	Tekoäly IT-konsultointi	25 min
H2	Tekoäly ja data-analytiikka johtaja	Vähittäiskauppa, maatalous, teollisuus, terveydenhuolto	34 min
H3	Yliopettaja / dosentti	Tekoälytutkimus	29 min
H4	Tekoäly ja data-analytiikka johtaja	Energia-ala	38 min
H5	Perustaja partneri	Data-analytiikka & tekoäly konsultointi	29 min
H6	Tekoälyn ja datan muutosjohtaja	Tekoäly ja data konsultointi	1h 7 min
H7	Johtava tekoäly- ja datakonsultti	Strateginen tekoäly konsultointi	44 min
H8	Kaupallinen johtaja	Media-ala	33 min

Taulukossa 4 on esitetty tutkimukseen osallistuneiden henkilöiden tunnus, asema organisaatiossa, toimiala tai tutkimusala ja haastattelun kesto. Haastattelun keston on sisällytetty ainoastaan tutkimuksen kannalta relevanttien aineistojen sisältävä aika, eli alun esittäytymiset ja muut tutkimukseen kuulumattomat keskustelut on rajattu pois.

Tässä tutkimuksessa haastateltiin kahdeksaa asiantuntijaa, joka riitti tuottamaan syvällistä tietoa aiheesta. Haastateltavien määrä perustuu laadullisen tutkimuksen luonteeseen, jossa suurta otosta tärkeämpää on kerätyn aineiston laadullinen syvyys ja toistuvuuden saavuttaminen vastauksissa (Eriksson – Kovalainen 2008).

3.6 Aineiston analysointi

Haastatteluiden pitämisen jälkeen tutkijalla on hallussaan paljon dataa. Seuraava vaihe on tehdä selkoa siitä, mitä kerätty data pitää sisällään, eli analysoida sitä. Ensimmäiseksi kerätty haastattelumateriaali muutettiin tekstimuotoon, eli litteroitiin hyödyntämällä Wordin Transcribe-toimintoa. Tämän jälkeen litterointi tarkistettiin huolellisesti varmistaen, että se vastasi alkuperäistä äänitallennetta.

Tässä tutkimuksessa haastatteluaineiston analyysi toteutettiin yhdistäen teorian ohjaamaa koodausta ja aineistolähtöistä temaattista analyysiä. Teorialähtöinen koodaustapa tarkoittaa, että koodauksessa käytettiin valmiita teemoja ja koodiluokkia, jotka perustuvat aiempaan teoriaan ja teoreettiseen viitekehykseen. (Fereday – Muir-Cochrane 2006, 82.) Litteroidut haastattelut siis jäseneltiin ennalta määriteltyihin analyysikategorioihin, jotka perustuivat kilpailija-analyysin viiteen prosessivaiheeseen. Näin ollen analyysin alkuvaiheessa käytetty koodaus pohjautui valmiisiin luokkiin, joiden avulla haastateltavien vastaukset jaoteltiin niihin kilpailija-analyysin prosessin vaiheisiin, joita ne käsittelevät. Tämä lähestymistapa mahdollisti systemaattisen tarkastelun siitä, miten tekoälytehostettu big data -analytiikka vaikuttaa kilpailija-analyysin prosessin eri vaiheisiin.

Teorialähtöisen koodauksen jälkeen aineistoa tarkasteltiin tarkemmin temaattisen analyysin avulla kunkin prosessivaiheen sisällä, jolloin aineistosta nousseita teemoja voitiin tunnistaa ja analysoida (Fereday – Muir-Cochrane 2006, 82). Teema määritellään aineistossa toistuvaksi kaavaksi, joka vähimmillään kuvaa ja jäsentää havaintoja ja laajimmillaan tarjoaa tulkinnan ilmiön eri ulottuvuuksista. Temaattisessa analyysissä siis pyritään löytämään näitä toistuvia kaavoja aineistossa. Temaattinen analyysi mahdollisti sen, että vaikka aineiston perusjäsenitys oli ennalta määritelty, aineistolähtöiset havainnot saivat myös tilaa, mikä on tyypillistä abduktiiviselle analyysille. Tämän analyysitavan tavoitteena on tunnistaa tutkimusaiheen kannalta merkittäviä toistuvia teemoja haastatteluaineistosta. Prosessi edellyttää aineiston huolellista lukemista kerta toisensa jälkeen, jotta teemat voidaan tunnistaa. (Fereday – Muir-Cochrane 2006, 82.)

3.7 Tutkimuksen luotettavuuden arviointi

Laadullisessa tutkimuksessa tutkimuksen luotettavuuden arviointi poikkeaa usein määrällisestä tutkimuksesta. Määrällisessä tutkimuksessa käytetään usein tutkimusta arvioitaessa kriteereitä, kuten yleistettävyyttä, reliabiliteettiä ja validiteettiä. Näitä on

kuitenkin usein haastavaa soveltaa laadulliseen tutkimukseen, koska arviointikriteerien on oltava yhteensopivia laadullisen tutkimuksen metodologian, tavoitteiden ja taustaoletusten kanssa, jotka poikkeavat määrällisestä tutkimuksesta (Eriksson – Kovalainen 2008, 291–292). Koska tämä tutkimus perustuu subjektivistiseen ontologiaan ja epistemologiaan, sen arvioinnissa käytetään arvioinnin kriteereinä *uskottavuutta, riippuvuutta, vahvistettavuutta ja siirrettävyyttä* (Eriksson – Kovalainen 2008, 294), sekä arvioidaan tutkimuksen *vilpittömyyttä, johdonmukaisuutta* (Tracy 2010, 849) ja *refleksiivisyyttä* (Cassell ym. 2018, 524–525). Tätä tutkimusta arvioidaan seuraavaksi siis yhteensä seitsemällä eri arviointikriteerillä.

Uskottavuudella tarkoitetaan sitä, että tutkimuksen on oltava riittävän syvälinen ja tutkijan tulee olla tarpeeksi perehtynyt aiheeseen. Tämän lisäksi tutkimuksessa käytetyn empiirisen aineiston tulee olla riittävä, jotta tutkimuksessa esitetyt johtopäätökset ovat perusteltuja. Tärkeää uskottavuutta arvioitaessa on pohtia, tulisiko joku muukin tutkija samoihin päätelmiin, mikäli tällä olisi sama aineisto käytössään. (Eriksson – Kovalainen 2008, 294). Lukijan on siis pystyttävä seuraamaan tutkijan päättelyä ja hyväksymään tutkimuksen johtopäätökset (Cassell ym. 2018, 523). Tutkielmassa on pyritty uskottavuuteen käyttämällä lähteinä kattavasti mahdollisimman tuoreita vertaisarvioituja tutkimuksia käsiteltävistä aiheista. Myös haastatteluaineistosta pyrittiin saamaan riittävän laaja ja monipuolinen, jotta tulokset eivät perustuisi vain muutaman henkilön näkemyksiin. Kun samat teemat alkoivat toistua haastatteluissa, todettiin, että aineisto oli tarpeeksi laaja ja riittävä.

Riippuvuudella tarkoitetaan sitä, että tutkija tarjoaa riittävästi tietoa tutkimuksen prosessista. Tutkimusprosessin tulisi olla looginen, jäljitettävä ja dokumentoitu. (Eriksson – Kovalainen 2008, 294.) Tässä tutkimuksessa riippuvuus varmistettiin kuvaamalla yksityiskohtaisesti tutkimuksen eri vaiheet, kuten aineistonkeruu, analyysimenetelmä ja tulosten muodostaminen. Haastatteluprosessi, haastateltavien valintaperusteet ja haastattelurungon muodostaminen on myös kuvattu läpinäkyvästi.

Vahvistettavuudella tarkoitetaan sitä, että tutkimuksen löydökset ja tulkinnat eivät perustu pelkästään tutkijan omiin ajatuksiin ja oletuksiin, vaan ne voidaan perustella aineiston avulla (Eriksson – Kovalainen 2008, 294). Tutkimuksessa tämä varmistettiin avaamalla analyysiprosessi yksityiskohtaisesti sekä esittämällä tutkimuksen keskeiset löydökset suoraan haastatteluaineistosta nousevien esimerkkien avulla, jotta se, miten

johtopäätöksiin päästiin, olisi jokaisen jäljitettävissä ja arvioitavissa. Myös tutkimuksessa hyödynnetty abduktiivinen lähestymistapa, eli teorian ja empirian vuoropuhelu, mahdollistaa johtopäätösten tekemisen siten, että ne pohjautuvat aiempaan tutkimukseen ja kirjallisuuteen, mikä vahvistaa tämän tutkimuksen luotettavuutta.

Siirrettävyys puolestaan tarkoittaa, että tutkijan tulee osoittaa, miten tutkimuksen tulokset suhteutuvat aiempiin tutkimuksiin ja voivat olla sovellettavissa muihin konteksteihin (Eriksson – Kovalainen 2008, 294). Siirrettävyyttä on pyritty vahvistamaan analysoimalla tutkimustulosten yhteyksiä aiempaan kirjallisuuteen ja tunnistamalla samankaltaisuuksia aiempien tutkimusten löydösten kanssa. Myös tutkimuksen analyysi toteutettiin siten, että se vastasi teoreettista viitekehystä.

Vilpittömyyden arvioinnilla tutkimuskontekstissa viitataan siihen, että tutkijan tulee olla rehellinen omista lähtökohdistaan, reflektoida subjektiivisia näkemyksiään ja olla avoin tutkimusprosessin haasteista (Tracy 2010, 841–842). Vilpittömyyden arviointiin liittyy siis vahvasti tutkijan *reflektiivisyys*, eli tutkijan tietoisuus omasta roolistaan tutkimuksessa ja siitä, miten hänen taustansa, asenteensa ja valintansa vaikuttavat tutkimusprosessiin ja tulkintoihin (Cassell ym. 2018, 524–525). Tutkijan rooli ja lähtökohdat tutkimusaiheen tarkasteluun on tunnistettu ja niitä on pyritty refleктоimaan koko tutkimusprosessin ajan. Myös tutkijan mahdolliset ennako-oletukset tekoälytehostetusta big data -analytiikasta ja sen hyödyistä kilpailija-analyysissa on tiedostettu, ja niiden vaikutusta tulkintoihin on pyritty minimoimaan kriittisen analyysin ja aineistolähtöisen tarkastelun avulla.

Vilpittömyyden takaamiseksi on välttämätöntä tuoda esiin myös tutkimuksen haasteita ja rajoitteita. Yksi tutkimuksen rajoitteista liittyy siihen, miten haastateltavat ymmärsivät tutkimuksen keskeisen käsitteen, tekoälytehostetun big data -analytiikan. Vaikka tutkimuksen alussa pyrittiin varmistamaan yhteinen ymmärrys tarjoamalla esitietomateriaalia, ja taustatyön avulla perehdyttiin haastateltavien asiantuntemukseen, osa vastaajista tarkasteli aihetta suppeammasta näkökulmasta kuin tutkimuksen viitekehyksessä tarkoitettiin. Haastatteluissa nousi esimerkiksi esiin tapauksia, joissa haastateltavien käyttämä data-analytiikka kytkeytyi pelkästään pienimuotoiseen data-analyysiin, eikä näin ollen varsinaisen big datan hyödyntämiseen.

Vaikka generatiivinen tekoäly voi olla osa tekoälytehostettua big data -analytiikkaa, tämän tutkimuksen tarkoituksena ei ollut myöskään keskittyä ainoastaan generatiivisen

tekoälyn, kuten ChatGPT:n, hyödyntämiseen kilpailija-analyysissa. Osa haastateltavista tarkasteli kuitenkin tekoälyn merkitystä ainoastaan generatiivisen tekoälyn näkökulmasta ilman yhteyttä laajempaan big data -analytiikkaan, jolloin heidän näkemyksensä eivät täysin liittyneet tutkimuksen pääasialliseen viitekehykseen. Tämä vaikutti haastattelun aineistoihin siten, että osa keskusteluista liittyi enemmän generatiivisen tekoälyn käyttöön kuin varsinaiseen tekoälytehostettuun big data -analyysiin kilpailija-analyysissa. Tämä saattaa vaikuttaa tutkimuksen tuloksiin siten, että osa havainnoista perustuu kapeampaan käsitykseen tekoälytehostetusta analytiikasta kuin mitä tutkimuksessa tavoiteltiin. Haastatteluaineiston analyysissä on kuitenkin pyritty huomioimaan nämä eroavaisuudet ja erottamaan löydökset sen mukaan, missä määrin ne liittyvät varsinaiseen tekoälytehosteiseen big data -analytiikkaan.

3.8 Tutkimuksen eettisyyden arviointi

Tutkimuksen luotettavuuden arvioimisen lisäksi on tärkeää tuoda avoimesti esille tutkimuksen eettisyyttä. Tutkimusetiikka viittaa niihin periaatteisiin ja käytäntöihin, joita tutkijan on seurattava varmistaakseen, että tutkimus on eettisesti kestävä ja luotettava. Eettisten säännösten tavoitteena on ehkäistä epäeettistä toimintaa, väärinkäytöksiä ja vahingon aiheuttamista sekä edistää rehellisyyttä, oikeudenmukaisuutta ja osallistujien kunnioittamista. Lisäksi tutkijan on oltava tietoinen mahdollisista eettisistä huolenaiheista ja riskeistä jo tutkimuksen suunnitteluvaiheessa ja pyrittävä minimoimaan ne. (Sauders ym. 2023, 253–254.) Eriksson ja Kovalainen (2008, 64) muistuttavat myös, että tutkimusetiikka on koko tutkimusprosessin mukana kulkeva, joka koskee koko tutkimusprosessia, alkaen tutkijan ja tutkimuskohteen suhteesta ja päättyen tutkimuksen raportointiin ja julkaisuun.

Sauders ym. (2023, 257) määrittelevät eettisen tutkimuksen arviointikriteereiksi seuraavat periaatteet: rehellisyys, oikeudenmukaisuus ja avoimuus, toisten kunnioittaminen, haittojen välttäminen, osallistujien yksityisyyden suojaaminen, vapaaehtoisuus ja oikeus vetäytyä tutkimuksesta, tietoon perustuva suostumus, tietojen luottamuksellisuus ja anonymiteetti, vastuu tiedon analysoinnista ja tulosten raportoinnista, tietojen hallinnan säännösten noudattaminen sekä tutkijan turvallisuuden varmistaminen.

Tässä tutkimuksessa on noudatettu näitä eettisiä arviointikriteerejä johdonmukaisesti. Rehellisyys, oikeudenmukaisuus ja avoimuus on varmistettu kuvaamalla

tutkimusprosessi ja tulokset tarkasti ilman harhaanjohtavia tai vääristeleviä tulkintoja. Osallistujien kunnioittaminen on toteutettu selkeällä ja avoimella tiedottamisella sekä varmistamalla, että osallistuminen perustuu täysin vapaaehtoisuuteen. Oikeudenmukaisuutta, vapaaehtoisuutta ja tietoon perustuvaa suostumusta on korostettu lähettämällä jokaiselle haastatteluun osallistuvalla henkilöllä suostumuslomake sähköpostitse ennen haastattelua (ks. liite 4). Lomakkeessa on korostettu, että osallistuminen on täysin vapaaehtoista, eikä ketään voida pakottaa osallistumaan. Osallistumisesta kieltäytymisellä ei ole negatiivisia seurauksia ja haastateltavilla on mahdollisuus keskeyttää tai perua osallistumisensa milloin tahansa ilman seuraamuksia. Suostumus on pyydetty joko kirjallisesti tai sähköpostitse.

Osallistujien yksityisyyden suojaaminen on varmistettu siten, että tutkija on käsitellyt haastatteluaineiston huolellisesti, ottaen huomioon luottamuksellisuuden ja osallistujien tietosuojan. Aineiston käsittelyssä on noudatettu Turun yliopiston tietosuojavaatimuksia, ja datan hallinnassa on noudatettu Turun yliopiston datan hallinnan vaatimuksia.

Tutkimukseen osallistujille on toimitettu myös tietosuojailmoitus (ks. liite 3), jossa on esitelty tutkimuksen tarkoitus, tietojen käsittelytavat, käytännön toteutus sekä aineiston säilytys ja hallinta. Tietojen luottamuksellisuus ja anonymiteetti on varmistettu poistamalla tunnistettavat tiedot ja säilyttämällä aineisto turvallisesti. Lisäksi organisaation anonymiteetti on taattu, jotta voidaan minimoida mahdolliset riskit ja suojella osallistujien yksityisyyttä.

Vastuu tiedon analysoinnista ja tulosten raportoinnista on toteutettu huolellisesti siten, että analyysi on suoritettu asianmukaisesti ja tulokset on esitetty rehellisesti ilman manipulointia tai vääristelyä. Tutkimuksessa on myös noudatettu tietojen hallinnan säännöksiä, kuten tietosuojalainsäädäntöä ja tutkimusaineiston säilytysohjeita.

4 Tekoälytehosteisen big data -analytiikan tuomat mahdollisuudet kilpailija-analyysissä

4.1 Tekoälytehosteinen big data -analytiikka kilpailijatiedon tarpeen määrittelyssä

4.1.1 Tekoälytehosteinen big data -analytiikka ideoinnin tukena tiedon tarvemäärittelyssä

Kilpailija-analyysin ensimmäinen vaihe, jossa hahmotetaan, mitä tietoa kilpailijoista kerätään, on keskeinen koko loppu analyysiprosessin kannalta. Tässä kilpailija-analyysin vaiheessa haastatteluissa ei niinkään noussut esiin big data -analytiikan merkitys, vaan haastateltavat korostivat, että erityisesti generatiivinen tekoäly, kuten ChatGPT tai Copilot, voivat tuoda tähän prosessin vaiheeseen useita hyötyjä. Generatiivinen tekoäly voi esimerkiksi toimia sparrauskaverina, näkökulmien laajentajana ja ideoinnin rikastajana. Haastateltavat 5 ja 2 kuvasivat tätä seuraavasti:

Generatiiviset chat botithan on hyviä just tuollaisessa ajatusten herättelyssä, että sieltä voi just kysyä vastauksia, vaikka just että, ketkä ovat meidän kilpailijoita ja sieltä voi omat ajatukset sitten lähteä lentää vähän paremmin. H5

Voisi ihan tietysti tekoälyn avulla määritellä niitä tarpeita, lähteä sparrailemaan, miettiä mitä siellä voitaisiin ehkä tarvita, siinä voisi olla Copilot tai ChatGPT apurina (...) Mutta jos haluaisi katsoa trendejä, mitkä vaikka 2025 olisi ajankohtaisempia, niin esimerkiksi Perplexity AI:lla, kun se tuo ne lähteet siihen niin hyvin mukaan. H2

Haastateltava 7 korosti, että big data ei tässä kohtaa ole hyödyksi, mutta korosti myös, että tekoäly voi olla ajatusten herättelijä ja apuna lähteiden kartoittamisessa.

Mä ehkä tässä nyt jättäisin sen big datan tästä kohtaa pois ja sanoisin, että tekoälyhän siihen kannattaa käyttää kyllä, että tavallaan vähintään semmoisena sparrailukaverina, että hakee vähän sitä, että mitäs keskeisiä asioita tällä toimialalla, vaikka on? (...) Jutella taas sparrailukaverina sen tekoälyn kanssa ikään kuin siitä, että mitkä ne voisi olla ne järkevät lähteet? Mistä sinä sitä tietoa lähdet keräämään? H7

Tässä analyysivaiheessa big data -analytiikan merkitys jäi vähäiseksi, mutta generatiivinen tekoäly nähtiin hyödyllisenä ideoinnin ja näkökulmien herättelijänä sekä tiedon tarpeiden jäsentäjänä.

4.1.2 Tiedon tarpeiden määrittelyn tehostuminen tekoälyn avulla

Ideoinnin helpottamisen lisäksi tekoälytehosteinen big data -analytiikka myös tehostaa tiedon tarpeiden määrittelyn vaihetta. Useat haastateltavat mainitsivat, että tekoäly on esimerkiksi selvästi nopeuttanut tarvemäärittelyä ja helpottanut analyysirungon hahmottamista. Haastateltava 4 kertoi, että tekoälyn avulla tarvemäärittely ja suunnitteluvaihe ovat tehostuneet. Hänen mukaansa heidän yrityksessään suositellaan jo tekoälyn hyödyntämistä erilaisissa aivoriivissä.

Ideointi, tarvemäärittely ja suunnitteluvaihe ovat tehostuneet todella paljon, kun Chat GPT:n kaltaiset työkalut ovat mukana prosessissa (...) Meillä kannustetaan, että vaikka erilaisissa brainstorming sessioissa aina olisi jo tekoälyä apuna siellä. On kyllä huomattu yleisemmällä tasolla sitä tehostumista just tällaisessa alkuvaiheen suunnittelussa. H4

Haastateltava 8 puolestaan korosti, että generatiiviset mallit mahdollistavat nopean perusrungon luomisen kilpailija-analyysille. Tämän lisäksi haastateltava 1 mainitsi, että tekoäly auttaa myös hahmottamaan kilpailijakenttää nopeasti.

Ensinnäkin, nykyään generatiivisen tekoälymallit mahdollistaa ainakin semmoisen perusrungon tekemisen hyvin nopeasti. Ja sitten tietenkin niitten eri tietolähteiden kartoittamisen nopeasti. H8

Mutta nyt, mitä tää ChatGPT esimerkiksi mahdollistaa, niin kyllä se antaa aika nopeasti sen tiedon, että mitkä on tän toimialan ja sektorin ”vendorit” esimerkiksi, jotka toimittaa jotain tekoälypalveluita, niin sitten siitä tulee hyvin nopea lista näistä yhtiöistä. H1

Myös haastateltava 6 mainitsee, että tekoäly helpottaa kilpailija-analyysikehikkojen luomista, mutta lisää kuitenkin, että tekoälyn tuottamaa geneeristä sisältöä ei nähdä lopullisena, vaan lähtökohtana, jota voidaan ja tulisi rikastaa organisaation omalla hiljaisella tiedolla. Haastateltava kuvaa, että kun tekoälyä opetetaan yrityksen kontekstuaalisissa asioissa ja sille syötetään dataa, keskustelusta voi tulla enemmän yrityksen kontekstiin sopivaa ja näin kilpailija-analyysin toteuttamisen kannalta hyödyllisempää.

Sen kanssa voi lähteä sparrailemaan että: ”OK mikä olisi semmoinen esimerkiksi hyvä markkina ja kilpailijatiedon keruun kehikko?” ja sitten sitä voi lähteä rikastamaan esimerkiksi sellaisen näkökulmilla, mitä meillä on siinä yrityksessä semmoista hiljaista tietoa, mitä me tiedetään, mutta tekoäly ei tiedä, koska se tieto mitä sillä tekoälyllä on, niin se on aina geneeristä. H6

Tässä kommentissa korostuu ihmisen rooli tekoälyn tukena tiedon tarpeen määrittelyssä. Mikäli käyttäjä auttaa tekoälyä syöttämällä tälle enemmän dataa, johon tekoälyllä ei itsellään ole pääsyä, tiedon tarpeen määrittely voi olla tehokkaampaa.

Tekoäly, erityisesti generatiiviset mallit, koettiin hyödylliseksi työkaluksi tiedon tarpeiden nopeassa jäsentämisessä ja analyysirunkojen suunnittelussa, mutta niiden tuottamaa sisältöä pidettiin enemmän lähtökohtana, jota on tärkeää täydentää organisaation omalla asiantuntemuksella.

4.1.3 Generatiivisen tekoälyn haasteet ja rajoitteet tiedon tarpeiden määrittelyssä

Vaikka haastatteluissa nousi esiin hyötyjä, joita tekoäly tuo kilpailija-analyysin alkuvaiheeseen, esiin nousi myös rajoitteita ja kriittisiä huomioita. Haastateltavat korostivat erityisesti, että vaikka generatiiviset tekoälytyökalut voivat tarjota nopeasti ideoita ja alustavan rungon kilpailija-analyysille, niiden tuottama tieto on usein liian geneeristä eikä välttämättä ota huomioon organisaation yksilöllistä liiketoimintakontekstia.

Tämä käy ilmi erityisesti haastateltavien 6 ja 8 kommentista, joissa he kritisivat tiedon tarpeen määrittelyn ulkoistamista tekoälylle. Heidän mukaansa on yhä ihmisen vastuulla ymmärtää, mitä tietoa yrityksessä tarvitaan ja miksi. Haastateltavan 6 mielestä on jopa laiskuutta, että organisaatio ei itse mieti bisneslogiikkansa näkökulmasta, mitä tietoa kilpailijoista kannattaisi kerätä juuri kyseisen yrityksen näkökulmasta.

Toki me voidaan mennä kaikki ChatGPT tai Geminiin ja kysellä että: minulla on tällainen yritys ja mitä minun olisi hyvä tietää, tutkia ja tietää kilpailijoistani? Ja sieltähän jo saadaan erinäisiä vastauksia, mutta ne on tosi geneerisiä aina. Että ehkä enemmän se on siitä, että mikä on yritykselle itselle ja strategista, että mistä sitä tietoa olisi hyvä kerätä? Koska se just helposti aina ulkoistetaan sillä tavalla että: tekoäly sitten kertoo meille kaiken tarvittavan. No ei se kerro. Että se on aikamoista laiskuutta, jos ei edes itse mietitä bisneslogiikan näkökulmasta, että: jotta minä voisin ohjata tätä minun toimintaa paremmin, jotta minä voisin tehdä parempia skenaarioanalyyssejä, niin mitä kaikkea tietoa minä siihen oikeastaan tarvitsisin? H6

Ehkä siinä on myös semmoinen riski, että sitten helposti menee vähän sieltä mistä aita on matalin, kun saa sen perusrungon niin helposti, niin sitten lähdetään, että no tähän riittää ja lähdet tekemään sitä. H8.

Nämä havainnot tuovat esiin tekoälyn hyödyntämiseen liittyvän mahdollisen riskin, eli vaaran siitä, että teknologian helppokäyttöisyys johtaa pinnalliseen analyysiin ilman

kriittistä reflektointia. Myös haastateltava 5 jakoi tämän näkemyksen ja korosti, että analyysin alkuvaihe on edelleen vahvasti ihmiskeskeistä työtä.

En tiedä onko tää nyt vanhakantaista ajattelua, mutta mä näkisin, että se on kuitenkin sellaista, ei ehkä tekoälyvetoista tekemistä, enemmän sellaista ihmisvetoista tekemistä vielä toi määrittely. H5

Lisäksi esiin nousi generatiivisten mallien rajoitteet tiedon ajantasaisuuden suhteen. Haastateltava 6 huomautti, että kielimallien tietopohja ei ole jatkuvasti päivittyvä, vaan rajoittuu niiden koulutusdataan.

Mutta yhtä kaikki siis se data on koottu. Se ei ole se dynaamisesti päivittyvää, eli silloin semmoinen ”cut off point” että mitä se tietää. H6

Eli vaikka tekoälyn hyödyntämisestä voidaan saada tiedon tarvemäärittelyssä hyötyjä, liittyy siihen myös erinäisiä riskejä. Haastateltavat nostivatkin näkökulmia myös siihen, mitkä tekijät vaikuttavat siihen, milloin lisäarvon määrä on suurempi. Lisäarvon suuruus riippuu muun muassa käyttäjän osaamistasosta ja aiemmasta kokemuksesta kilpailija-analyysin toteuttamisesta. Tilanteissa, joissa analyysiä ei ole aiemmin tehty, tekoäly voi tarjota erityisen merkittäviä hyötyjä ohjauksen ja rakenteen muodossa.

Jos lähtisi ehkä täysin uudelta alalta liikkeelle, niin voisi tekoälyn avulla määrittellä niitä tarpeita, lähteä sparrailemaan. H2

Jos olet alkuvaiheessa tällaisissa asioissa, niin se on semmoinen 'it gets you up to speed'. H6

Sen sijaan syvää toimialaosaamista omaaville käyttäjille tekoäly voi toimia enemmän avustavana työkaluna.

Jos pyöritetään konsulttipumppua, joka tekee kilpailija-analyysia geneerisesti eri toimialoille, niin siitä voi olla iso hyöty sparrailukaverina (...). Mutta jos olet sen toimialan syvä asiantuntija, niin se on enemmän tekstin tuottamista tai ajatusten laittamista paperille. H7

Vaikka tekoäly nähtiin hyödyllisenä osana tarvemäärittelyä, haastatteluissa korostui myös sen tuottaman tiedon geneerisyys, ajantasaisuuden rajallisuus ja riski pinnalliseen analyysiin, ellei käyttäjä hyödynnä omaa asiantuntemustaan ja liiketoimintalogiikkaa kriittisesti. Saadun lisäarvon suuruus riippuu muun muassa käyttäjän osaamistasosta ja aiemmasta kokemuksesta kilpailija-analyysin toteuttamisesta.

4.2 Tekoälytehosteinen big data -analytiikka kilpailijatiedon keräämisessä

4.2.1 Kilpailijatiedon keräämisen tehostuminen tekoälytehostetun big data -analytiikan avulla

Kilpailija-analyysin seuraava vaihe tiedon tarpeen määrittelyn jälkeen on kilpailijatiedon kerääminen. Haastateltavien mukaan tekoälytehostettu big data -analytiikka mahdollistaa kilpailijatiedon keruun tehostumisen, sillä sen avulla voidaan käsitellä valtavia määriä sekä rakenteista että rakenteetonta dataa aiempaa tehokkaammin ja monipuolisemmin. Datan keräämisen yhteydessä haastatteluissa toistunut teema oli manuaalisen työn väheneminen ja automaation lisääntyminen, mitkä lisäävät prosessin tehokkuutta. Haastateltava 3 kommentoi, että tekoäly pystyy keräämään ja käsittelemään laajoja tietomassoja, jotka ennen olivat käyttäjälle työläästi saavutettavissa.

Eli tavallaan kyllä se on mahdollistanut sen, että pystytään koko globaali internetissä oleva data ottaa haltuun huomattavasti helpommin kuin ennen. (...) Aikaisemmin tämä vaati paljon käsityötä, mutta nyt tekoälyllä saadaan nopeasti raakadataa, jota ihminen voi jalostaa. H3

Haastateltavien mukaan perinteinen verkkoharavointi (engl. web scraping) nähtiin edelleen käyttökelpoisena keinona, mutta generatiivisen tekoälyn yhdistäminen tähän prosessiin mahdollistaa sen tehostumisen. Kilpailijatiedon keräämisen prosessi ei enää heidän mukaansa ole yhtä manuaalinen kuin ennen. Haastateltava 7 avaa tätä muutosta seuraavasti:

Jos kilpailijatietoa täytyisi jotenkin kerätä silleen massana, ja sä haluaisit kerätä yli toimialojen tai jonkun tyyppistä rekisteriä nyt kilpailijoista, niihin liittyvistä attribuuteista, niin nyt on päässyt vähän ”webskreippaamalla”, mikä voisi löyhästi olla big dataan liittyvä asia, niin sillä tavallaan pystyt seulomaan sieltä netistä asioita. Oot pystynyt jo ennen tekoälyä, mutta mä uskon, että jonkun ChatGPT apin tai muun vastaavan käyttäminen sen kanssa yhdessä niin varmasti tuottaa taas uudenlaisia mahdollisuuksia, että aikaisemmin se netin skreippaaminen on semmoista aika manuaalisia intensiivistä hommaa kuitenkin, että jokainen nettisivu on vähän erilainen ja ne rakenteet on erilaisia ja sun täytyi aikaisemmin se asia vähän kustomoida helposti silleen sivu toiselle, mutta se että, nyt sinä voit sitten helpommin raapia jonnekin sun tietokantaan niitä nettisivuja vähän sillä tavalla, ei rakenteellisessa muodossa ja sitten ehkä käyttää tekoälyä sitten sen tiedon... Se menee vähän tuon tiedonkeräykseen ja oleellisen tiedon seulontaan. H7

Haastateltava 3 toi esiin myös, että tekoälymallit mahdollistavat kielimallien avulla myös erikielisen datan hyödyntämisen, mikä on mahdollistanut yhä kattavamman kilpailija-

analyysin tekemisen yli maantieteellisten rajojen, missä ennen kielimuuri saattoi olla iso ongelma.

Pystytään tekemään analyysiä helpommin, vaikka Suomesta Aasiaan tai vaikka Etelä-Amerikkaan, että pystytään globaalisti tarkastelemaan tai mitä me ollaan tehty big data -analyysseja, niin siinä on tullut haasteena se kieli, mutta mun mielestä laajojen kielimallien tulo on helpottanut sitä puolta, koska ne pystyy sitten ymmärtämään ja kääntämään hyvinkin erilaisia kieliä.
H3

Useiden haastateltavien mukaan tekoäly- ja big data -analytiikkateknologiat tehostavat kilpailijatiedon keräämistä erityisesti nopeuden, skaalaavuuden ja työvaiheiden automatisoitavuuden näkökulmasta.

Jos ajatellaan, että jossain ajassa, että paljonko sä nyt vaikka tunnissa saat aikaiseksi, niin kyllähän varmasti enemmän saa tunnissa aikaiseksi. Ja ihan siitä lähtien, että sä voit käyttää sitä myös koodaus apurina siinä, että se tavallaan määrittelet sen ”skreipperin” sinne verkkoon sen tekoälyn avulla.
H7

Kyllä se se sitten isoin hyöty on se nopeus. Hyvin nopeasti saat sen perussetin tietyllä tavalla kasaan ja ei mene aikaa siihen, että sä lähdet kaivamaan sitä tietoa jostain vaan se tieto tietolähteet on tosi nopeasti käsillä. H8

Ehdottomasti skaalautunut, että meilläkin on pieni porukka talossa, jonka vastuulla on tällainen toimintaympäristön ja kilpailijoiden seuranta ja aikaisemmin, kun se on ollut manuaalista työtä, niin se on pystynyt tietyn määrän selailemaan kaikenlaisia tiedotteita ja muita ja sitten tekemään tiivistelmiä, niin kyllä se skaalautuminen on siinä tuonut tuommoista tehoa, että pystyy valtavia määriä sitten tiivistämään sen henkilön tueksi, joka sitä käyttää tai jonka vastuulla se on. H4

Haastateltavan 2 näkemyksen mukaan tekoälytehostettu big data -analytiikka ei ole sinänsä ratkaisevasti lisännyt organisaation mahdollisuutta kerätä määrällisesti enemmän dataa kuin ennen, sillä tekninen infrastruktuuri, kuten pilvipalvelut ja skaalautuvat järjestelmät, ovat olleet olemassa jo aiemmin:

No varsinaisesti siihen, että kuinka paljon sitä mahdollista kerätä, niin siinä ei ole ollut enenkään rajoitteita, koska meillä on ollut kuitenkin kaikkia pilviä aikaisemminkin ja skaalautuvuutta. H2

Tämä huomio osoittaa, että monissa organisaatioissa tekninen kapasiteetti ei välttämättä ole ollut tiedonkeruun pullonkaula, mutta monet haastateltavista korostivat, että kerääminen on tehostunut siten, että saadaan laajempi kokonaisuus erilaisista datalähteistä ja eri muotoista dataa taloudellisemmin ja helpommin:

Varsinkin jos olisi sellainen tilanne, että kilpailijakenttä on tosi laaja, että yksittäisten kilpailijoiden sijaan vaan seurata, että mitä jollain koko isolla toimialalla tapahtuu, niin nyt pystytään katsomaan varmaan tosi paljon laajempaa kokonaisuutta niitä datalähteitä, vaikka ne olisi ollut aiemminkin saatavissa, mutta ei ollut ehkä taloudellisesti mielekäästä sitten hyödyntää niitä, ainakaan isolla frekvenssillä. H5

Joo, siinä mekin ollaan tehty ehkä tällaisella vähän esitekoälymallilla, (...) ne ei ole vielä hirveän tekoälyvetoisia, enemmänkin perinteistä ”webskreippausta” (verkkoharavointia), niin mekin tietysti kerätään kilpailijoista tietoa, vaikka hintakilpailutusportaali tai sellainen, mistä me haetaan kilpailijoiden hintatietoja, niin tää kielimalli tekoäly tuo siihen sellaisen lisän, että myös laajemmasta valikoimasta lähteitä saadaan sitä vaikka hintasignaaleja tai signaalia siitä kilpailijoiden tilanteesta, että ei tarvitse niin tarkasti olla rakenteista tietoa, vaan sitten tekoäly avulla voidaan poimia sellaista strukturoimatonta tietoa helpommin. H4

Tekoälytehosteisen big data -analytiikan hyödyntäminen kilpailijatiedon keräämisessä on erityisen tehokasta silloin, kun tekoäly pystyy hyödyntämään rajapintoja ja integroimaan eri lähteistä tulevan datan yhteen analyysiprosessiin. Rajapinnat tarjoavat yhteisen perustan, jonka avulla ohjelmistot voivat keskustella keskenään. Niiden avulla tietojen ja toiminnallisuuksien jakaminen esimerkiksi applikaatioiden ja sivustojen välillä on mahdollista ilman, että ohjelmoijan tarvitsee tuntea toisen ohjelman sisäistä rakennetta. (Stylos – Myers, 2007, 50–52.). Haastateltavat toivat esiin, että generatiivinen tekoäly voi madaltaa teknistä kynnystä käyttää näitä rajapintoja. Haastateltava 2 kuvasi esimerkin, jossa ChatGPT:n avulla rakennettiin Python-malli Tilastokeskuksen datan hakemiseen:

ChatGPT:n ohjeiden avulla rakennettiin Python-malli, joka haki Tilastokeskuksen datat API:n kautta. H2

Lisäksi H2 totesi, että tekoäly on mahdollistanut pääsyn myös sellaisiin rajapintoihin, joita ei aiemmin teknisesti saavutettu:

Jos on vaikka vaikeampia rajapintayhteyksiä (...), niin kielimallien kanssa on pystytty saamaan ne käyttöön. H2

Kilpailijatiedon keruu on tehostunut merkittävästi tekoälytehostetun big data -analytiikan ansiosta. Haastateltavat korostivat tämän taustalla erityisesti manuaalisen työn vähenemistä, nopeutta, automaatiota ja rajapintojen helpompaa hyödyntämistä.

4.2.2 Tekoälytehosteisen big data -analytiikan mahdollistama automaatio kilpailijatiedon keruussa

Tiedonkeruun tehostumisen taustalla on tiedonkeruun automaatio. Haastattelujen perusteella tekoälytehosteisen big data -analytiikan mahdollistama tiedonkeruun automaatio nähtiin yhtenä merkittävimmistä kilpailija-analyysin tiedon keräysvaiheeseen vaikuttavista muutoksista. Useat haastateltavat kokivat, että tekoälypohjaiset työkalut, kuten tekoälyagentit, mahdollistavat kilpailijatiedon keräämisen huomattavasti pienemmällä manuaalisella työpanoksella kuin aiemmin. Perinteisesti kilpailija-analyysi on ollut aikaa vievä prosessi, jossa analyttikot ovat manuaalisesti keränneet ja seuloneet tietoa eri lähteistä. Nyt tekoälypohjaiset työkalut voivat automaattisesti skannata ja kerätä dataa mikä tehostaa prosessia.

Mahdollisuus just kerätä dataa niin monipuolisesti, pienellä vaivannäöllä, koska me saadaan hirvittävän paljon asioita automatisoitua, just sekä ohjelmallisesti että tekoälyä avusteisesti. H6

Haastateltava 7 tuo esiin merkittävän muutoksen kilpailijatiedon keruun teknisessä toteutuksessa. Hänen mukaansa tekoäly vähentää tarvetta manuaaliselle ohjelmoinnille. Käyttäjän täytyy ainoastaan määrittää tiedonkeruun lähteet ja tekoäly pystyy ilman ihmisen manuaalista työtä kerätä dataa näistä lähteistä.

Ehkä se isoin on varmaan se, että sinun ei tarvitse välttämättä koodata sitä kaikkea logiikkaa tietyllä lailla alusta asti. Että, sinulla voi ehkä olla se yksi vaihe, jossa sinä tavalla tai toisella hankit sen listan niistä lähteistä, että: suunnilleen nämä sivustot ja näin edespäin ja sen jälkeen sulla olisi voi olla se vaihe, missä sulla on niiden sivustojen läpikäynti tekoälyavusteisesti. Niin kyllä se, vaikka nyt tää netin raapiminen on tuttua vanhasta elämästä, niin varmasti sitä vaihetta tehostaa kyllä merkittävästi. H7

Haastateltava 5 uskoo jopa, että yksinkertaistettu kilpailija-analyysi voidaan jatkossa kokonaan automatisoida näiden teknologioiden avulla, eikä näin ollen tällaisia yksinkertaista kilpailija-analyysiä tuottavia yrityksiä enää tarvita:

Onhan tuossa aste-eroja, että niille tasolle mennään, mutta sellainen, että saadaan signaaleja siitä, että mitä kilpailijalla tapahtuu, niin minä vähän tosiaan tälleen poleemisesti väitän, että sellaista liiketoimintaa ei 2-3 vuoden päästä enää ole, koska sen voi hyvin helposti jo automatisoida sen osuuden, että se täytyy olla sitten, jos sitä haluaa myydä asiakkaille kilpailijaseurantaa, niin sen täytyy mennä ihan merkittävästi syvemmälle, kun se, että julkisia tietolähteitä seurataan ja niistä koostetaan summauksia. H5

Kiinnostava havainto, jonka haastateltava 4 puolestaan nosti esiin, oli, että tekoölyn avulla voidaan hakea tietoa myös ennalta määrittämättömistä lähteistä ilman, että käyttäjän tarvitsee tietää tarkalleen, mistä tietoa etsitään tai miltä se näyttää. Haastateltava 4 kuvasi, kuinka tekoölyagentille voidaan antaa laajoja hakuohjeita internetistä, jolloin tekoöly toimii paitsi hakijana myös suodattajana ja analyysin esivaiheen toteuttajana. Näin automaatio laajentaa sekä tiedon määrällistä kattavuutta että vähentää tarvetta ohjaukselle.

Sellaista ennalta määrittämättömistä ja ennalta tunnistamattomista paikoista on sitten helpompi kerätä sitä tietoa, että ei tarvitse etukäteen määritellä, että ”tästä kohdasta tätä sivua löytyy tämä tieto”, vaan pystyy antamaan isompia API:ja sitten että: etsi kaikesta pohja tiedoista tuolta internet hauista ja sitten suodata. H4

Tiedonkeruun tehostumisen taustalla on siis ennen kaikkea automaatio. Tekoöly mahdollistaa kilpailijatiedon keräämisen aiempaa laajemmista lähteistä huomattavasti vähemmällä manuaalisella työpanoksella, mikä muuttaa koko keruuvaiheen luonnetta.

4.2.3 Tekoölyn ja big data -analytiikan mahdollistama multimodaalisuus kilpailijatiedon keruussa

Yksi haastatteluissa noussut teema tekoölytehosteisen big data -analytiikan hyödyntämisessä kilpailijatiedon keräämisen yhteydessä oli sen mahdollistama multimodaalisuus. Multimodaalisuudella tarkoitetaan erimuotoisten datalähteiden, kuten videoiden, kuvien ja erilaisten tiedostomuotojen, yhä laajempaa hyödyntämistä. (Giannakos ym. 2019). Tietoa kilpailijoista voidaan nyt kerätä ja analysoida esimerkiksi verkkosivuilta ja rakenteettomista tietokannoista aiempaa tehokkaammin ja helpommin. Myös esimerkiksi videot, podcastit, PDF-tiedostot ja muut ei-strukturoidut materiaalit voidaan nyt helpommin saada osaksi kilpailija-analyysiä kuin ennen.

Vaikka monimuotoisia tietolähteitä on aiemminkin ollut olemassa ja näin ollen niitä olisi käytännössä ollut mahdollista hyödyntää kilpailija-analyysissä, se on kuitenkin ollut hyvin työlästä eikä tehokasta analyysin kannalta. Uutta onkin se, että moninaisten tietolähteiden hyödyntäminen on tekoölyn kehittymisen myötä tullut merkittävästi helpommaksi, nopeammaksi ja skaalautuvammaksi, jolloin niiden hyödyntäminen myös kilpailija-analyysissä on helpottunut. Haastateltava 6 kuvaa tätä seuraavasti:

Sitten just tällainen multimodaalisuus, eli tarkoitan just näitä eri muotoisia dataja, että me voidaan just yhtäkkiä tehdä hirvittävän helposti myöskin

videosta analyysiä, se voidaan kääntää, sekä kuviksi että se voidaan kääntää transkriptioksi, jota taas sitten voidaan syöttää tavalla ja toisella sitten tekoälylle ja sitä kautta jälleen kerran rikastaa sitä kokonaisdatan vaihtelevuutta ja kattavuutta mitä me kerätään. Ja just tässäkin se helppous ja sen nopeus ja se vaivannäön pienuus versus se, että minkälaista mahdollista hyötyä se tuo. H6

Haastateltava 6 antoi käytännön esimerkin, miten videoita oli käytetty kilpailija-analyysissä:

Esimerkiksi mitä me ollaan tehty eräälle rakennusyhtiölle, liittyy myös kilpailija- ja markkinatiedon keruuseen, että minkälaisia rakennushankkeita tai tontteja mahdollisesti esimerkiksi tulee vapaaksi. Niin me ollaan, tosiaan otettu näitä kunnan valtuustokokouksia, niistä tehtyjä videoita ja ajettu niistä transkriptio ja ne kanssa sitten just toimi polttoaineena tämmöisille sitten tekoälyn ratkaisuille, jossa sitten pyritään siihen parempaan analyysiin, että mitä tämä kaikki meille tarkoittaa. H6

Hänen kuvauksensa tuo esiin epäsuoran tavan seurata kilpailijoiden toimintaa valtuustokokousvideoiden kautta, joissa käsitellään kunnan rakennushankkeita ja tonttien jakautumista. Näiden avulla voidaan tunnistaa, mihin hankkeisiin kilpailijat ovat osallistumassa, mitä alueita ne tavoittelevat ja millaisia strategisia liikkeitä ne tekevät. Haastateltava 6 korostaakin, että nykyään kilpailijatiedon keräämisessä mielikuviutus on lähinnä rajana ja toisaalta mielikuviutuksen puute voi haitata kattavan kilpailija-analyysin tekemistä.

Myös haastateltavat 7 ja 3 nostivat esiin multimodaalisuuden ja eri muotoisten tietojen keräämisen helpottumisen, mikä laajentaa kilpailija-analyysin tekemistä. Haastateltavat kuvasivat, kuinka tekoäly voi esimerkiksi litteroida YouTube-videoita, tulkita PDF- ja Word-dokumentteja sekä yhdistää eri lähteistä peräisin olevaa dataa, vaikka se ei olisi yhteismitallista. Tämä vähentää aiempia työpanoksia, kun verkkoharavointityökalujen säätämistä ei tarvita, ja helpottaa tiedon keräämistä, kun formaattiriippuvuus ei ole enää hidasteena.

Tämmöinen multimodaalisuus, mitä tämä nykyinen tekoäly mahdollistaa, niin toki sitä kautta varmaan sä voisit, vaikka jotain kilpailijoiden YouTube - videoita, vaikka litteroida, ja sieltä kerätä jotain tietoa, mitä sä et ikään kuin verkosta sinällään löydä. Ja sitten toki se, mikä aikaisemmin on ollut haastavaa, nykyään ehkä vähemmän haastavaa, tämmöisten PDF-, Word - muotoisten tai taulukkomuotoisten aineistojen lukeminen sieltä verkkosivuilta, että se on ollut aikaisemmin vähän semmoista, että se on tavallaan pitänyt ”konfata” siihen ”skreippiin”, että: ”hei tällä sivulla on tämmöinen taulukko ja siinä on tämmöiset ”hederit” ja käy hakemassa sieltä

nämä solut ja tallennat tänne näin”, niin sen parsiminen varmasti on vähän yksinkertaisempaa nykyään. H7

Varsinkin tää multimodaalisuus on erittäin monikeinonen analytiikka, mikä tarkoittaa, että yhdistetään eri datalähteitä. Että mun mielestä siinä on se yksi iso juttu että, nää uudet big data -mallit ja muut, niin pystyy yhdistää ja vertailemaan isoja data aineistoja keskenään, vaikka ne ei olisi yhteismitallisia. H3

Tilinpäätösanalyysit ovat keskeinen osa kilpailija-analyysiä (Järvenpää ym. 2017, 447) ja tästä syystä haastatteluissa noussut havainto tilinpäätösmateriaalien keräämisen helpottumisesta on tärkeä havainto. Toinen tärkeä havainto on, että kilpailija-analyysissä hyödynnettävä kilpailijatieto ei aina ole suoraan kilpailijan itsensä tuottamaa, vaan kyse voi olla myös toissijaisista lähteistä, kuten analyysitalojen tai muiden ulkopuolisten tahojen julkaisuista. Tämä laajentaa ymmärrystä kilpailijatiedon keräämisen lähteistä. Tiedon, jota analysoidaan, ei nimittäin tarvitse olla varsinaista kilpailijayrityksen viestintää, vaan myös erilaiset kolmannen osapuolen lähteet, kuten asiantuntijapuheet, analyysit ja jopa podcast-keskustelut, voivat tarjota kilpailijoista olennaista tietoa.

Mitä ikinä tavallaan materiaaleja siellä voi olla kilpailijalla, jotka ei ole verkkosivu HTML-pohjaisia, vaan just vaikka nyt, ehkä toi tilinpäätös matskut on hyvä esimerkki siitä, mitä monesti saattaa olla PDF:nä. Tai sitten voit nyt sitten videoista tai podcasteista tai muista vastaavista, niin siellä voi olla sitten jotain muunlaisia keskusteluja tai analyysseja. Ei välttämättä ole edes sen firman omia, vaan voi olla, vaikka Inderesin tai jonkun muun vastaavan jotain tällaisia osavuosikatsausanalyysseja tai jotain muita, niin semmoisten litterointia. H7

Multimodaalisuus mahdollistaa myös sosiaalisen median sisältöjen hyödyntämisen kilpailija-analyysissä. Sosiaalisessa mediassa oleva data on myös hyvin monimuotoista, kuten kommentteja, kuvia, videoita ja reaktioita, ja ne voivat tarjota arvokkaita signaaleja kilpailijoiden toimista ja markkinakehityksestä. Haastateltava 4 kertoo, että hän on aiemmalla urallaan ollut mukana projektissa, jossa sosiaalisen median postauksista kerättiin sentimenttejä tai avainsanoja, joita verrattiin kilpailijoiden vastaaviin tietoihin. Sentimenttianalyysi on prosessi, jossa pyritään ennustamaan tunteita sanoissa, lauseissa tai tekstikokoelmissa (Li ym. 2018, 6940–6941). Tekoälyn tulon myötä tämä prosessi on hänen mukaansa myös helpottunut:

Aikaisemmalla uralla jo vähän tehtiinkin, elikkä some postauksista haettiin sentimentteja tai tällaisia avainsanoja omiin tuotteisiin ja kilpailijoiden tuotteisiin liittyen... mutta sekin on näiden uusien mallien myötä helpottunut tosi paljon. H4

Tekoölyn ja big data -analytiikan mahdollistama multimodaalisuus on laajentanut kilpailijatiedon keräämisen mahdollisuuksia merkittävästi. Nyt myös aiemmin vaikeasti hyödynnettävissä olleet lähteet, kuten videot, podcastit, PDF-tiedostot ja sosiaalisen median sisällöt, voidaan analysoida tehokkaasti osana kilpailija-analyysia.

4.3 Tekoölytehosteinen big data -analytiikka oleellisen kilpailijatiedon seulonnassa

4.3.1 Tekoölytehosteisen big data -analytiikan tuomat mahdollisuudet kilpailijatiedon seulonnassa

Kilpailijatiedon seulonta on tiedon keräämisen jälkeen seuraava vaihe kilpailija-analyysissä. Perinteisesti tämä vaihe on vaatinut merkittävää manuaalista työtä ja asiantuntijoiden kykyä lukea, tulkita ja tiivistää suuria määriä informaatiota. Haastatteluista käy ilmi, että tekoöly, erityisesti suuret kielimallit, ovat kuitenkin muuttaneet tätä prosessia. Mallit mahdollistavat suurten strukturoimattomien tietomassojen automaattisen läpikäynnin, luokittelun ja tiivistämisen sekä olennaisten teemojen esiin nostamisen. Tämä tehostaa datan seulomista ja vapauttaa asiantuntijoiden aikaa myöhempisiin kilpailija-analyysin prosessin vaiheisiin, eli analyysiin ja tulkintaan.

Just nämä erityyppiset strukturoimattomat datalähteet, että aikaisemminhan tässä on ollut just tosi paljon sitä, että ihmiset lukevat, ihmiset tekee niitä johtopäätöksiä jne. Koska monet asiat on kuitenkin ollut semmoisessa eikvantifioitavassa muodossa, mutta nyt me voidaan näitä tosi paljon käsitellä näillä isoilla kielimalleilla. Jotka on yllättävänkin hyviä kyllä hahmottamaan just sellaisia nousevia kokonaisuuksia semmoisesta massasta, mitä ihminenkin saattaa hahmottaa, mutta meillä menee sen läpikäymiseen ihan tuhottoman paljon aikaa. (...) Meillä on enemmän mahdollisuuksia kuin koskaan nimenomaan leikkiä tän strukturoidun ja strukturoimattoman välillä ja muuntaa siitä strukturoimattomasta just strukturoitua. H6

Nämä uusimmat mallit, on jo todella hyviä siinä, että pystyy tosi isoja, pitkiä tekstimassoja, erilaisia kerättyjä materiaaleja ja dokumentteja tiivistämään ja sitten tietysti luokittelemaan. (...) Tehostaa sitä, että jos ajattelee toimintaympäristöä, niin se on tosi paljon laajempikin vielä, kun se kilpailijakenttä niin, sitä materiaalia on vaan niin paljon, että se läpikäyminen ja tiivistäminen, niin se tehostaminen siinä ja tuominen sellaisiksi päivätiivistelmiksi tai viikkotiivistelmiksi niin helpottaa sitten asiantuntijoiden työtä, kun saa, saa vähän pureskeltua, analysoitua materiaalia sitten eri lähteistä. H4

Erytyisesti generatiiviset tekoölymallit ovat parantaneet kilpailijatiedon seulontaa siten, että niiden avulla voidaan kyselypohjaisesti promptaamalla kohdentaa haku olennaisiin tietoihin, mikä poistaa tarpeen käydä manuaalisesti läpi laajoja aineistoja. Lisäksi

tekoälyn hyödyntäminen mahdollistaa laajan tietomassan skannaamisen ilman tarkkaa ennakkotietoa siitä, mitä yksityiskohtia etsitään. Perinteisesti datan analysointi on vaatinut jonkinlaisen hypoteesin tai ennalta määrätyn lähestymistavan, mutta tekoälyn avulla voidaan lähteä kartoittamaan laajempaa ilmiökenttää, tunnistamaan trendaavia aiheita ja nostamaan esiin tärkeitä signaaleja kilpailijamarkkinoilta.

No siihen (seulomiseen) tuo tosi paljon, koska just tossa aikaisemmin viittasin, että jos on tosi isoja tauluja vaikka, niin se on ihan mahdoton niitä kahlata läpi tai pitäisi olla aika lailla jo tietoa mitä sieltä etsii. Niin nyt sitten voi vaan olla se iso tietomassa siinä taustalla ja lähteä kyselemään chatin kautta että: ”tuo kaikkia eri pointteja esille ja eri kulmista ja mitkä kilpailijat, vaikka mitkä tuotteet tässä tietyssä kategoriassa nousee/kehittyy?” Ja saadaan semmoista isoa kuvaa nopeasti ja sitä seulomista (...) Tehdä vaikka perinteisiä koneoppimismalleja siihen päälle. H2

Tekoäly on aika hyvä luomaan struktuuria sellaiseen struktuurimattomaan. Eli esimerkiksi se, että: onko tämä trendi? vai eikö tämä ole trendi? H6

Haastateltava 5 tuo esiin, että perinteinen data-analyysin prosessi muuttuu ja nopeutuu, kun datan mallinnusvaihe jää kokonaan pois prosessista, kun tekoäly kykenee ohittamaan mallinnusvaiheen ja tuottamaan analyttistä tietoa pelkän kysymyksen promptin avulla:

Mutta että kyllä se on ilman muuta se toinen massiivinen etu, että puhutaan tällaisesta neulan löytämisestä heinäsuovasta, että kyllä nää tekoälymallit löytää ihan vaikka olisi iso vuosiraportti, niin ne löytää ihan yksittäisiä asioita, jos sä tiedät mitä kysyä sieltä. Ja tällä lailla sä pystyt läpikäymään strukturoimaan ja summaroimaan hyvin isoja tietoaineistoja löytääksesi sieltä sitten juuri sen asian, mitä sä haluat tietää, mitä markkinassa tapahtuu tai mitä kilpailijoilla tapahtuu? Ja tosiaan että se on ollut ennen vaivalloista työtä, joko ihan tehty käsin tai sitten data analytiikalla aiemmin ratkaisu, niin se on ollut hyvin vaivalloista. Nyt se koko mallinnus osa jää tavallaan väliin, että voi vaan promptaamalla tehdä, niin se on kyllä aivan massiivinen muutos. H5

Tekoäly, erityisesti suuret kielimallit, ovat tehostaneet kilpailijatiedon seulontaa automatisoimalla strukturoimattoman datan läpikäyntiä, tiivistämistä ja luokittelua. Tämä vähentää manuaalista työtä, nopeuttaa olennaisten signaalien tunnistamista ja mahdollistaa tiedon kohdennetun tarkastelun ilman valmiita hypoteeseja.

4.3.2 Haasteet kilpailijatiedon seulomisessa tekoälytehosteisesti

Tekoälymallit kehittyvät jatkuvasti, mutta eivät ole vielä täydellisiä. Jotkut haastateltavat olivat sitä mieltä, että suuren massadatan seulonnassa on vielä myös toistaiseksi joitain puutteita. Haastateltava 4 nostaa esiin, että tekoälymallit eivät nykyisellään esimerkiksi kykene aina luotettavasti arvioimaan lähteiden ajallista relevanssia. Tämä voi johtaa

tilanteisiin, joissa analyysin pohjaksi nousee vanhentunutta tai kontekstiltaan epäolennaista aineistoa. Näistä syistä haastateltavan 4 mielestä ihmistä yhä tarvitaan arvioimaan tekoälyn seulomia aineistoja.

Se, mikä on vielä vähän heikko nykymalleissa, mitä ollaan huomattu niin, datamassat kun on niin isot ja nuo tekoälymallit ei ole vielä hirveän hyviä siinä ajallisessa ulottuvuudessa, että mallit saattaa nostaa jotain tosi vanhoja dokumentteja oleelliseksi, mutta ne ei osaa arvioida sitä, että tämä on vaikka 6 vuotta vanha artikkeli tai sitten me ollaan ehkä käytetty väriä malleja ja muuta mutta... myös sellainen riski siinä on... Mä en sano että hallusinointi on enää niin hirveän iso ongelma, mut ehkä just tällainen ajantasaisuuden ja oleellisuuden arviointi, niin siellä on vielä kyllä haasteita niissä tekoälymalleissa. Joo siis no ehkä se hyöty siitä, että pystyy seulomaan isoja massoja, on ehkä vielä isompi, että sitten ihmistä tarvitaan kuitenkin siellä vielä arvioimaan, että onko se seulottu materiaali vielä miten validia. H4

Toisaalta haastateltava 6 kuitenkin kertoo, että heidän kehittämässään kilpailija-analyysiin keskittyvässä tekoälymallissaan tekoäly analysoi dataa reaaliaikaisesti ottaen huomioon kommenttien aikaleimat, jolloin äkillisesti nousevat aiheet tulevat nopeammin esiin massasta. Eli hänen esimerkissään aikaleimat olivat toimineet ja toivat siis lisäarvoa analyysille.

Just tää ruokavalmistaja esimerkki on hyvä, että kun me ei olla määritelty niitä seurattavia topiikkeja, sen kyseisen langan sisällä erikseen, mutta sitten kun meillä on siellä nämä aikaleimat, että milloin on mitään keskusteltu ja että millaiset asiat on tullut nyt esille, että siellä on voinut olla vaikka, että yhtäkkiä kilpailija on voinut lanseerata jonkun uuden tuotteen ja koska sitä tuotetta ei ole aikaisemmin ollut, eikä me olla oltu siitä tietoisia, niin me ei aiemmin olla voitu mitenkään määritellä, mutta yhtäkkiä se sitten alkaakin saada pöhinää ja ihmiset keskustele ja sehän on taas meille tosi kiinnostavaa. H6

Kaikki mallit eivät siis vielä ole samalla tasolla oleellisen tiedon seulonnan funktiossa. Haastattelussa nousi esiin myös kysymys siitä, kykenevätkö tekoälymallit tunnistamaan poikkeavia tai epätyypillisiä tietoja laajoista aineistoista. Tällä tarkoitettiin virheitä, jotka saattaisivat jäädä manuaalisessa analyysissä huomaamatta. Haastateltavan 4 mielestä ihminen on parempi huomaamaan poikkeuksia datasta.

Ihminen on vielä ehkä parempi huomaamaan sitten niitä poikkeuksia sieltä. H4

Haastateltava 7 on myös sitä mieltä, että seulottujen tietojen validointi on kriittistä, koska tekoäly ei aina osaa arvioida, mikä on oikeasti tärkeää ja miksi. Haastateltava korostaakin, että tekoäly pystyy tiivistämään laajoja aineistoja, mutta sen tuloksiin ei voi ainakaan

vielä sokeasti luottaa, vaan käyttäjän on pohdittava, miksi jokin asia on tekoälyn mielestä olennainen ja millaisilla perusteilla tekoäly tekee valintojaan. Mutta haastateltavan mielestä tässä mennään toistaiseksi tekoälyn opetusdatan ja kykyjen rajoilla.

Mahdollistaa vaikka sen, että se jostain isommasta tekstiaineistosta pyydät sitä kertomaan sen mielestä tietyltä kantilta olennaiset asiat, mutta sitten siinä kohtaa se kysymys tavallaan tulee, että luotatko sä siihen? Ja millä perusteella ja miksi ne on olennaisia sen mielestä? toki sä voit käydä sen mallin kanssa tällaista keskustelua siinä, että se perustelee itseensä, mutta tuon tyyppisessä keskustelussa helposti mennään vähän sen opetusdatan rajoilla tai kykyjen rajoilla, että sitten se saattaa ruveta perustelevaan niitä asioita vähän typerillä pointeilla. H7

Haastateltavan 7 mielestä oleellisen tiedon seulonnassa ihmistä ei kokonaan tulisi poistaa yhtälöstä, kun kilpailijadataa seulotaan, sillä sen oleellisen tiedon seulominen on niin kriittinen vaihe tuleville prosessin vaiheille, joissa dataa lähdetään analysoimaan.

Koska tää oleellisen tiedon seulonta on seuraavia vaiheita varten niin oleellinen juttu, että edellisissä vaiheissa sulla voi olla sitä tietoa paljon. Ja sitten sieltä, kun pikkuvirheet saattaa hävitä sinne tavallaan keskiarvoihin ja muihin sen massan kautta. Mutta sitten tässä kohtaa, kun se ehkä tiivistää, sitä jotenkin sitä tietoa ja aggregoit ja koitat miettiä, että mikä sitä olennaista on, niin kyllä siinä käyttäisin ehkä nyt aika vahvasti kuitenkin sitä ihmisen ajattelukykyä jotenkin sitten validoimaan sitä, että mitä se tekoäly ehdottaa. H7

Vaikka tekoälymallit tehostavat kilpailijatiedon seulontaa, haastateltavat korostivat niiden rajoitteita erityisesti tiedon ajantasaisuuden, poikkeamien tunnistamisen ja oleellisuuden arvioinnin osalta. Tästä syystä ihmisen rooli tiedon validoinnissa ja tulkinnassa nähtiin edelleen hyödyllisenä.

4.4 Tekoälytehosteinen big data -analytiikka kilpailijatiedon analysoinnissa

4.4.1 Tekoälytehosteisen big data -analytiikan tuoma tehostushyöty kilpailijoiden analysoinnissa

Kun kilpailijatietoa on kerätty ja seulottu, seuraava vaihe kilpailija-analyysiprosessissa on kerätyn tiedon analysointi. Perinteisesti kilpailijatiedon analysointi on vaatinut paljon manuaalista työtä ja syvällistä osaamista toimialasta ja datan rakenteistamisesta. Haastatteluaineiston perusteella tekoälytehostettu big data -analytiikka on tuonut tähän prosessiin merkittäviä muutoksia. Kilpailijoiden analysointi voidaan nykyisin toteuttaa aiempaa nopeammin, skaalautuvammin ja monipuolisemmin. Haastatteluiden mukaan

tekoäly mahdollistaa kerättyjen suurten tietomassojen nopean läpikäynnin. Tämän jälkeen tekoälytehostetut big data -analytiikkaohjelmat pystyvät tiivistämään ja tekemään raportteja, sanallisia tulkintoja ja visuaalisia esityksiä datasta, jonka avulla tietojen analysointi on helpompaa ja tehokkaampaa. Erityisesti aloilla, joilla kilpailijoita on paljon tai toimiala on laaja, tekoälytehosteinen big data -analytiikka mahdollistaa markkinan tilan hahmottamisen nopeammin. Haastateltavan 6 mielestä tekoälyn suurin potentiaali korostuu eniten juuri tässä kilpailija-analyysin vaiheessa:

Just se analyysi on varmaan ehkä se isoin, missä saadaan nää tekoälyn hyödyt. Just se kielimalli tulee niin valtavan suureen rooliin muokkaamaan strukturoimattomasta datasta strukturoitua ja voidaan yhdistellä sitä lähdeagnostisesti. Se on ehkä se isoin muutos, että just se tehostaminen, mistä puhuttiin ja sitten myöskin se laadullinen tarkkuus, koska vaikka tän kaiken tekisi ihmiset niin, me ei vaan olla niin hyviä siinä. H6

Myös haastateltavat 2 ja 3 alleviivaavat näiden teknologioiden tuomaa tehostumista kilpailijoiden analysoinnissa. Haastateltava 3 korostaa eritoten aloja, joissa kilpailijoita voi olla paljon:

Nopeutta tehokkuutta ja sitten jossain tapauksessa myös laatua lisää, että aikaisemmin on ehkä pystynyt silleen vaan tarjoamaan luvut jollekin, että tuossa, mutta nyt kun sä pystyt tekee siihen lukujen päälle myös sen ne syvemmät analyysit ja selitykset ja kaikki, niin se tuo myös tavallaan sitä laatua ja helppolukuisuutta sille loppukäyttäjälle. H2

Ennenhän käytettiin paljon konsultteja tai sitten tehtiin itse, googlattiin ja haettiin tietoa. Mutta nyt tekoäly on automatisoinut sitä paljon. (...) Tietysti riippuu markkinasta, jos ollaan B2B-markkinassa, vaikka hissimarkkinassa, Kone Suomessa, niin kyllä se tuntee aika hyvin kilpailijansa. Mutta sitten jos mennään B2C-markkinaan tai johonkin, missä voi olla satoja ja tuhansia kilpailijoita, niin tällaisessa se kyllä auttaa paljon. H3

Haastateltava 5 kertoo, että varsinkin kapeassa, hyvin ennalta määritellyssä kilpailijavertailussa, jossa käytetään strukturoimatonta dataa, analyysiprosessin tehostuspotentiaali on korkea.

Kyllä mä näkisin, että vaikka otettaisiin se benchmarkkaus hyvin kapeasti, että ollaan etukäteen tunnistettu dimensiot, joissa halutaan vertailla omaa yritystä pääkilpailijoihin. Niin kauan, kun ne dimensiot on sellaisia, että se data on ei-rakenteisessa muodossa ja se pitää jotenkin parsia kasaan, niin onhan se tehostuspotentiaali aivan massiivinen. H5

Haastateltava 7 korostaa, että varsinkin tilanteissa, jossa yritys pyrkii keräämään ja vertailemaan kilpailijatietoa säännöllisesti, mahdollisesti muodostaen historiallisen vertailuaineiston, tekoäly pystyy nopeuttamaan ja skaalaamaan tätä prosessia:

Vai teetkö sä sitä silleen ammattimaisesti, että sä vaikka kerran neljänneksessä käyt läpi sun kilpailijoiden uusimmat kotkotukset ja sitten sä haluat rakenteistaa ne jonkinlaiseen muotoon ja sulla on ehkä joku tietokanta niistä sitten olemassa, mitä sä voit katsoa jatkossa tai joku semmoinen raporttikokonaisuus, mikä tavallaan vaikka vertailee edellisiin kausiin tai muuta vastaavaa, niin tuommoiset asiatkin tietenkin nopeutuu. H7

Haastateltava 4 tuo puolestaan esiin, kuinka tekoälytehosteinen big data -analytiikka mahdollistaa kilpailijoiden hintatiedon jopa reaaliaikaisen keräämisen:

Se on ehkä tällaista edistynyttä analytiikkaa, en kutsuisi sitä tekoälyksi, koska siellä ei nyt välttämättä ihan mitään kielimalleja ole, mutta sitten kun me ollaan kerätty sitä dataa tuolta, skreipattu vaikka hintatietoja ja muita niin, siellähän pyörii ihan jatkuvasti useita kertoja päivässä se prosessi, että analysoidaan niitä poikkeamia ja miten kilpailijat on muuttanut hintoja. H4

Kilpailija-analyysin automaatio oli myös merkittävä teema, jonka nähtiin vaikuttavan tehostuspotentiaaliin. Useampi haastateltava toi esiin, että lähes koko kilpailija-analyysi pystytään nyt ja tulevaisuudessa hyvin pitkälti automatisoimaan. Haastateltava 5 näki, että varsinkin yksinkertainen kilpailija-analyysin prosessi voidaan automatisoida näiden edistyneiden teknologioiden avulla jopa kokonaan.

Sellainen yksioikoinen kilpailija-analyysi niin sehän voidaan automatisoida jo nyt kokonaan ja varsinkin vuoden parin päästä, että jos kilpailija analyysi on enemmän sellaista, että kerätään julkisista tietolähteistä tietoa ja yhteismitallistetaan ja summaroidaan se, niin toihan on yksi suoraviivaisimmista tietotyön tehtävistä automatisoida tekoälyagenteilla nykyisin. H5

Myös haastateltavan 4 mukaan näiden teknologioiden mahdollistama automaatio on muuttanut analyytikoiden työnkuvaa. Aiemmin suuri osa kilpailija-analyytikkojen työstä keskittyi raporttien kirjoittamiseen ja tiedon koostamiseen, mutta nyt tekoälyn avulla tiivistelmät ja analyysit voidaan automatisoida, jolloin analytikoilla on enemmän aikaa keskittyä strategiseen ajatteluun.

Jos nyt ajattelee menee vaikka 10 vuotta taaksepäin niin mitä tuollaiset kilpailija analytytikot, niin varmasti käyttivät ison osan kuukaudesta siihen, että ne kirjoitti, sitten niitä raportteja sisäiseen käyttöön, niin nyt se on pitkälti automatisoitua jo se tiivistelmien tekeminen ja sitten nämä analytytikot pystyy

keskittymään sitten oleellisiin asioihin ja ehkä miettimään erilaisia kilpailijastrategioita ja muita sitten siinä ajallaan. H4

Kilpailijoiden analysointi on tehostunut merkittävästi tekoälytehostetun big data -analytiikan ansiosta. Prosessit ovat aiempaa nopeampia, skaalautuvampia ja osin automatisoitavissa, mikä vapauttaa asiantuntijoiden aikaa strategiseen työskentelyyn ja mahdollistaa syvällisempien analyysien toteuttamisen suuremmista tietomassoista.

4.4.2 Esimerkki reaaliaikaisesta kilpailijatietotaulusta

Automaation ja reaaliaikaisen analyysin tuomat mahdollisuudet kilpailija-analyysille näyttäytyivät haastateltavan 6 tuoman konkreettisen esimerkin kautta. Haastateltava 6 esitti asiakkaalle räätälöidyn kilpailija- ja markkinatietotietotaulun, joka havainnollistaa tekoälytehostetun big data -analytiikan potentiaalia käytännössä. Esimerkissä hajanaisista ja osin epästrukturoimattomista lähteistä kerätty tieto jalostettiin reaaliaikaisesti päivittyväksi kokonaisuudeksi, jonka avulla asiakasyritys pystyi seurata markkinatrendejä ja kilpailijoiden toimintaa yhdestä paikasta. Esimerkkitapauksessa kyseessä oli barbeque-tuotteita valmistava elintarvikealan yritys. Lopullisessa taulussa haastateltava 6 kertoi olleen muun muassa kommentteja keskusteluforumilta, toimialaraportteja, tilinpäätösdataa, uutisia ja dataa kilpailijoiden omista julkaisuista. Kuvassa 10 näkyvässä taulussa on kuitenkin vasta Reddit-keskustelujen data.



Kuvio 10 Haastateltavan 6 esittämä markkinatietotaulu

Alkuvaiheessa hyödynnettiin erityisesti Reddit-keskusteluja, joiden avulla seurattiin kuluttajien mielipiteitä ja keskustelunaiheita markkinoilla:

Tämä oli semmoinen ensiversio, jossa me tuotiin tänne sisään dataa tällaisesta Redditiin foorumista ja langasta, jossa on tosi aktiivinen sellainen barbecue -keskustelu. (...) Skreipattu heidän API:n kautta. H6

Tämä on merkittävä esimerkki siitä, miten big data -analytiikka mahdollistaa myös epästrukturoitujen ja käyttäjälähtöisten tietolähteiden hyödyntämisen kilpailija-analyysissä. Haastateltava 6 jatkaa, että kilpailijatauluun lisättiin myöhemmin muitakin tietolähteitä, kuten toimialaraportteja ja kilpailijatietoa:

Tänne on sitten myöhemmin lisätty muitakin datalähteitä, että esimerkiksi just tällaisia mitä mä mainitsin: toimialaraportteja, tietoa kilpailijoista, kaikenlaista. (...) kilpailijoiden nämä osavuositarkastukset ja tilinpäätökset ja esimerkiksi 'capital markets day' -materiaalit. H6

Lisäksi haastateltava 6 mainitsee, että monet yritykset hyödyntävät usein myös kaupallisia mediatietopalveluita, jotka kokoavat ja analysoivat uutisia ja markkinatietoa, ja niistä saatavat tiedot voidaan myös integroida osaksi kilpailijataulua:

Tällä asiakkaalla, ruokatoimittajalla, oli just tällainen Valonan palvelu, niin se kytkettiin tuohon mukaan osaksi, että periaatteessa ne uutishetit tuli sinne, ja ne periaatteessa tarjosi semmoisen ylimääräisen datan lähteen, joka sitten pystyttiin "plugaamaan" tähän isoon kielimalliin ja tähän kokonaisratkaisuun kiinni. H6

Tämän lisäksi haastateltava 6 mainitsi, että heidän taulunsa ei ollut pelkästään perinteinen tietojen visualisointityökalu, vaan siihen oli integroitu myös tekoälychat -toiminto, jonka avulla käyttäjät pystyivät vuorovaikutteisesti kysymään syvällisempää analyysiä taulusta ja sen datasta. Tämä tarkoittaa, että pelkän raportoinnin sijaan analyttikot tai päätöksentekijät saattoivat esittää tarkentavia kysymyksiä, suodattaa tietoa ja saada lisäanalyysiä tekoälyn avulla.

Me pystymme vielä lähteä kyselemään siltä lisäkysymyksiä, että nyt tuli tämmöinen analyysi, vähän niin kuin siinä ruoanvalmistajan casessa, että sitten meillä on kyky myöskin keskustella niiden analyysitulosten kanssa. Entä: 'entäpä tämä näkökulma, entäpä tuo? Entä jos olisi näin? H6

Haastateltavan 6 kuvaama tapaus osoittaa, kuinka tekoälyn ja big data -analytiikan avulla on mahdollista tehdä aiempaa kattavampi, automatisoitu ja reaaliaikainen markkina- ja kilpailija-analyysi. Taulu yhdisti perinteiset taloudelliset raportit, kilpailijoiden

strategiamateriaalit, uutispalvelut sekä sosiaalisen median datan yhdeksi reaaliaikaisesti päivittyväksi kokonaisuudeksi.

4.4.3 Tekoälytehosteinen big data -analytiikka kilpailijoiden markkinakäyttäytymisen analysoinnissa

Perinteisesti kilpailijoiden trendien ja liikkeiden tunnistaminen ja analysointi on vaatinut ihmisen asiantuntemusta ja pitkälle kehitettyjä analyysimenetelmiä, mutta tekoälyn ja big data -analytiikan hyödyntäminen tässä prosessissa on mahdollistanut suurien tietomäärien jäsentämisen, toistuvuuksien ja muutosilmiöiden tunnistamisen kilpailijoiden liikkeistä aiempaa tehokkaammin ja automatisoidummin. Tekoälyn avulla suuresta massasta dataa on mahdollista helpommin erottaa esimerkiksi toistuvia ilmiöitä, jotka voivat kertoa nousevasta trendistä. Nämä trendit voi olla mitä vain muutoksia datassa, esimerkiksi muutoksia kilpailijoiden liikkeissä markkinassa.

Tekoälyn keskeinen etu kilpailijatiedon analysoinnissa on sen kyky käsitellä rakenteetonta dataa tehokkaasti. Tämä on merkittävä muutos verrattuna aikaisempiin luonnollisen kielen menetelmiin, joissa analyysin kohteena olevat aihealueet piti määritellä staattisesti etukäteen. Tämä johti rajoituksiin, koska analyysi ei kyennyt tunnistamaan uusia, yllättäviä ilmiöitä, joita ei ollut etukäteen ohjelmoitu. Suurten kielimallien myötä analyysi voi nyt mukautua dynaamisesti muuttuvaan sisältöön, mahdollistaen trendien ja signaalien havaitsemisen ilman staattisia luokitteluja. Tämä helpottaa myös kilpailija-analyysiä, sillä kilpailijadatassa ilmenevät muutokset auttavat yrittystä huomamaan kilpailijoiden liikkeet entistä nopeammin.

Ja nämä on superhyviä siinä mielessä, koska kun aikaisemmin tällaisia tehtiin, niin me määriteltiin staattisesti näitä tällaisia ”topiikkeja”(aiheita/teemoja). Esimerkiksi me tehtiin sellaista NLP pohjaista analyysiä, että me yritettiin atribuoida jokin tietty kommentti tai lauseen osa vaikka johonkin tiettyyn staattisesti ennalta määriteltyyn ”topiikkiin”. Mutta sehän oli tosi kökköä koska just se, että ei me vaikka tiedetä, että voi olla että ehkä huomenna tulee joku uusi trendi, joku uusi mauste, mitä me ei olla keksitty tuonne määritellä ja sitten ne ei tuolla ollenkaan näy seurannassa. Mutta nythän kaikki tää on dynaamista, nämä ”topiikit” on dynaamisia. Tää analyysi on dynaamista. H6

Mitä me tehdään tällä hetkellä, on sellainen vähän laajemmin sen koko toimintaympäristön, varsinkin tällaisten uusien tuotteiden, uusien energiatuotteiden suhteen käytetään tällaista tekoälypalvelua siihen, että koostetaan, tällaisia trendaavia ilmiöitä ja hälytyksiä, vaikka uusista tutkimuksista johonkin tiettyyn vetyteknologiaan tai muuhun liittyen. Siellä tietysti monesti esiintyy kilpailijoita sillä tavalla, että joku kilpailija on

kehittänyt jotain uutta ja se tulee sitä kautta tietoon, kun seurataan niitä teknologioita ja toimintaympäristöä, niin tulee se kilpailija siinä samalla. H4

Ja sitten toki sä voit pyytää sitä tekoälyä, vaikka siitä tekstuaalisesta datasta etsimään semmoisia tavallaan teemoja, jotka toistuu vaikka eri artikkeleissa tai eri lähteissä ja tän tyyppistä, mikä ei ole siis ollut aiemmin mahdollista. H7

Haastateltavan 2 kommentti alleviivaa tekoälyn merkitystä erityisesti kilpailijoiden määrän kasvaessa ja markkinoiden muuttuessa nopeasti, jolloin perinteinen seuranta ei riitä.

Joo pystyy kyllä niitä hyvin havaitsemaan ja semmoista että vaikka, että kun ihan kaikkea mitä tapahtuu markkinassa, kun niitä kilpailijoita on niin paljon ja tuotteita on niin paljon, niin se seuraaminen on tosi vaikeata. Tommoinen olisi mahdollista, että seuraat näitä kaikkia kilpailijoita näitä kaikkia tuotteita, mitä siellä tapahtuu. H2

Tekoälytehostettu big data -analytiikka mahdollistaa kilpailijoiden markkinakäyttäytymisen analysoinnin aiempaa dynaamisemmin, tunnistuen toistuvuuksia, trendejä ja muutoksia suurista tietomassoista, mikä tekee kilpailijoiden liikkeiden seuraamisesta reaaliaikaisempaa ja kattavampaa.

4.4.4 Tekoälytehosteinen big data -analytiikka kilpailijoiden liikkeiden ennakoinnin ja ennustamisen tukena

Merkittävä havainto, jonka haastateltavat toivat esiin, oli se, että tekoälytehosteinen big data -analytiikka muuttaa kilpailija-analyysin luonnetta reaktiivisesta ennakoivaksi. Koska tekoäly on hyvä trendien havaitsemisessa datasta, se voi myös parhaimmillaan mahdollistaa kilpailijoiden strategisten muutosten ennakoinnin. Haastateltavan 4 kommentti korostaa, että yritykset voivat ennakoida markkinamuutoksia ja kilpailijoiden liikkeitä jo ennen niiden toteutumista. Seurannan kautta havaitaan myös kilpailijoiden kehitystyötä, sillä kun tarkastellaan teknologisia muutoksia ja alan kehitystä, kilpailijoiden innovaatiot tulevat samalla näkyviin.

Kyllähän se mahdollistaa sen, että me löydetään sellaisia hiljaisia signaaleja ja vasta nousemassa olevia trendejä jo ennen, kun joku kilpailija isosti, vaikka lyö läpi (...) vaikka jonkun kilpailijan taloudellinen menestyminen, mihin suuntaan se tulee menemään. Tai indikaatioita siitä, että mihin markkinoihin joku firma tähtää, jos rupeaa tulemaan, vaikka uusia hankesuunnitelmia tai viestiä siitä, että joku kilpailija on lähdössä jollekin uudelle liiketoimintalueelle, niin tuon tyyppisiä varmasti helpommin tekoälykin myös löytää. H4

Haastateltava 6 tuo puolestaan esille, että kilpailijoiden strategisten muutosten havaitsemisessa voidaan hyödyntää erilaisia rakenteellisia indikaattoreita, kuten patenttihakemuksia, osavuosikatsauksia, rekrytointeja ja R&D-panostuksia. Tekoäly voi nopeuttaa ja tehostaa näiden analysointia automatisoimalla suurten tietomassojen analysointia ja tunnistamalla sieltä toistuvuuksia ja poikkeavuuksia, jotka voivat viitata strategiaan muutoksiin.

Mitä me itse ajatellaan, että mitkä voisi olla niitä asioita, mitkä kertoo siitä strategisesta muutoksesta. Useinhan ne on just, esimerkiksi just patentit, jos on yritys siellä patentteja, jos siellä on esimerkiksi isoja muutoksia vaikka jos on julkinen yhtiö, niin esimerkiksi, että mitä he raportoivat osavuosikatsauksessaan tai tilinpäätöksissään, vuosikertomuksissaan just esimerkiksi suhteessa, vaikka R&D panostuksiin tai ylipäänsä, että miten tase on rakentunut. Sitten myös rekrytoinnit, jos lähdetään ihan uusia profiileja hakemaan, niin näähän on sellaisia alueita, jotka usein kertoo, että jotain saattaa olla muuttumassa (...) Totta kai niitä voidaan nyt vielä paremmin ja nopeammin kerätä robotiikka- ja tekoälyavusteisesti. H6

Haastateltavan 5 mielestä tämä kilpailijoiden liikkeiden ennakoiminen ja ennustaminen voidaan pian suorittaa lähes kokonaan tekoälyn avulla.

Kyllä mä näkisin, että toikin on sellaista, että kyllä ajan myötä ne perus foresighting insightit ja raportit pystytään hyvin pitkälle tekoälyllä luomaan, ja sitten siihen varmaan ihminen ottaa näkemystä. Näin uskon, että tulee käymään, että nehän on luonteeltaan tietysti epävarmoja nämä tällaiset foresight väittämät. Mutta kyllä mä näkisin, että se tekoäly tulee pystymään siinä ihan samaan tasoon, kuin sen joku seniori asiantuntijakin. H5

Haastateltava 6 kuvasi käytännön esimerkin, jossa ennustava analytiikka ja generatiivinen tekoäly yhdistettiin kilpailija-analyysin kontekstissa konkreettiseksi liiketoimintaratkaisuksi. Hän painotti, että vaikka suuret kielimallit ovat erittäin kykeneviä käsittelemään monipuolista dataa, ne eivät yksin riitä tarkkaan ennustamiseen:

Jos mä sanoisin että: ”tässä on mulla tämmöinen iso datamassa, voitko luoda tästä jonkun ennustemallin, että mitä tapahtuu 2 vuoden päästä”, niin eihän se osaa sellaista luoda. H6

Tämän vuoksi haastateltava 6 painottaa, että organisaatioissa tarvitaan yhä data-analytikoita ja datatieteilijöitä rakentamaan ennustemalleja. Hän painottikin, että generatiivinen tekoäly ja perinteinen ennustava analytiikka tulee nähdä toisiaan täydentävinä, ei erillisinä. Haastateltava 6 kuvasi myös käytännön asiakasesimerkin ennustamisesta globaalista B2B-yrityksestä, joka käy vuosittain monimutkaisia sopimusneuvotteluja asiakkaidensa kanssa. Ennustamisen kohteena oli tämän B2B-

yrittäjien asiakkaalle ehdotettavat hintatasot ja siirtyminen korkeamman katteen tuotteisiin, mitä varten yritys tarvitsi tietoa omista kustannuksistaan, asiakkaistaan, asiakkaiden asiakkaista ja kilpailijoista.

Mitä me tiedetään niistä asiakkaista ja asiakkaan asiakkaista ja sitten meidän kilpailijoista ja ehkä asiakkaan kilpailijoista ja toimialasta? H6

Haastateltava 6 kertoo, että kerätyn datan pohjalta rakennettiin ennustemalli ja hintajoustomalli, jotka perustuivat yhtiön tietovarastossa olevaan rakenteelliseen dataan, kuten tuotantokustannuksiin, historiallisiin katteisiin ja strukturoimattomaan kilpailijatietoon, jota oli kerätty esimerkiksi verkkosisällöistä ja markkinaraporteista. Lopullinen ratkaisu, jonka he B2B-yritykselle räätälöivät, oli työkalu, jossa tekoäly laski suositeltuja hintatasoja eri volyymeille ja tuotti myös perusteluja, joilla asiakasta voitiin vakuuttaa:

Se tekoälyratkaisu ehdottaa ihan konkreettisia hintatasoja sitten tietyille volyymeille, vielä lisättynä argumenteilla, että millä voidaan sitten lähteä suosittelemaan ja voitelemaan sitä asiakasta. H6

Tämä esimerkki osoittaa, kuinka ennustava mallintaminen yhdistettynä tekoälyyn voi tukea kilpailija-analyysiä osana päätöksentekoa. Esimerkissä ennustettiin omaa kysyntää, tuotantoon liittyviä muuttujia ja kilpailijoiden ja asiakkaiden käyttäytymistä. Tekoälylle syötettiin suuri määrä dataa, jonka pohjalta tekoälymalli muodosti näkemyksen siitä, miten markkina ja kilpailijat todennäköisesti toimivat tulevaisuudessa. Nämä ennusteet sitten integroitiin osaksi hinnoittelumallia, joka tuotti konkreettisia suosituksia hintatasoista.

Tekoälytehosteinen big data -analytiikka muuttaa kilpailija-analyysin luonnetta reaktiivisesta ennakoivaksi. Koska tekoäly on hyvä trendien havaitsemisessa datasta, se voi myös parhaimmillaan mahdollistaa kilpailijoiden strategisten muutosten ennakoinnin. Myös kilpailijoiden liikkeiden ennustaminen koneiden avulla mahdollistuu.

4.4.5 Tekoälytehosteinen big data -analytiikka kilpailijoiden heikkouksien ja vahvuuksien analysoinnissa

Yksi kilpailija-analyysin keskeisistä tavoitteista on tunnistaa kilpailijoiden vahvuuksia ja heikkouksia, jotta yritys voi hyödyntää tätä tietoa strategisessa päätöksenteossa (Porter 1980, 63–67). Haastateltavien mukaan tekoälytehosteinen big data -analytiikka voi tukea

tätä prosessia. Haastateltava 2 esimerkiksi kuvailee prosessia, jossa uudelle markkinoille tulleiden brändien menestystä voidaan analysoida:

Varsinkin jos on tullut jotain uusia vaikka brändejä markkinoille, niin sitten voi katsoa että ensin datasta ihan määrällinen data kertoo, että: ”hei tällöisen tuotteen myynti kasvaa” ja sen jälkeen pystyy tekemään sitä lisäanalyysiä etsiä, että ”OK. No miksihän ne kasvaa? Et mitäköhän nää tehnyt eri tavalla? ja sitten siihen joko lähtee etsimään ihan julkisista lähteistä tai sitten lähtee promptaamaan tekoälyn kanssa, että mitä ne on mahdollisesti... mitä siellä voisi olla kilpailussa erottautumistekijöitä ja sitten on myös erilaisia tutkimusportaaleita, esimerkiksi internetissä lanseerataan kaikki tai tuodaan kaikki tuotteet julki mitä on lanseerattu meidän kategorioissa ja pystyy vähän sitä kautta lähteä katsomaan. H2

Myös haastateltavat 4 ja 5 näkivät tekoälyn potentiaalinen kilpailijoiden vahvuuksien ja heikkouksien tunnistamisessa, kunhan datapohja on riittävän kattava:

Näen, että kyllä tuosta varmasti löytyy hyötyjä just, jos on kerännyt sitä dataa ja sitten haluaa lähteä analysoimaan sitten niitä vahvuuksia, heikkouksia. Luulisin, että tekoäly löytää kyllä niitä jo aika hyvin. H4

Kyllähän se ihan promptaamalla niin varmaan tulee uusia näkökulmia, että kun kysyy että tässä on nyt mitä me tiedetään yrityksestä A ja ehkä vähän esimerkkejä, että minkälaisia asioita lasketaan, vahvuuksia ja heikkouksia, niin kyllä sieltä varmasti löytyy uusia näkökulmia. Vähintään ajatuksia herättelemään ja uskoisin kyllä, että toikin on sellaista asiaa, joka kyllä lähitulevaisuudessa sitten pystytään pitkälle automatisoimaan. H5

Haastateltava 5 kuitenkin muistutti, että kyse ei ole täysin objektiivisesta prosessista.

Vahvuuksien ja heikkouksien arviointi perustuu aina tulkintaan ja näkökulmiin:

Mutta sehän ei ole ihan faktapohjainen kysymys, että mitkä on yrityksen heikkoudet ja vahvuudet, että siinä on aina jonkun näkemyksestä kyse, että siinä ei varmaan ihminen ihan nyt heti katoa kuitenkaan. H5

Tämä alleviivaa sitä, että vaikka tekoäly pystyy kokoamaan ja analysoimaan dataa, lopullinen arvio kilpailijan vahvuuksista ja heikkouksista vaatii vielä inhimillistä tulkintaa ja liiketoimintaymmärrystä. Haastateltava 6 korosti myös samaan sävyyn, että heikkoudet ja vahvuudet voivat olla usein hiljaista, koneiden saavuttamattomissa olevaa tietoa.

Tällaisia asioita, joita ei välttämättä mitenkään oikein saada pintautettua, tekoäly ei pääse näihin asioihin käsiksi. Mutta jos me mietitään tätä kilpailukyvyn rakennetta, että siellä sinullahan voi olla yliverainen tuote tai sinulla voi olla mielettömät verkostot vaikka, mutta sitten toisaalta just tämä toimintatapa voi olla myös ihan hirvittävän suuri vahvuus ja se on ehkä

vaikein saada sitten datan avulla esiin koska se on sitä yrityksen hiljaista ja sisäistä dataa. H6

Yhteenvedona tekoälytehostettu big data -analytiikka voi tukea kilpailijoiden vahvuuksien ja heikkouksien tunnistamista tuottamalla uusia näkökulmia ja jäsentämällä laajoja tietomassoja, mutta lopullinen arviointi vaatii edelleen ihmisen tulkintaa.

4.5 Tekoälytehosteisen big data -analytiikan tuomat hyödyt kilpailijatiedon tulkinnassa ja johtopäätöksissä

Tulkinta- ja johtopäätösvaihe on tässä tutkimuksessa käsiteltävän kilpailija-analyysin viimeinen vaihe, jossa analysoidusta datasta jalostetaan liiketoimintaa hyödyttäviä tulkintoja ja johtopäätöksiä. Haastattelujen perusteella tekoäly, erityisesti suuret kielimallit, voivat tukea tätä vaihetta useilla tavoilla, kuten tiivistämällä analysoitua dataa, ehdottamalla johtopäätöksiä, visualisoimalla tuloksia ja jopa tarjoamalla toimintakehotuksia. Haastateltava 2 kuvailee, että tekoälytehosteinen big data -analytiikka tukee kilpailija-analyysin tulkinta- ja johtopäätösvaihetta tehokkuuden ja työtyytyväisyyden näkökulmasta. Haastateltava korostaa erityisesti, että tekoäly voi nopeuttaa tulkintaa ja johtopäätösten tekemistä sekä helpottaa raporttien laatimista.

No se on ihan selkeä, että kun tiedot on kerätty analysoitu niin se, miten se pystyy tosi kätevästi tekemään, varsinkin sekä tulkintaa että johtopäätöksiä (...) että: 'tässä on nämä tiedot, käytä näitä tietoja ja tee näistä nyt yhteenvedot ja johtopäätökset', niin tosi kätevästi ja sitten kun se tulee se sanallinen vastaus sieltä, niin sitten jos pitää kirjoittaa jotain raportteja tai muita, niin ne tulee kyllä tosi nopeasti. (...) Niin ei tuu sitä tyhjän paperin syndroomaa ja nyt on päässyt nopeammin liikkeelle, että vaikka kerralla ei tulisi lopullista vastausta niistä, niin pääsee aika pitkälle (...), mitä ollaan tutkittu sisäisesti niin myös työtyytyväisyyttä lisäännä, koska aikaisemmin ollaan pystytty automatisoimaan semmoisia tylsiä rutiinihommia. H2

Tämä kuvastaa sitä, miten tekoäly voi vähentää manuaalisen työn määrää tässäkin prosessin vaiheessa ja nopeuttaa raportointia ja päätöksentekoa. Myös haastateltavan 4 kommentti tuo esiin tekoällyn kyvyn tiivistää analysoitua tietoa ja muodostaa johtopäätöksiä, mikä voi tehostaa kilpailija-analyysiä ja päätöksentekoa. Haastateltava 4 korostaa, että kielimallit ovat tehokkaita yhteenvedojen tekemisessä ja analyysin jalostamisessa johtopäätöksiksi. Tämä havainnollistaa, että tekoäly voi toimia tehokkaana tukityökaluna analytikoille, sillä sen avulla voidaan pienentää informaatiokuormaa ja esittää kilpailija-analyysin lopputuloksia selkeämpinä johtopäätöksinä.

Joo sama oikeastaan kuin tuossa seulonnassa, niin kaikki kielimallit on mun mielestä tosi hyviä tekemään tällaisia summauksia ja johtopäätöksiä tiivistämään sitä analysoitua tietoa, että jos siinä analysointivaiheessa laittaisi taas enemmän, että ”analysoi todella monelta kantilta tätä asiaa”, niin sitten taas se johtopäätösvaihe niin siellä tekoäly voi, tai sillä on jo nyt kyvykkyys taas tiivistää sitä selvemmiksi johtopäätöksiksi ja ehkä tällaisiksi toimintaohjeistuksiksi jopa. Eli mitä tulisi tehdä näiden tietojen perusteella?
H4

Yhteenvetona tekoälytehostettu big data -analytiikka voi tukea kilpailija-analyysin tulkinta- ja johtopäätösvaihetta erityisesti tiivistämällä analysoitua tietoa ja muodostamalla selkeitä yhteenvetoja, mikä tehostaa raportointia ja nopeuttaa päätöksentekoa.

4.5.1 Tekoälytehosteisen big data -analytiikan mahdollistaman visualisoinnin hyöty tulkintoja tehdessä

Tekoäly voi tukea kilpailijatiedon tulkintaa myös sanallisten tiivistelmien ja erilaisten visuaalisten esitysten avulla. Haastateltavien mukaan tekoäly pystyy tuottamaan laskelmia ja visualisoimaan analyysin tuloksia nopeasti, mikä helpottaa analyysin ymmärtämistä ja viestintää. Haastateltava 2 toi esiin, että tekoälyltä voi pyytää suoraan esimerkiksi markkinaosuuksien laskentaa ja niiden kehityksen esittämistä graafisesti:

Se pystyy tehdä sitä laskentaa, pystyy visualisoimaan tuloksia, tekee graafeja. H2

Myös haastateltava 4 mainitsi, että analytiikkatyökaluissa tekoäly voi tuottaa ensimmäiset luonnokset visualisoinneista, mikä madaltaa kynnystä analyysin aloittamiseen ja vähentää manuaalista työtä:

Tekoäly tekee ne ensimmäiset graafit (...), helpottaa sitä tyhjän paperin syndroomaa. H4

Haastateltava 3 puolestaan toi esiin, että vaikka sama data ja visualisointi olisi tarjolla eri sidosryhmille, se ei välttämättä johda yhteiseen näkemykseen. Tässäkin tekoäly voi auttaa:

Joo se auttaa visualisoimaan ja sitten se auttaa sanoittamaan niitä tuloksia, mutta siitä en ole ihan varma, että paljonko se siihen päätöksentekoon dynamiikkaan edelleen vaikuttaa, että edelleenkin saattaa olla, että myyntijohtajalla ja markkinointijohtajalla on erilainen käsitys asioista, vaikka ne saa saman datan, mutta toinen uskoo ja toinen ei usko niihin tuloksiin.

Mutta voi se auttaa ehkä siinä vähän sitten tulkitsemaan toisenkin näkökulmasta sitä, että se voi auttaa toista ymmärtämään toista paremmin. H3

Haastateltavan 7 mukaan tekoälyllä tuotetut visualisoinnit eivät kuitenkaan vielä ole laadultaan hyviä, mutta kehityssuunta on selkeä:

Pitää tehdä nyt tästä aineistosta Powerpoint ja siihen pitää olla nää asiat kerrottuna, niin tee mulle ne slaidit ja tee ne kuvat sinne slaideille. Tämähän ei tällä hetkellä vielä tapahdu, vaan ainakin siis oman kokemuksen mukaan lopputulokset on aika silleen huonoja. (...) Tavallaan tuon tyyppisen asian automatisointi kyllä varmasti on mahdollisuus, jota moni odottaa kieli pitkällä. H7

Haastateltava 6 korostaa myös tekoälymallien käytettävyyttä luonnollisen kielen analyysissä ja datan visualisoinnissa.

Jos mä saisin sen visuaalisesti, niin sehän on yleensä paljon vaikutusvaltaisempaa. Niin tällaisiahan mahdollisuuksia meille on tullut nyt tosi paljon, että me voidaan periaatteessa tekoälystä ja strukturoimattomasta datasta kääntää sitä koodiksi, jota voidaan visualisoida. Jos esimerkiksi oot kokeillut Chat GPT:ssä Code Interpreterä, niin sille voi antaa vaikka artikkelin ja sanoa, että tee tällaisia asioita ja sille voi antaa tämän luonnollisen kielen komennon että ”use code interpreter ja tee minulle dashboard, että onko tässä vaikka jotain tiettyä toistoa tai trendiä?” Niin, kyllähän se tekee sen. H6

Tekoäly voi siis tukea kilpailijatiedon tulkintaa tuottamalla nopeasti laskelmia ja visualisointeja, jotka helpottavat analyysin hahmottamista ja viestintää, vaikka niiden laatu ja päätöksentekoa edistävä vaikutus koetaan vielä vaihtelevaksi.

4.5.2 Tekoälyn luomat toimintakehotukset kilpailija-analyysissä

Johtopäätösten ja visualisointien lisäksi tekoäly voi myös ehdottaa toimintakehotuksia tai päätösehdotuksia analysoidun datan perusteella. Haastateltava 2 kuvasi esimerkiksi, miten tekoäly voi analysoida kilpailijoista ja tuotteista kerättyä dataa ja antaa toimintakehotuksen sen pohjalta, millainen tuote voisi menestyä tietyllä markkinalla:

Me ollaan tekemässä vaikka uutta lanseerausta, niin tämän kaiken tiedon valossa, minkälaisen tuotteen lähtisit tekemään (...)? Se tarjoaa semmoisen tuotteen, että nää ominaisuudet voisi olla siinä. Niin sehän on tietynlainen toimintakehotus. H2

Tämä havainnollistaa generatiivisen tekoälyn kykyä muodostaa hypoteeseja ja vaihtoehtoisia ratkaisuehdotuksia, joita analyttikot pystyvät sitten hyödyntämään

kilpailija-analyysin loppupäässä. Myös haastateltava 4 sanoo, että tekoäly on kykeneväinen antamaan toimintakehotuksia.

Tekoäly voi, tai sillä on jo nyt kyvykkyys taas tiivistää sitä selvemmillä johtopäätöksiksi ja ehkä tällaisiksi toimintaohjeistuksiksi jopa. Eli mitä tulisi tehdä näiden tietojen perusteella? H4

Haastateltava 5 huomautti kuitenkin, että toimintasuosituksukset jäävät usein geneerisiksi, ellei tekoälylle anneta selkeää kontekstia. Tämän vuoksi asiantuntijan tulkinta ja liiketoimintaymmärrys ovat yhä välttämättömiä:

Toimintakehotukset kyllä auttaa näkemään mahdollisuuksien kirjjon (...), mutta ne ei aina sovi kontekstiin, jos ei sitä ole annettu mallille. H5

Haastateltavan 3 mukaan tarkempi, toimialakohtainen koulutus tai räätälöity kielimalli voi lisätä suositusten osuvuutta, mutta tällä hetkellä tekoäly toimii parhaiten päätöksenteon tukena, ei sen korvaajana. Myös haastateltava 8 korosti, että suositukset toimivat lähinnä ajattelun virikkeinä, eivät suorina ohjeina:

Ne on lähinnä semmoisia omaa ajatusta ohjaavia (...), mutta ei vielä sellaisia, että suoraan ohjaisivat valintoja.(...) mutta mä en ainakaan luota niihin vielä, että kyllä mä ajattelen että ne on aika pintapuolisia. H8

Tekoäly voi siis tukea kilpailija-analyysin päätöksentekoa ehdottamalla toimintakehotuksia analysoidun datan perusteella, mutta niiden hyödyllisyys riippuu vahvasti annetusta kontekstista ja siksi niitä pidettiin ennen kaikkea ajattelua ohjaavina ehdotuksina, ei suoraan sovellettavina ratkaisuuina.

4.6 Kilpailija-analyysin prosessin muutos dynaamiseksi kokonaisuudeksi

Aineistosta nousi myös merkittävä havainto koskien kilpailija-analyysin perinteistä, prosessimaista luonnetta. Haastateltavien mukaan edistyneet teknologiat ovat muuttaneet tätä perinteistä, lineaarisesti etenevää kilpailija-analyysin prosessia aiempaa dynaamisemmaksi, jolloin perinteinen, lineaarinen vaiheittainen eteneminen ei enää kuvaa prosessia, vaan esimerkiksi tiedon keruu ja seulonta voivat tapahtua rinnakkain tai lomittain. Haastateltava 5 kuvailee, että perinteisesti data-analytiikan prosessiin kuuluu tiedon kerääminen, varastointi ja tallennus, mutta hänen mukaansa nämä vaiheet voidaan automatisoida niin, että tekoälyagentti kykenee suoraan hakemaan, seulomaan ja analysoimaan tietoa ilman, että yrityksen tarvitsee itse ensin kerätä tai ylläpitää laajoja

tietovarantoja. Tämä lomittaisi tiedon keruun, seulomisen ja analysoinnin yhdeksi dynaamiseksi vaiheeksi.

Yksinkertaisimmillaan niin, se voidaan kokonaan automatisoida jopa niin, ettei ole varsinaisesti edes tarvetta, ainakaan kohta, kerätä ja varastoida mitään isoja strukturoimattomia tietomassoja, vaan sen voi AI -agentille antaa tehtäväksi aika suoraan sitten, että: ”tuota päivitys, mitä kilpailijalla kuuluu näissä ja näissä asioissa, näitä ja näitä asioita painottaen” että tuossa näkisin kyllä massiivisen automaation mahdollisuuden. H5

Haastateltava 5 ei ollut ainoa, jonka mukaan kilpailija-analyysin prosessimalli on teknologioiden tulon myötä muuttunut, sillä myös haastateltava 6 kyseenalaisti perinteisen kilpailija-analyysin vaiheittaisen prosessimallin, sillä tekoälyratkaisuisissa analyysi, tulkinta ja hyödyntäminen limittyvät toisiinsa:

Ei se enää mene in very orderly fashion, step by step, vaan vaiheet lomittuvat (...) analyysi, tulosten tulkinta ja jakaminen tapahtuvat rinnakkain, ja se on tosi hyvä asia. H6

Näiden havaintojen perusteella kilpailija-analyysin perinteiset vaiheet: tiedon kerääminen, seulonta, analyysi ja tulkinta, sulautuvat yhdeksi dynaamiseksi kokonaisuudeksi, joka päivittyy jatkuvasti. Ne eivät siis ole enää tiukasti erottuvia vaihteita, vaan prosessi on jatkuvaa, limittäistä ja samanaikaista toimintaa. Ainoastaan tiedon tarpeen määrittely säilyy aineiston perusteella omana vaiheenaan, jossa tarvitaan ihmisen selkeää liiketoimintakontekstin hahmottamista ja ohjaamista. Haastateltavan 6 tuoma reaaliaikainen kilpailijatietotaulu esimerkki (ks. luku 4.4.2) havainnollistaa tätä käytännönläheisesti, sillä sen avulla tiedot kilpailijasta on kerätty, seulottu ja analysoitu reaaliaikaisesti tekoälyn toimesta.

4.7 Tulosten tarkastelu

Tutkimuksen tulokset osoittavat, että tekoälytehostettu big data -analytiikka tuo merkittäviä hyötyjä kilpailija-analyysin eri vaiheisiin. Tulokset on avattu taulukossa 5.

Taulukko 5 Tutkimuksen empiirisen osuuden tulokset

Tiedon tarpeen määrittely	<ul style="list-style-type: none"> • Ajattelun rikastaminen ja ideoinnin tukeminen generatiivisilla malleilla • Analyysirunkojen esimerkit ja nopeampi hahmottaminen • Tiedon lähteiden kartoittaminen • Kilpailijakentän hahmottaminen
Tiedon kerääminen	<ul style="list-style-type: none"> • Datan keruun tehostuminen ja automatisointi • Multimodaalisuus: tietoa voidaan kerätä eri muodoissa (esim. video, PDF, sosiaalinen media) ja laajentaa näin kilpailija-analyysiä • Manuaalisen työn väheneminen • Monikielinen analyysi • Rajapintojen tehokkaampi hyödyntäminen
Oleellisen tiedon seulonta	<ul style="list-style-type: none"> • Strukturoimattoman datan jäsentely, luokittelu ja tiivistäminen • Promptipohjainen seulonta • Merkittävien teemojen automaattinen tunnistus: relevanttien tietojen kohdentaminen. Myös ilman hypoteesia.
Tiedon analysointi	<ul style="list-style-type: none"> • Analysointinopeuden ja skaalautuvuuden parantuminen • Strukturoimattoman datan muuntaminen rakenteiseksi • Kilpailijoiden vahvuksien ja heikkouksien kartoittaminen • Kilpailijoiden liikkeiden tunnistaminen ja ennakointi; analyysistä reaktiivisen sijaan proaktiivista • Ennustusten tekeminen kilpailijoiden strategioista ja markkinamuutoksista.
Tiedon tulkinta ja johtopäätösten tekeminen	<ul style="list-style-type: none"> • Tulosten tiivistäminen ja raportointi, sanallisten yhteenvedojen ja visualisointien teko • Toimintakehotusten ehdottaminen • Päätöksenteon tukeminen; vuorovaikutteinen tulkinta datasta.
Prosessin muutos	<ul style="list-style-type: none"> • Kilpailija-analyysin vaiheet sulautuvat dynaamiseksi, jatkuvasti päivittyväksi prosessiksi, jossa tiedon keruu, seulonta, analyysi ja tulkinta tapahtuvat limittäin. Ainoastaan tiedon tarpeen määrittely säilyy erillisenä vaiheena, sillä se vaatii ihmisen liiketoimintaymmärrystä ja ohjausta.

Ensimmäisessä vaiheessa, tiedon tarpeen määrittelyssä, generatiivinen tekoäly toimii ajattelun rikastuttajana ja ideoinnin tukena erityisesti kokemattomille käyttäjille tai uusilla toimialoilla. Se nopeuttaa analyysirungon suunnittelua ja auttaa kilpailijatiedon lähteiden kartoittamisessa. Haasteena ensimmäisessä vaiheessa tekoälyn tuottama geneerinen sisältö, joka vaatii käyttäjän ohjausta ja rikastuttamista yrityksen hiljaisella tiedolla.

Kilpailijatiedon keräysvaiheessa tekoäly mahdollistaa nopean ja laaja-alaisen kilpailijatiedon keruun automatisoinnin. Tiedon lähteet voivat olla myös multimodaalisia, kuten videoita, PDF-tiedostoja tai sosiaalisen median sisältöä, mikä laajentaa huomattavasti kilpailija-analyysin tietopohjaa. Kolmannessa vaiheessa tekoäly nopeuttaa

tiedon seulontaa ja oleellisten tietojen löytämistä. Tietoa voidaan seuloa promptipohjaisesti.

Analyysivaiheessa tekoäly mahdollistaa kilpailijoiden vahvuuksien ja heikkouksien kartoittamisen, markkinakäyttäytymisen analysoinnin ja ennakoinnin sekä strategisten liikkeiden ennustamisen. Analyysistä tulee proaktiivista. Johtopäätös- ja tulkintavaiheessa tekoäly tukee johtopäätösten muodostamista tarjoamalla tiivistelmiä, visualisointeja ja toimintakehotuksia, vaikka lopulliset päätökset vaativat edelleen inhimillistä tulkintaa ja kontekstin ymmärrystä. Tärkeä huomio oli myös, että kilpailija-analyysin prosessin luonne muuttuu aiemmasta vaiheittaisesta entistä dynaamisemmaksi ja automatisoidummaksi.

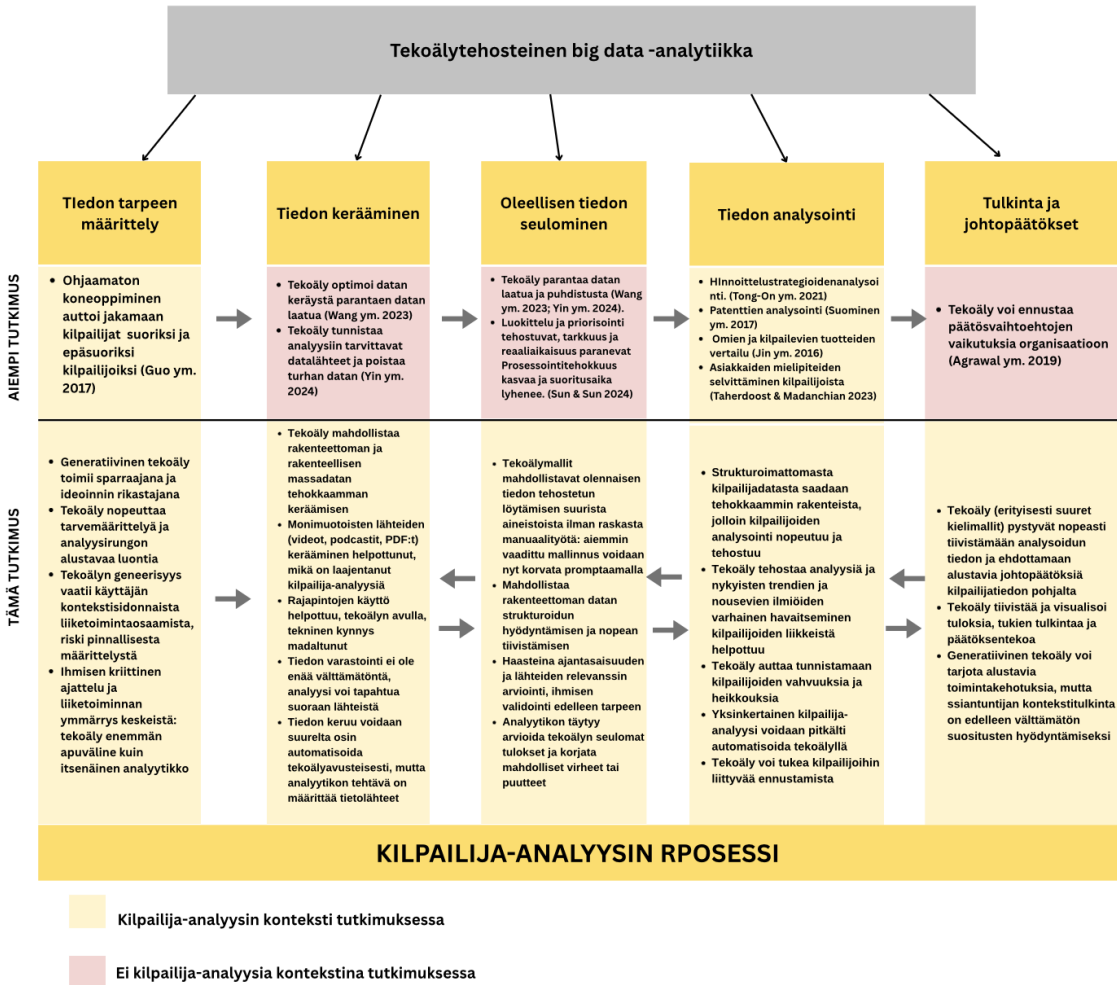
5 Johtopäätökset

5.1 Teoreettiset johtopäätökset

Tämän pro gradu -tutkielman tarkoituksena oli selvittää, millaisia mahdollisuuksia tekoälytehosteinen big data -analytiikka tuo kilpailija-analyysiin. Tähän tutkimuskysymykseen pyrittiin vastaamaan kahden osaongelman kautta: (1) millaisia mahdollisuuksia tekoälytehosteinen big data -analytiikka tuo kilpailijatiedon tarpeiden määrittelyyn, keräämiseen ja seulontaan ja (2) millaisia mahdollisuuksia tekoälytehosteinen big data -analytiikka tuo kilpailijatiedon analysointiin ja tulkintaan.

Täten tämä tutkimus tuo teoreettista kontribuutiota strategisen markkinoinnin kilpailija-analyysin tutkimukseen avaamalla, miten edistyneet teknologiat tukevat kilpailija-analyysiä ja sen eri vaiheita. Samalla tutkimus luo uutta teoriaa kilpailija-analyysin prosessimalliin (Pirttilä 2000) osoittamalla, miten tekoälyn ja big data -analytiikan hyödyntäminen ovat muokanneet perinteistä lineaarista kilpailija-analyysin prosessia yhä dynaamisemmaksi ja automatisoidummaksi. Samalla tutkimus vahvistaa yleistä tietotyöläisten työn muutoksen trendiä osoittamalla, että myös kilpailija-analyytikon työ muuttuu vähemmän manuaaliseksi.

Aiempi kirjallisuus on keskittynyt tekoälyn ja big data -analytiikan hyödyntämiseen yksittäisissä, rajatuissa kilpailija-analyysin osa-alueissa, kuten patenttien (Suominen ym. 2017), kilpailijoiden tuotteiden vertailun (Taherdoost – Madanchian 2023) tai hinnoitteluseurannan (Tong-On ym. 2021) konteksteissa. Kattavaa teoreettista ymmärrystä siitä, miten tekoälytehostettu big data -analytiikka voi tukea koko kilpailija-analyysin prosessia ja muuttaa sitä, on kuitenkin puuttunut. Tämä tutkimus vastaa tähän tutkimusaukkoon tarkastelemalla kilpailija-analyysiä viiden vaiheen kokonaisuutena. Alla on päivitetty teoreettinen viitekehyskuvio, jossa teoreettiset kontribuutiot on esitetty.



Kuvio 11 Tekoölytehosteisen big data -analytiikan tuomat mahdollisuudet kilpailija-analyysiin

Kuvio on jaettu alkuperäiseen ja tämän tutkimuksen kautta syntyneeseen täydennettyyn kuvioon, josta näkyvät sekä uudet tekoölyn ja big data -analytiikan tuomat hyödyt kilpailija-analyysissä että prosessin luonteen muutos nuolilla. Ensimmäisenä vasemmalla käsitellään ensimmäiseen osaongelmaan sisältyvää tiedon tarpeen määrittelyä. Tutkimuksen tulokset osoittavat, että kilpailijatiedon tarpeiden määrittelyssä generatiiviset kielimallit toimivat tehokkaina ajattelun rikastuttajina ja ideoinnin tukena erityisesti silloin, kun analyysia tekevä henkilö on kokematon tai toimii uudella toimialalla. Mallit myös helpottavat kilpailija-analyysin perusrungon suunnittelua, kilpailijoiden hahmottamista ja kilpailijatiedon lähteiden löytämistä. Tutkimuksen perusteella tekoöly voi myös tukea kilpailijakentän hahmottamista. Tämä havainto on linjassa Guon ym. (2017) tutkimuksen kanssa, jossa tekoölyllä rakennettu järjestelmä mahdollisti kilpailijoiden tunnistamisen ja luokittelun suoriin ja epäsuoriin kilpailijoihin.

Näiden hyötyjen ohella tässä tutkimuksessa ilmeni myös haasteita tiedon tarpeiden määrittelyssä. Generatiivisen tekoälyn tuottama tieto on useimmiten yleisluontoista eikä kykene itsenäisesti huomioimaan organisaatiokohtaista strategista kontekstia tai hiljaista tietoa, jotka ovat keskeisiä tekijöitä määriteltäessä, millaista kilpailijatietoa yritys tarvitsee. Tämä havainto ei ole uusi, sillä myös aiemmissa tutkimuksissa on korostettu tekoälyjärjestelmien geneerisyyttä ja kontekstuaalisen herkkyyden puutteen aiheuttamia haasteita (Lopez-Solis 2024, 12–14). Tässä tutkimuksessa ilmeni kuitenkin, että syöttämällä tekoälylle organisaation hiljaista tietoa ja kontekstuaalista informaatiota, on mahdollista muokata analyysiä kilpailija-analyysin kontekstille sopivammaksi.

Aiemmissa tutkimuksissa on osoitettu, että tekoäly ja big data -analytiikka tehostavat yleisesti datan keruuta, kuten relevanttien lähteiden tunnistamista ja keruun automatisointia (Wang ym. 2023; Yin ym. 2024). Tässä tutkimuksessa esiin nousut edistyneiden teknologioiden mahdollistama multimodaalisuus ei myöskään ole ilmiönä uusi (mm. Giannakos ym. 2019), mutta sitä, mitä hyötyä se tuo kilpailija-analyysiin, ei ole vielä tunnistettu aiemmissa tutkimuksissa. Tämä tutkimus laajentaakin aiempaa ymmärrystä osoittamalla, miten nämä teknologiat muuttavat nimenomaan kilpailijatiedon keruuta ja siten kilpailija-analyysiä. Kilpailijoista voidaan nyt kerätä tietoa entistä tehokkaammin laajemmista ja monimuotoisemmista lähteistä, kuten videoista, podcasteista ja sosiaalisesta mediasta, ja analysoida sitä monikielisesti ilman kieliosaamista. Aiemmin käsitys siitä, mitä tietoa kilpailijoista kannattaisi kerätä, oli paljon suppeampi, eikä näitä tietolähteitä ole hyödynnetty aiemmin kilpailija-analyysissä.

Kilpailijatiedon kerääminen pystytään myös lähes kokonaan automatisoimaan, jonka seurauksena kilpailija-analyttikoiden manuaalinen työ vähenee ja he voivat keskittyä kilpailija-analyysin myöhempisiin vaiheisiin, strategiseen tulkintaan ja varsinaiseen analysointiin. Nämä tekijät muuttavat kilpailija-analyysin edellytyksiä kokonaisvaltaisesti, sillä analyysin tietopohja laajenee huomattavasti, sen ajantasaisuus paranee, ihmisen tekemä manuaalinen työ vähenee automaation myötä ja kilpailija-analyysi voi perustua entistä suurempaan ja moninaisempaan dataan. Analyttikkojen työnkuvan muutos on tunnistettu muissakin tutkimuksissa. Manuaalisen datan käsittelyn väheneminen on tunnistettu muuttavan analyttikkojen työtä (Hamadaqa ym. 2024, 20) ja tämä koskee myös kilpailija-analyysiä tekeviä.

Tiedon seulonnassa tekoälytehosteisen big data -analytiikan avulla voidaan tiivistää ja priorisoida suuria määriä monimuotoista dataa. Erityisesti suuret kielimallit mahdollistavat tietomassojen automaattisen jäsentelyn, tiivistämisen ja kohdistetun tiedonhaun ilman perinteistä manuaalista työtä tai ennakkohypoteeseja. Nämä huomiot ovat linjassa aiemman tutkimuksen kanssa (ks. Wang ym. 2023; Yin ym. 2024; Sun – Sun 2024). Tekoälytehosteinen big data -analytiikka pystyy siis tämän tutkimuksen mukaan myös tehostamaan tätä vaihetta kilpailija-analyysin prosessissa. Samaan aikaan tulokset kuitenkin osoittavat, että tässä vaiheessa prosessia voi olla sama riski kuin tiedon tarpeen määrittelyssä, eli että generatiivinen tekoäly ei aina kykene itsenäisesti määrittämään, mikä on strategisesti merkityksellistä analyysin kannalta. Tämän lisäksi ajallisen relevanssin arvioiminen on osoittautunut osittaiseksi haasteeksi.

Toiseen osaongelmaan puolestaan sisältyi kilpailijatiedon analysointi, tulkinta ja johtopäätökset. Analyysivaiheessa tekoälytehosteinen big data -analytiikka tuo uusia mahdollisuuksia erityisesti kilpailijoiden strategisten liikkeiden tunnistamiseen ja ennakoivaan analyysiin. Löydökset viittaavat siihen, että analyysi muuttuu reaktiivisesta enemmän proaktiiviseksi, mikä avartaa aiemman kilpailija-analyysin teorian näkökulmaa, jossa kilpailija-analyysi nähtiin lähinnä takautuvana tarkasteluna (ks. Wright ym. 2009). Aiempi tutkimus on käsitellyt kilpailijoiden liikkeiden ennakoitua lähinnä patenttien analysoinnin näkökulmasta (Suominen ym. 2017), mutta tämä tutkimus osoittaa, että tekoälytehostettu big data -analytiikka mahdollistaa hiljaisten kilpailijoihin liittyvien signaalien havaitsemisen, trendien tunnistamisen ja jopa strategisten liikkeiden ennustamisen kilpailija-analyysin kontekstissa. Tekoäly tukee myös kilpailija-analyysin keskiössä olevaa kilpailijoiden vahvuuksien ja heikkouksien kartoittamista.

Kilpailijatiedon tulkinta- ja johtopäätösvaiheessa tekoälytehostettu big data -analytiikka tukee prosessia tarjoamalla johtopäätösvaihtoehtoja, nopeuttamalla raportointia, visualisoimalla analyysien tulokset sekä tuottamalla alustavia toimintasuosituksia. Nämä yhdessä tehostavat kilpailija-analyysin viimeistä vaihetta. Samoin kuin Yin ym. (2024) toteavat, myös tämän tutkimuksen mukaan tekoäly voi automatisoida datan visualisointia analysoimalla käyttäjän tarpeet ja datan ominaisuudet, ehdottamalla sopivia esitystapoja ja mahdollistamalla vuorovaikutteisen datan tutkimisen. Tässä tutkimuksessa havaittiin, että kilpailija-analyysin johtopäätösten tekeminen tehostuu näiden teknologioiden avulla, mikä on linjassa aiemman tutkimuksen kanssa, joka käsittelee yleistä yrityksen

päätöksentekoa tekoälytehosteisesti (Mohamed 2025). Davenport ja Ronanki (2018) korostavat, että tekoäly ei korvaa päätöksentekoa, vaan siirtää ihmisen roolia kysymysten asettajaksi ja merkitysten tulkitsijaksi.

Tämä tutkimus kehittää myös aiempaa kilpailija-analyysin prosessimallia (Pirttilä 2000). Perinteinen kilpailija-analyysin prosessimalli etenee hyvin vahvasti eri vaiheissa, mutta tutkimuksesta selvisi, että tämä ei kuvaa nykyistä kilpailija-analyysia, jossa tekoälyä ja big data -analytiikkaa hyödynnetään. Tänä päivänä prosessi on dynaamista, eri prosessin vaiheet tapahtuvat usein rinnakkain ja lomittain ja tieto päivittyy jatkuvasti. Tätä kuvaavat kuviossa 11 edestakaiset nuolet tiedon keräämisen, seulomisen, analyysin ja tulkinnan ja johtopäätösten välissä. Ainoastaan alun tiedon tarpeen määrittely nähtiin selvästi erillisenä vaiheena prosessissa, ja tästä syystä sen ja keräämisen välissä on yhdensuuntainen nuoli. Tutkimus ehdottaa, että kilpailija-analyysiä tulisi tarkastella yhä enemmän dynaamisena ja reaaliaikaisena prosessina.

Tämä tutkimus tarjoaa markkinoinnin tutkimukselle uutta ymmärrystä siitä, miten tekoäly ja big data -analytiikka voivat muuttaa perinteistä kilpailija-analyysiä. Se osoittaa, että tekoäly ja big data -analytiikka eivät pelkästään tehosta ja laajenna kilpailija-analyysiä, vaan myös uudelleenmuotoilevat koko kilpailija-analyysin prosessia sekä kilpailija-analyttikon roolia.

5.2 Liikkeenjohdolliset suositukset

Tutkimus tarjoaa liikkeenjohdolle näkökulmia siihen, miten tekoälytehostettu big data -analytiikka voi muuttaa ja tehostaa kilpailija-analyysin prosessia sekä yritykset voivat myös ymmärtää, mitä hyötyjä he voivat saada, mikäli integroivat edistyneitä teknologioita osaksi kilpailija-analyysensä. Kilpailija-analyysin käsittely kattavana prosessina mahdollistaa sen, että liikkeenjohto saa selkeän ja jäsennellyn kokonaiskuvan kaikista prosessin vaiheista ja siitä, miten jokainen vaihe on merkityksellinen seuraavan kannalta. Näin varmistetaan, että mikään vaihe ei jää epäselväksi tai irralliseksi, vaan organisaatiolla on valmiudet toteuttaa kilpailija-analyysi johdonmukaisesti alusta loppuun.

Ensimmäinen liikkeenjohdolle merkittävä huomio on tiedon tarpeen määrittelyn helpottuminen. Kilpailija-analyysi voidaan tekoälytehostetun big data -analytiikka avulla aloittaa nopeasti ja tehokkaasti. Generatiiviset mallit tukevat analyysin alkuvaiheen

ideointia ja tiedontarpeen määrittelyä esimerkiksi tarjoamalla ehdotuksia, auttamalla analyysirunkojen rakentamisessa ja laajentamalla näkökulmia. Tämä nopeuttaa analyysin suunnittelua ja madaltaa kynnystä aloittaa kilpailijaseuranta, eritoten pienemmissä organisaatioissa ja yrityksissä, joissa kilpailija-analyysiä ei ennen ole tehty. Myös kokemattomat kilpailija-analyytikot voivat tehdä yhä laadukkaampaa tiedon tarpeen määrittelyä tekoälyn avulla.

Tekoälyn ja big data -analytiikan avulla kilpailija-analyysistä on mahdollista saada aiempaa syvällisempää ja kattavampaa. Teknologiat mahdollistavat esimerkiksi multimodaalisen datan hyödyntämisen, kuten videoiden, podcastien ja PDF-tiedostojen, tuoden analyysiin mukaan uusia tietolähteitä, joiden käyttö on aiemmin ollut työlästä tai mahdotonta, mutta jotka nyt tarjoavat merkittävää lisäarvoa kilpailijaseurantaan. Kilpailija-analyysistä voidaan multimodaalisuuden avulla saada kattavampaa kuin ennen, ja suosituksena onkin, että yritykset käyttäisivät enemmän luovuutta siinä, mistä kaikkialta dataa kilpailijoista voitaisiin kerätä. Näiden mahdollisuuksien myötä organisaatioiden kannattaa aktiivisesti korvata manuaalisia tiedonkeruun ja -seulonnan vaiheita tekoäly- ja big data -ratkaisuilla tehokkuuden parantamiseksi.

Analyysivaiheessa puolestaan tekoäly tarjoaa huomattavia hyötyjä analyysin nopeudessa ja laadussa. Suuret kielimallit pystyvät tekemään itsenäisiä analyysejä ja tukemaan visuaalista esittämistä, mikä parantaa tulosten ymmärrettävyyttä ja päätöksentekoa. Lisäksi tekoäly mahdollistaa kilpailijoiden strategisten liikkeiden entistä tehokkaamman havainnoinnin ja jopa niiden ennakoinnin. Myös kilpailijoiden tulevaisuuden liikkeistä on mahdollista tehdä ennustuksia. Nämä mahdollisuudet tekevät kilpailija-analyysistä proaktiivista eikä pelkästään reaktiivista toimintaa. Tämä mahdollistaa yrityksille sen, että ne voivat pyrkiä olemaan askelta edellä kilpailijoitaan.

Merkittävä huomio liikkeenjohdolle on manuaalisen työn vähentyminen ja automaation lisääntyminen koko kilpailija-analyysin prosessissa. Tämä tarkoittaa, että yritysten on näiden teknologioiden avulla mahdollista saavuttaa kattavampi ja laadukkaampi kilpailija-analyysi aiempaa kustannus- ja resurssitehokkaammin. Kun tekoäly mahdollistaa tietojen automatisoidun keräämisen, jäsentelyn ja analyysin, kilpailija-analyytikkojen tehtävät siirtyvät kohti strategista tulkintaa ja päätöksentekoa. Tämä muutos edellyttää uudenlaista osaamista ja roolien määrittelyä organisaatioissa.

Tekoälytehosteinen big data -analytiikka hyödyttää organisaatioita tehostamalla koko kilpailija-analyysin prosessia vähentämällä manuaalista työtä, laajentamalla analyysin kattavuutta ja datalähteitä sekä siirtymällä ennakoivaan ja proaktiiviseen kilpailijaseurantaan. Näin ollen liikkeenjohdolle suositellaan tekoälyn ja big data -analytiikan hyödyntämistä kilpailija-analyysiä tehtäessä.

5.3 Tutkimuksen arviointi

Tutkimus tarjoaa uutta teoreettista ymmärrystä ja käytännönläheisiä havaintoja, mutta samalla työ kohtaa myös tiettyjä rajoitteita, jotka on huomioitava tuloksia tulkittaessa. Ensinnäkin tutkimuksen aineisto perustuu kahdeksaan asiantuntijahaastatteluun, ja vaikka haastateltavat edustivat laajasti eri toimialoja ja teknologiaosaamista, näyte on silti rajallinen. Ulkoisen validiteetin näkökulmasta tutkimuksen tulokset ovat kontekstisidonnaisia, sillä haastatellut edustivat Suomessa toimivia yrityksiä, ja tutkimus ajoittui ajankohtaan, jolloin tekoälyn ja big data -analytiikan hyödyntäminen oli monilla toimialoilla vielä kehitysvaiheessa. Muualla, eri maantieteellisissä ja teknologisissa konteksteissa, näiden teknologioiden kehityksen aste ja niistä saadut hyödyt, voivat poiketa merkittävästi, mikä voi rajoittaa tulosten yleistettävyyttä.

Luvussa 3.7 on pyritty arvioimaan kattavasti tutkimuksen luotettavuutta, mutta laadullisen tutkijan arvioidessa tutkimustaan, arviointi ei ole ikinä täysin objektiivista (Eskola – Suoranta 1998) ja tutkijan puolueettomuutta voi olla vaikea täysin poistaa (Tuomi – Sarajärvi 2018). Laadulliset haastattelut tarjoavat syvällistä tietoa, mutta haastattelijan rooli ja kysymyksenasettelu voivat myös vaikuttaa vastauksiin (ns. haastattelijan vaikutus). Lisäksi tutkijan oma tulkinta vaikuttaa aineiston analyysiin.

Lopuksi, koska tutkimus painottui asiantuntijoiden näkemyksiin, se ei sisällä käytännön näyttöä tekoälyn käytön vaikutuksista kilpailija-analyysiin. Lisäksi on huomioitava, että osa haastateltavista saattoi pohjata näkemyksensä enemmän tekoälyn ja big data -analytiikan potentiaaliin kuin omiin vakiintuneisiin käytänteisiin.

5.4 Jatkotutkimusmahdollisuudet

Tämä tutkimus on tarjonnut ymmärrystä siitä, miten tekoälytehostettu big data -analytiikka voi tukea kilpailija-analyysin eri vaiheita. Tutkimuksen löydökset herättävät kuitenkin jatkokysymyksiä, joita voidaan tarkastella tulevassa tutkimuksessa.

Yksi jatkotutkimuksen suunta voisi olla kvantitatiivinen tutkimus, jossa tutkittaisiin luvun 4.2.2 kaltaista reaaliaikaista tekoälyä ja big dataa hyödyntävää kilpailijatietotaulua ja sen vaikuttavuutta yrityksen strategiseen päätöksentekoon. Tutkimuksessa voitaisiin verrata aikaa ennen työkalun käyttöönottoa käyttöönoton jälkeiseen aikaan ja tutkia esimerkiksi päätöksenteon nopeutta, laatua ja reagointikykyä markkinamuutoksiin. Myös vastaavanlainen vertailu voitaisiin suorittaa mitaten suorituskykyä. Tällainen tutkimus olisi hyödyllistä, jotta nähtäisiin, tuoko kilpailija-analyysin tehostaminen edistyneillä teknologioilla yrityksille parempaa suorituskykyä.

Toinen jatkotutkimuskohde voisi olla tutkimus suomalaisten yritysten tämänhetkisestä kilpailija-analyysin toteutuksesta ja siitä, missä vaiheessa teknologista siirtymää yritykset ovat kilpailija-analyysin osalta. Tämä antaisi yleiskuvaa siitä, kuinka laajasti edistyneitä teknologioita on käytössä kilpailijoiden seuraamisessa ja analysoinnissa.

Tutkimuksesta nousi myös viitteitä, että kilpailijoiden liikkeiden ennakointi ja ennustaminen on helpottunut tekoälyn ja big data -analytiikan hyödyntämisen myötä. Olisikin hyödyllistä tutkia tätä aihetta enemmän ja syvällisemmin, että millä tarkkuudella nämä teknologiat pystyvät ennakoimaan kilpailijoiden liikkeitä ja tekemään ennustuksia ja niiden pohjalta toimintakehotuksia. Myös toimintakehotusten hyödyllisyys voisi olla mittarina, eli ohjaako tekoäly oikeasti oikeisiin hypoteeseihin ja toimintaan.

Lopuksi, koska kilpailija-analyysin tutkimus ei ole ollut kovinkaan pinnalla viime vuosina, olisi myös mielenkiintoista tutkia, että onko edistyneiden teknologioiden mahdollistama kilpailija-analyysin tehostus ja laadun parannus tehnyt kilpailija-analyysistä entistä merkittävämmän yrityksen toiminnan ja päätöksenteon kannalta.

6 Yhteenveto

Tämän tutkimuksen tavoitteena oli selvittää, mitä mahdollisuuksia tekoälytehosteinen big data -analytiikka tuo kilpailija-analyysiin ja sen prosessin eri vaiheisiin. Kilpailija-analyysi on merkittävä osa markkinointitutkimusta ja tärkeä työkalu yrityksen strategisessa johtamisessa, jotta yritykset pystyvät tavoittelemaan kilpailuetua suhteessa kilpailijoihinsa. Edistyneet teknologiat ovat mullistaneet jo hyvin laajalti yritysten strategista johtamista ja käytänteitä, joten on hedelmällistä tutkia, miten ne voivat tukea myös kilpailija-analyysiä.

Tutkimuskysymykseen vastattiin kahden osaongelman kautta, joista ensimmäinen tutki tekoälytehosteisen big data -analytiikan tuomia mahdollisuuksia kilpailija-analyysin alkuosuuteen eli tiedon tarpeenmäärittelyyn, keräämiseen ja seulontaan. Toinen osaongelma puolestaan kattoi prosessin loppuosan eli kilpailijoiden analysoinnin, tulkinnat ja johtopäätökset. Tutkimuksen teoreettinen viitekehys pohjautui kilpailija-analyysin viisivaiheisen prosessimalliin ja tekoälyn ja big data -analytiikan kirjallisuuteen, yhdistäen nämä kaksi osa-aluetta. Tekoälytehosteista big data -analytiikkaa ei ole aiemmin tutkittu kilpailija-analyysin prosessimallin yhteydessä, joten tämä tutkimus vastasi löydöksillään tähän tutkimusaukkoon.

Tutkimuksen tulokset osoittivat, että tekoälytehosteinen big data -analytiikka tehostaa, laajentaa ja automatisoi koko kilpailija-analyysin prosessia. Tiedon tarpeenmäärittelyssä tekoäly auttaa ideoinnissa, kehysten luonnissa ja tiedon lähteiden hahmottamisessa. Myös kilpailukentän hahmottaminen helpottuu. Tämä prosessivaihe vaatii silti yhä ihmisen omaa liiketoimintalähtöistä ajattelua ja kontekstitietoisuutta. Loppuvaiheet kilpailija-analyysin prosessissa, eli tiedon kerääminen, seulonta, analysointi, tulkinta ja johtopäätösten tekeminen, nähtiin sulautuvan yhdeksi dynaamiseksi kokonaisuudeksi aiemman vaiheittaisen etenemisen sijaan. Tekoäly ja big data -analytiikka mahdollistavat näiden vaiheiden koneellisen automaattisen suorittamisen samanaikaisesti, mikä haastaa aiemman vaiheittaisen prosessimallin luonteen. Kilpailija-analyysin tiedon kerääminen voidaan tehdä reaaliaikaisesti automaattisesti seuloen ja muodostaen siitä helposti jäseneltäviä visuaalisia ja tekstuaalisia analyysikokonaisuuksia. Tämä tehostaa kilpailija-analyysiä ja muuttaa kilpailija-analyttikkojen työnkuvaa, sillä manuaalisen työn määrä vähenee ja panokset voidaan siirtää sitä enemmän validoijan ja strategisen tulkinnan osa-alueille.

Tehokkaan kilpailija-analyysin lisäksi kilpailijoista on näiden teknologioiden avulla myös mahdollista kerätä entistä laajemmin dataa sekä muodoltaan että määrältään. Näiden edistyneiden teknologioiden mahdollistaman multimodaalisuuden ansiosta, kilpailijoita pystytään helpommin analysoimaan myös esimerkiksi video-, kuva-, podcast- ja tiedostomuotojen kautta, jolloin kilpailija-analyysistä saadaan aiempaa kattavampaa ja vaikuttavampaa.

Myös varsinainen analyysi muuttuu syvällisemmäksi ja hyödyllisemmäksi. Tekoälytehosteinen big data -analytiikka mahdollistaa entistä helpommin kilpailijoiden käyttäytymistrendien tunnistamisen, ja ennakkoinnin ja jopa ennustamisen. Tekoäly voi myös auttaa kilpailijoiden vahvuuksien ja heikkouksien kartoittamisessa.

Nämä havainnot tuovat uutta ymmärrystä markkinointitieteelle, jossa kilpailija-analyysin prosessia ei näiden edistyneiden teknologioiden kontekstissa ole vielä tutkittu. Tulokset laajentavat ymmärrystä kilpailija-analyysin prosessista ja kuinka siitä voidaan saada entistä tehokkaampaa ja vaikuttavampaa yrityksen päätöksenteon näkökulmasta. Myös kilpailija-analyysin prosessin ja kilpailija-analyttikkojen työnkuvan havaittiin muuttuneen näiden teknologioiden integroinnin myötä ja aiempi lineaarinen vaiheittainen prosessimalli haastettiin. Nämä löydökset ovat sekä teoreettisesti että käytännön liikkeenjohdolle merkittäviä löydöksiä. Liikkeenjohto pystyy peilaamaan omaa tarvettaan kilpailija-analyysin kehittämiseksi näiden teknologioiden avulla, kun positiiviset mahdollisuudet ovat tiedossa.

Lähteet

- Adom, A. Y. – Nyarko, I. K. – Kumi Som, G. N. (2016) Competitor analysis in strategic management: Is it a worthwhile managerial practice in contemporary times? *Journal of Resources Development and Management*, Vol. 24, 116–127.
- Agarwal, S. – Ahmad, N. – Jamali, D. – Ahmad, N. – Chauhan, P. (2024) AI and big data in contemporary marketing. *Computer*, Vol. 57 (4), 137–142.
- Agrawal, A. – Gans, J. S. – Goldfarb, A. (2019) Exploring the impact of artificial Intelligence: Prediction versus judgment. *Information Economics and Policy*, Vol. 47, 1–6.
- Ali, Z. – Bhaskar, S. B. (2016) Basic statistical tools in research and data analysis. *Indian Journal of Anaesthesia*, Vol. 60 (9), 662–669.
- Anand, G. – Kodali, R. (2008) Benchmarking the benchmarking models. *Benchmarking: An International Journal*, Vol. 15(3), 257–291.
- Bakator, M. – Vukoja, M. – Manestar, D. (2023) Achieving competitiveness with marketing 5.0 in new business conditions. *UTMS Journal of Economics*, Vol. 14 (1), 63–73.
- Banh, L. – Strobel, G. (2023) Generative artificial intelligence. *Electronic Markets*, Vol. 33 (1), 63.
- Barney, J. (1991) Firm resources and sustained competitive advantage. *Journal of Management*, Vol. 17 (1), 99–120.
- Basu, R. – Lim, W. M. – Kumar, A. – Kumar, S. (2023) Marketing analytics: The bridge between customer psychology and marketing decision-making. *Psychology & Marketing*, Vol. 40 (12), 2588–2611.
- Bennett, R. (2003) Competitor analysis practices of British charities. *Marketing Intelligence & Planning*, Vol. 21 (6), 335–45.
- Bergen, M. – Peteraf, M. A. (2002) Competitor identification and competitor analysis: A broad-based managerial approach. *Managerial and Decision Economics*, Vol. 23 (4–5), 157–169.
- Bernhardt, D. C. (1994) “I want it fast, factual, actionable” - Tailoring competitive intelligence to executives' needs. *Long Range Planning*, Vol. 27 (1), 12.
- Calof, J.– Arcos, R. – Sewdass, N. (2017) Competitive intelligence practices of European firms. *Technology Analysis & Strategic Management*, Vol. 29 (6), 639–672.

- Cambria, E. – White, B. (2014). Jumping NLP curves: A review of natural language processing research. *IEEE computational intelligence magazine*, Vol. 9 (2), 48–57.
- Cao, G. – Duan, Y. – Edwards, J. S. – Dwivedi, Y. K. (2021) Understanding managers' attitudes and behavioral intentions towards using artificial intelligence for organizational decision-making. *Technovation*, Vol. 106, 102312.
- Rheinhardt, A. – Kreiner, G. E. – Gioia, D. A. – Corley, K. G. (2018). Conducting and publishing rigorous qualitative research. Teoksessa Cassell, C. – Cunliffe, A. L. – Grandy, G. (toim.) *The SAGE Handbook of Qualitative Business and Management Research Methods: History and Traditions*, 515–532., SAGE Publications Ltd. London.
- Chakraborti, S. – Dey, S. (2019) Analysis of competitor intelligence in the era of big data: An integrated system using text summarization based on global optimization. *Business & Information Systems Engineering*, Vol. 61 (3), 345–355.
- Chiu, M.-C. – Huang, J.-H. – Gupta, S. – Akman, G. (2021) Developing a personalized recommendation system in a smart product service system based on unsupervised learning model. *Computers in Industry*, Vol. 128, 103421.
- Creswell, J. W. (2007) *Qualitative inquiry and research design: Choosing among five approaches*, 2. p. SAGE Publications, Inc., Los Angeles.
- Creswell, J. W. – Poth, C. N. (2016) *Qualitative Inquiry and research design: Choosing Among Five Approaches*. 4. p. SAGE Publications, Los Angeles.
- Davenport, T. H. – Ronanki, R. (2018) Artificial intelligence for the real world. *Harvard Business Review*, Vol. 96 (1), 108–116.
- Davenport, T. H. (2018) From analytics to artificial intelligence. *Journal of Business Analytics*, Vol. 1 (2), 73–80.
- De Mauro, A. – Greco, M. – Grimaldi, M. (2015) What is big data? A consensual definition and a review of key research topics. Teoksessa: *AIP conference proceedings*, Vol. 1644 (1), 97-104).
- Dean, J. – Ghemawat, S. (2008). MapReduce: Simplified data processing on large clusters. *Communications of the ACM*, Vol. 51(1), 107–113.
- Devi, R. S. – Siddharth, R. K. – Manjula, D. – Xu, Z. (2015) An efficient approach for web indexing of big data through hyperlinks in web crawling. *The Scientific World*, 1–9.

- Dijcks, J.-P. (2012) *Oracle: big data for the enterprise*. Oracle White Paper.
- Duan, Y. –Edwards, J. S. – Dwivedi, Y. K. (2019) Artificial intelligence for decision making in the era of big data – evolution, challenges and research agenda. *International Journal of Information Management*, Vol. 48, 63–71.
- Eriksson, P. – Kovalainen, A. (2008) *Qualitative methods in business research*. SAGE Publications Ltd, London.
- Eskola, J. – Suoranta, J. (1998) *Johdatus laadulliseen tutkimukseen*. 3. uud. p. Vastapaino, Tampere.
- Fereday, J. – Muir-Cochrane, E. (2006) Demonstrating rigor using thematic analysis: A hybrid approach of inductive and deductive coding and theme development. *International Journal of Qualitative Methods*, Vol. 5 (1), 80–92.
- Feuerriegel, S. –Hartmann, J. – Janiesch, C. – Zschech, P. (2024) Generative AI. *Business & Information Systems Engineering*, Vol. 66 (1), 111–126.
- Fisher, C. (2010) *Researching and writing a dissertation: An essential guide for business students*. Pearson Education, Limited.
- Fitz-enz, J. – Mattox, J. R. (2014) *Predictive Analytics for Human Resources*. John Wiley & Sons, Hoboken, New Jersey.
- Fleisher, C. S. – Bensoussan, B. E. (2007). *Business and Competitive Analysis: Effective Application of New and Classic Methods*. Saddle River, NJ: FT Press.
- Giannakos, M. N. – Sharma, K. – Pappas, I. O. – Kostakos, V. –Velloso, E. (2019) Multimodal data as a means to understand the learning experience. *International Journal of Information Management*, Vol. 48, 108–119.
- Gilligan, C. – Wilson, R. M. S. (2009) *Strategic marketing planning*. 2. p. Elsevier/Butterworth-Heinemann, Amsterdam.
- Gilligan, C. – Wilson, R. M. S. – Richard M. S. (2003) *Strategic marketing planning*. Butterworth-Heinemann, Amsterdam.
- Gonçalves, C. T. – Gonçalves, M. J. A. – Campante, M. I. (2023) Developing integrated performance dashboards visualisations using power BI as a platform. *Information (Basel)*, Vol. 14 (11), 13–14.
- Goodfellow, I. – Bengio, Y. – Courville, A. (2016) *Deep learning*. The MIT Press.
- Grix, J. (2002) Introducing students to the generic terminology of social research. *Politics*, Vol. 22 (3), 175.

- Guo, L. – Sharma, R. – Yin, L. – Lu, R. – Rong, K. (2017) Automated competitor analysis using big data analytics. *Business Process Management Journal*, Vol. 23(3), 735–762.
- Gupta, S. – Kar, A. K. – Baabdullah, A. – Al-Khowaiter, W. A. A. (2018) Big data with cognitive computing: A review for the future. *International Journal of Information Management*, Vol. 42, 78–89.
- Haenlein, M. – Kaplan, A. (2019a) A brief history of artificial intelligence: On the past, present and future of artificial intelligence. *California Management Review*, Vol. 61 (4), 5–14.
- Hair, J. F. Jr. (2007). Knowledge creation in marketing: The role of predictive analytics. *European Business Review*, Vol. 19(4), 303–315.
- Hamadaqa, M. H. M. – Alnajjar, M. – Ayyad, M. N. – Al-Nakhal, M. A. – Abunasser, B. S. – Abu-Naser, S. S. (2024) Leveraging artificial intelligence for strategic business decision-making: Opportunities and challenges. *International Journal of Academic Information Systems Research*, Vol. 8 (8), 16 –23.
- Hatzijordanou, N. – Bohn, N. – Terzidis, O. (2019) A systematic literature review on competitor analysis: status quo and start-up specifics. *Management Review Quarterly*, Vol. 69, 415–458.
- Heinen, C. – Hoffjan, A. (2005) The strategic relevance of competitor cost assessment- An empirical study of competitor accounting. *Journal of Applied Management Accounting Research*, Vol. 3(1), 17–33.
- Hinds, J. – Joinson, A. N. (2018) What demographic attributes do our digital footprints reveal? A systematic review. *PloS One*, Vol. 13 (11), 17–18.
- Hirsjärvi, S. –Hurme, H. – Gaudeamus oy, kustantaja. (2022) *Tutkimushaastattelun teemahaastattelun teoria ja käytäntö*. 2. p. Gaudeamus.
- Ho, J. C. – Lee, C.-S. (2008). The DNA of industrial competitors. *Research Technology Management*, Vol. 51 (4), 17–20.
- Huang, M.-H. – Rust, R. –Maksimovic, V. (2019) The feeling economy: Managing in the next generation of artificial intelligence (AI). *California Management Review*, Vol. 61 (4), 43–65.
- Huang, M.-H. – Rust, R. T. (2021) A strategic framework for artificial intelligence in marketing. *Journal of the Academy of Marketing Science*, Vol. 49 (1), 30–50.
- Ivanov, D. – Tang, C. S. – Dolgui, A. – Battini, D. – Das, A. (2021) Researchers’ perspectives on Industry 4.0: Multi-disciplinary analysis and opportunities for

- operations management. *International Journal of Production Research*, Vol. 59 (7), 2055–2078.
- Järvenpää, M. – Lämsiluoto, A. – Partanen, V. – Pellinen, J. (2017) *Talousohjaus ja kustannuslaskenta* 2.–4. p. Sanoma Pro Oy.
- Jin, J. – Ji, P. – Gu, R. (2016) Identifying comparative customer requirements from product online reviews for competitor analysis. *Engineering Applications of Artificial Intelligence*, Vol. 49, 61–73.
- Kamkankaew, P. – Phattarowas, V. – Khumwongpin, S. – Limpiaongkhanan, P.–Sribenjachot, S. (2022) Increasing competitive environment dynamics and the need of hyper-competition for businesses. *International Journal of Sociologies and Anthropologies Science Reviews*, Vol. 2 (5), 9–20.
- Kaplan, A. – Haenlein, M. (2019b) Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, Vol. 62 (1), 15–25.
- Knudsen, E. S. – Lien, L. B. – Timmermans, B. – Belik, I. – Pandey, S. (2021) Stability in turbulent times? The effect of digitalization on the sustainability of competitive advantage. *Journal of Business Research*, Vol. 128, 360–369.
- Kreutzer, R. T. – Sirrenberg, M. (2020) *Understanding artificial intelligence: Fundamentals, use cases and methods for a corporate AI journey*. Springer International Publishing.
- Kumar, V.D.A. – Varadarajan, V. – Gupta, M.K. – Rodrigues, J.J.P.C. – Janu, N. (2022) AI empowered big data analytics for industrial applications. *Journal of Universal Computer Science*, Vol. 28 (9), 877–881.
- Laney, D. (2001) "3D data management: Controlling data volume velocity and variety", *META Group Res. Note*, vol. 6 (70), 1.
- Li, Z. – Fan, Y. – Jiang, B. – Lei, T. – Liu, W. (2019) A survey on sentiment analysis and opinion mining for social multimedia. *Multimedia Tools and Applications*, Vol. 78 (6), 6939–6967.
- Lian, X. – Zhang, X. (2022) Learning-based data storage [Vision] (Technical Report).
- López-Solís, O. – Luzuriaga-Jaramillo, A. – Bedoya-Jara, M. – Naranjo-Santamaría, J. Bonilla-Jurado, D. – Acosta-Vargas, P. (2025) Effect of generative artificial intelligence on strategic decision-making in entrepreneurial business initiatives: A systematic literature review. *Administrative Sciences*, Vol. 15 (2), 66.

- Lukoianova, T. – Rubin, V. (2014) Veracity roadmap: Is big data objective, truthful and credible? *Advances In Classification Research Online*, Vol. 24 (1).
- Madden, S. (2012) From databases to big data. *IEEE Internet Computing*, Vol. 16 (3), 4–6.
- Mohamed, G. (2025) ‘Comparative analysis of AI-driven decision support systems and traditional spreadsheets: Evaluating accuracy and consistency in business intelligence’. *Journal of Science and Technology*, Vol. 30 (4).
- Mohammed, M. – Khan, M.B. – Bashier Mohammed, B.E. (2016) *Machine learning: algorithms and applications*. CRC Press.
- Monino, J.-L. (2021) Data value, big data analytics, and decision-making. *Journal of the Knowledge Economy*, Vol. 12 (1), 256–267.
- Nag, R.– Hambrick, D. C.– Chen, M.-J. (2007) What is strategic management, really? Inductive derivation of a consensus definition of the field. *Strategic Management Journal*, Vol. 28 (9), 935–955.
- Nguyen, D. K. – Sermpinis, G. – Stasinakis, C. (2023) Big data, artificial intelligence and machine learning: A transformative symbiosis in favour of financial technology. *European Financial Management: The Journal of the European Financial Management Association*, Vol. 29 (2), 517–548.
- Pellissier, R. – Nenzhelele, T. E. (2013) Towards a universal competitive intelligence process model. *South African Journal of Information Management*, Vol. 15 (2), 1–7.
- Petrescu, M.– Krishen, A. S. (2023) Mapping 2022 in Journal of Marketing Analytics: what lies ahead? *Journal of Marketing Analytics*, Vol. 11 (1), 1–4.
- Pirttilä, A. (2000) *Kilpailijaseuranta*. WSOY, Helsinki.
- Porter, M.E. (1980) *Competitive strategy: techniques for analyzing industries and competitors*. Free Press, New York
- Porter, M.E. (1985) *Competitive advantage: creating and sustaining superior performance*. Free Press, New York
- Porter, M. E. (2004) *Competitive Strategy: Techniques for Analyzing Industries and Competitors*. Free Press, New York.
- Prasnikar, J. – Debeljak, Z. – Ahcan, A. (2005) Benchmarking as a tool of strategic management. *Total Quality Management & Business Excellence*, Vol. 16 (2), 257–275.

- Puusa, A. – Juuti, P. (2020) *Laadullisen tutkimuksen näkökulmat ja menetelmät*. Gaudeamus, Helsinki.
- Rahmani, A. M. – Azhir, E. – Ali, S. – Mohammadi, M. – Ahmed, O. H. – Ghafour, M. Y. – Ahmed, S. H. – Hosseinzadeh, M. (2021) Artificial intelligence approaches and mechanisms for big data analytics: A systematic study. *Computer Science*, Vol. 7, 1–28.
- Riemer, K. – Peter, S. (2024) Conceptualizing generative AI as style engines: Application archetypes and implications. *International Journal of Information Management*, Vol. 79, 3–4.
- Robinson, C. V. – Ahmad, F. – Simmons, J. E. L. (2021) Consolidation and fragmentation in environmental scanning: A review and research agenda. *Long Range Planning*, Vol. 54 (3), 1–2.
- Rouach, D. – Santi, P. (Oct 2001) Competitive intelligence adds value: Five intelligence attitudes. *European Management Journal*, Vol. 19 (5), 552–559.
- Russel, S. – Norvig, P. (2016) *Artificial Intelligence: A Modern Approach Fourth Edition*. Prentice Hall Series in Artificial Intelligence. Prentice Hall.
- Saggi, M. K. – Jain, S. (2018) A survey towards an integration of big data analytics to big insights for value-creation. *Information Processing & Management*, Vol. 54 (5), 758–790.
- Saldana, J. (2011) *Fundamentals of qualitative research*. Oxford University Press, Incorporated.
- Sarker, I. H. (2021) Machine learning: Algorithms, real-world applications and research directions. *SN Computer Science*, Vol. 2 (3), 160–160.
- Sarker, I.H. – Kayes, A.S.M. – Badsha, S. – Alqahtani, H. – Watters, P. – Ng, A. (2020) Cybersecurity data science: An overview from machine learning perspective. *Journal of Big Data*, Vol. 7 (1), 12–14.
- Saunders, M. – Lewis, P. – Thornhill, A. (2007) *Research methods for business students*. 4. p. Financial Times Prentice Hall, Edinburgh Gate, Harlow.
- Saunders, M. N. – Lewis, P. – Thornhill, A. (2019) *Research methods for business students*. 8. p. Pearson Education Limited, Harlow.
- Schmidhuber, J. (2015) Deep learning in neural networks: An overview. *Neural Networks*, Vol. 61, 85–117.

- Schroeck, M.– Shockley, R.– Smart, J.– Romero-Morales, D.– Tufano, P. (2012) Analytics: The real-world use of big data: How innovative enterprises extract value from uncertain data. IBM Institute for Business Value.
- Srinivas, P. V. V. S. – Gayathri, K. – Bhavitha, K. – Sarath, K. D. (2023) BLIP-NLP model for sentiment analysis. *Teoksessa: 2023 2nd International Conference on Edge Computing and Applications (ICECAA)*, 468–475. IEEE.
- Stylos, J. – Myers, B. – Hosking, J. – Cox, P. (2007) Mapping the space of API design decisions. *IEEE Symposium on Visual Languages and Human Centric*, 50–60.
- Sun, M. – Sun, L. (2024) Optimization of artificial intelligence in localized big data real-time query processing task scheduling algorithm. *Frontiers in Physics*, Vol. 12.
- Sun, Z. – Sun, L. – Strang, K. (2019) Big data analytics services for enhancing business intelligence. *The Journal of Computer Information Systems*, Vol. 58 (2), 162–169.
- Suominen, A. – Toivanen, H. – Seppänen, M. (2017) Firms' knowledge profiles: Mapping patent data with unsupervised learning. *Technological Forecasting & Social Change*, Vol. 115, 131–142.
- Suthaharan, S. (2014). Big data classification: Problems and challenges in network intrusion prediction with machine learning. *Performance Evaluation Review*, Vol. 41 (4), 70–73.
- Taherdoost, H. – Madanchian, M. (2023) Artificial intelligence and sentiment analysis: A review in competitive research. *Computers (Basel)*, Vol. 12 (2), 37.
- Tan, F. – Zhang, Q. – Mehrotra, A. – Attri, R. – Tiwari, H. (2024) Unlocking venture growth: Synergizing big data analytics, artificial intelligence, new product development practices, and inter-organizational digital capability. *Technological Forecasting & Social Change*, Vol. 200, 123174.
- Tong-On – Panithan – Siripipatthanakul – Supaprawat – Phayaphrom – Bordin (2021) The implementation of business intelligence using data analytics and its effects towards on performance in the hotel industry in Thailand. *International Journal of Behavioral Analytics*, Vol. 1 (2), No. 9, 1–16.
- Tracy, S. J. (2010) Qualitative quality: Eight “big-tent” criteria for excellent qualitative research. *Qualitative Inquiry*, Vol. 16 (10), 837–851.

- Trunk, A.– Birkel, H. – Hartmann, E. (2020) On the current state of combining human and artificial intelligence for strategic organizational decision making. *Business Research (Göttingen)*, Vol. 13 (3), 875–919.
- Tuomi, J. – Sarajärvi, A. (2018) *Laadullinen tutkimus ja sisällönanalyysi*. Kustannusosakeyhtiö Tammi, Helsinki.
- Wang, P. (2019) On defining artificial intelligence. *Journal of Artificial General Intelligence*, Vol. 10 (2), 1–37.
- Wang, Y. – Zhang, B. – Ma, J. – Jin, Q. (2023) Artificial intelligence of things (AIoT) data acquisition based on graph neural networks: A systematical review. *Concurrency Computat Pract Exper*, Vol. 35 (23), 1–24.
- Wilson, L. (2019) *Data-driven marketing content: A practical guide*. Emerald Publishing Limited, Bingley.
- Witten, I. H. –Frank, E. – Hall, M. A. (2011) *Data mining: Practical machine learning tools and techniques*. 3. p. Elsevier/Morgan Kaufmann.
- Wittmann, C. M. (2024) Resource-advantage theory, market segmentation and competitor analysis. *Journal of Marketing Management*, Vol. 40 (13–14), 1286–1299.
- Wright, E. W. – Hillon, Y. C. – Garrido-Lopez, M. – Fowler, D. (2019) A new scorecard for strategic planning. *Journal of Business Strategy*, Vol. 40 (2), 50–58.
- Wu, C. – Buyya, R. – Ramamohanarao, K. (2016) Big Data analytics=machine learning+ cloud computing, arXiv preprint arXiv:1601.03115.
- Yaneva, D. (2020) Importance of the marketing environment analysis in the process of strategic marketing decision making. *Economics & Law*, Vol. 2(2), 24–32.
- Yazti, D. Z. – Krishnaswamy, S. (2014) Mobile big data analytics: Research, practice, and opportunities. Teoksessa: *Proceedings of IEEE 15th international conference on mobile data management (MDM)*, 1–2.
- Yin, S. – Li, H. –Sun, Y. – Teng, L. (2024) Data visualization analysis based on explainable artificial intelligence: A survey. *IJLAI Transactions on Science and Engineering*, Vol. 2 (2), 13–20.
- Young, T. –Hazarika, D. –Poria, S. – Cambria, E. (2018) Recent trends in deep learning based natural language processing. *IEEE Computational Intelligence Magazine*, vol. 13 (3), 55–75.

- Zaripova, R. – Kosulin, V. – Shkinderov, M. – Rakhmatullin, I. – Pushkarev, M. (2023) Unlocking the potential of artificial intelligence for big data analytics. *E3S Web of Conferences*, Vol. 460, 04011.
- Zong, Z. – Guan, Y. (2025) AI-driven intelligent data analytics and predictive analysis in Industry 4.0: Transforming knowledge, innovation, and efficiency. *Journal of the Knowledge Economy*, Vol. 16, 864-903.
- Özdemir, V. – Hekim, N. (2018) Birth of industry 5.0: Making sense of big data with artificial intelligence, “the internet of things” and next-generation technology policy. *Omics*, Vol. 22 (1), 65–76.

Liitteet

Liite 1. Haastattelukysymykset

Tiedon tarpeiden määrittely

Teema: Tekoälyn ja big data -analytiikan käyttämisen merkitys tiedon tarpeiden määrittelyyn kilpailija-analyysissä.

1. Mitä hyötyjä/mahdollisuuksia tekoälypohjainen Big Data -analytiikka on tuonut kilpailija-analyysin alkuvaiheeseen, eli tiedon tarpeiden määrittelyyn? (Eli siihen vaiheeseen, kun pohditaan, mitä tietoa tarvitsemme tehdäksemme kilpailija-analyysin)
 - Esimerkit: Onko se auttanut tunnistamaan uusia kilpailijoita, kenestä pitäisi dataa kerätä? Auttaako se näyttämään suuntaa, mitä tietoa kannattaa kerätä, jotta yrityksen liiketoimintatavoitteet saavutettaisiin?
2. Kuinka tekoäly ja big data ovat **tehostaneet** tätä prosessin vaihetta, jossa yrityksenne määrittelee, mitä tietoa kilpailija-analyysissä tarvitaan? (nopeus, kustannustehokkuus etc.)

Tiedon kerääminen

Teema: Big data -analytiikan ja tekoälyn merkitys tiedonkeruuprosessiin.

Kysymykset:

3. Mitä uusia mahdollisuuksia tekoäly & big data – analytiikka on tuonut kilpailijatiedon keräämiseen?
 - Esimerkiksi uusia tiedon lähteitä, määrällisesti enemmän tietoa kuin ennen jne?
4. Onko tekoäly & big data -analytiikka tehostanut kilpailijatiedon keräämistä?
 - Onko tiedonkeruun prosessi esimerkiksi nopeutunut, vaatiiko se vähemmän resursseja, onko se laadultaan tarkempaa ja oikeellisempaa?
 - Onko siitä tullut reaaliaikaista?

Oleellisen tiedon seulominen

5. Millaisia uusia mahdollisuuksia tekoäly ja Big Data -analytiikka ovat tuoneet kilpailijatiedon seulontaan? *Eli oleellisen tiedon löytämiseen datasta*
 - *kykyä havaita poikkeamia, epätyypillisiä trendejä tai signaaleja, joita manuaalinen analyysi ei ehkä havaitsisi?*
6. Miten tekoäly ja Big Data -analytiikka ovat **tehostaneet** kilpailijatiedon seulontaa, esimerkiksi nopeuden, tarkkuuden tai olennaisen tiedon löytämisen näkökulmasta?

Tiedon analysointi

7. Miten tekoälypohjainen analytiikka on muuttanut kilpailijatiedon analysointiprosessia vs perinteinen analyysiprosessi?
 - Esimerkiksi benchmarkingissa
8. Mitä uusia mahdollisuuksia tai oivalluksia tekoäly on tuonut kilpailijoiden analysointiin?
9. Millaisia uusia mahdollisuuksia tekoäly on tuonut kilpailijoiden vahvuuksien ja heikkouksien tunnistamiseen analyysivaiheessa?
10. Onko tekoäly mahdollistanut kilpailijoiden strategisten muutosten ennakkoinnin? Jos kyllä, niin miten?
11. Kuinka tekoäly on auttanut tunnistamaan piilotettuja trendejä tai kaavoja kilpailijoiden markkinakäyttäytymisessä?
12. Miten tekoäly ja big data ovat vaikuttaneet kilpailijoiden tulevaisuuden suunnitelmien hahmottamiseen? Onko se mahdollistanut jotain uutta?
13. Onko tekoäly ja big data -analytiikka **tehostanut** tiedon analysoinnin prosessia?

Tulkinta ja johtopäätökset.

14. Millaisia uusia mahdollisuuksia tekoäly on tuottanut kilpailija-analyysin johtopäätös- ja tulkintavaiheessa?
 - Onko se esimerkiksi auttanut tiedon visualisoinnissa, jotta sitä on helpompi tulkita, tai auttanut ymmärtämään analyyseja muuten?
 - Onko se auttanut tekemään johtopäätöksiä tai antanut toimintakehotuksia, ja jos kyllä, niin millaisia?
15. Onko tekoäly ja big data -analytiikka **tehostanut** tulkintaa ja johtopäätösten tekemistä?

Liite 2. Esitietomateriaali

Pro gradu -tutkimuksen aiheet ja haastattelussa käsiteltävät teemat

Tutkimuksen aihe: Millaisia uusia mahdollisuuksia edistynyt tekoälytehosteinen big data -analytiikka tuo kilpailija-analyysin prosessin eri vaiheisiin?

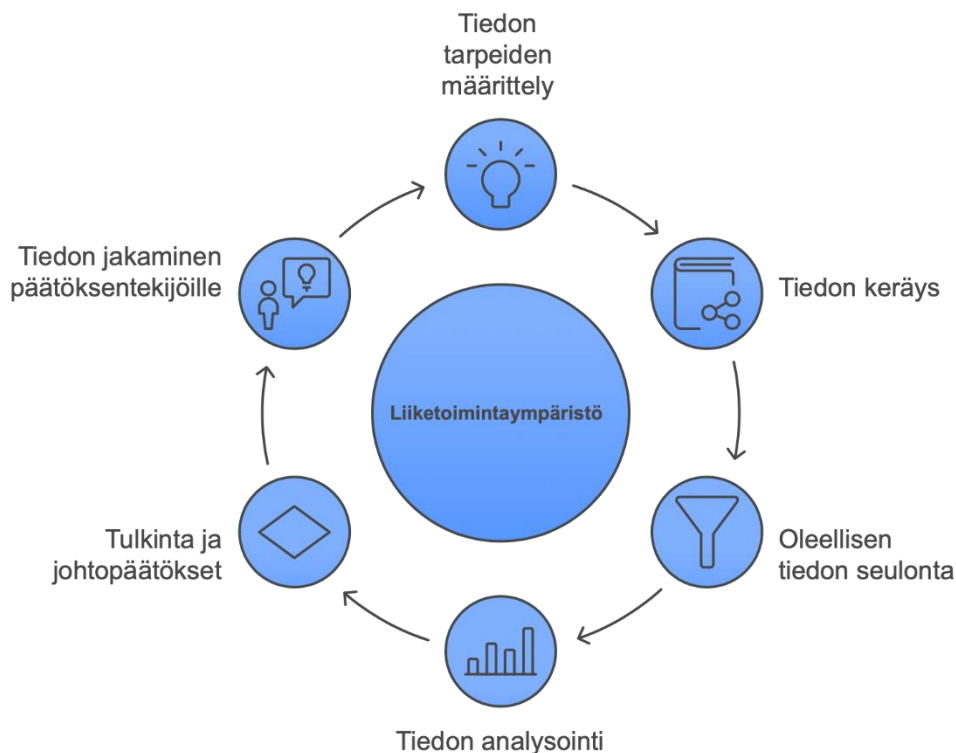
Haastattelun osa-alueet avattuna:

Tekoälytehosteinen Big Data -analytiikka

- Big Data -analytiikan tavoitteena on luoda suurista datamääristä uusia oivalluksia, jotka auttavat yrityksiä ymmärtämään markkinoita, kilpailijoita ja asiakkaita. Tämä prosessi on nykyään useimmiten tekoälypohjaista, mikä tarkoittaa, että big data -analytiikan prosessin eri vaiheissa hyödynnetään tekoälyalgoritmeja ja koneoppimismenetelmiä. Tekoäly on vienyt big data-analytiikkaa eteenpäin, sillä tekoäly tehostaa datan analysointia ja käsittelyä.

Kilpailija-analyysi

- Tutkimuksessa käsiteltävä kilpailija-analyysi tarkoittaa tiedon keräämistä ja analysointia kilpailijoista sekä niiden toimista, jotta yritys voi kehittää omia kilpailustrategioitaan ja saavuttaa kilpailuetua markkinoilla.
- Tutkimuksessani hyödynnetään Pirttilän (2000) kilpailija-analyysin kuusivaiheista prosessimallia (ks. kuva alla), keskittyen viiteen vaiheeseen eli tiedon tarpeiden määrittely, tiedon keräys, oleellisen tiedon seulonta, tiedon analysointi ja tulkinta ja johtopäätökset. Tiedon jakaminen ei sisälly tutkimukseni keskiöön.
- **Eli tekoälytehosteisen big data -analytiikan hyötyjä tarkastellaan seuraavissa prosessin eri vaiheissa vaihe kerrallaan:**



Tutkimus on teemahaastattelu, ja sen teemoja ovat:

1. Tekoälytehosteisen big data -analytiikan hyödyt ja mahdollisuudet yrityksen alustavassa tiedon tarpeen määrittelyssä, kilpailija-analyysiä tehtäessä

2. Tekoälytehosteisen big data -analytiikan hyödyt ja mahdollisuudet kilpailijatiedon keräämisessä
3. Tekoälytehosteisen big data -analytiikan hyödyt ja mahdollisuudet oleellisen kilpailijatiedon seulomisessa
4. Tekoälytehosteisen big data -analytiikan hyödyt ja mahdollisuudet kilpailijatiedon analysoimisessa
5. Tekoälytehosteisen big data -analytiikan hyödyt ja mahdollisuudet kilpailija-analyysin johtopäätösten ja tulkintojen tekemisessä

Tämän haastattelun tarkoituksena on kerätä asiantuntijanäkemyksiä tekoälytehosteisen big data -analytiikan hyödyntämisestä kilpailija-analyysissä. Kerättyjä vastauksia käytetään pro gradu -tutkielmani aineistona, ja ne käsitellään täysin luottamuksellisesti tutkimuksen eettisten periaatteiden mukaisesti. Kiitos ajastasi ja arvokkaasta panoksestasi tutkimukseeni!

Liite 3. Tietosuojalomake



Tietosuojailmoitus

1. Rekisterin nimi	Asiantuntijoiden näkemyksiä tekoälytehosteisen big data -analytiikan tuomista mahdollisuuksista kilpailija-analysissä
2. Rekisterinpitäjä	
3. Vastuuhenkilön yhteystiedot	
4.. Henkilötietojen käsittelyn tarkoitukset ja käsittelyn oikeusperuste	<p>Tutkimuksessa kerätään haastatteluita, joissa kysytään asiantuntijoiden kokemuksia ja käsityksiä tekoälytehosteisen big data -analytiikan tuomista mahdollisuuksista kilpailija-analysistä tehtäessä. Sähköpostiosoitteita käytetään haastattelukutsun lähettämiseen. Haastatteluissa kerätään tietoa asiantuntijoiden kokemuksista ja käsityksistä esimerkiksi kilpailija-analyysin prosessin eri vaiheista, ja miten kussakin vaiheessa tekoälypohjaisesta big data -analytiikasta on hyötyä.</p> <p>Henkilötietojen EU:n yleisen tietosuojasetuksen 6 artiklan mukaisena käsittelyperusteena on</p> <p><input checked="" type="checkbox"/> käsittely on tarpeen tieteellistä tutkimusta varten (yleinen etu 6 art. 1 a-kohta)</p> <p><input checked="" type="checkbox"/> rekisteröity on antanut suostumuksensa henkilötietojen käsittelyyn (suostumus 6 art. 1 e-kohta)</p> <p><input type="checkbox"/> muu mikä _____</p>
5. Käsiteltävät henkilötietoryhmät	<p>Rekisteriin talletetaan rekisteröidystä seuraavia tietoja: asema organisaatiossa ja toimiala.</p> <p>Yhteydenottoa varten kerätään sähköpostiosoitteet, mutta niitä ei tallenneta tutkimuksen aineistoon eikä käytetä muihin tarkoituksiin.</p>
6. Henkilötietojen vastaanottajat ja vastaanottajaryhmät.	Tietoja ei siirretä eikä luovuteta ulkopuolisille.
7. Tiedot tietojen siirrosta kolmansiin maihin	Henkilötietoja ei luovuteta EU:n tai Euroopan talousalueen ulkopuolelle.

8. Henkilötietojen säilyttämisaika tai sen määrittämisen kriteerit	Haastattelunauhoitteista kirjoitetaan tekstitiedostot ja nauhoitteet tuhotaan. Samalla tutkimusaineistosta poistetaan suorat tunnistetiedot. Tietoja säilytetään enintään 1.6.2030 asti, jonka jälkeen aineisto hävitetään tietoturvallisesti.
9. Rekisteröidyn oikeudet	Rekisteröidyllä on oikeus pyytää pääsy häntä itseään koskeviin henkilötietoihin sekä oikeus pyytää tietojensa oikaisemista tai poistamista taikka käsittelyn rajoittamista tai vastustaa niiden käsittelyä. Oikeutta henkilötietojen poistamiseen ei sovelleta tieteellisessä tai historiallisessa tutkimustarkoituksessa silloin, kun poisto-oikeus todennäköisesti estää käsittelyn tai vaikeuttaa sitä suuresti. Poisto-oikeuden toteuttamista arvioidaan tapauskohtaisesti. Rekisteröidyllä on oikeus tehdä valitus valvontaviranomaiselle.
10. Tiedot siitä, mistä henkilötiedot on saatu	Haastattelukutsujen lähettämiseksi pyydetään sähköpostiosoitteita. Muut tiedot kerätään suoraan haastattelututkimukseen osallisuilta.
11. Tiedot automaattisen päätöksenteon ml. profiloinnin olemas-saolosta	Tietoja ei käytetä automaattiseen päätöksentekoon tai profiloinnin tekemiseen.

Liite 4. Suostumuslomake



PRO GRADU-TUTKIMUS

SUOSTUMUS TUTKIMUKSEEN

Tämä on kutsu osallistua tutkimushaastatteluun, minkä vuoksi toivomme teidän lukevan alla olevan hankekuvauksen ja antavan suostumuksenne haastatteluihin.

Tutkimuksen nimi ja toimijat: Tekoälytehosteisen big data -analytiikan mahdollisuudet kilpailija-analyysissä

Taustaa tutkimukselle: Tutkielman tarkoituksena on selvittää, mitä mahdollisuuksia tekoälytehosteinen big data -analytiikka tuo kilpailija-analyysin prosessin eri vaiheisiin.

Pyyntö osallistua tutkimukseen: Jotta voimme selvittää, onko tekoälytehosteisesta big data -analytiikasta hyötyä kilpailija-analyysissä, tarvitsemme asiantuntijoiden näkökulmia aiheesta. Tätä selvitämme haastattelemalla ihmisiä liittyen mahdollisuuksiin, joita tekoälytehosteinen big data -analytiikka tuo kilpailija-analyysin prosessin eri vaiheissa. Osallistuminen on täysin vapaaehtoista, ja siitä on mahdollisuus vetäytyä missä vaiheessa haastattelua tahansa. Tutkimuksen tuloksia hyödynnetään Pro-gradu tutkielmassa. Lisäksi tutkimustuloksia voidaan hyödyntää akateemisissa julkaisuissa.

Aineiston hankintamenetelmät ja sen hallinta: Haastattelemme yrityksiä työntekijöitä heidän näkemyksistään tekoälytehosteisen big data -analytiikan tuomista mahdollisuuksista kilpailija-analyysissä. Kahdenkeskiset keskustelumme ovat pituudeltaan noin 30-60 minuuttia.

Haastattelut äänitetään suostumuksellanne. Aineiston käsittelyssä noudatamme Suomen Akatemian eettisiä sääntöjä ja aineisto on luottamuksellista.

Haastatteluaineiston litteroinnissa käytämme koodeja, joten yksittäiset henkilöt eivät ole tunnistettavissa eikä aineisto ole yksilöitävissä.

Litteroinnin jälkeen äänitykset hävitetään. Litteroinnin koodiavain on vain graduntekijällä. Haastatteluaineisto tallennetaan vain graduntekijän käytössä olevalle suljetulle Turun yliopiston verkkolevyllä. Tutkielman valmistuttua ei-yksilöitävissä oleva, litteroitu haastatteluaineisto, joka säilytetään tietoturvalisesti yliopiston suosituksen mukaisesti 5 vuotta, jonka jälkeen se hävitetään.

Tutkimuksen henkilötietojen käsittely perustuu yleiseen etuun, jolloin tutkittavilla ei ole niiden poisto-oikeutta.

Lisätietoja hankkeesta: Gradun tekijä: Salla Nurmi salla.matilda.nurmi@gmail.com

Gradun ohjaaja: KTT Helena Rusanen helena.rusanen@utu.fi

Lisätietoja Turun yliopiston tietosuojasta: tietosuoja@utu.fi.

Minua on pyydetty osallistumaan tutkimukseen, jonka tarkoituksena on selvittää, onko tekoälytehosteisesta big data -analytiikasta hyötyä kilpailija-analyysissä. Olen perehtynyt edellä olevaan selvitykseen ja saanut riittävästi tietoa tutkimuksesta ja sen yhteydessä suoritettavasta tietojen keräämisestä, käsittelystä ja tallentamisesta. Tutkimuksen sisältö on kerrottu minulle ja olen saanut riittävän vastauksen kaikkiin

tutkimusta koskeviin kysymyksiini. Selvityksen antoi Salla Nurmi. Minulla on ollut riittävästi aikaa harkita tutkimukseen osallistumista.

Ymmärrän, että tähän haastatteluun osallistuminen on vapaaehtoista. Minulla on oikeus milloin tahansa ja syytä ilmoittamatta keskeyttää haastatteluun osallistuminen.

Allekirjoituksellani vahvistan, että osallistun tässä asiakirjassa kuvattuun tutkimukseen ja suostun haastateltavaksi.

Allekirjoitus

Päiväys

Nimen selvennys

Suostumus vastaanotettu

Suostumuksen vastaanottajan allekirjoitus

Päiväys

Nimen selvennys

Asiakirjasta on tehty kaksi identtistä kappaletta, toinen opinnäytetyön tekijän arkistoon ja toinen haastateltavalle.