

Data-analytiikka rahanpesun estämisessä

Vilma Saarinen

Tekijä(t) Vilma Saarinen	
Koulutusohjelma Liiketalouden koulutusohjelma	
Raportin/Opinnäytetyön nimi Data-analytiikka rahanpesun estämisessä	Sivu- ja liitesivumäärä 54 + 1
Rahanpesu on nykypäivänä merkittävä rikosoikeudellinen ongelma, jonka estämiseen etenkin pankkien tulee kiinnittää huomiota. Pankeissa hyödynnetään yhä enenevässä määrin data-analytiikkaa rahanpesutapausten havaitsemisessa ja estämisessä. Opinnäytetyössä perehdytään data-analytiikan hyödyntämiseen ja kannattavuuteen rahanpesun estämisessä Suomessa toimivien pankkien näkökulmasta. Työn tutkimustavoitteena oli selvittää, miten pankit voisivat hyödyntää data-analytiikkaa tehokkaasti ja kannattavasti rahanpesun estämisessä. Opinnäytetyön tietoperusta tarkastelee dataa, data-analytiikkaa, sen menetelmiä ja prosessia, rahanpesun estämistä sekä lisäarvoja ja haasteita, joita data-analytiikan hyödyntäminen rahanpesun estämisessä tuo pankeille. Tietoperustan aineistona on käytetty aiheeseen liittyvää kirjallisuutta, lakeja, artikkeleita sekä erilaisia tutkimuksia ja raportteja. Tietoperustaa seuraa tutkimusosio, jossa esitellään tutkimuksen toteutus ja tulokset. Tutkimus toteutettiin kvalitatiivisena eli laadullisena tutkimuksena syksyn ja talven 2020 aikana. Tutkimusaineisto kerättiin puolistrukturoidun haastattelun avulla neljältä eri Suomessa toimivalta pankilta. Haastatteluista kolme toteutettiin Microsoft Teams -alustan välityksellä ja yksi sähköpostitse. Tutkimusaineiston avulla pyrittiin luomaan kokonaisvaltainen kuva tutkittavasta ilmiöstä. Tutkimuksesta selvisi, että pankit ovat hyvin eri vaiheissa data-analytiikan hyödyntämisessä. Hyödyntämisen skaala vaihteli pankkien välillä suunnitelmista lisätä ja kehittää järjestelmiä aina jo olemassa oleviin hyvin nykyaikaisiin menetelmiin. Tutkimuksen perusteella rahanpesua ehkäistään pankeissa eri vaiheissa eri tavalla, aina asiakkaan tuntemisesta lähtien data-analytiikkaa apuna käyttäen. Tutkimuksen perusteella data-analytiikan hyödyntämiselle on pankeissa suuri tarve ja se koetaan erittäin kannattavana, vaikka kehitettävääkin olisi. Data-analytiikan hyödyntäminen on tutkimuksen mukaan pankeille suuri investointi, mutta sen tuomat hyödyt ovat myös merkittävät. Tutkimustulosten perusteella dataa kertyy pankeille valtavia määriä koosta riippumatta. Esille nousi kuitenkin, että pankeissa olisi tarvetta etenkin verrannolliselle datalle, jotta data-analytiikkaa voitaisiin hyödyntää tehokkaammin. Tutkimuksen tulosten perusteella suurin pankkeja kiinnostava asia, joka tehostaisi data-analytiikan hyödyntämistä ja tekisi siitä siten kannattavampaa, olisi pankkien välinen yhteistyö.	
Asiasanat Rahanpesu, data, analytiikka, pankki	

Sisällys

1	Johdanto	1
2	Data ja data-analytiikka	4
2.1	Data	4
2.1.1	Pankki- ja maksutilien valvontajärjestelmä	6
2.1.2	Tietosuoja-asetus.....	6
2.2	Data-analytiikka	7
2.3	Data-analytiikka finanssialalla	7
2.4	Datan analysointiprosessi	8
2.5	Big data.....	10
2.6	Koneoppiminen	11
2.6.1	Koneoppimisalgoritmien luokittelu	12
2.6.2	Koneoppimisen malleja	13
3	Rahanpesun estäminen	15
3.1	Rahanpesuprosessi	16
3.2	Säätely.....	16
3.3	Rahanpesun estämisen prosessi	17
3.4	Rahanpesun estämisen haasteita	19
4	Data-analytiikan hyödyntäminen rahanpesun estämisessä	20
4.1	Asiakasdatan jakaminen pankkien kesken	21
4.2	Data-analytiikan tuoma lisäarvo	21
4.3	Data-analytiikan mukana tulevat haasteet.....	22
5	Data-analytiikan hyödyntäminen rahanpesun estämisessä pankkien näkökulmasta ...	24
5.1	Tutkimuksen toteutus	24
5.2	Tutkimuksen tulokset	28
5.2.1	Rahanpesun estämisen menetelmät ja data-analytiikan hyödyntäminen..	29
5.2.2	Rahanpesun estämiseen liittyvä data	33
5.2.3	Data-analytiikan hyödyntämisen kannattavuus.....	35
5.2.4	Tulevaisuuden mahdollisuudet.....	37
6	Pohdinta.....	40
6.1	Tulosten tarkastelu.....	40
6.2	Johtopäätökset sekä kehittämis- ja jatkotutkimusehdotukset.....	44
6.3	Tutkimuksen luotettavuus ja eettiset näkökohdat	45
6.4	Opinnäytetyöprosessin ja oman oppimisen arviointi.....	47
	Lähteet	50
	Liitteet.....	55
	Liite 1. Haastattelurunko.....	55

1 Johdanto

Rahanpesu on nykypäivänä suuri rikosoikeudellinen ongelma, jonka estämiseen etenkin pankkien tulee kiinnittää huomiota. Rahanpesun estämisen rooli on kasvanut viime aikoina huomattavasti. Rahan liikkuminen maasta toiseen on helpompaa ja nopeampaa kuin koskaan ennen teknologian kehityksestä ja globalisaatiosta johtuen. Vanhanaikaiset menetelmät ja manuaaliset prosessit eivät pysy yhä tiukentuvien vaatimusten ja datan määrän kasvun perässä, joten parannukset data-analytiikan hyödyntämisessä ovat tarpeen. (Chen ym. 2018, 245; EY 2019.)

Rahanpesun estäminen on osa talousrikollisuuden torjuntaa, joka pitää lisäksi sisällään esimerkiksi terrorismin rahoittamisen, veronkierron ja korruption estämistä. Rahanpesun estämiseksi laissa määritellyn ilmoitusvelvollisen on tehtävä ilmoitus rahanpesun selvittelykeskukselle havaitessaan epäilyttäviä liiketoimia. Ilmoitusvelvollisia ovat esimerkiksi pankit, vakuutusyhtiöt, kiinteistönvälitysliikkeet sekä veroneuvontaa tarjoavat liike- ja ammattiyrittäjät. Opinnäytetyö rajataan koskemaan Suomessa toimivia pankkeja, koska heillä on merkittävä rooli rahanpesun ja terrorismin rahoittamisen estämisessä. (Laki rahanpesun ja terrorismin rahoittamisen estämisestä 28.6.2017/444.)

Data-analytiikka on puolestaan osa laajempaa analysointia tukevaa datatieteen alaa, joka käsittää lisäksi esimerkiksi algoritmit, tilastotieteen, tiedonhallinnan ja sovelluskehityksen (Ambientia s.a.; Hyvärinen 28.4.2020). Opinnäytetyössä tutkitaan data-analytiikan keskeisiä aihealueita, sen sijaan, että syvennyttäisiin vain yhteen aihealueeseen. Data-analytiikan järjestelmät ovat yrityksille erittäin mittava investointi, ja siitä aiheutuu monenlaisia kustannuksia aina hankkimisesta jatkuvaan järjestelmien päivittämiseen ja henkilöstön koulutukseen. Sen avulla yritykset voivat kuitenkin esimerkiksi tehostaa toimintaansa saaden kilpailuetua kilpailijoihin nähden. Data-analytiikan hyödyntämisestä on paljon apua myös pankeille, jotka pyrkivät vaativan sääntelyn alla havaitsemaan rahanpesutapauksia ja estämään niitä, toimiakseen yhteiskunnallisesti vastuullisena toimijana ja välttääkseen suuria sakkoja. Tästä syystä opinnäytetyössä on haluttu lähteä tutkimaan aihetta kannattavuuden näkökulmasta. (Rouse 2020.)

Opinnäytetyölle valitut tavoitteet tukevat työlle valittua rajausta ja näkökulmaa sekä ovat yhteiskunnallisesti tärkeitä. Alalla toimijat saavat työstä tärkeää tietoa toimiakseen tehokkaasti ja kannattavasti tällä tarkkaan säännellyllä alalla. Opinnäytetyön päätavoitteena on selvittää, miten pankit voisivat hyödyntää data-analytiikkaa tehokkaasti ja kannattavasti rahanpesun estämisessä. Tähän saadaan vastaus alatavoitteiden kautta, jotka ovat seuraavat:

- Miten pankit hyödyntävät data-analytiikkaa rahanpesun estämiseksi?
- Miten kannattavana data-analytiikan hyödyntäminen rahanpesun estämisessä koetaan pankeissa?

Taulukossa 1 on esitetty opinnäytetyön alatavoitteiden yhteys teoreettiseen viitekehykseen, haastattelun kysymyksiin ja tutkimustuloksiin peittomatriisin avulla.

Taulukko 1. Peittomatriisi

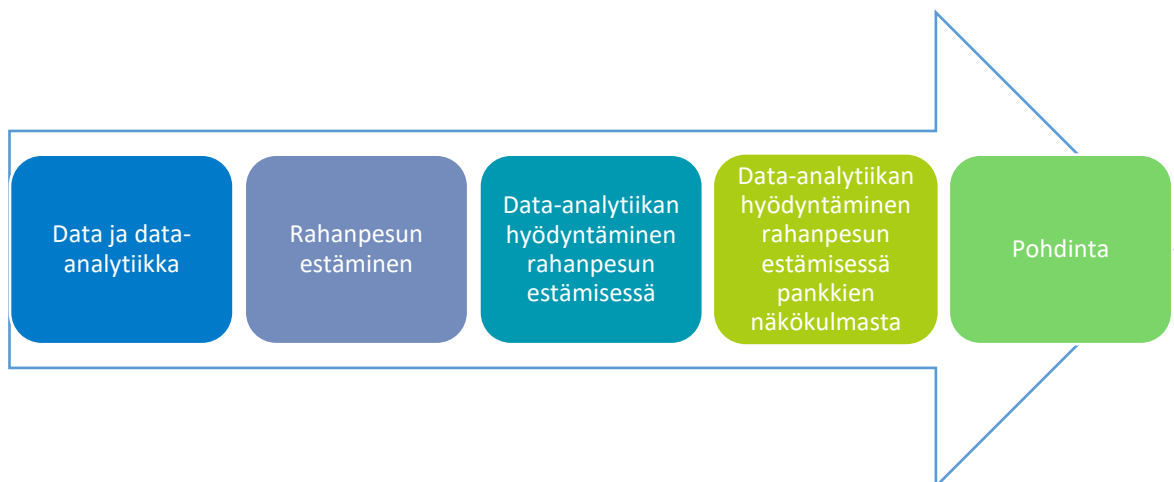
Opinnäytetyön alatavoitteet	Teoreettinen viitekehys	Haastattelun kysymykset	Tulokset
Alatavoite 1: Miten pankit hyödyntävät data-analytiikkaa rahanpesun estämiseksi?	2.1, 2.1.1, 2.1.2, 2.2, 2.3, 2.4, 2.5, 2.6, 2.6.1, 2.6.2, 3, 3.3, 3.4, 4, 4.1, 4.2 ja 4.3	3, 4, 5, 6 ja 7	5.2.1, 5.2.2 ja 6.1
Alatavoite 2: Miten kannattavana data-analytiikan hyödyntäminen rahanpesun estämisessä koetaan pankeissa?	2.1.1, 2.1.2, 2.2, 2.3, 2.4, 2.6, 2.6.1, 2.6.2, 3.4, 4, 4.1, 4.2 ja 4.3	8, 9, 10, 11 ja 12	5.2.3, 5.2.4 ja 6.1

Idea opinnäytetyöhön syntyi kiinnostuksesta data-analytiikkaan ja rahanpesun estämiseen liittyvään sääntelyyn. Aihe tukee ja syventää omia opintojani ja asiantuntemusta itseäni kiinnostavalla alalla sekä tukee työllistymismahdollisuksiani tällä alalla lisäten substanssiosaamista ja kertoen motivaatiostani alaa kohtaan. Aiheesta tutkimisen arvoisen tekee sen ajankohtaisuus ja mielenkiintoisuus sekä se, että sitä ei ole aikaisemmin tutkittu opinnäytetöissä. Aiheen ajankohtaisuudesta kertoo erityisesti parin viimeisen vuoden aikana esille nousseet uutiset koskien asiakkaan tuntemista ja pankkien saamia sakkoja puutteellisen rahanpesun estämisen takia. Pankit ovat myös viime aikoina yhä enenevässä määrin hyödyntäneet ennakoivaa ja reaaliaikaista analytiikkaa eri toiminnoissaan, kuten myös rahanpesun ja petosten valvonnassa ja selvittelyssä (Lehto 27.5.2019). Tällä opinnäytetyöllä kannustetaan pankkeja hyödyntämään data-analytiikkaa entistä tehokkaammin rahanpesun estämisessä sekä kehittämään rahanpesun estämistä kohti tulevaisuutta.

Opinnäytetyö toteutetaan kvalitatiivisena eli laadullisena tutkimuksena ilman ulkopuolista toimeksiantajaa. Tutkimusaineisto kerätään haastattelemalla neljää Suomessa toimivaa pankkia. Kolmea pankkia haastatellaan Microsoft Teams -alustan välityksellä ja yhtä sähköpostitse. Haastattelut toteutetaan loka- ja marraskuun aikana 2020. Haastattelut toteutetaan puolistrukturoituna haastatteluna. Tämä näkyy siten, että haastattelurungossa (liite

1) on valmiit kysymykset, jotka esitetään jokaiselle haastateltavalle. Kysymyksiin ei ole kuitenkaan olemassa valmiita vastausvaihtoehtoja, joten haastateltavat pystyvät antamaan tarkan kuvauksen omista näkökulmistaan.

Opinnäytetyö etenee Haaga-Helia ammattikorkeakoulun perinteisen opinnäytetyörakenteen mukaisesti. Opinnäytetyön kulkua on esitetty kuviossa 1. Työ etenee siten, että johdannon jälkeen siirrytään tietoperustaan, jossa tutustutaan ensin tarkemmin data-analytiikkaan ja sen prosessiin, dataan sekä big dataan ja koneoppimisen menetelmiin. Tämän jälkeen käydään läpi rahanpesua, sen estämistä ja siihen liittyvää sääntelyä. Viimeisimmässä tietoperustan luvussa keskitytään siihen, miten data-analytiikkaa on mahdollista hyödyntää rahanpesun estämisessä sekä tuodaan ilmi sen tuomia haasteita ja lisäarvoa. Tietoperustan jälkeen seuraa työn empiirinen osuus, jossa esitellään tutkimuksen toteutus ja tulokset. Lopuksi analysoidaan tutkimuksen tuloksia sekä pohditaan muun muassa tutkimuksen luotettavuutta, johtopäätöksiä sekä arvioidaan omaa oppimista. (Kuvio 1.)



Kuvio 1. Työn eteneminen

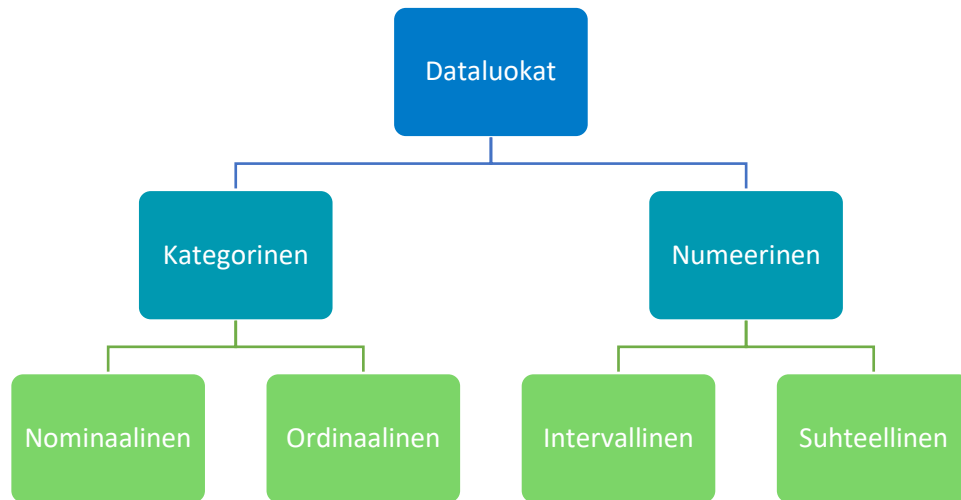
2 Data ja data-analytiikka

Päätöksenteosta on tullut yhä monimutkaisempi prosessi yritysmaailmassa vuosien varrella. Strukturoitujen ja strukturoimattomien datojen määrä ja moninaisuus kasvaa päivittäin valtavalla nopeudella. Monilla organisaatioilla on käytössään analyyttisiä työkaluja suurten tietomäärien käsittelyyn, mutta datan määrän kasvaessa, suurien tietomäärien hallitseminen vaikeutuu huomattavasti, eivätkä aiemmat työkalut enää riitä. Tässä luvussa tutustutaan tarkemmin data-analytiikan määritelmään ja sen prosessiin, dataan, big dataan ja koneoppimiseen sekä otetaan katsaus data-analytiikan hyödyntämiseen finanssialalla. (Chattopadhyay 2016.)

2.1 Data

Data on halutusta kohteesta kerättyä ja tallennettua aineistoa, kuten asiakkaiden nimiä ja tilitapahtumia, jota voidaan kutsua myös raakadatakseksi. Tällaisesta tietokantoihin kerätystä raakadatasta ei ole sellaisenaan kovin paljon hyötyä päätöksenteon näkökulmasta. Dataa on nykypäivänä valtavia määriä. Marr kertoo vuonna 2018 kirjoittamassaan artikkelissa, että sen hetkellä tahdilla dataa syntyi 2,5 kvintiilitavua päivässä ja silloin 90 % datasta oli syntynyt edellisten kahden vuoden aikana. Tällä hetkellä dataa syntyy vielä nopeammin ja vauhti kiihtyy entisestään. (Marr 2018; Martinsuo & Kärri 2017, 19.)

Data voidaan jakaa neljään luokkaan, joita ovat nominaalinen, ordinaalinen, intervallinen ja suhteellinen. Kuviossa 2 on havainnollistettu datan jakautumista luokkiin. Datan jakaminen luokkiin on tärkeää analytiikan ammattilaisille, sillä se mahdollistaa datan, jota ei voi helposti laskea, systemaattisen mittaamisen ja analysoinnin sekä ohjaa valitsemaan oikeat tekniikat datan käsittelyyn. Nominaalisia ja ordinaalisia dataa voidaan pitää kategorisina datoina, kun taas intervalliset ja suhteelliset datat voidaan ryhmitellä numeerisiksi datoina. (Gupta 2016, luku 2; kuvio 2.)



Kuvio 2. Datan luokittelu (mukaillen Gupta 2016)

Nominaaliasteikkoa käytetään muuttujiin, joissa eri kategoriat on määritelty selkeästi, esimerkiksi käteisen talletuksiin ja nostoihin. Näiden tulee olla toisiaan poissulkevia ja tyhjentäviä, esimerkiksi ihmisten luokittelu naisiin ja miehiin. Nominaaliasteikolla kategoriat eivät noudata suhteellista järjestystä. Miesten nimeäminen "1" ja naiset "2" ei tarkoita, että naiset olisivat kaksi kertaa jotain kuin miehet tai, että 1 olisi parempi kuin 2. Numerot ovat nominaaliasteikolla siis vain tunnistamista ja kategorisointia varten. (Gupta 2016, luku 2.)

Ordinaaliasteikkoa käytetään datan kategorioiden järjestämiseen, esimerkiksi asiakkaan riskisyyden mukaan. Ordinaaliasteikko mahdollistaa kategorioiden vertailun antamatta kuitenkaan mitään tietoa eron suuruudesta. Asiakas, jolla on tason 2 riskiluokitus, ei välttämättä tarkoita kaksi kertaa riskisempää kuin asiakas, jonka riskiluokitus on tasolla 4. (Gupta 2016, luku 2.)

Intervalliasteikkoa voidaan käyttää datan kategorioiden järjestämiseen kuten ordinaaliasteikkoakin, mutta intervalliasteikolla tiedetään tarkasti havaintojen erojen suuruudet. Esimerkiksi 20 °C on enemmän kuin 5 °C. Yhteen- ja jakolaskuja intervalliasteikolla ei kuitenkaan voi suorittaa. Toinen määrittelevä tekijä on, ettei intervalliasteikolla ole todellista nollaa. Vaikka lämpötila olisi nolla, ei se tarkoita, ettei lämpötilaa olisi. (Gupta 2016, luku 2.)

Suhdelukuasteikko on hyvin samankaltainen kuin intervalliasteikkokin, mutta suhdeluasteikolle on määritelty nollapiste. Asteikon nollapiste ei ole vain käytäntö, vaan määrään todellinen nolla. Ikä on esimerkiksi määre, jota voidaan tarkastella suhdeluasteikolla. Iälle on määritelty nollapiste, sillä ihminen ei voi olla alle 0-vuotias. Suhdelukuasteikolla on kaikkien muiden kolmen edellä mainitun asteikon ominaisuuksia, mikä tekee siitä kaikkein jalostetuimman. (Gupta 2016, luku 2.)

2.1.1 Pankki- ja maksutilien valvontajärjestelmä

Viranomaisten sähköistä tiedonsaantia pankki- ja maksutileistä syyskuusta 2020 alkaen edistää laki pankki- ja maksutilien valvontajärjestelmästä. Lain mukaan muun muassa rahanpesun selvittelykeskuksella sekä finanssivalvonnalla on oikeus käyttää pankki- ja maksutilien valvontajärjestelmää, mikäli se on välttämätöntä rahanpesun estämiseksi, paljastamiseksi ja selvittämiseksi. Laki mahdollistaa paremman yhteistyön viranomaisten kesken rahanpesun estämiseksi. Finanssiala ry:n mukaan sähköinen järjestelmä nopeuttaa tiedonsaantia ja mahdollistaa entistä laajemmat kyselyt. Tiedot tehdyistä kyselyistä on kirjattuna ylös kymmenen vuoden ajan, joka varmistaa oikeusturvan, tietoturvallisuuden sekä henkilötietojen suojan. (Finanssiala ry 2019; Laki pankki- ja maksutilien valvontajärjestelmästä 571/2019; Valtiovarainministeriö 2018.)

2.1.2 Tietosuoja-asetus

Yrityksille ja organisaatioille on asetettu tarkat vaatimukset koskien henkilötietojen keräämistä, säilytystä ja hallinnointia yleisessä tietosuoja-asetuksessa (GDPR). Yleistä tietosuoja-asetusta täsmennetään ja täydennetään tietosuojalaissa (5.12.2018/1050). EU:n yleinen tietosuoja-asetus ja tietosuojalaki on tehty luonnollisten henkilöiden suojelemiseksi. Henkilötiedoilla tarkoitetaan esimerkiksi nimeä, tuloja tai kulttuurista profiilia. Yritys ei saa käsitellä henkilökohtaisia tietoja, jotka koskevat esimerkiksi henkilön etnistä alkuperää, poliittisia mielipiteitä tai uskonnollista vakaumusta. Yleinen tietosuoja-asetus koskee yrityksiä, jotka käsittelevät henkilötietoja ja sijaitsevat EU:ssa sekä niitä yrityksiä, jotka sijaitsevat EU:n ulkopuolella, mutta käsittelevät henkilötietoja, jotka liittyvät tavaroiden tai palveluiden tarjoamiseen henkilöille EU:ssa, tai yritys seuraa yksilöiden käyttäytymistä EU:ssa. (Asetus luonnollisten henkilöiden suojelusta henkilötietojen käsittelyssä sekä näiden tietojen vapaasta liikkuvuudesta direktiivin 95/46/EY kumoamisesta 2016/679/EU.)

Reese tuo konsulttiyrityksen blogikirjoituksessa esille tietosuoja-asetuksen tuomia haasteita rahanpesun estämiseen. Rahanpesun estämisen -säännökset edellyttävät suuren määrän henkilötietojen keräämistä, käsittelyä ja analysointia, mutta toisaalta EU:n tietosuoja-asetus rajoittaa asiakkaasta kerättäviä tietoja ja niiden käyttöä. Tietosuoja-asetus edellyttää, että asiakkaasta kerättyjä tietoja käsitellään asianmukaisesti ja lainmukaisesti. Tietoja saa käsitellä vain tiettyä ja laillista tarkoitusta varten, ja vain kyseistä tarkoitusta varten tarvittavia tietoja saa käsitellä. Pankin on siis huolehdittava, että rahanpesulainsäädännön nojalla kerätyt tiedot eivät riko tietosuoja-asetuksen vaatimuksia. Luvussa 4.1 perehdytään asiakastietojen jakamiseen pankkien kesken, jossa on otettava tarkkaan huomioon tietosuoja-asetus. (Reese 24.5.2018; tietosuoja-asetus.)

2.2 Data-analytiikka

Data-analytiikka on prosessi, jossa tutkitaan aineistoja, data-analytiikkaan erikoistuneiden järjestelmien ja ohjelmistojen avulla, trendien ja halutun informaation löytämiseksi sekä johtopäätösten tueksi. Se on datan ja matematiikan yhdistämistä, jonka avulla saadaan vastauksia liiketoiminnan kysymyksiin, löydetään yhteyksiä, ennustetaan lopputuloksia sekä automatisoidaan päätöksentekoa. Tavoitteena data-analytiikan hyödyntämisessä on siis sovelletun matematiikan, tilastojen, ennustavan mallintamisen ja koneoppimistekniikoiden avulla löytää merkityksellisiä malleja datasta sekä löytää uutta tietoa ja täten parantaa liiketoiminnan suorituskykyä. (Rouse 2020; SAS s.a.)

Olennainen osa data-analytiikkaa ovat koneoppiminen ja tekoäly, sillä niiden avulla suuristakin datamäärästä on mahdollista tehdä nopeitakin analyyseja (Empirica s.a). Kaikki kolme ovat omia käsitteitään, mutta ovat kuitenkin jossain määrin päällekkäisiä ja osa-alueet liippaavat usein toinen toisiaan (Hyvärinen 28.4.2020).

Data-analytiikka voidaan jakaa kvantitatiiviseen ja kvalitatiiviseen data-analytiikkaan. Kvantitatiivisessa data-analytiikassa data esiintyy numeerisessa muodossa, jota pystytään vertailemaan ja mittaamaan tilastollisesti. Kvalitatiivinen lähestymistapa on puolestaan enemmän tulkitsevaa ja selittävää. Kvalitatiivisessa analyysissä keskitytään ymmärtämään ei-numeerisia tietoja, kuten ääntä, kuvaa, tekstiä ja videoita. (Rouse 2020.)

Yritykset voivat hyödyntää data-analytiikkaa esimerkiksi tehostamaan toimintaa, kehittääkseen asiakaspalvelutoimia tai reagoidakseen nopeasti markkinoilla nouseviin trendeihin saaden kilpailuetua kilpailijoihin nähden. Data-analytiikasta on hyötyä monenlaisille liiketoiminnoille. Esimerkiksi verkkokaupat ja markkinointipalvelujentarjoajat pystyvät tunnistamaan kävijät, jotka todennäköisemmin ostavat tietyn tuotteen tai palvelun navigoinnin ja sivujen katselumallien perusteella. Myös pankit pystyvät havaitsemaan rahanpesu ja henkilöllisyysvarkautapauksia sekä estämään niitä data-analytiikan avulla. (Rouse 2020.)

2.3 Data-analytiikka finanssialalla

Finanssiala on yksi varhaisimpia data-analytiikan käyttöönottajia. Capital One, Yhdysvaltojen suurimpiin pankkeihin lukeutuva teknologiakeskeinen pankki, saavutti markkinajohtajuuden luottokorttialalla käyttäessään ensimmäisenä data-analytiikkaa erottautuakseen kilpailijoistaan. (Gupta 2016, luku 1.)

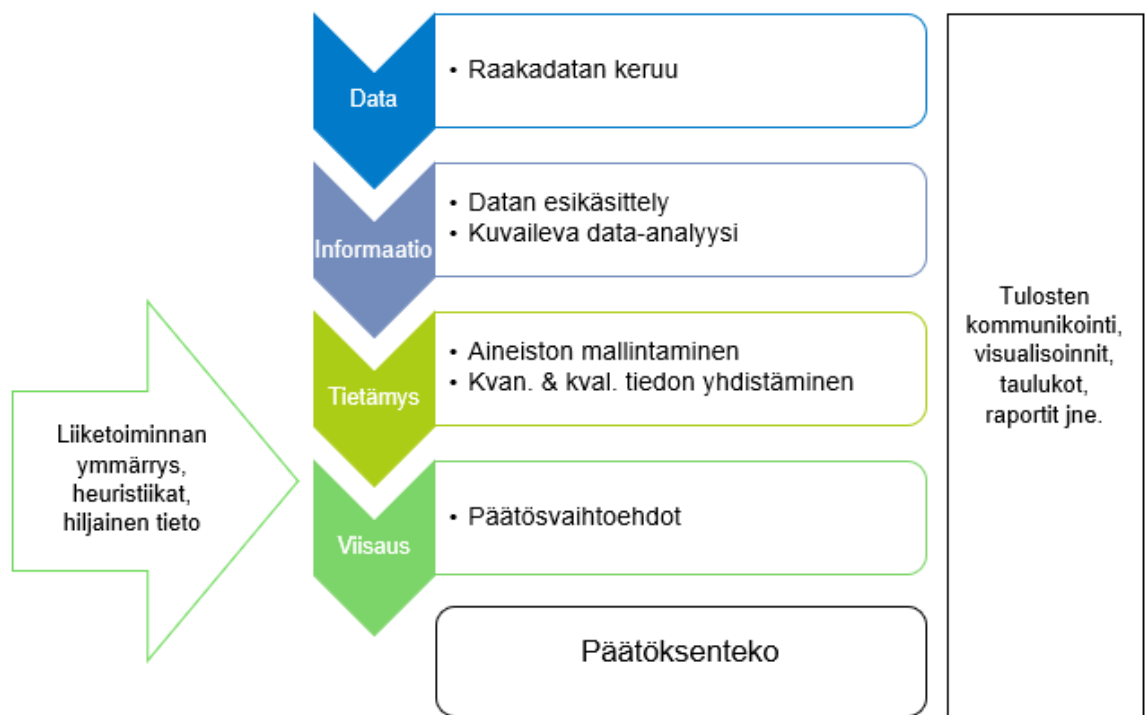
Harvard Business Review kirjoittaa tutkimuksesta, jossa haastateltiin yli 3 000 johtohenkilöä 14 eri sektorilta, 10 eri maasta ja selvitettiin, kuinka yritykset voisivat menestyä tekoälyn avulla. Tutkimuksessa nousee myös esille kaksi toimialaa, tele- ja finanssitoimiala, jotka ottavat johtoasemaa ja ovat ensimmäisten joukossa hyödyntämässä data-analytiikkaa. Tele- ja finanssitoimialalla investointeja tekoälyyn kasvatetaan yli 15 % vuodessa, joka on seitsemän prosenttiyksikköä enemmän kuin kaikilla toimialoilla keskimäärin seuraavien kolmen vuoden aikana. (Bughin, McCarthy, Chui 2017.)

McKinseyn tutkimuksessa havaittiin, että lähes jokainen pankki mainitsee edistyneen analytiikan viiden tärkeimmän prioriteettinsa joukossa. Saman tutkimuksen mukaan johtavassa asemassa olevat pankit ovat käyttäneet satoja miljoonia euroja dataan, erityisesti riskidataan sekä vaatimusten noudattamiseen. Tutkimus tuo ilmi, että edistyneen analytiikan hyödyntäminen on tuottanut johtavassa asemassa oleville pankeille 300 miljoonaa euroa. (Garg, Grande, Miranda, Sporleder & Windhagen 2017.)

Data-analytiikan hyödyntäminen tulee lisääntymään McAuleyn mukaan etenkin juuri rahanpesun, verkkopankkihuijausten ja vakuutuspetosten ehkäisemisessä ja selvittelyssä. Oppivan ja sääntöpohjaisen analytiikan avulla voidaan automatisoida rutiinin omaiset työtehtävät vakuutuskorvausten käsittelyssä, ja nostaa epäilyttäviä tapauksia tarkempaan selvitykseen. (Lehto 27.5.2019.)

2.4 Datan analysointiprosessi

Data-analytiikka prosessin vaiheet jaetaan ja määritellään kirjallisuudessa usealla eri tavalla. Kaikkien pohjalla on kuitenkin ajatus siitä, että datasta saataisiin tärkeää informaatiota päätöksenteon tueksi. SAS jakaa data-analytiikan kolmeen vaiheeseen – dataan, havaintoon ja käyttöönottoon (SAS s.a.). Martinsuo ja Kärri jakavat analysointiprosessin pienempiin osiin, jopa viiteen vaiheeseen – datan keräämiseen, käsittelyyn, mallintamiseen, vaihtoehtojen vertailuun ja päätöksentekoon (Martinsuo & Kärri 2017, 20). Taanila tuo puolestaan menetelmäblogissaan esille, että datasta saadaan vastauksia datan valmistelun, kuvailevan analytiikan, selittävän analytiikan ja ennakoivan analytiikan kautta (Taanila 26.8.2020). Kuviossa 3 on esitetty data-analytiikan prosessi Martinsuon ja Kärrien esittelemää prosessia mukaillen.



Kuvio 3. Datan analysointiprosessi (mukaillen Martinsuo & Kärri 2017)

Ensimmäisessä vaiheessa kerätään raakadataa, jota analysointiprosessissa lähdetään jalostamaan. Kuten kappaleessa 2.1 todettiin, data on halutusta kohteesta kerättyä aineistoa, joko numeerista tai ei-numeerista, joka ei sellaisenaan vielä tue päätöksentekoa. Analytiikkaa varten dataa on mahdollista poimia erilaisista lähteistä. Se, minkälaista dataa tarvitaan, riippuu siitä, mihin halutaan vastauksia. Dataa on mahdollista kerätä esimerkiksi yrityksen omista tietokannoista, kuten CRM- ja ERP-järjestelmistä, erilaisilta tiedontuottajilta, jotka tarjoavat sekä ilmaista että maksullista dataa, kyselytutkimuksilla tai kokeellisilla tutkimuksilla, jotka on toteutettu varta vasten analytiikkaa ajatellen tai erilaisista sensoreista tai mittalaitteista. (Kuvio 3; Martinsuo & Kärri 2017, 19–20; Taanila 26.8.2020.)

Jotta datasta saataisiin hyödyllistä informaatiota päätöksenteon tueksi, seuraavana vaiheena on datan esikäsittely ja jalostus ihmisaivoille ymmärrettävään muotoon. Toisessa eli informaatiovaiheessa datasta muodostetaan esimerkiksi trendikuvaajia, keskiarvoja ja muita mielekkäitä tunnuslukuja ja kuvaajia. Kuitenkin ennen tätä, data tulee esikäsitellä. Esikäsitellyllä tarkoitetaan datan yhdistelyä, siihen tutustumista, siivoamista ja sen muunnoksia. Tämä tarkoittaa esimerkiksi datan taulukkomuotoisuuden tarkistamista, päivämäärien muuntamista päivämääräksi tunnistettavaan muotoon, puuttuvien arvojen käsittelyä, arvojen luokittelua ja uudelleenkoodaamista sekä monia muita vaiheita. Suuret määrät monimutkaista dataa tekevät datan esikäsitteystä hyvin aikaa vievää. Datan keräämisen ja esikäsitteilyn arvioidaan vievän jopa 80 % analysointiprosessista (SAS s.a.). Datan ja informaation tuottaminen ovat pääosin teknisiä kysymyksiä ja näitä aikaa vieviä tehtäviä

pystytään automatisoimaan tekoälyn avulla. (Kuvio 3; Martinsuo & Kärri 2017, 19–21; Taanila 26.8.2020.)

Analysointiprosessin kolmannessa eli tietämysvaiheessa hyödynnetään ihmisaivoille ymmärrettävään muotoon jalostettua dataa aineiston mallintamisessa sekä kvantitatiivisen ja kvalitatiivisen tiedon yhdistämisessä. Kolmas vaihe on siis informaation tulkitsemista ja mahdollisesti tarvittavien toimenpiteiden tunnistamista. Tietämys vaihe vaatii Martinsuo ja Kärriin mukaan laajempaa ymmärrystä tarkasteltavasta kohteesta ja parempia valmiuksia analysoida aineistoa, tulkita tuloksia yrityksen näkökulmasta sekä tunnistaa muutostarpeita. Martinsuo ja Kärri tuovat kirjassaan esille, että analysointimenetelmät riippuvat käytössä olevasta datasta ja valinta kvalitatiivisten ja kvantitatiivisten menetelmien välillä tehdään aina tapauskohtaisesti. (Kuvio 3; Martinsuo & Kärri 2017, 19–21.)

Neljännessä eli viisauksivaiheessa yhdistetään tietoja eri lähteistä ja hyödynnetään aiempia kokemuksia vaihtoehtoisten toimintatapojen tunnistamiseen sekä vertaillaan eri vaihtoehtojen hyviä ja huonoja puolia. Viisauksivaihe vaatii edellisen vaiheen tapaan laajempaa ymmärrystä tarkasteltavasta kohteesta, jotta pystytään tuottamaan niin pitkälle jalostettua tietoa, että se on sellaisenaan ilman lisätyötä hyödynnettävissä päätöksenteossa. Mikäli data-analytiikkaan käytetään apuna palveluntarjoajia, tarkoittaa tämä hyvin kiinteää yhteistyötä yrityksen ja palveluntarjoajan välillä. (Kuvio 3; Martinsuo & Kärri 2017, 20–21.)

Näiden kaikkien edellä mainittujen vaiheiden kautta saadaan mielekästä informaatiota päätöksenteon tueksi, joka onkin viimeisen eli viidennen vaiheen tehtävänä. (Kuvio 3; Martinsuo & Kärri 2017, 19–21.)

2.5 Big data

Kalyvas ja Overly määrittelevät big datan prosessiksi, jonka avulla saadaan päätöksenteon oivalluksia. Prosessi käyttää ihmisiä ja tekniikkaa analysoimaan nopeasti suuria määriä erityyppisiä tietoja useista lähteistä tuottamaan käyttökelpoista dataa päätöksenteon tueksi. Big data ei terminä viittaa vain tietomäärään tai -tyyppiin vaan nimenomaan prosessiin. Salo puolestaan määrittelee big datan kirjassaan käsitteeksi, jolla viitataan kahteen asiaan. Ensimmäinen on nopeasti kasvavan ja monipuolistuvan datan määrä, joka luo haasteita organisaatioille ja yhteiskunnalle, ja toinen on tähän haasteeseen tarjottavat ratkaisut. Salokin korostaa kirjassaan, ettei big datan käsite viittaa vain määrään ja siihen, että sitä tulee jatkuvasti lisää. Big dataa kuvaillaan kirjallisuudessa usein kolmen V:n mallilla, jossa big data määritellään kolmen piirteen avulla: volyymi (volume), vauhti (velocity) ja vaihtelevuus (variety). Volyymillä viitataan siihen, että datan määrä kasvaa maailmassa eksponentiaalisesti. Vauhdilla puolestaan tarkoitetaan kiihtyvää nopeutta,

jolla dataa syötetään tietojärjestelmiin ja jolla sitä halutaan saada käyttöön. Kolmas piirre, vaihtelevuus, kuvaa datan muuttumista yhä heterogeenisemmaksi sen lähteiden monipuolistuessa. (Kalyvas & Overly 2014, luku 1.1–1.1.1; Salo 2013, 10.)

Big data tarjoaa mahdollisuuksia kaikenkokoisille toimijoille. Aloittaville yrityksille se tarjoaa mahdollisuuksia esimerkiksi tuote- ja palveluliiketoimintaan. Salo tuo kirjassaan esille, että mitä aikaisemmin investoi aikaa ja osaamista projektiin perehtymiseen ja osajien rekrytoimiseen, sitä paremmin voi tulevaisuudessa menestyä. Keskisuurille ja suurille yrityksille big data puolestaan tarjoaa mahdollisuuden erottautua kilpailijoista ja hankkia näin kilpailuetua. Salo tuo kirjassaan esille, että vaikka dataa jaettaisiinkin kilpailijoiden kesken, on silti mahdollisuus luoda ja säilyttää kestävä kilpailuetu. Hän tuo esille, että mitä suurempi osa yrityksen datan varaan rakentamasta lisäarvosta perustuu analytiikkaan, sitä harmittomampaa raakadatan luovuttaminen on kilpailullisesti. Salo mainitsee kirjassaan, että kestävä kilpailuedun lähde pitkälle tulevaisuuteen tulee olemaan tallennettujen tai muuten saatavilla olevien datamassojen päälle rakennettu analytiikkakerros, joka muodostuu niin osaamisesta ja ymmärryksestä kuin konkreettisista teknologioista ja käytössä olevista palveluista. (Salo 2013, 90.)

2.6 Koneoppiminen

Ennakoivaan analytiikkaan käytetty koneoppiminen on yksi tekoälyn osa-alueista. Koneoppiminen ei ole mikään uusi ilmiö, vaan tekniikoiden periaatteet on kehitetty jo 50–60-luvuilla. Sen sijaan, että toiminta olisi ohjelmoitu valmiiksi, koneoppiminen käyttää dataa oppimiseen ja luokitteluun. Algoritmit, joita koneoppiminen käyttää, oppivat käytössä olevasta datasta askel askeleelta, ja mahdollistaa näin koneoppimismallin jatkuvan kehittymisen. Kehittymisen ansiosta koneoppimismalli pystyy entistä paremmin kuvailemaan käytettävissä olevaa dataa ja täten ennustamaan lopputuloksia. Koneoppiminen toimii sitä paremmin ja luotettavammin, mitä enemmän dataa on käytettävissä. Koneoppimisen arvo perustuukin juuri siihen, että se pystyy jatkuvasti ottamaan uutta dataa huomioon ja siten tarjoamaan ennustuksia tulevast. Data jaetaan koneoppimisessa opetusdataan ja testidataan. Tämän ideana on, että malli opetetaan opetusdatan avulla ennustamaan jotain haluttua lopputulosta ja testidata sitten puolestaan selvittää, miten hyvin opetus on onnistunut. (Kananen & Puolitaival 2019, 109; Merilehto 2018, 27–31.)

Yrityksissä käytetään koneoppimista yhä enenevässä määrin. Sillä täydennetään tai jopa korvataan asiantuntija-arvioita ja manuaalista päätöksentekoa. Koneoppimisen mallit ovat nopeita ja edullisia käyttää, joka varmasti osaltaan lisää sen hyödyntämistä yrityksissä. Isosta datamäärästä voidaan ennustaa nopeasti asioita päätöksenteon tueksi. Samankaltaisten ennusteiden tekeminen vaatisi muuten runsaasti enemmän aikaa. Kehitysvaiheen

jälkeen koneoppimisen ohjelmien käyttökustannukset ovat yleensä alhaiset, joka tekee niiden hyödyntämisestä suhteellisen edullista. (Kananen & Puolitaival 2019, 111.)

Koneoppimisen malleja voidaan hyödyntää tarkasti rajattuun tehtävään tai ongelmaan. Koneoppimista voidaan hyödyntää muun muassa ruokasuositusten luomiseen, liikennevalojen ohjaamiseen, henkilön tunnistamiseen tai vaikka juuri petosten estoon. Malleihin voidaan lisätä hälytyksiä, jolloin muutoksiin pystytään reagoimaan nopeasti. (Kananen & Puolitaival 2019, 111; Merilehto 2018, 29–33.)

2.6.1 Koneoppimisalgoritmien luokittelu

Koneoppimisen algoritmit voidaan luokitella esimerkiksi algoritmin opetustavan mukaan ohjattuun oppimiseen (supervised learning), ohjaamattomaan oppimiseen (unsupervised learning) ja vahvistusoppimiseen (reinforcement learning) tai algoritmien käyttötarkoituksen mukaan. (Kananen & Puolitaival, 112; Taanila 26.8.2020.)

Ohjatussa oppimisessa algoritmi opetetaan opetusdatalla eli jo valmiiksi luokitellulla datalla. Koneoppimisen nykyisistä sovelluksista suurin osa hyödyntää ohjattua oppimista. Siinä ennalta haluttu lopputulos on tiedossa ja se annetaan mallille opetusdatan syötteenä, esimerkiksi luottokorttiosoite ja vastena, esimerkiksi petos vai ei. Esimerkiksi luottokorttipetosten havaitsemiseksi sovellus opetetaan luottokorttiosoitodatalla, jossa on erilaisia tietoja kustakin ostosta sekä tieto siitä, oliko luottokorttiosoite petos vai ei. Tämän datan perusteella muodostuu sitten malli, jota käyttäen tulevista luottokorttiosoitista voidaan tunnistaa luottokorttipetokset. Ohjatun oppimisen algoritmeja ovat esimerkiksi logistinen regressio ja päätöspuut. (Merilehto 2018, 28; Taanila 26.8.2020.)

Ohjaamattomassa oppimisessa algoritmi puolestaan oppii luokittelemattoman datan pohjalta. Koneelle ei ole annettu ennakoitavaa muuttujaa, vaan se löytää datassa olevat säännönmukaisuudet ja suhteet, ja päättää asioita niiden pohjalta. Ohjaamattomaan oppimiseen paljon käytettyjä algoritmeja ovat esimerkiksi k-means clustering sekä pääkomponenttianalyysi. (Merilehto 2018, 19; Taanila 26.8.2020.)

Vahvistusoppimista voidaan pitää edellä mainittujen välimuotona. Vahvistusoppimisessa algoritmilta annetaan palautetta sen suorittamista tehtävistä antamalla kuitenkin oikeita vastauksia. Algoritmi oppii näistä sille annetuista palautteista. Algoritmi on suunniteltu löytämään keinot saada mahdollisimman paljon positiivisia pisteitä. Vahvistusoppimisen hyviä puolia ovat esimerkiksi, ettei sen käyttämiseen tarvita paljoa historiadataa eikä oikeita data-vastaus-pareja. Sen hyödyntäminen käytännössä on kuitenkin ollut haastavampaa kuin ohjatussa ja ohjaamattomassa oppimisessa. Merilehto tuo kirjassaan esille, että

useimmat algoritmit oppivat onnistuneesti vain silloin, kun säännöt, olosuhteet ja tavoitteet ovat muuttumattomat. Vahvistusoppimisen käyttökohteena on esimerkiksi itseohjautuvat autot ja dynaaminen hinnoittelu. (Merilehto, 2018, 19; Kananen & Puolitaival, 2019, 112–159; Taanila 26.8.2020.)

Ohjattu ja ohjaamaton oppiminen voidaan jakaa edelleen alatyyppeihin, joita on esitetty seuraavassa alaluvussa.

2.6.2 Koneoppimisen malleja

Merilehto (2018, 34) mainitsee kirjassaan viisi koneoppimisen mallia. Terminologia ei ole täysin vakiintunutta ja siksi käsitteet voivat olla suomennettu muualla kirjallisuudessa toisin. Koneoppimisen malleja ovat Merilehdon mukaan luokittelu (classification), ryhmittely (clustering), regressio (regression), suosittelu (recommendation) sekä poikkeamien etsiminen (anomaly detection). (Merilehto 2018, 34.)

Luokittelussa, jota voidaan pitää ohjatun oppimisen alatyypinä, yksittäinen havainto eli datapiste luokitellaan ennalta määrättyihin kahteen tai useampaan kategoriaan. Luokittelua suorittavat algoritmit tulevat kyseeseen, kun ennakoitava asia on epäjatkuva. Luokittelua voidaan käyttää esimerkiksi kohdennetussa markkinoinnissa asiakastyypin perusteella tai kuvan tunnistamisessa. Luokittelumallilta voidaan esimerkiksi kysyä, onko asiakas suuren, kohtalaisen vai pienen riskin asiakas. (Kananen & Puolitaival, 114–115; Merilehto 2018, 34; Taanila 26.8.2020.)

Ryhmittelyssä, joka on puolestaan ohjaamattoman oppimisen alatyyppejä, algoritmi etsii ryhmiä analysoimalla ilman ennalta määriteltyjä luokkia. Algoritmi ei vaadi ihmisen valmiiksi merkkejä ennakkotietoja havaintojen samankaltaisuuksien löytämiseksi, vaan se löytää ne itsenäisesti. Ryhmittelyllä saadaan vastauksia muun muassa siihen, millaisia asiakas-segmenttejä on löydettävissä tai onko eri ryhmien rahankäytössä eroa. (Kananen & Puolitaival, 116; Merilehto 2018, 34; Taanila 26.8.2020.)

Regressio on ohjatun oppimisen alatyyppejä, jossa ennakoitava asia on jatkuva. Regressiolla ennustetaan numeerista arvoa. Helpoimmin kahden muuttujan välistä suhdetta voidaan tarkastella x-y-koordinaatistolla, jossa pisteet ovat numeerisia havaintoja ilmiöstä, esimerkiksi testituloksia. Regression tavoitteena on löytää malli siihen, miten ilmiötä voidaan kuvata ja ymmärtää. Sen käyttökohteita ovat esimerkiksi kannattavuuden kehityksen ja markkinoinnin tuottojen ennustaminen. (Kananen & Puolitaival, 113; Merilehto 2018, 34; Taanila 26.8.2020.)

Suosittelun tavoitteena on esimerkiksi arvioida mistä tuotteista asiakas pitää, ja täten suositella niitä. Tässä algoritmit näyttävät asian tai havainnon suhteellisen tärkeyden muihin verrattuna. Suosittelevia algoritmeja käytetään esimerkiksi internetin hakukoneoptimoinneissa. (Kananen & Puolitaival, 116; Merilehto 2018, 34.)

Poikkeamien etsimisen tarkoituksena on nimensä mukaisesti löytää aineistosta selkeästi normaalista poikkeava data. Tämän käyttökohteita ovat esimerkiksi laitteistojen lokien tarkkailu ja toimenpidesuosituksien tuottaminen tai luottokorttien väärinkäytön havaitseminen. (Merilehto 2018, 34.)

3 Rahanpesun estäminen

Rahanpesu on vaikuttanut maailmantalouteen monien vuosien ajan. Rahaa pestään suuria määriä vuosittain ja se uhkaa maailmantaloutta ja sen turvallisuutta. YK:n huumeiden ja rikollisuuden toimiston tekemän tutkimuksen mukaan arvio rahanpesun arvosta on maailmanlaajuisesti 2–5 prosenttia koko maailman bruttokansantuotteesta eli 1,6–4 biljoonaa dollaria vuodessa. Suomessa vuonna 2018 pelkästään pankkien rahanpesutapauksia paljastui arvion mukaan 9 000 kappaletta. Rahanpesutapaukset lisääntyvät dramaattisesti teknologian kehityksestä ja globalisaatiosta johtuen. Rahanpesu määritellään rikoslain 32 luvun 6–10 §:ssä toiminnaksi, jossa henkilö peittää tai häivyttää rikoksella hankittua hyötyä tai omaisuutta, jotta omaisuus saataisiin näyttämään laillisesti hankitulta (rikoslaki 19.12.1889/39). Talousrikosten havaitsemisella ja estämisellä on tärkeä rooli pankeissa. Rahanpesun havaitseminen vaatii sekä taloudellisten tietojen turvaamista että jakamista, ja rikoksella hankitun omaisuuden polkujen jäljittämistä. Tässä luvussa tutustutaan tarkemmin rahanpesuun ja sen estämisen prosessiin, sääntelyyn sekä haasteisiin. (Chen ym. 2018, 245; FATF s.a.; Lehto 27.5.2019; United Nations Office on Drugs and Crime s.a.; Weeks-Brown 2018.)

OECD:n yhteydessä toimiva FATF (Financial Action Task Force on Money Laundering), joka on hallitusten välinen rahanpesun ja terrorismin rahoittamisen vastainen toimintaryhmä, ja joka perustettiin vuonna 1989 pidetyssä G7-huippukokouksessa, arvioi vuonna 2019 huhtikuun puolessa välin julkaistussa Suomea koskevassa maaraportissaan, että Suomen rahanpesun ja terrorismin rahoittamisen torjuntatoimet ovat melko hyvin kunnossa, mutta erityisesti valvonnassa olisi parannettavaa. Suomessa on raportin mukaan riittävä tietämys rahanpesun ja terrorismin rahoittamisen riskeistä ja niitä estetään kattavasti ja koordinoitusti. Raportin mukaan Suomen viranomaiset tekevät säännöllisesti yhteistyötä ulkomaisten kollegoidensa kanssa, jakaen oikea-aikaisia ja tarkkoja tietoja sekä suorittaen monimutkaisia ja kansainvälisiä rahanpesututkimuksia. Suomen tulisi maaraportin mukaan kuitenkin parantaa rahanpesun selvittelykeskuksen tuottamien analyysien käyttöä sekä tosiallisia edunsaajia koskevien tietojen saantia. FATF tuo raportissaan lisäksi ilmi, että rahoitusalan yritykset eivät ilmoita epäilyttävistä liiketoimista tarpeeksi usein lisätutkimuksia varten. Raportin mukaan valvojien tulisi siis parantaa riskiperusteista valvontaa näillä aloilla ja vaatia resurssien merkittävää lisäämistä. Finanssivalvonta on sittemmin ilmoittanut etenkin kyseisen raportin myötä syventävänsä rahanpesun estämisen valvonnan riskiperusteisuutta ja käyttää valvontaa tehostavia rangaistuksia entistä tehokkaammin sekä lisäävänsä valmiuttaan tehdä tarkastuksia kaksinkertaistamalla valvonnan henkilöstöään. Finanssivalvonta korostaa lisäksi lehdistötiedotteessaan eurooppalaisen yhteisen valvonnan vahvistamista rahanpesun ja terrorismin kansainvälisen luonteen

vuoksi. Suunnitteilla on lisätä kansallisten rahanpesun estämisen valvojien ja Euroopan keskuspankin välistä tietojenvaihtoa ja yhteistyötä sekä Euroopan pankkiviranomaisten valtuuksia valvonnan tehostamiseksi. (FATF 2019; Finanssivalvonta 16.4.2019; Lehto 27.5.2020.)

3.1 Rahanpesuprosessi

Rahanpesu on ollut Suomessa rangaistava teko vuodesta 1994 lähtien. Se, millaisen rangaistuksen tekijä saa, riippuu siitä, onko tekijä tehnyt myös niin sanotun alkurikoksen, esimerkiksi huumausainerikoksen, vai osallistunut vain rahanpesuun. Mikäli alkurikoksentekijä osallistuu myös rahanpesuun, häntä ei pääsääntöisesti tuomita erikseen rahanpesusta, vaan rangaistus tulee alkurikoksen tekemisestä. Mikäli toinen henkilö auttaa rikoksentekijää peittämään tai häivyttämään rikoksella saatujen varojen alkuperää, mutta ei ole osallistunut itse alkurikokseen, voidaan hänet tuomita rikoslain 32 luvun 6–1 §:n mukaisesti rahanpesurikoksesta rangaistukseen. Suomessa pestävästä rahasta valtaosa on talousrikollisuudella hankittua hyötyä. (Poliisi s.a.)

Rahanpesu voi tapahtua esimerkiksi kierrättämällä varoja pankkien kautta tai ostamalla tai sijoittamalla omaisuutta. Rahanpesuprosessi voidaan jakaa kolmeen vaiheeseen: sijoitusvaiheeseen (placement), harhautusvaiheeseen (layering) ja palautusvaiheeseen (integration). Sijoitusvaiheessa laittomat varat sijoitetaan lailliseen rahoitusjärjestelmään. Tällöin esimerkiksi suuret määrät käteistä voidaan jakaa pienemmiksi summiksi, ja tallettaa suoraan pankkitilille. Varat sijoitetaan usein eri tilien ja jopa eri pankkien kesken. Pienempiin osiin jakamisella pyritään herättämään vähemmän huomiota. Harhautusvaiheessa laittomat varat muutetaan toiseen muotoon, esimerkiksi ostamalla sijoitusinstrumentteja, jotta niiden alkuperää olisi vaikeampi selvittää. Varat voidaan myös siirtää useiden pankkien tilien kautta ympäri maailmaa, ja täten peitellä niiden alkuperää. Kolmannessa vaiheessa, eli palautusvaiheessa toiseen muotoon muutetut varat sijoitetaan lailliseen liiketoimintaan, jonka jälkeen niiden alkuperää on lähes mahdotonta selvittää. Tällöin rikoksentekijän palautuvat varat voivat näyttää laillisesti hankitulta toimeentulolta. (FATF s.a.; Poliisi s.a.)

3.2 Sääntely

Lain rahanpesun ja terrorismin rahoittamisen estämisestä tavoitteena on estää rahanpesua, edistää sen paljastamista ja selvittämistä sekä tehostaa rikoksella hankitun omaisuuden jäljittämistä ja takaisinsaantia. Rahanpesun estäminen perustuu kansainvälisiin standardeihin. Globaaleilla finanssimarkkinoilla sääntelyn pyrkimyksenä on yhtenäisten asiakkaiden tuntemista koskevien menettelytapojen noudattaminen. Tärkeässä roolissa rahanpesun estämisestä on maailmanlaajuinen rahanpesun ja terrorismin rahoittamisen vastainen

toimintaryhmä FATF. EU:n rahanpesudirektiivit perustuvatkin FATF:in suosituksiin. Suomessa valtiovarainministeriö vastaa rahanpesun vastaisen lainsäädännön kehittämisestä. Laissa rahanpesun ja terrorismin rahoittamisen estämisestä on määriteltynä tahot, jotka ovat velvollisia ilmoittamaan rahanpesun selvittelykeskukselle epäilyttävistä liiketoimista. Lain määrittelemiä ilmoitusvelvollisia ovat esimerkiksi pankit ja vakuutusyhtiöt. Finanssivalvonnan valvottavilla, kuten juuri pankeilla, on velvollisuus tuntea asiakkaidensa toimintaa, sitä kautta havaita ja selvittää epäilyttäviä liiketoimia sekä ilmoittaa epäilyistään rahanpesun selvittelykeskukselle, joka toimii Keskusrikospoliisin yhteydessä. (FATF s.a.; Finanssivalvonta s.a.; RapeL 28.6.2017/444.)

Lain rahanpesun ja terrorismin rahoittamisen estämisestä 2 luvussa säädetään riskiarvion laatimisesta. Kansalliset viranomaiset, sisäministeriö ja valtiovarainministeriö, sovittavat yhteen kansallisen rahanpesun ja terrorismin rahoittamisen riskiarvion laadintaa. Rahanpesun rahoittamisen riskiarvion laadinnasta vastaa valtiovarainministeriö. Riskiarviossa on tunnistettava ja arvioitava Suomen rahanpesun rahoittamisen riskejä sekä otettava huomioon Euroopan komission laatima EU:n laajuinen rahanpesun rahoittamisen riskiarvio. Riskiarvion tavoitteisiin kuuluu muun muassa yksilöidä riskit toimialoilla sekä tukea ja tehostaa torjuntaa ja voimavarojen kohdentamista. Kansallisen riskiarvion lisäksi toimivaltaisen valvontaviranomaisen on laadittava lain mukaan riskiarvio sen valvonnan piiriin kuuluvien ilmoitusvelvollisten rahanpesun rahoittamisen riskeistä. (RapeL 28.6.2017/444.)

3.3 Rahanpesun estämisen prosessi

Ilmoitusvelvollisen on varauduttava rahanpesun estämiseen jo ennen asiakassuhteen aloittamista. Kuviossa 4 on havainnollistettu rahanpesun estämisen prosessia riskiarvion laadinnasta työntekijöiden suojelemiseen lain näkökulmasta. Pankkien ja muiden ilmoitusvelvollisten on lain rahanpesun ja terrorismin rahoittamisen estämisestä 2 luvun 3 §:n mukaan laadittava ja päivitettävä säännöllisesti riskiarvio rahanpesun riskien tunnistamiseksi ja arvioimiseksi. Riskiarvion laadinnassa on otettava huomioon ilmoitusvelvollisen toiminnan luonne, koko ja laajuus. Ilmoitusvelvollisella on oltava edellä mainitut tekijät huomioon ottaen riittävät toimintaperiaatteet, menettelytavat ja valvonta rahanpesun ja terrorismin rahoittamisen riskien vähentämiseksi ja tehokkaaksi hallitsemiseksi. (RapeL 28.6.2017/444.)



Kuvio 4. Rahapesun estämisen prosessi (RapeL 28.6.2017/444)

Lain mukaisesti vakituista asiakassuhdetta perustettaessa ilmoitusvelvollisen on tunnistettava asiakkaansa ja todennettava tämän henkilöllisyys. Kaikki asiakkaan tuntemista ja liiketoimia koskevat asiakirjat ja tiedot on pidettävä ajantasaisina ja olennaisina sekä säilytettävä luotettavalla tavalla viiden vuoden ajan vakituisen asiakassuhteen päättymisestä. Pankit arvioivat asiakassuhteeseen liittyvää talousrikollisuuden riskiä esittämällä asiakkaalle kysymyksiä asiakkaasta ja siitä, miten asiakas aikoo käyttää pankin tuotteita ja palveluita sekä pyytämällä asiakirjoja jo asiakassuhteen alkaessa. Pankkien on muun muassa selvitettävä, mistä heidän asiakkaidensa varat ovat peräisin ja miten asiakkaat aikovat käyttää tilejään. Näin pankit voivat tunnistaa ja oppia tuntemaan asiakkaansa sekä heidän liiketoimintansa, ja sitä kautta pystyvät havaitsemaan epäilyttäviä liiketoimia. (Nordea s.a.; RapeL 28.6.2017/444.)

Talousrikollisuuden riskin arviointi vaikuttaa merkittävästi siihen, miten pankki seuraa asiakkaidensa tilejä myöhemmin. Laki velvoittaa ilmoitusvelvollisia hankkimaan ja seuraamaan jatkuvasti tietoja asiakkaidensa toiminnasta, laadusta ja laajuudesta. Nordean rahapesun estämistä käsittelevältä sivulta selviää, että heillä pankin henkilöstö ja IT-järjestelmät valvovat jatkuvasti asiakkaidensa toimia. Rahapesutapausten havaitsemiseksi on kiinnitettävä erityisesti huomiota seuraavan kaltaisiin tapahtumiin: rahojen alkuperä ja kohde, varojen nopea siirto, pitkään epäaktiivisen asiakkaan tilin aktivoituminen uudelleen tai usein toistuvat käteisen rahan tapahtumat. Perinteisesti rahoituslaitokset ovat luottaneet voimakkaasti manuaalisiin, ihmisen tekemiin toimiin rahapesun estämiseksi. Useat eri välikädet tarkastelevat yksityiskohtia fyysisesti ja kirjoittavat selostuksia ennen kuin epäilyttävästä toiminnasta ilmoitetaan sääntelyviranomaisille. (Ey 2019; Nordea s.a.; Lehto 27.5.2019; RapeL 28.6.2017/444.)

Lain rahanpesun ja terrorismin rahoittamisen estämisestä 4 luvussa säädetään ilmoitus- ja salassapitovelvollisuudesta sekä liiketoimen keskeyttämisestä. Ilmoitusvelvollisen on viipymättä ilmoitettava rahanpesun selvittelykeskukselle epäilyttävästä liiketoimesta. Nordeassa toimitaan heidän nettisivujensa mukaan rahanpesun estämiseksi lähettämällä ilmoitus valvontajärjestelmään tai pankkineuvojan käsiteltäväksi, jos asiakkaan tilitapahtumissa huomataan jotain epäilyttävää. Tilitapahtumat tarkistetaan muiden tavanomaisesta poikkeavien tapahtumien varalta. Mikäli pankki ei löydä poikkeaville tapahtumille selitystä ja epäilyksiä on edelleen, siirretään asia jollekin Nordean asiantuntijatiimeistä tutkittavaksi yksityiskohtaisesti. Mikäli yksityiskohtainen tutkimus vahvistaa pankin epäilyksen epäilyttävästä toimesta, ilmoitetaan asiasta viranomaisille. Pankin on syytä avustaa viranomaisia lisätutkimuksissa. Lain mukaan ilmoitusvelvollisen on keskeytettävä liiketoimi lisäselvityksiä varten. Epäilyttävän tapahtuman ilmetessä ja vahvistuessa pankki harkitsee asiakassuhteen lopettamista. (Nordea s.a.; RapeL 28.6.2017/444.)

Ilmoitusvelvollisen keskeisiin velvollisuuksiin kuuluu lisäksi tärkeänä osana työntekijöiden koulutus ja suojeleminen. Pankin on huolehdittava, että sen työntekijät saavat koulutuksen lain rahanpesun ja terrorismin rahoittamisen estämisestä ja sen nojalla annettujen säännösten noudattamisen varmistamiseksi. Pankin tulee lisäksi toteuttaa toimenpiteitä niiden työntekijöiden suojelemiseksi, jotka tekevät ilmoituksen epäilyttävästä liiketoimesta tai terrorismin rahoittamisen epäilystä. (RapeL 28.6.2017/444.)

3.4 Rahanpesun estämisen haasteita

Breslow ym. (2017) tuovat artikkelissaan ilmi haasteita, joita rahoituslaitokset kohtaavat rahanpesua estävien ohjelmien käytössä. Haasteita tuo huonolaatuinen data, jota joudutaan hankkimaan useista hajanaisista lähteistä, joka vaikeuttaa tietojen yhdistämistä. Monet pankit joutuvat edelleen soittamaan kalliita asiakaspuheluja päivittääkseen virheellisiä tai puutteellisia asiakkaan tuntemistietoja tietokannoissaan, joka lisää myös ylimääräisiä henkilöstökustannuksia. Toisena haasteena artikkelissa tuodaan esille väärin positiivisten hälytysten määrä, joka lisääntyy usein uusien kalibrointityökalujen ja raja-arvojen lisäämisestä. Suuret määrät väärä positiivisia hälytyksiä lisäävät merkittävästi henkilötyötunteja turhien hälytysten selvittämiseen. Haasteena havaittiin myös hajautetut järjestelmät ja alustat, jotka rajoittavat transaktioiden seuraamisen automatisointia ja asiakkaan tuntemisvelvollisuuden täyttämistä. Tutkijat käyttävät suurimman osan ajastaan tietojen keräämiseen sen sijaan, että aika käytettäisiin tutkimustyöhön, mikä johtuu osaltaan hajautetuista järjestelmistä. Näihin haasteisiin ratkaisua tuo osaltaan data-analytiikka, johon perehdytään tarkemmin seuraavassa luvussa. (Breslow, Hagstroem, Mikkelsen & Robu 7.11.2017.)

4 Data-analytiikan hyödyntäminen rahanpesun estämisessä

Yhä tiukemmat vaatimukset ja valtava datan määrän kasvu tekevät talousrikollisuuden torjunnasta entistä vaikeampaa pankeille, manuaalisten prosessien ja vanhanaikaisten tekniikoiden takia. Tutkimuksista ja artikkeleista käy ilmi, että rahanpesun estämisen toimi-joille on määrätty viime vuosikymmenen aikana suuria määriä sakkoja rahanpesun torjunnan, asiakkaan tuntemisen ja pakotemääräysten noudattamatta jättämisestä, joten muutokset ovat tarpeen. Lehto tuokin Talouselämän artikkelissa esille, että pankit hyödyntävät yhä enenevässä määrin ennakoivaa ja reaaliaikaista analytiikkaa rahanpesun ja petosten valvonnassa ja selvittelyssä. (Chen ym. 2018, 245—246; EY 2019; Lehto 27.5.2019.)

Institute of International Financen (IIF) vuonna 2018 tekemän tutkimuksen mukaan, jossa selvitettiin koneoppimistekniikoiden soveltamista rahanpesun torjunnassa, suurin osa (35 %) kyselyyn vastanneista käyttää jo koneoppimismenetelmiä. Kysely tehtiin 59 rahoituslaitokselle, jotka koostuivat 54 pankista ja 5 vakuutusyhtiöstä. Kyselyyn valitut rahoituslaitokset olivat eri kokoisia ja valittujen joukossa oli rahoituslaitoksia kaikilta eri mantereilta. Tutkimuksen mukaan 34 prosentilla kyselyyn vastanneista koneoppimismenetelmät ovat ko-keilussa ja 29 % vastanneista ilmoitti ottavansa uusia analytiikka tekniikoita käyttöönsä lähitulevaisuudessa. Koneoppimisen käyttötapaukset kyselyyn vastanneiden joukossa vaihtelevat edistyneen analytiikan soveltamisesta virheellisten positiivisten hälytysten vähentämiseksi jo olemassa oleviin seurantajärjestelmiin asetetusta lisäsuodattimesta aina perinteisten lähestymistapojen yhdistämisestä graafiteoriaan ja ohjattuun oppimiseen suorittamaan seuranta itse. Pääasiallinen tavoite kaikilla käyttötapauksilla on kuitenkin vapauttaa yrityksen analyytikoiden resursseja, ja käyttää heidän keskittymisensä ensi sijassa päätöksentekoon. Tutkimuksen mukaan rahoituslaitokset eivät tällä hetkellä halua automatisoida päätöksentekoprosessia, mutta tukevat analyytikoidensa työtä entistä tehokkaammilla tekniikoilla ja automatisoimalla manuaalisempia prosessivaiheita. (Delle-Case, Bailey, Carr & Ekberg 2018, 1–3.)

Breslow ym. (2017) kokemuksen mukaan data-analytiikan hyödyntämisessä on paljon mahdollisuuksia, mutta he nostavat kolme aluetta, joihin pankkien tulisi ensi sijassa sijoittaa. Näitä ovat tehokkaat tietojen yhdistämisalustat, edistynyt tilastollinen mallintaminen, kuten koneoppimiseen perustuva riskipisteytys ja prosessien automatisointi. (Breslow ym. 7.11.2017.)

Chen ym. (2018) tuovat ilmi tutkimuksessaan kattavan katsauksen koneoppimisen algoritmeihin ja menetelmiin, joilla voidaan havaita epäilyttäviä tapahtumia. Tutkimuksesta käy ilmi neljä tärkeää ominaisuutta, jotka rahanpesun estämiseksi käytetyillä algoritmeilla tulee

olla kunnossa, jotta sillä saataisiin vaadittavia hyötyjä: datan laatu, havaitsemisen tarkkuus, skaalautuvuus ja reaktio aika. Monet rahanpesun estämisen järjestelmät koostuvat useista eri algoritmeista ja tekevät näin niiden hyödyntämisestä tehokkaampaa. (Chen ym. 2018 248–251.)

4.1 Asiakasdatan jakaminen pankkien kesken

Burton-Taylor Internationalin raportin mukaan rahanpesun estämisestä, asiakkaan tuntemisesta ja niihin liittyvistä talousrikoksiin ja sääntöjen noudattamiseen liittyvistä toimista aiheutuvat kustannukset ovat maailmanlaajuisesti kasvaneet 17,5 % vuonna 2019. Tämä tarkoittaa ennätyselliset 905 miljoonaa dollaria. Raportin mukaan uuden asiakkaan rekisteröinti ja tunnistaminen vie pankeilta valtavasti aikaa ja resursseja, jotta siihen liittyvät sääntelyt tulee noudatettua. (Griboval 12.1.2020.)

Puutteellisen ja rajallisen datan analysointi ei ole paras tapa saada laajaa kuvaa transaktioista ja havaita mahdollisia rahanpesutapauksia. Institute of International Finance toteaa raportissaan, että merkittävä tekijä, joka estää pankkien rahanpesun estämistä, on juuri riittämätön datan jakaminen pankkien välillä epäilyttävien tilien ja toimintojen ilmoittamiseksi. Siten asiakkaisiin liittyvien tietojen jakaminen muiden pankkien kanssa näyttäisi olevan tarpeellista. Erilaisten BI-järjestelmien sekä koneoppimis- ja automaatoratkaisujen käyttäminen tietojen jakamiseen pankkien välillä kansainvälisellä tasolla parantaisi selvästi nykyisiä tapoja toimia sekä kasvattaisi jäädytettyjen varojen määrää vuosittain. Monet pankit eivät kuitenkaan halua jakaa asiakasdataansa kilpailijoilleen, ja se rikkoisikin asiakkaan yksityisyyden suojaa. Yhä tiukentuvat tietosuojaa käsittelevät säännökset asettavat tiukat rajoitukset tietojen käsittelylle ja jakamiselle ja näin hidastavat ja rajoittavat asiakkaan tuntemista ja rahanpesun estämistä. Yhdistämällä koneoppimisen ja tietosuojasetuksen rajoissa tiedon jakamisen pankkien välillä ja jopa kansainvälisten rajojen yli, rahoituslaitoksilla on mahdollisuus olla huomattavasti tehokkaampia yhä kalliimmassa rahanpesun ja terrorismin rahoittamisen estämisessä. (Delle-Case ym. 2018, 6–7; Griboval 12.1.2020; Shainski 20.11.2019.)

4.2 Data-analytiikan tuoma lisäarvo

Onnistunut havaitsemisjärjestelmä tunnistaa epäilyttävän toiminnan trendit tai mallit nopeasti luomalla hälytyksen, joka osoittaa petoksen todennäköisyyden. Kun pistemäärä ylittää tietyn raja-arvon, tapaus otetaan tarkempaan selvittelyyn. Nykypäivänä tärkeää rahanpesun estämisessä data-analytiikan avulla on juuri reaaliaikaisuus. Pankin edustaja kommentoi Talouselämän artikkelissa, että nopeaan muistinvaraiseen laskentaan perustuva

analytiikka voi tutkia maksuliikennettä millisekunneissa eli sekunnin tuhannesosissa. Prosessista on saatu huomattavasti nopeampi aikaisempaan verrattuna, sillä nämä operaatiot on ajettu aikaisemmin pankeissa yön yli. McAuley, SAS Institutun petosten ja talousrikosten torjunnan yksikön johtaja kertoo, että analytiikan avulla suurestakin datamassasta on mahdollista tunnistaa ihmistä nopeammin epätavallista rahaliikennettä tai muita poikkeustilanteita, jotka sitten päätyvät tarkempiin tutkimuksiin. (Chen ym. 2018, 247; Lehto 27.5.2019.)

Kehittyneitä data-analytiikka menetelmiä, kuten tekoälyä, koneoppimista, luonnollisen kielien käsittelyä ja kognitiivista automaatiota, voidaan käyttää nopeuttamaan tai automatisoimaan merkittävä osa henkilötöitä vaativasta työstä vähentämällä operationaalisia kustannuksia ja antamalla ihmisille vapauden keskittyä ennaltaehkäiseviin toimenpiteisiin. Koneoppimisen avulla pankit voivat vähentää transaktioiden valvomisessa vääriä positiivisia hälytyksiä, joka vähentää merkittävästi manuaalisen työn tarvetta. Tämän takia tutkijat pystyvät keskittymään tarkemmin tutkivaan ja päätöksentekoa vaativaan työhön. Tämä vähentää myös säännösten noudattamatta jättämiseen liittyviä sakkoja, kun oikeat tapaukset huomataan herkemmin ja niihin voidaan käyttää enemmän aikaa. Breslow ym. (2017) kokemuksen mukaan koneoppimista hyödyntämällä saadaan vähennettyä 20–30 prosenttia vääriä hälytyksiä, jonka avulla manuaalisen työn tarve pienenee jopa 50 prosenttia. (Breslow ym. 7.11.2017; EY 2019.)

Älykkäät tietojärjestelmät käyttävät koneoppimista tai sumeaa logiikkaa strukturoimattomaan tili ja transaktio dataan luodakseen 360-asteen näkymän rahanpesuepäilyistä. Näiden avulla pankit voivat automaattisesti vahvistaa enemmän asiakkaan tietoja, tunnistaa tosialliset edunsaajat nopeammin ja kartoittaa tiettyjen asiakkaiden yhteyksiä muihin henkilöihin, etenkin suuremman riskin henkilöihin. Data-analytiikan avulla on lisäksi mahdollista havaita yhteyksiä nopeasti monien eri asiakkaiden tekemien pienten käteistalletusten ja näiden asiakkaiden samalle vastaanottajalle lähettämien tilisiirtojen välillä. Data-analytiikkaa voidaan käyttää myös ennaltaehkäisevissä toimissa, kuten parantamaan asiakkaan tuntemisprosessia, joka parantaa osaltaan asiakaskokemusta, pakoteseurannan tehostamisessa ja asiakastapahtumien seurannassa, mikä auttaa tunnistamaan ennalta riskit ja mahdollisuudet. (Breslow ym. 7.11.2017; EY 2019.)

4.3 Data-analytiikan mukana tulevat haasteet

Data-analytiikka järjestelmät ovat mittava investointi pankeille. McKinseyn artikkelissa tuodaan ilmi, että pankit ovat käyttäneet jopa satoja miljoona dollareita asiakkaantuntemisen ja rahanpesun estämisen vaatimusten täyttämiseksi erilaisiin prosesseihin ja järjestelmiin. (Breslow ym. 7.11.2017.)

Chen ym. (2018, 247) tuovat tutkimuksessaan esille rahanpesun estämiseen liittyvän datan analysoinnin haasteita. Haasteet voidaan luokitella viiteen ryhmään: datan määrä ja monimutkaisuus, luokkien epätasapaino, käsitteen muuntautuvuus, luokkien päällekkäisyys sekä virheellisyys. Datan muuntaminen esikäsittelyvaiheessa vaatii paljon vaivannäköä, ollen erittäin haastava ja aikaa vievä tehtävä. Suurimman osan kokonaisajasta rahanpesun estämisprosessissa vie datan käsittely, sillä rahanpesun estämiseen liittyvää dataa on valtavia määriä. Esimerkiksi Nordean tilien kautta kulkee vuosittain noin kaksi miljardia tilisiirtoa. (Chen ym. 2018, 248; Nordea s.a.)

5 Data-analytiikan hyödyntäminen rahanpesun estämisessä pankkien näkökulmasta

Tässä luvussa esitellään opinnäytetyön tutkimuksen toteutus perusteluineen, käyden läpi vaiheet tutkimustavoitteiden asettamisesta tutkimusaineiston analyysiin, sekä haastattelujen tulokset sitaatein havainnollistaen.

5.1 Tutkimuksen toteutus

Opinnäytetyölle valittujen tavoitteiden pyrkimyksenä oli tukea työlle valittua rajausta ja näkökulmaa sekä olla yhteiskunnallisesti tärkeitä. Työn tarkoituksena oli, että alalla toimijat saisivat työstä tärkeää tietoa toimiakseen tehokkaasti ja kannattavasti tällä tarkkaan säännellyllä alalla sekä kannustaa pankkeja hyödyntämään data-analytiikkaa entistä tehokkaammin rahanpesun estämisessä ja kehittämään rahanpesun estämistä kohti tulevaisuutta. Tämän opinnäytetyön päätavoitteena oli selvittää, miten pankit voisivat hyödyntää data-analytiikkaa tehokkaasti ja kannattavasti rahanpesun estämisessä. Päätavoitteeseen saatiin vastaus alatavoitteiden kautta. Opinnäytetyön alatavoitteet olivat seuraavat: Miten pankit hyödyntävät data-analytiikkaa rahanpesun estämiseksi sekä miten kannattavana data-analytiikan hyödyntäminen rahanpesun estämisessä koetaan pankeissa.

Tutkimusmenetelmät voidaan jakaa kvantitatiiviseen eli määrälliseen ja kvalitatiiviseen eli laadulliseen tutkimukseen. Tämän opinnäytetyön empiirinen osa toteutettiin kvalitatiivisena tutkimuksena. Kirjallisuudessa on todettu, että kvalitatiivisessa tutkimuksessa ei pyritä tilastollisiin yleistyksiin, kuten kvantitatiivisessa tutkimuksessa, vaan sen avulla pyritään saamaan syvälinen näkemys tutkittavasta ilmiöstä. Kvalitatiivinen tutkimus sopii siksi ilmiöiden, joita ei entuudestaan juuri tunneta, tutkimiseen. Näillä perusteluilla kvalitatiivinen tutkimus oli perustellumpaa tutkimukselle asetettujen tavoitteiden takia kuin kvantitatiivinen menetelmä. Näiden lisäksi tässä kyseisessä tutkimuksessa tutkimuskohteiden joukko oli lukumäärältään melko pieni, joten tästäkään syystä määrällinen yleistäminen ei olisi ollut soveltuva tälle tutkimukselle. (Kananen 2017, 32–33; Tuomi & Sarajärvi 2018, 98.)

Tutkimustavoitteet ratkaistaan kerätyllä aineistolla. Yleisimpiä kvalitatiivisen tutkimuksen aineistontuottamistapoja ovat dokumentit, haastattelut ja havainnointi. Tähän opinnäytetyöhön aineiston tuottamistavaksi valittiin haastattelu ja tarkemmin puolistrukturoitu haastattelu, eli kaikille haastateltaville esitettiin samat kysymykset noudattaen ennalta luotua haastattelurunkoa, mutta kysymyksiin ei ollut olemassa valmiita vastausvaihtoehtoja, kuten esimerkiksi kyselylomakkeella on. Tutkimusta olisi voinut lähteä toteuttamaan myös dokumenttien pohjalta, mikäli Suomessa toimivista pankeista olisi dokumentaatiota, joka

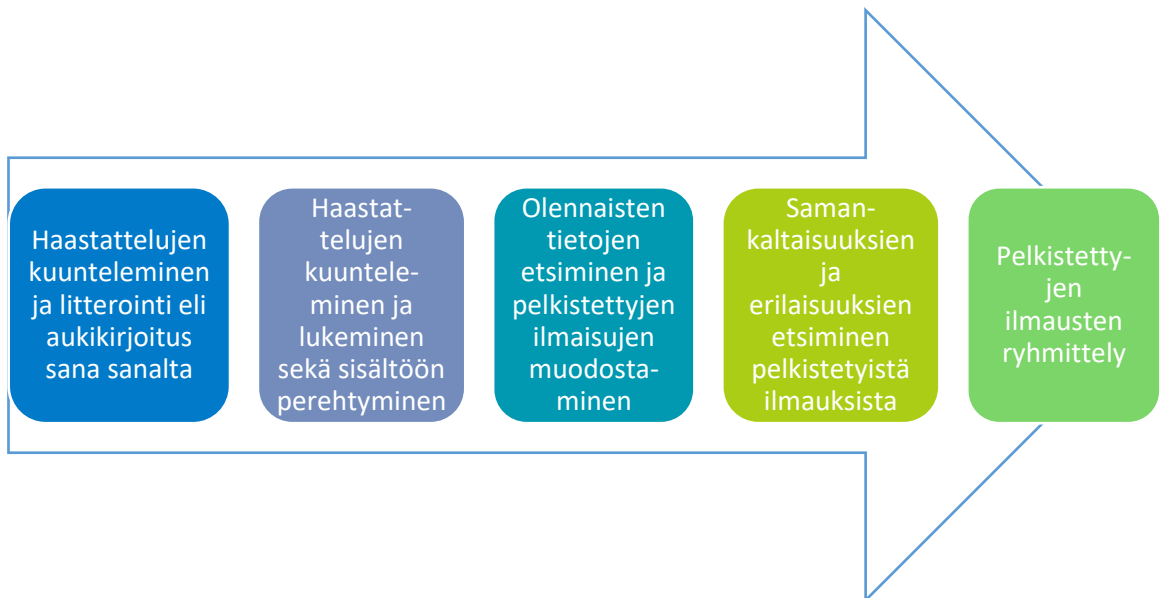
hyödyttäisi tutkimusta, näin ei kuitenkaan ollut. Aineistontuottamistavaksi valikoitui haastattelu, koska se oli kaikkein luontevin tutkimuksen tavoitteiden kannalta. Haastattelu sopiikin Tuomi ja Sarajärven kirjan mukaan sellaisiin tapauksiin, kun halutaan tietää, mitä joku ajattelee tai miksi hän toimii niin kuin toimii. Saman kirjan mukaan haastattelussa tärkeintä on saada mahdollisimman paljon tietoa halutusta ilmiöstä. Tutkimuksen haastattelukysymykset olivat avoimia kysymyksiä, jotta haastateltavien oli mahdollista avata omia näkemyksiään ja käytäntöjään mahdollisimman tarkkaan. Haastattelurunko (liite 1) koostettiin 12 haastattelukysymyksestä, joiden avulla oli mahdollista saada vastaus tutkimustavoitteisiin. Haastattelukysymykset määriteltiin tutkimustavoitteiden ja teoreettisen viitekehysten pohjalta. Tällä pyrittiin varmistamaan, että haastattelurungossa olevilla kysymyksillä kysyttäisiin tutkimuksen aiheen kannalta oikeita asioita ja tutkimuksen päätavoite saavutettaisiin. Peittomatriisin avulla varmistuttiin siitä, että haastattelurungossa oli riittävästi kysymyksiä molempiin tutkimuksen alatavoitteisiin liittyen, joiden avulla saataisiin oleelliset vastaukset myös tutkimuksen päätavoitetta ajatellen. Haastattelurungon kysymyksiin saatiin tukea menetelmäohjaajalta ja kysymykset hyväksytettiin opinnäytetyön ohjaajalla vielä ennen itse haastatteluiden toteuttamista. Haastattelut toteutettiin vuonna 2020 lokamarraskuun aikana. (Kananen 2017, 52–67; liite 1; Tuomi & Sarajärvi 2018, 83–85.)

Koska kvalitatiivisen tutkimuksen avulla pyritään saamaan syvälinen ymmärrys tutkittavasta ilmiöstä, on kvalitatiivisissa tutkimuksissa tärkeää, että haastateltavat tietävät tutkitavasta ilmiöstä mahdollisimman paljon tai heillä on kokemusta asiasta. Tuomi ja Sarajärvi korostavat, että haastateltavien valinta tästä syystä tulee olla harkittua ja tarkoitukseen sopivaa. Kvalitatiivisessa tutkimuksessa käytetään siis suppeaa, harkinnanvaraisesti ja tarkoituksenmukaisesti koottua näytettä. Tutkimusta varten haastateltiin muutamia alan ammattilaisia, jotka työskentelevät data-analytiikan ja rahanpesun estämisen parissa, jotta saataisiin muodostettua mahdollisimman kattava ja kokonaisvaltainen ymmärrys tutkittavasta ilmiöstä. Haastateltaviksi haluttiin eri kokoisia pankkeja, jotta saataisiin mahdollisimman laaja käsitys menetelmistä ja kokemuksista pankkisektorilla, ja näin haastateltavat edustaisivat kohdejoukon näkökulmia aiheeseen. Haastateltavaksi valittiin neljä pankkia, joista kolmea pankkia haastateltiin Microsoft Teams -alustan välityksellä ja yhdeltä pankilta vastaukset kysymyksiin saatiin sähköpostitse. Kyseisiin ratkaisuihin päädyttiin vallitsevan poikkeustilanteen takia, jonka COVID-19 (koronavirustauti 2019) on aiheuttanut sekä haastateltavien kiireiden takia. Se, onko valittujen pankkien määrä riittävä tutkimuksen kannalta, selviää saturaatiolla eli kylläntymisellä. Tuomi ja Sarajärvi kertovat kirjassaan, että saturaatiolla tarkoitetaan tilannetta, jossa aineisto alkaa toistaa itseään. Tämä tarkoittaa sitä, että haastateltavat eivät tuota enää tutkimustavoitteen kannalta uutta tietoa. Täten voidaan todeta, että neljä pankkia olivat riittävä tutkimuksen kannalta, sillä

haastateltavilla ei ollut enää antaa uutta tietoa tutkimustavoitteen kannalta. Haastateltavien suhteellisen pienen joukon riittävyttä tukee se, että haastateltavat joutuvat toimimaan rahanpesun estämisen parissa sääntelyn alla, jonka takia erityisesti rahanpesun estämisen menetelmät ovat hyvin samankaltaisia. (Tuomi & Sarajärvi 2018, 98–99.)

Haastateltavat saivat haastattelukysymykset jo etukäteen nähtäväksi, jotta heidän oli mahdollista tutkia kysymyksiä etukäteen ja antaa siten syvällisempiä vastauksia. Tällä varmistuttiin myös siltä, ettei itse haastattelutilanteessa tule esille, ettei haastateltava voi vastata kysymykseen lainkaan kysymyksen sensitiivisyyden takia. Tuomi ja Sarajärvi mainitsevat myös kirjassaan, että haastattelukysymykset on perusteltua antaa haastateltaville hyvissä ajoin etukäteen, jotta halutusta asiasta saataisiin mahdollisimman paljon tietoa ja siten tukisi haastattelun onnistumista. Kahdesta pankista haastatteluun osallistui kaksi pankin edustajaa ja kahdesta yksi. Haastattelut nauhoitettiin, jonka mahdollisuutta tiedusteltiin pankeista etukäteen, jotta se ei tulisi haastattelutilanteessa yllätyksenä. Haastattelujen pituudet vaihtelivat 36 minuutista 52 minuuttiin. Haastateltaville annettiin mahdollisuus anonymisointiin eli henkilötietojen häivyttämiseen aiheen sensitiivisyyden vuoksi. Haastateltavat anonymisointiin myös sen takia, että haastateltavan tai yrityksen nimellä ei ollut keskeistä merkitystä tutkimuksen tulosten esittämisen kannalta. Haastatellut pankit anonymisoitiin nimeämällä pankit sattumanvaraisessa järjestyksessä seuraavanlaisesti: pankki A, pankki B, pankki C sekä pankki D. (Tuomi & Sarajärvi 2018, 85-86.)

Kvalitatiivisten tutkimusaineistojen analyysitapoja on olemassa useita. Näitä analyysimenetelmiä ovat esimerkiksi teemoittelu, tyypittely, sisällönanalyysi sekä diskursiiviset menetelmät. Tämän opinnäytetyön tutkimusaineiston analyysimenetelmäksi valittiin sisällönanalyysi. Sen pyrkimyksenä on järjestellä tutkimusaineisto tiiviiseen ja selkeään muotoon kadottamatta sen sisältämää informaatiota. Sisällönanalyysi valikoitui tutkimusaineiston analyysimenetelmäksi, koska sen avulla oli mahdollista esitellä aineiston keskeiset asiat tiivistetysti sanallisessa muodossa. Tutkimusaineistoa kertyi haastatteluista paljon avoimien kysymysten takia, joten koin, että sisällönanalyysi sopisi tutkimukselle erinomaisesti. Sisällönanalyysin avulla tutkimusaineisto saadaan kuitenkin vain järjestetyksi johtopäätösten tekoa varten. Sisällönanalyysillä järjestetyn tutkimusaineiston pohjalta muodostetaan johtopäätökset. Tutkimuksen johtopäätökset on esitetty luvussa 6. Kuviossa 5 on esitelty aineistolähtöisen sisällönanalyysin eteneminen, jota tätä tutkimusta analysoitaessa on noudatettu. (Eskola & Suoranta 1998, 161, Kuvio 5; Tuomi & Sarajärvi 2018, 117–122.)



Kuvio 5. Aineistolähtöisen sisällönanalyysin eteneminen (mukaillen Tuomi & Sarajärvi 2018, 123)

Ensimmäinen ja tärkeä vaihe haastatteluaineiston analyysissä on aineiston litterointi eli tekstiksi muuttaminen. Litteroinnin avulla tutkimusentekijä tutustuu saatuun aineistoon. Nämä tekstiksi puretut äänitallenteet muodostavat haastattelututkimuksen tutkimusaineiston. Litteroinnin tarkkuustason määrittää tutkimustavoite sekä analyysitapa. Opinnäytetyössä ei tutkita puheen tuottamista, joten puhettavan, esimerkiksi puheessa olevien taukojen ja äänenpainojen litterointi ei ole tarpeen. Sen sijaan puheen sisältöä opinnäytetyössä tutkitaan, joten nauhoitukset litteroidaan huomioiden sanat ja sisältö. Tässä tutkimuksessa nauhoitetut haastattelut kuunneltiin ensin kokonaisuudessaan läpi, jonka jälkeen haastattelut kirjoitettiin kirjalliseen muotoon sanat ja sisältö huomioiden. Litteroinnin jälkeen aineisto kuunneltiin sekä luettiin vielä läpi useita kertoja, jotta saatiin käsitys siitä, mitä on kasassa. (Hyvärinen, Nikander & Ruusuvuori 2017, 427–437.)

Kolmannessa vaiheessa litteroidusta aineistosta etsittiin tutkimustavoitteiden kannalta tärkein ja olennaisin tieto. Tämän jälkeen alkuperäisistä ilmauksista pyrittiin muodostamaan pelkistetyt ilmaisut tiivistämällä pitkät alkuperäiset lauseet omin sanoin. Tämän jälkeen näistä pelkistetyistä ilmauksista pyrittiin löytämään samankaltaisuuksia ja erilaisuuksia, jonka jälkeen pelkistetyt ilmaukset vielä ryhmiteltiin.

Tässä opinnäytetyössä pelkistetyt ilmaukset ryhmiteltiin neljään osioon, joita ovat: Rahanpesun estämisen menetelmät ja data-analytiikan hyödyntäminen, Rahanpesun estämiseen liittyvä data, Data-analytiikan hyödyntämisen kannattavuus sekä Tulevaisuuden mahdollisuudet. Näistä kaksi ensimmäistä vastaavat ensimmäiseen alatavoitteeseen ja

kaksi jälkimmäistä toiseen alatavoitteeseen. Tutkimusaineistosta pyrittiin löytämään teksti-
katkelmat ja luokittelemaan ne omiin osioihinsa. Jako neljään osioon tuki ja selvensi tutki-
mustavoitteiden saavuttamista. Seuraavaksi esitellään tutkimuksen tulokset ryhmiteltynä
näihin neljään osioon.

5.2 Tutkimuksen tulokset

Tässä alaluvussa esitellään tutkimuksen tuloksia, joita havainnollistetaan sitaatein. Haas-
tateltavat on esitelty taulukossa 2. Haastateltavat toimivat pankeissa vaihtelevasti datan ja
rahanpesun estämisen parissa. Haastateltaville kokemusta rahanpesun estämisen pa-
rissa toimimiseen on kertynyt neljästä vuodesta jopa 15 vuoteen.

Taulukko 2. Haastateltavien esittely

	Haastateltavat	Kokemus ja työnkuva
Pankki A	Head of analytics and process management Tekoälyn osaamiskes- kuksen vetäjä	Yksikön johtaminen, kyvykkyyksien rakentami- nen ja niiden toimintojen varmistaminen. Toimi- nut 10 vuotta talousrikollisuuden torjunnan pa- rissa ja kyseisessä työyhteisössä 2 vuotta. Oikeiden ihmisten ja käyttötapauksien löytämi- nen, joka parantaa kilpailukykyä: myynnin tuke- minen/asiakkaan parempi ymmärtäminen, ope- ratiivisen tehokkuuden parantaminen sekä riskin ymmärtäminen. Noin 10 vuoden kokemus alalta ja 4 vuotta kyseisessä työyhteisössä
Pankki B	AML Compliance Officer	Asiakkaan tunteminen, rahanpesun ja terroris- min rahoittamisen estäminen, compliance näkö- kulma (sääntelyn noudattaminen), liiketoiminnan ja pankkien tuki. AML-tehtävissä toiminut 15 vuotta ja kyseisessä työyhteisössä 3 vuotta.
Pankki C	Head of Regulatory Ma- nagement	Rahanpesun ja terrorismin rahoittamisen estä- misen prosessien kehittäminen sekä MiFIDD II liittyvä sijoittajansuoja ja siihen liittyvät soveltu- vuusarviointi ja asianmukaisuusprosessit. Ra- hanpesun estämisen tehtäviä takana 4 vuotta ja työkokemusta kaiken kaikkiaan 6 vuotta kysei- sessä pankissa.

Pankki D	GFCP Senior Expert	Toimii operatiivisessa ryhmässä, joka tutkii kompleksisia kokonaisuuksia liittyen mahdolliseen rahanpesuun tai terrorismin rahoittamiseen. Toiminut alalla 4 vuotta ja kyseisessä työyhteisössä noin 1,5 vuotta.
	GFCP Master Expert	Toimii rahanpesun ja terrorismin rahoittamisen eston kehitysyksikössä, jossa työtehtävät ja-kaantuvat data engineer / scientist rooleihin. Toiminut alalla 4,5 vuotta ja nykyisessä yksikössään 3 vuotta.

Seuraavaksi esitellään tutkimuksen tulokset ja niitä havainnollistetaan sitaatein. Tulokset on ryhmitelty neljään osioon: Rahanpesun estämisen menetelmät ja data-analytiikan hyödyntäminen, Rahanpesun estämiseen liittyvä data, Data-analytiikan hyödyntämisen kannattavuus sekä Tulevaisuuden mahdollisuudet.

5.2.1 Rahanpesun estämisen menetelmät ja data-analytiikan hyödyntäminen

Pankit toivat haastatteluissa ilmi, että rahanpesua ehkäistään eri vaiheissa eri tavalla, aina asiakkaan tuntemisesta lähtien ja rahanpesun estämiseen liittyy useita toimintoja. Rahanpesun estämiseen liittyy pankki B:tä edustaneen AML Compliance Officerin mukaan erityisesti lakisäätteisiä menetelmiä, joita pankeilla tulee olla käytössä. Pankki C:tä haastattelussa edustanut Head of Regulatory Management kertoi, että rahanpesun estämisen peruspilarit koostuvat asiakkaan tuntemisesta sekä jatkuvasta tuntemisesta ja seurannasta. Myös muut haastateltavana olleet pankkien edustajat toivat saman esille.

Asiakkaan tuntemiseen liittyy pankeissa erilaisia menetelmiä ja vaihteita. Nämä menetelmät vaihtelevat hieman pankkien kesken, mutta samankaltaisuuksiakin on havaittavissa. Pankki B:stä kerrottiin, että heillä asiakkaat monitoroidaan heti uutta asiakkuutta avatessa, mutta myös koko asiakaskanta monitoroidaan päivittäin pakotelistoja ja PEP-tietoja (Politically exposed person) vasten. Pankki C:n Head of Regulatory Management totesi haastattelussa, että kaikki lähtee liikkeelle asiakkaan tuntemisesta. Eli asiakkaan hakiessaan heidän asiakkaakseen, pankki C:ssä haetaan tietoa siitä, kuka asiakas on. Hän kertoi myös, että tämän lisäksi heidän on tunnistettava, miksi asiakas haluaa heidän asiakkaakseen, millaisia liiketoimia asiakas on suorittamassa sekä millainen hänen taloudellinen asemansa on suhteessa näihin liiketoimiin. Myös muut pankit toivat haastattelussa esille, että asiakkailta on kysyttävä tietoja, jotta heidän tunnistamisensa olisi mahdollista. Pankki

C:stä kerrottiin myös, että asiakkaan tuntemiseen kuuluu, että laki velvoittaa pankkia tunnistamaan asiakkaan riskit. Pankki C:ssä tunnistetaan ja luokitellaan asiakkaan riski perustuen asiakkaan antamiin tietoihin ja tuotevalikoimaan, jonka asiakas on valinnut. Korkean riskin asiakkailta pankki C kerää haastattelun mukaan tehostetusti tietoja asiakkaasta:

”Me kysytään niin kun tarkempia tietoja, että mihin hän käyttää jotain tuotetta tai jos hänellä on jotain varoja tulossa tilillensä, että mikä näiden varojen alkuperä on tai jos hän suorittaa jotain liiketoimia korkeariskisiin maihin, niin mikä niiden liiketoimintojen tarkoitus on”. (Pankki C.)

Pankki C:n Head of Regulatory Management kertoi, että yhtenä oleellisena asiana asiakkaan tuntemisessa on myös tarkistaa asiakkaan poliittinen vaikutusvaltaisuus. ”Jos asiakas on poliittisesti vaikutusvaltainen henkilö, niin silloin myös laki edellyttää meiltä, että pitää tiedustella tarkempia tietoja varojen alkuperästä”, pankki C:tä edustanut kertoi. Tähän he käyttävät apuna myös kaupallisia listoja sekä kysyvät itse asiakkailta, tunnistavatko he itsensä poliittisesti vaikutusvaltaiseksi henkilöksi. Näiden perusteella pankki C:ssä tehdään päätös asiakkaan poliittisesta vaikutusvaltaisuudesta. Pankki D:ssä rahanpesun estämisen tehtävät ovat selkeästi jaettuna eri tiimien kesken. Pankki D:n GFCP Senior Expert kertoi, että heidän pankissaan KYC-tiimin (Know your customer) tehtävänä on selvittää pankin asiakkaiden taustoja ja varmistaa, että pankki ymmärtää ketä heillä on asiakkaana.

Asiakkaan tuntemisen lisäksi toinen rahanpesun estämisen peruspilareista on jatkuva tunteminen ja seuranta, kuten pankki C:n edustaja haastattelussa totesi. Rahanpesun estämisessä olennaista on myös sekä pankki A:n että pankki B:n mukaan ymmärtää asiakasta ja asiakkaalle normaalia ja epänormaalia käyttäytymistä. Kun asiakkaat ja heidän toimensa tunnetaan, pankit pystyvät havaitsemaan poikkeavia liiketoimia. Pankki A:sta kerrottiin, että he tiedustelevat asiakkailta varojen alkuperää heidän tallettaessaan käteistä. Pankki B:ssä yhtenä rahanpesun estämisen keinona on, että asiakkaat antavat arvion omasta maksuliikenteestään, esimerkiksi arviot siitä, miten käyttävät käteistään tai onko asiakkaalla maksuja ulkomaille. Tätä asiakkaan arviota verrataan sitten asiakkaan toteutuneeseen maksuliikenteeseen, minkä tavoitteena on löytää epätavanomaisia liiketoimia. ”Epätavanomaisten liiketoimien seurantaan eli siihen itse rahanpesun estämisen seurantaan on tietynlaiset raportit ja osittain automatisoitu seuranta, mutta se on aika manuaalista vielä tällä hetkellä”, pankki B:stä kerrottiin. Myös pankki C:ssä asiakkaan antamia tietoja verrataan asiakkaan tekemiin liiketoimiin ja pyritään siten havaitsemaan poikkeavuuksia. ”Me pidetään asiakastiedot ajan tasalla muun muassa sillä tavalla, että me saadaan

säännöllisiä päivityksiä väestörekisterikeskuksesta, mutta sen lisäksi me myös monitoroidaan meidän asiakaskäyttäytymistä mahdollisten poikkeamien varalta”, pankki C:n Head of Regulatory Management kertoi. Hän jatkoi vaadittavilla lisätoimilla, mikäli asiakaskäyttäytymisessä havaitaan poikkeavuuksia:

”Jos me havaitaan asiakaskäyttäytymisessä jotain poikkeavuutta, niin silloin se tarkoittaa, että meidän pitää joko asiakkaalta pyytää lisätietoja näihin liiketoimiin liittyen tai sitten jos me pidetään näitä tapahtumia hyvinkin epäilyttävänä, niin sitten meillä on velvoite joko pysäyttää nämä liiketoimet tai sitten tehdä ilmoitus Keskusrikospoliisin rahanpesun selvittelykeskukseen”. (Pankki C.)

Pankki D:ssä jatkuva tunteminen ja seuranta on jaettu eri tiimien kesken seuraavanlaisesti: Transaktio ja sanktio –monitorointi ja tutkinta tutkii ja raportoi tiettyjen raja-arvojen ja sääntöjen mukaan selvittelyyn nostettuja tapauksia ja FCIA-yksikkö (Financial Crime Intelligence and Analytics), jossa haastateltavat itse työskentelevät, selvittää ja toteuttaa analytiikkaan ja temaattisiin kokonaisuuksiin liittyviä projekteja, joiden avulla esille nostetaan mahdollisesti riskisiä asiakkaita, joita yksikön sisällä oleva kompleksinen tutkinnan ryhmä tutkii ja tarvittaessa raportoi.

Haastatteluista tuli esille, että data-analytiikan hyödyntämiselle on pankeissa suuri tarve.

Pankki A:n edustajat toivat ilmi data-analytiikan tarpeen rahanpesun estämisessä:

”Transaktio volyymien ollessa niin suuria, dataa ja järjestelmiä on hyödynnettävä, sillä ihminen ei pystyisi millään käymään kaikkia läpi” ja ”Suurista asiakasmääristä johtuen, ei ole realistista, että pystytään kaikkea käsiparein seuraamaan, jolloin tietysti täytyy olla järjestelmät tukemassa tätä toimintaa”. Samalla kannalla oli myös pankki C: ”Data ja analytiikka on niin kun aivan olennainen osa tätä työtä nykypäivänä ja ilman tätä data-analytiikkaa me ei käytännössä tehokkaasti pystyttäis rahanpesua estämään”. Pankki B:ssä rahanpesun estäminen on toistaiseksi vielä melko manuaalista, mutta hekin ovat kuitenkin huomanneet tarpeen data-analytiikan lisäämiselle ja ovatkin lähitulevaisuudessa kehittämässä järjestelmiään. ”Ollaan kehittämässä näitä järjestelmiä ja meillä on iso järjestelmähanke meneillään”, pankki B:n AML Compliance Officer mainitsi.

Se, missä määrin pankit hyödyntävät data-analytiikkaa rahanpesun estämisessä, vaihtelee haastateltujen pankkien välillä paljonkin. Pankki A:lla on käytössään laajasti hyvin erilaisia data-analytiikka menetelmiä. ”Transaktion osapuoliin, asiakkaaseen ja kontekstiin liittyä kaikkiin omat menetelmänsä ja ne muodostavat loogisen kokonaisuuden, jossa on asiakas, transaktio ja sen osapuolet ja kaikki siihen liittyvä, data, aika jne.”, pankki A:n tekoälyn osaamiskeskuksen vetäjä kuvaili. Pankki A:ssa menetelmien valinta riippuu ensisi-

jaisesti käyttötarkoituksesta, selitettävyydestä ja siitä, mitä menetelmällä halutaan saavuttaa. Pankki A:n tekoälyn osaamiskeskuksen vetäjä kuvaili datan, automaation ja ihmisen työnjakoa siten, että data auttaa ihmistä päätöksenteossa tehden monimutkaisista asioista ja transaktioista ymmärrettäviä, jolloin tutkijan aika on käytettävissä itse päätöksentekoon ja johtopäätöksiin. "Malli tai tekoäly ei itsessään välttämättä päätä moniakaan asioita, vaan ihminen tekee edelleen aika monessa asiassa päätöksen, mutta sitten taas järjestelmän pitäisi tehdä ihmisen työ jotenkin siedettäväksi ja tehokkaaksi", pankki A:n tekoälyn osaamiskeskuksen vetäjä kommentoi. Pankki A:sta kerrottiin, että heillä tutkijat käyvät läpi liiketoimet yksitellen, joista järjestelmä antaa hälytyksen ja tarkastaa, onko liiketoimessa jotain epäilyttävää. "Kone pystyy tekemään tiettyyn pisteeseen olettamukset, että tietyt kriteerit täyttyvät, jolloin epäilyttävän oloinen tai asiakkaan normaaliin toimintaan poikkeava liiketoimi olisi kyseessä, mutta se voi tietysti olla täysin luonnollinen ja asiaankuuluva toiminto", pankki A:n analytiikka ja prosessien hallinta yksikön päällikkö huomautti. Isot hankinnat, kuten asuntokaupat, jolloin rahaa liikkuu huomattavasti enemmän kuin normaalisti aiheuttavat järjestelmiin niin sanottuja virheellisiä hälytyksiä, joiden seasta tutkijoiden on löydettävä oikeat rahanpesuepäilykset.

Pankki B:ssä rahanpesun estäminen on puolestaan vielä melko manuaalista, mutta heillä data-analytiikkaa hyödynnetään esimerkiksi epätavanomaisten liiketoimien havaitsemiseen jossain määrin. Heillä data-analytiikkaa käytetään erityisesti asiakasmonitoroinneissa, jossa etenkin PEP-asiakkaat ovat iso haaste pankille.

Pankki C:n Head of Regulatory Management kertoi, että heillä data-analytiikkaa käytetään esimerkiksi asiakkaan riskiluokitteluun ja poikkeamien havaitsemiseen asiakaskäyttäytymisessä. Heillä ei kuitenkaan varsinaisesti vielä ole esimerkiksi koneoppimisen menetelmiä, joilla pystyisivät automatisoimaan prosesseja, koska kokevat sen toistaiseksi riskialttiiksi. He ovat kuitenkin rakentaneet sille valmiuksia ja seuraavat sen mahdollisuuksia. Myös pankki C:stä kerrottiin, että kaikkea rahanpesun estämisen prosessissa ei voida automatisoida, ennen kaikkea asiakkaan antamien selvitysten käsittely ja arviointi tehdään C:ssä ihmisen toimesta.

Pankki D:ssä on puolestaan käytössä nykyaikaiset työkalut rahanpesun estämiseen. Pankki D:n GFCP Master Expert kertoi, että heidän pankissaan on käytössä Hadoop-sateenvarjon alla olevat teknologiat, esimerkiksi Apache Spark, jolla pankki tekee niin sanottu preprossesing-vaiheet. Tämän päällä analyttikot voivat käyttää esimerkiksi R/Python -kieliä ja niiden kirjastoja. "Lähtökohtaisesti analytiikan tarkoitus on tuottaa tarkempaa ymmärrystä epäilyttävästä rahaliikenteestä ja komplementoida olemassa olevia monitorointijärjestelmiä", pankki D:n GFCP Master Expert kertoi.

5.2.2 Rahanpesun estämiseen liittyvä data

Haastatteluista tuli ilmi, että jokaisella pankilla, koosta riippumatta, asiakkaaseen ja rahanpesun estämiseen liittyvää dataa kertyy suuret määrät. Pankit kertoivat, että dataa kertyy asiakkaista, transaktioista, korttitapahtumista, KYC-tiedoista ja verkkopankkikäynneistä. ”Tämän lisäksi rahanpesuun liittyvässä analytiikassa voidaan hyödyntää myös kansainvälisiä yritystietokantoja ja muita mahdollisia rekistereitä”, pankki D:n GFCP Senior Expert kertoi. Pankki A:sta kerrottiin myös, että pankin omien tietojen lisäksi asiakkaan tietoja on mahdollista saada ulkopuolisilta tarjoajilta, esimerkiksi erilaisilta henkilötietorekistereiltä. Tällaiset ulkopuoliset tarjoajat saattavat tarjota tietoja esimerkiksi asiakkaan luottotiedoista tai yritykseen liittyvistä omistussuhteista. Pankki C:n Head of Regulatory Management toi lisäksi haastattelussa esille, että on olemassa konsulttiyrityksiä ja muita palveluntoimittajia, joilta on mahdollista ostaa verrannollista dataa, mutta näkee tässä ongelmia. ”Usein siinä on ongelma se, että se verrannollinen data tulee esimerkiksi Keski-Euroopasta, ja Keski-Euroopan asiakaskäyttäytyminen ei ole verrannollista Suomen asiakaskäyttäytymiseen, koska siellä käytetään esimerkiksi käteistä todella paljon enemmän”, hän kertoi. Näitä kaikkia tietoja pankit voivat sitten hyödyntää rahanpesuun ja terrorismin rahoittamiseen liittyvässä analytiikassa sekä esimerkiksi erilaisiin raportteihin, tilastoihin ja keskiarvopoikkeamiin.

Pankit toivat haastatteluissa ilmi, että asiakkaasta saataviin tietoihin liittyy jonkin verran käyttörajoituksia, mutta laki antaa kuitenkin verrattane laajan mahdollisuuden käyttää dataa rahanpesun estämiseen. Pankki D:stä nostettiin esille, että asiakkaan tietojen käyttöön liittyy lain määrittämiä käyttörajoituksia muun muassa GDPR ja oikeus käyttää rikostuomiotietoja. Silloin kun pankista tutkitaan ja epäillään rikollista toimintaa, niin silloin on laajempi keinovalikoima datan hyödyntämisessä käytössä. ”Jos on syytä epäillä tai halutaan sulkea jokin epäilyttävä pois, niin silloin meillä on keinot käytettävissä ja se on täysin luvalista tällaisessa tilanteessa käsitellä asiakkaan tietoja, koska silloin ollaan niin sanotusti oikealla asialla”, pankki A:n analytiikka ja prosessien hallinta yksikön päällikkö kommentoi pankin keinoja käyttää asiakkaasta saatavia tietoja. Pankki B:n AML Compliance Officer huomautti, että laki säätelee osaa tiedoista, joita ei voida käyttää kuin vain juuri rahanpesun estämisen tarkoitukseen. AML Compliance Officer antoi esimerkiksi PEP-tiedon, joka voidaan kysyä vain lain määäämissä puitteissa. Pankki B:n AML Compliance Officer kommentoi:

”Asiakkaan tuntemistietoja ja muita henkilötietoja voidaan käyttää rahanpesun ja terrorismin rahoittamisen estämiseen, paljastamiseen, selvittämiseen ja tutkintaan

saattamista varten. Jos nämä tiedot on hankittu ainoastaan rahanpesun ja terrorismin rahoittamisen estämiseksi ja paljastamiseksi, niitä ei käytetä muuhun näiden kanssa yhteen sopimattomaan tarkoitukseen, esimerkiksi markkinointiin.” (Pankki B.)

Pankki C:n Head of Regulatory Management nosti haastattelussa esiin esimerkkejä, joita ei voida käyttää edes rahanpesun estämistarkoituksessa. ”Tiedot liittyen asiakkaan terveydentilaan on sellaisia, joita me ei voida käsitellä edes tässä tarkoituksessa” ja ”toinen on sitten taas tällaista diskriminoivaa tietoa. Tietenkään mitään etnistä profilointia me ei saada tehdä, koska se on laitonta”.

Pankki A:sta kerrottiin, että asiakkaasta saatavia tietoja käytetään tarpeen mukaan ja yritetään saada kokonaiskuva niin asiakkaasta kuin transaktiostakin. Kuten data-analytiikka menetelmän valinnassakin, hyödynnettävän datan valinta riippuu pankki A:ssa yksittäisen rahanpesun estämisen käyttötapauksesta ja siitä, mitä halutaan etsiä. Pankki B:n AML Compliance Officer totesi, että oleellisin data rahanpesun estämisen seurannassa on asiakkaiden tilitapahtumat, jotka ovat aina saatavilla. Pankki B:llä on käytössään Data warehouse -raportointijärjestelmä, jonne tiedot kerätään ja sitä kautta tekevät analytiikkaa. Pankki B:n data on ulkopuolisessa IT-talossa, joka toimittaa datan pankki B:lle heidän sitä tarvitessaan. Tämä hieman hankaloittaa pankki B:ssä datan saantia, sillä esimerkiksi historiatietoja pitää hakea erillisiltä raporteilta, jos dataa tarvitaan pidemmältä ajalta. Ulkopuoliselta IT-talolta pitkän ajan dataa pankki B:ssä joudutaan tällä hetkellä hakemaan datapöimintatilausten kautta, joka ei ole kovin kustannustehokasta eikä aina niin nopea prosessi. Myös datan käsittelyn ja raportoinnin osalta pankissa B on laajempi projekti meneillään. Pankki D:n GFCP Master Expert kertoi datan hyödyntämisestä rahanpesun estämiseksi:

”Yleensä asetelma on seuraavanlainen: ensiksi muodostetaan teorialason ymmärrys tutkittavasta ilmiöstä ja sitä pyritään validoida datalla. Sama asia voidaan tehdä myös käänteisesti, jolloin data luokitellaan ensiksi ja annetaan algoritmien oppia itse datasta. Tämän jälkeen tulokset pyritään selittämään ihmisjärjen avulla”. (Pankki D.)

Pankki D:stä nostettiin myös esille, kuten pankki A:stakin, että datan hyödyntämisessä on tärkeää hyödyntää useita eri työkaluja, jotta ilmiöitä olisi mahdollista ymmärtää.

5.2.3 Data-analytiikan hyödyntämisen kannattavuus

Haastatteluista tuli ilmi, että data-analytiikan hyödyntäminen on iso kustannuserä pankeille ja siitä aiheutuu monenlaisia kustannuksia aina hankintakustannuksista henkilöstökustannuksiin. Data-analytiikasta aiheutuu pankeille yksittäisiä kertakustannuksia, kuten hankintakustannukset, mutta myös jatkuvia kustannuksia, kuten lisenssikustannukset. Sekä pankki A:n edustajat että pankki C:n edustaja toivat haastatteluissa esille, että tiedon tallentamisen ja prosessoinnin kustannukset ovat iso kuluerä pankeille. Tämä sen takia, koska asiakasta ja transaktiota on ymmärrettävä, jolloin myös datan määrä on valtava. Koska pankki B:ssä IT-talo on pankin ulkopuolinen taho, kokevat he data-analytiikan hyödyntämisen hyvinkin kalliiksi, erityisesti datan keräämisen osalta:

”Tämä on itseasiassa aika kallista, kun meillä IT-yhtiö on ulkopuolella, jonka takia tavallaan ostetaan omaa dataa muualta ja tietenkin kun meillä ei itsellään varsinaisesti ole data-analytiikkaa ainakaan tälle osa-alueelle tällä hetkellä, niin joudutaan ostamaan näitä palveluita muualta”. (Pankki B.)

Pankki C:n Head of Regulatory Management puolestaan kertoi, että henkilöstökustannukset ovat yksi suurimmista kuluista, kun analytikot ovat omasta takaa. ”Meillä pitää olla henkilöitä, jotka rakentaa sekä meidän analytiikka malleja, että sitten analysoi niitä analytiikan lopputuotoksia... ja on ihmisiä, jotka käsittelee esimerkiksi näitä asiakkailta tulevia vastauksia ja arvioi niiden luotettavuutta”, pankki C:n Head of Regulatory Management kertoi.

Erityisesti pankki C:n Head of Regulatory management nosti haastattelussa esille toisen kannattamattomuuteen vaikuttavan seikan kustannusten lisäksi. Vaikka dataa kertyy pankeille valtavia määriä, koosta riippumatta, kokevat he, että dataa voisi olla vielä enemmän. Pienenä toimijana pankki C toivoisi erityisesti, että heillä olisi enemmän dataa yrityspuolella, jossa heidän otantansa on suhteellisen pieni. C:n edustaja toi esille, että ennen kaikkea tekoälyn ja koneoppimisen hyödyntäminen tehokkaasti vaatisi paljon dataa - mitä enemmän dataa on käytettävissä niin sitä paremmin ne toimivat. Koska dataa ei ole riittävästi saatavilla, kokevat he juuri esimerkiksi koneoppimisen ja big datan hyödyntämisen kannattamattomaksi.

Kaikista haastatteluista kävi ilmi, että data-analytiikka tuo valtavasti lisäarvoa rahanpesun estämiseen. Pankit kokevat, että data-analytiikan hyödyntäminen tuo pankeille etenkin tehokkuutta ja kustannussäästöjä rahanpesun estämisen parissa toimimiseen. Data-analytiikan avulla on mahdollista vähentää henkilöstökustannuksia, mutta ala vaatii kuitenkin edelleen jossain määrin henkilötyövoimaa, sillä kaikkea data-analytiikankaan avulla ei

pystytään automatisoimaan. Pankki A:n tekoälyn osaamiskeskuksen vetäjä kommentoi data-analytiikan tuomia hyötyjä suhteessa henkilöstökustannuksiin: "Mitä paremmin ymmärrät asiakasta ja transaktiota, sitä vähemmän tarvitaan ihmisen tekemää tarkastustyötä". Pankki A:n analytiikan ja prosessien hallinta yksikön päällikkö jatkoi data-analytiikasta saatavilla lisäarvoilla: "Data-analytiikka rajoittaa sitä määrää merkittävästi, mihin me voidaan keskittyä, jolloin ihmisen työvoima voidaan käyttää sellaiseen, missä ihmisen päätöksentekoa tarvitaan". Data-analytiikka mahdollistaa siis myös ihmisten osaamisen kehittymisen, kun työtehtävät ovat entistä mielekkäämpiä ja vaihtelevampia koneiden tehdessä rutiininomaiset työt. Vaikka data-analytiikan hyödyntäminen on pankki B:ssä vielä melko vähäistä, kokevat he myös data-analytiikan vähentäneen työtunteja niissä toiminnoissa, joissa sitä on jo hyödynnetty, kuten asiakas- ja pakotemonitoroinneissa. Pankki B:n edustaja kommentoi: "Tietenkin sitä työaikaa ja henkilöiden resursseja vapautuu muuhun ja jää sellainen turhien asioiden manuaalinen tekeminen, jolloin prosessit tehostuvat". Pankki D:stä kerrottiin, että analytiikan ideana on lähtökohtaisesti tehdä asiat skaalautuvasti, joka tuo tehokkuutta verrattuna manuaaliseen asiakastasolla tehtävään käsittelyyn. Analytiikka lisää pankeissa tehokkuutta, väärin positiivisten ilmoitusten vähentyessä, ja sitä kautta tuo pankeille kustannussäästöjä, kun toimintoihin vaaditaan vähemmän henkilöstöä. Pankki D:stä korostettiin, että tästä huolimatta pankin rahanpesun estäminen on korkeatasoista ja lain vaatimalla tasolla. Pankki A:sta kommentoitiin samaa:

"Kun prosessissa pystytään jotakin automatisoimaan, niin siinä tulee tietenkin luonnollisesti kustannussäästöjä, mutta toinen aspekti siihen on myös se, että asia mikä me halutaan saavuttaa, saavutetaan jotenkin automatisoituna, tavasta riippumatta, niin se tulee tasalaatuisena. Eli siihen ei tule sellaista inhimillisen virhetekijän mahdollisuutta tai yksittäisen tutkijan osaamisesta riippuvaa laatupoikkeamaa". (Pankki A.)

Tehokkuuden ja kustannussäästöjen lisäksi pankki D:ssä nähdään data-analytiikan hyödyiksi se, että se mahdollistaa erilaisien mallien kehittämisen ja hyödyntämisen rahanpesun torjunnassa. He kertoivat, että rahanpesun ja terrorismin rahoittaminen ovat jatkuvassa muutoksessa ja muutokseen vastaamisessa tulee pystyä toimimaan työkaluilla, jotka ovat muokattavissa muutoksen mukaan. "Data-analytiikka tuo mahdollisuuden ketterämpään muutokseen ja syvempään ymmärrykseen epäilyttävästä toiminnasta", pankki D:n GFCP Senior Expert mainitsi. Myös pankki C:n Head of Regulatory Management kertoi, että data-analytiikka on helpottanut datan jalostamista ja sitä kautta helpottaa datan käyttöä erilaisissa analytiikka malleissa. Pankki A:sta tuotiin myös esille, että datan analysoinnin avulla tiedetään entistä enemmän tapahtumista, nähdään erilaisia käyttäytymismalleja asiakkailta ja maksuissa, jonka avulla puolestaan pystytään säätämään erilaiset

automaatiikat ja järjestelmät tarkemmalle tasolle, jotta pankki saisi positiivisia hälytyksiä vähemmän ja saatavat hälytykset olisivat mahdollisimman laadukkaita. Pankki A:n tekoälyn osaamiskeskuksen vetäjä toi esille myös, että rahanpesun estämisen tulisi näkyä mahdollisimman vähän asiakkaan arjessa, silloin kun asiat ovat kunnossa. Data-analytiikka mahdollistaa asiakkaiden mahdollisimman vähäisen häirinnän rahanpesun estämiseen liittyvillä asioilla. Data-analytiikan avulla pankki A kertoi lisäksi täyttävänsä muun muassa viranomaisten vaateet, mutta myös saavuttavansa pankin omat asettamat tasot.

Jokainen haastateltu pankki kokee data-analytiikan hyödyntämisen rahanpesun estämisessä erittäin kannattavana. Tästä kertoo myös pankki B:n suunnitelmat lisätä data-analytiikkaa rahanpesun estämisessä. "Til tapahtumia on miljoonia, niin niiden analysointi on mahdotonta ilman analytiikkaa", pankki B:n AML Compliance Officer kommentoi. "Teknologiset investoinnit ja työvoimakustannukset ovat suuria, mutta samalla pankille tuotettava lisäarvo on merkittävä ilmiöiden selittämisessä ja lakisääteisen monitoroinnin tehostamisessa", pankki D:n GFCP Master Expert kertoi. Pankki C:n Head of Regulatory Management kommentoi rahanpesun estämisen kannattavuutta:

"Rahanpesun estäminen on sellainen, joka ei suoraa tuo euroja sisään, et se on periaatteessa kustannus pankeille, mutta toki niin kun käytännössä se on todella tärkeä osa meidän yhteiskuntavastuuta ja todella tärkeä osa meidän mainetta ja todella tärkeä osa myös sitä, et me ei jouduta maksamaan mahdollisia sakkoja...data-analytiikka on edellytys siihen, että meillä on tehokkaat rahanpesun estämisen menetelmät niin se on myös kannattava investointi meille". (Pankki C.)

Pankki A kokee data-analytiikan hyödyntämisen jopa pakollisena nykypäivänä ja tulevaisuudessa. Samalla kannalla on myös pankki C:n edustaja: "Se on mun mielestä melkein pakko käyttää data-analytiikkaa, jotta rahanpesua pystyttäisiin tehokkaasti estämään". Pankki A:n edustajat kommentoivat data-analytiikan kannattavuudesta seuraavanlaisesti:

"Ei tää ilmiönä enää tuu menemään pois. Ihmisen tekemä tarkastustyö ei suinkaan ole sellainen evoluution huipentuma tässä, vaan niin kun näiden kahden symbioosi. Ihmisen tekemä tarkastelu ja arviointi ja sitten taas sellaisien työtehtävien, joita kukaan ei kaipaa, niiden mahdollisimman iso automaatio". (Pankki A.)

5.2.4 Tulevaisuuden mahdollisuudet

"Datan laatu, skaalautuvuus ja eri tietolähteiden yhdistäminen on suurin haaste ja samalla mahdollisuus. Tämä ongelma on käytännössä identtinen kaikilla toimialoilla tehtävässä analytiikassa", pankki D:n GFCP Master Expert kertoi data-analytiikan mahdollisuuksista. Pankki B toi haastattelussa puolestaan pitkän listan data-analytiikan mahdollisuuksista,

joita ovatkin joutuneet miettimään lähiaikoina paljonkin. Esille nousi esimerkiksi todennäköisyyksien helpompi laskettavuus, koneoppimisen ja fuzzy logicin hyödyntäminen suuremmissa määrin, suuremman datamäärän käsittely ja siitä informaation jalostaminen nopeammin manuaalitoimintaan verrattuna, hyödyntäminen asiakkaiden riskiluokitteluun ja profilointiin sekä analysoinnin käyttö testaukseen ja kehitykseen.

Sekä pankki A:n, B:n että C:n kanssa käydyissä haastatteluissa yhdeksi isoksi pankkien näkökulmasta kiinnostavaksi muutokseksi nousi pankkien välinen yhteistyö, esimerkiksi vertailu koko Suomen pankkisektorin dataan nähdään mielenkiintoisena. "Olisi loistavaa, jos voisi jakaa dataa tietysti nimettömänä tai anonyyminä eri pankkien kesken, koska tähän koskee kaikkia" ja "Kuitenkin kun kaikilla on sitä omaa dataa, mutta jos olisi vaikka joku instanssi, joka voisi koota dataa, jota kaikki voisivat käyttää esimerkiksi testaukseen ja tällaiseen, ja hyödyntää sitä sitten", pankki B:n AML Compliance Officer ajatteli. "Se on yks tällanen seikka, joka auttaisi tätä toimintaa, että pystyttäisiin keskenään sitä tietoa jakamaan ja pystyttäis sitä big data hyötyä saamaan... uskon, että se auttaisi kaikkia tämän alan toimijoita", pankki C:n Head of Regulatory Management kertoi. Hän jatkoi jakamisen tuomilla hyödyillä:

"Toinen on sitten tosiaan se, että se myös helpottais näitä koneoppimismallien käyttöä, koska ennen kaikkea tällanen tekoälyn ja koneoppimisen käyttäminen tehokkaasti vaatii paljon dataa, eli mitä enemmän sitä on käytettävissä niin sitä parempi ja sitä luotettavampaa siitä saadaan". (Pankki C.)

Tämän haasteena on kuitenkin pankkien mielestä sääntely ja pankkien haluttomuus paljastaa muille omia ideoitaan ja menetelmiään. Se, missä määrin pankit pystyvät tietoa tietystä asioista jakamaan pankkien kesken on hyvin vähäistä. Pankit voivat esimerkiksi kysyä kilpailevalta pankilta tietoja yksittäisestä maksusta tai transaktiosta. "Esimerkiksi jos saajassa tai lähettäjässä, joka on toisen pankin asiakas, on jotakin sellaista, saatetaan kysyä mikä on maksun tarkoitus tai taustoja, onko tämä asiakas tuo ja niin edespäin", pankki A:n analytiikka ja prosessien hallinta yksikön päällikkö kertoi tämänhetkisestä pankkien välisestä yhteistyöstä. Mitään niin sanottua mustaa listaa korkean riskin asiakkaista pankit eivät siis voi jakaa keskenään, mikä voisikin hyödyttää pankkeja esimerkiksi uutta asiakkuutta avatessa, jos tällainen olisi mahdollista. Pankki A:n analytiikka ja prosessien hallinta yksikön päällikkö kommentoi:

"Kyllähän sitä on tällä hetkellä erilaisia hankkeita menossa, esimerkiksi Pohjoismaissa, mitkä tähtäävät juuri siihen, että esimerkiksi tiettyjä asiakkaita tuntemisen

asioita pysyttäisiin tekemään yhteistyössä, mutta se tietenkin vaatii merkittävää panostusta asiakkaan tietosuojan näkökulmasta ja laillisesta näkökulmasta”. (Pankki A.)

Pankki A:n analytiikka ja prosessien hallinta yksikön päällikkö arvioikin, että tämä on yksi mahdollinen suunta, johon oltaisiin pikkuhiljaa tulevaisuudessa menossa: “Kukaan ei enää pärjää yksin, varmasti enenevissä määrin halutaan tehdä yhteistyötä ja tämä alue on kaikille haaste ja maksaa paljon rahaa kaikille finanssilaitoksille hoitaa se velvollisuus ja tehdä asiat kunnolla, siinä mielessä yhteistyö on yksi ratkaiseva tekijä tässä”. Pankki C:n edustaja ei näe lainsäädännön sallivan tätä kuitenkaan ihan heti vaan näkee muutokset pidemmällä horisontissa.

Pankki B:n kanssa käydyssä haastattelussa toinen iso esille noussut asia on tarve saada enemmän tietoa rahanpesun selvittelykeskukselta, joka helpottaisi myös analysointia. AML Compliance Officer pankki B:stä kommentoi rahanpesun selvittelykeskukselta saamia tietoja:

“Ei saada yksittäisten asioiden kohdalla tietää, mitkä ilmoitukset ovat johtaneet johonkin. Toki saadaan raportteja palautteina, että minkälainen prosenttiosuus on johtanut johonkin tutkittavaan juttuun tai tämän tyyppistä, mutta se ei kohdennu esimerkiksi tietyn tyyppisiin tapauksiin, niin niistä on aika vaikea tietää, mikä on niin sanotusti ollut sellainen onnistunut ilmoitus”. (Pankki B.)

Pankki B toi ilmi, että mikäli tietoja saataisiin tarkemmalla tasolla, esimerkiksi, mikäli saataisiin tietää, missä ilmoituksista on ollut kyse rahanpesusta, helpottaisi se data-analytiikkaa ja rahanpesun estämistä. Sen avulla pankki pystyisi profiloimaan asiakkaitaan tarkemmin: “Voitaisiin helpommin laskea todennäköisyyksiä, että jos on tietynlainen profiili jollakin henkilöllä juuri niiden meillä olevien tietojen ja tilitapahtumien osalta, että kuinka todennäköisesti tällaisen henkilön tekemä toimi tulisi johtamaan ilmoitukseen.”

6 Pohdinta

Tämän opinnäytetyön viimeisessä luvussa tarkastellaan tutkimuksen tuloksia tietoperustaan peilaten ja esitellään niistä tehdyt johtopäätökset. Tämän jälkeen pohditaan tutkimuksen luotettavuutta ja eettisiä näkökohtia ja lopuksi arvioidaan opinnäytetyöprosessia ja omaa oppimista opinnäytetyöprosessin aikana.

6.1 Tulosten tarkastelu

Tämän alaluvun tarkoituksena on tarkastella tutkimuksen tuloksia ensin tutkimukselle määriteltyjen alatavoitteiden kannalta ja lopuksi johtaa näistä alatavoitteisiin saaduista vastauksista vastaus itse päätavoitteeseen. Opinnäytetyön päätavoitteena oli selvittää, miten pankit voisivat hyödyntää data-analytiikkaa tehokkaasti ja kannattavasti rahanpesun estämisessä. Päätavoitteesta johdettiin alatavoitteet, joiden tavoitteena oli selvittää, miten pankit hyödyntävät data-analytiikkaa rahanpesun estämiseksi ja miten kannattavana data-analytiikan hyödyntäminen rahanpesun estämisessä koetaan pankeissa. Tutkimuksen tuloksia tarkastellaan peilaten tietoperustaan ja pohditaan muun muassa tulosten merkitystä ja yllättävyyttä.

- Miten pankit hyödyntävät data-analytiikkaa rahanpesun estämiseksi?

Tutkimuksen perusteella pankkeja yhdistävänä tekijänä on suuri tarve data-analytiikan hyödyntämiselle. Sen hyödyntämistä pidetään pankeissa jopa edellytyksenä rahanpesun estämisen parissa toimimiseen. Tässä kuitenkin yllättää se, että pankit ovat hyvin eri vaiheissa sen hyödyntämisessä, vaikka data-analytiikan hyödyntäminen koetaankin hyvin tarpeellisena. Tähän voi olla syynä esimerkiksi pankin koko tai se, että aihe on vielä hyvin tuore, eivätkä kaikki pankit ole täten vielä ehtineet ottaa data-analytiikkaa käyttöön täysimääräisesti. Tutkimuksessa nousee esille, että rahanpesun estämisen prosessissa moni toiminnoista on vielä manuaalista. Erityisesti päätöksenteko on vielä sellainen, jota pankeissa ei ole lähdetty automatisoimaan. Sama todettiin myös Institute of International Financen vuonna 2018 tekemässä tutkimuksessa, jota esiteltiin luvussa 4.

Tutkimuksessa nousee esille rahanpesun estämisen peruspilarit, jotka ovat asiakkaan tunteminen sekä jatkuva tunteminen ja seuranta. Tutkimuksesta selviää, että pankeissa jokaiseen rahanpesun estämiseen liittyvään toimintoon liittyy omat data-analytiikan menetelmänsä ja valinta menetelmien välillä riippuu käyttötarkoituksesta, selitettävyydestä ja siitä, mitä sillä halutaan saavuttaa. Samankaltaisia huomioita nousi esille myös tietoperustassa. Lisäksi sekä tämän opinnäytetyön tutkimuksen että tietoperustassa esitellyn tutki-

muksen perusteella useamman menetelmän hyödyntäminen tekee rahanpesun estämisestä tehokkaampaa. Tutkimuksen perusteella asiakkaan tuntemiseen liittyy pankeissa erilaisia menetelmiä ja vaihteita. Nämä menetelmät vaihtelevat hieman pankkien kesken, mutta samankaltaisuuksiakin on havaittavissa. Myös jatkuvaan tuntemiseen ja seurantaan liittyvät menetelmät ovat pankeissa osittain samankaltaisia, kuten epäilyttävien liiketoimien havaitseminen vertaamalla liiketoimia asiakkaan antamiin arvioihin. Haastatteluiden perusteella data-analytiikkaa käytetään enemmän jatkuvaan tuntemiseen ja seurantaan kuin asiakkaan tuntemiseen, johon liittyy vielä paljon manuaalisia työtehtäviä, kuten asiakkaan tietojen tiedustelu. Pankkien kertoma rahanpesun estämisen prosessi noudattelee pääpiirteittäin lakia rahanpesun ja terrorismin rahoittamisen estämisestä. Pankit ovatkin varmasti tehostaneet rahanpesun estämistään julkisuudessa olleiden uutisten ja pankeille myönnettyjen sakkojen painottamana.

Tutkimuksessa nousee esille samankaltaisia käyttökohteita data-analytiikan hyödyntämiselle, kuten tietoperustassa esitellyissä tutkimuksissa ja artikkeleissakin. Tutkimuksesta voidaan havaita, että pankit hyödyntävät data-analytiikkaa esimerkiksi asiakkaiden riskiluokitteluun, epäilyttävien liiketoimien havaitsemiseen ja täten myös virheellisten positiivisten hälytysten vähentämiseen sekä tekemään automatisoitavissa olevat työtehtävät ihmisten puolesta, tehden ihmisten työn tehokkaaksi ja siedettäväksi, kohdistuen työtehtävät itse päätöksentekoon. Data-analytiikkaa käytetään siis tutkimuksen perusteella pankeissa päätöksenteon tukena, kuten tietoperustassakin kerrottiin. Tutkimuksen tuloksissa yllättää kuitenkin se, että tietoperustaa varten tutustumissani tutkimuksissa ja artikkeleissa, jotka käsittelivät data-analytiikan hyödyntämistä rahanpesun estämisessä, data-analytiikan käyttökohteita on laajemmin ja monipuolisemmin. Tietenkin voi olla, että nämä eivät vain syystä tai toisesta nousseet haastatteluissa esille, vaikka data-analytiikkaa hyödynnettäisiinkin pankeissa vielä laajemmin. Data-analytiikalla on kuitenkin olemassa hyvin laajasti erilaisia käyttökohteita, joten niiden kaikkien luetteleminen olisi lähes mahdotonta. Tutkimuksen tuloksissa yllättää myös se, miten vähän koneoppimista hyödynnetään pankeissa ja kuinka riskialttiina se koetaan. Tietoperustaa kootessani sain kuvan, että se toisi paljon lisäarvoa pankeille. Mikäli pankit hyödyntäisivät koneoppimista enemmän, tekisi se data-analytiikan hyödyntämisestä rahanpesun estämisessä kannattavammaksi tietoperustassa esitelyjen lähteiden perusteella.

Tutkimuksen tulosten perusteella kaikille pankeille, koosta riippumatta, dataa kertyy suuria määriä ja sitä on mahdollista saada myös ulkopuolisilta tarjoajilta. Haastatteluista nousi kuitenkin esille, että dataa voisi olla vielä enemmän, erityisesti verrannollista dataa. Datan laajempi saatavuus helpottaisi data-analytiikan hyödyntämistä, erityisesti koneoppimisen ja big datan osalta, joka osassa pankeista oli vielä hyödyntämättä tai oli hyvin vähäistä.

Samankaltaisia huomioita nousi myös esille erilaisissa tutkimuksissa ja artikkeleissa, joita tietoperustassa on esitelty.

- Miten kannattavana data-analytiikan hyödyntäminen rahanpesun estämisessä koetaan pankeissa?

Tutkimuksen tulosten perusteella pankit pitävät data-analytiikan hyödyntämistä rahanpesun estämisessä erittäin kannattavana ja jopa edellytyksenä rahanpesun estämiseen liittyvän sääntelyn noudattamiseksi. Tutkimustuloksissa esille nousee kuitenkin, että data-analytiikan hyödyntämisessä rahanpesun estämisessä on edelleen asioita, joiden kehittäminen lisäisi sen kannattavuutta.

Tutkimuksen perusteella data-analytiikan hyödyntäminen on kaiken kaikkiaan pankeille erittäin mittava investointi. Sen hyödyntämisestä aiheutuu monenlaisia kustannuksia aina hankintakustannuksista henkilöstökustannuksiin. Data-analytiikasta aiheutuu yksittäisiä kertakustannuksia, kuten hankintakustannukset, mutta myös jatkuvia kustannuksia, kuten lisenssikustannukset.

Toinen kannattamattomuuteen vaikuttava seikka, joka tutkimuksessa nousee esille, on se, että dataa voisi olla vielä enemmän, erityisesti verrannollista dataa, vaikka sitä kertyykin pankeille valtavia määriä. Tutkimustulosten perusteella suurempi määrä dataa mahdollistaisi pankeille tehokkaamman ja luotettavamman big datan ja koneoppimisen hyödyntämisen. Tutkimuksessa nousee esille, että ennen kaikkea tekoälyn ja koneoppimisen käyttäminen tehokkaasti vaatisi paljon dataa. Mitä enemmän dataa on käytettävissä, sitä luotettavampaa koneoppimisen hyödyntäminen on. Erityisesti luvussa 2.6 on nostettu esille samankaltaisia huomioita. Tutkimustulosten perusteella pankeissa toivottiin erityisesti verrannollista dataa.

Tutkimuksessa nousee esille, että pankit kokevat data-analytiikan hyödyntämisen tuovan pankeille etenkin tehokkuutta ja kustannussäästöjä rahanpesun estämisen parissa toimimiseen. Analytiikka lisää pankeissa tehokkuutta, väärin positiivisten ilmoitusten vähentyessä, ja sitä kautta tuo pankeille kustannussäästöjä, kun toimintoihin vaaditaan vähemmän henkilöstöä. Ihmisen työvoima voidaan siten käyttää sellaiseen, missä ihmisen päätöksentekoa tarvitaan. Data-analytiikka mahdollistaa siis myös ihmisten osaamisen kehittymisen, kun työtehtävät ovat entistä mielekkäämpiä ja vaihtelevimpia koneiden tehdessä manuaaliset työt. Tehokkuuden ja kustannussäästöjen lisäksi data-analytiikka mahdollistaa erilaisten mallien kehittämisen ja hyödyntämisen, jota kautta tulee mahdollisuuksia ketterämpään muutokseen ja syvällisempään ymmärrykseen epäilyttävästä toiminnasta.

- Miten pankit voisivat hyödyntää data-analytiikkaa tehokkaasti ja kannattavasti rahanpesun estämisessä?

Opinnäytetyölle asetettujen alatavoitteiden vastauksista voidaan johtaa, että vaikka data-analytiikkaa hyödynnetään tällä hetkellä Suomessa toimivissa pankeissa rahanpesun estämisessä kannattavasti pankkien mielestä, on sitä mahdollista vielä tehostaa ja kehittää kannattavammaksi.

Tutkimuksen perusteella pankit kaipaavat entistä enemmän dataa, erityisesti verrannollista dataa, vaikka jokaiselle pankille, koosta riippumatta, dataa kertyykin valtavia määriä. Se, missä määrin pankit pystyvät dataa jakamaan tällä hetkellä keskenään, on hyvin vähäistä. Tämän estävät tietosuoja-asetus ja muut asiakkaiden tietojen käsittelyä koskevat lait, kuten myös tietoperustan luvuissa 2.1.2 ja 4.1 todettiin. Tutkimuksesta nousee myös esille, että pankit kokevat, etteivät muut pankit halua jakaa omia ideoitaan ja menetelmiään muille pankeille.

Kuten teoriaosuuden luvussa 4.1 kerrottiin, Burton-Taylor Internationalin raportin mukaan rahanpesun estämisestä, asiakkaan tuntemisesta ja niihin liittyvistä taloudelliseen rikollisuuteen ja sääntöjen noudattamiseen liittyvistä toimista aiheutuvat kustannukset ovat maailmanlaajuisesti kasvaneet 17,5 % vuonna 2019. Raportin mukaan uuden asiakkaan rekisteröinti ja tunnistaminen vie pankeilta valtavasti aikaa ja resursseja, jotta siihen liittyvät sääntelyt tulee noudatettua. Mikäli pankit voisivat jakaa dataa keskenään, tulisi se vähentämään etenkin pankkien kustannuksia valtavasti. (Griboval 12.1.2020.)

Tutkimuksen tuloksista nouseekin myös esille kiinnostus jakaa dataa pankkien kesken. Tutkimustulosten perusteella rahanpesun estäminen on kaikille haaste ja maksaa paljon rahaa eivätkä pankit enää pärjää yksin. Tutkimuksen perusteella datan jakaminen pankkien kesken auttaisi kaikkia. Datan jakaminen pankkien kesken helpottaisi pankeja suorittamaan data-analytiikkaa kannattavammin suuremman määrän ja verrannollisen datan ansiosta. Jakaminen mahdollistaisi vertailun koko Suomen pankkisektorin dataan, jota olisi mahdollista käyttää esimerkiksi testaukseen. Datan laajempi saatavuus helpottaisi myös data-analytiikan hyödyntämistä, erityisesti koneoppimisen, tekoälyn ja big datan osalta. Ne tarvitsevat valtavat määrät dataa toimiakseen tehokkaasti ja luotettavasti. Lisäksi, mikäli dataa voitaisiin jakaa pankkien kesken, olisi jo heti asiakkuutta avatessa pankeilla enemmän tietoa asiakkaasta. Lisäksi usein rahanpesutapauksissa rikoksella saadut varat sijoitetaan eri pankkien kesken, joten datan jakaminen helpottaisi rahanpesun estämistä huomattavasti. Tutkimuksessa nousi esille myös kiinnostus saada enemmän tietoa rahanpesun selvittelykeskukselle annetuista ilmoituksista, joka sekin auttaisi osaltaan rahanpesun estämistä.

6.2 Johtopäätökset sekä kehittämis- ja jatkotutkimusehdotukset

Johtopäätöksenä voidaan todeta, että tehokas keino kehittää data-analytiikkaa rahanpesun estämisessä ja samalla itse rahanpesun estämisen prosessia tehokkaammaksi ja kannattavammaksi olisi muuttaa lainsäädäntöä siten, että datan jakaminen pankkien kesken olisi mahdollista. Tietosuoja-asetuksen ja lainsäädännön mukauttaminen rahanpesun estämistä tukevaksi olisi tärkeää niin pankeille kuin koko yhteiskunnallekin. Yhteistyö-pankkien kesken vähentäisi pankkien kustannuksia ja tekisi data-analytiikan hyödyntämisestä ja rahanpesun estämisestä tehokkaampaa ja luotettavampaa. Se mahdollistaisi data-analytiikkamenetelmien laajemman käyttömahdollisuuden, koska useat menetelmät, esimerkiksi koneoppimisen menetelmät toimivat sitä paremmin ja luotettavammin mitä enemmän dataa on saatavilla. Rahanpesutapaukset, etenkin tapauksissa, joissa varat si-joitetaan useiden pankkien kesken, tulisivat myös varmasti oleellisesti vähenemään, kun niiden havaitseminen olisi helpompaa yhteistyön ansiosta.

Edellä esitetyn tärkeän kehitysehdotuksen lisäksi uskon opinnäytetyön tuovan kokonaisuudessaan alalla toimijoille tärkeää tietoa toimiakseen tehokkaasti ja kannattavasti tällä tarkkaan säännellyllä alalla sekä kannustavan pankkeja hyödyntämään data-analytiikkaa entistä tehokkaammin rahanpesun estämisessä ja kehittämään rahanpesun estämistä kohti tulevaisuutta, koska työ tarjoaa laajasti erilaisia data-analytiikan mahdollisuuksia niin tietoperustassa kuin tutkimusosiossakin. Työstä on hyötyä niille toimijoille, jotka eivät ole vielä ottaneet data-analytiikkaa laajemmassa mittakaavassa käyttöönsä, mutta myös niille, jotka ovat jo pidemmällä data-analytiikan hyödyntämisessä. Varhaisemmassa vaiheessa olevat saavat työstä arvokkaita ideoita ja näkökulmia data-analytiikan lisäämiseen toiminnassaan ja edelläkävijät saavat vinkkejä toimintansa tehostamiseen.

Opinnäytetyötä tehdessä mieleeni nousi muutamia eri jatkotutkimusehdotuksia. Erityisesti olisi mielenkiintoista lähteä selvittämään ja tutkimaan, miten lainsäädännön muuttaminen datan jakamisen kannalta olisi mahdollista ja kuinka pian sellaista voisi odottaa tapahtuvan. Olisiko esimerkiksi mahdollista toteuttaa erillinen tietokanta, jota pankeilla ja kehittäjillä olisi mahdollisuus hyödyntää rahanpesun estämisessä ja uusia rahanpesun estämi-seen käytettäviä järjestelmiä suunnitellessa ja testatessa. Toinen mieleeni noussut asia, jota voisi lähteä tutkimaan, on tutkia tarkemmin koneoppimisen mahdollisuuksia ja kannattavuutta rahanpesun estämisessä. Tässä tutkimuksessa ei keskitytty koneoppimiseen, mutta kuitenkin huomattiin siihen liittyviä mahdollisuuksia ja mielipiteitä sen kannattavuudesta. Molemmat jatkotutkimusehdotukset tukisivat tätä tutkimusta ja olisivat myös yhtä lailla yhteiskunnallisesti tärkeitä kehittäen data-analytiikan hyödyntämistä rahanpesun estämisessä tehden rahanpesutapausten havaitsemisesta tehokkaampaa ja edullisempaa.

6.3 Tutkimuksen luotettavuus ja eettiset näkökohdat

Tutkimus koostuu sekä isoista että pienistä päätöksistä, joita tehdessä tulee tehdä luotettavuutta ja eettisyyttä koskevia ratkaisuja. Luotettavuus ja eettisyys tulee näkyä läpi opinnäytetyön. Niiden kannalta on olennaista noudattaa hyvää tieteellistä käytäntöä. Epäeettisellä ja epärehellisellä toiminnalla eli hyvän tieteellisen käytännön loukkauksilla, voidaan vahingoittaa tieteellistä tutkimusta ja pahimmillaan mitätöidä sen tulokset (Tutkimuseettinen neuvottelukunta 2012, 8). Opinnäytetyössä pyrittiin noudattamaan hyvää tieteellistä käytäntöä ja Haaga-Helian opinnäytetyön raportointiohjeita. Tässä alaluvussa esitellään ja arvioidaan tämän opinnäytetyön luotettavuutta ja eettisiä näkökohtia.

Opinnäytetyön luotettavuutta ja eettisiä näkökohtia pohdittiin jo opinnäytetyön suunnitteluvaiheessa. Suunnitteluvaiheessa tutustuttiin Haaga-Helian opinnäytetyön raportointiohjeisiin sekä opinnäytetyöprosessin etenemiseen. Näiden pohjalta luotettavuutta ja eettisiä näkökohtia pohdittiin esimerkiksi aiheen valinnassa, jossa pohdittiin, onko sitä mahdollista tutkia valituilla menetelmillä. Lisäksi suunnitteluvaiheessa pohdittiin sitä, että onko aihetta perusteltua tutkia aiheen sensitiivisyyden ja arkaluonteisuuden takia.

Opinnäytetyön hyvää eettistä periaatetta ja luotettavuutta tukee lähteiden valinta ja niihin viittaaminen. Tässä opinnäytetyössä käytettiin runsaasti ja monipuolisesti lähteitä, jotka olivat korkealaatuisia ja ajantasaisia. Tärkeää lähteiden valinnassa oli kirjoittajan asiantuntemus sekä lähteen tuoreus ja kansainvälisyys. Käytin työssäni esimerkiksi Harvard Business Review:n artikkelia, jota voidaan pitää yhtenä arvokkaimpana lähteenä. Työssä hyödynnettiin myös paljon lakitekstiä, jota voidaan myös pitää korkealaatuisena. Opinnäytetyössä pyrittiin löytämään ja hyödyntämään samaa aihetta käsitteleviä lähteitä, joka tukisi lähteiden luotettavuutta. Tutkimuksessa alkuperäinen lähde pyrittiin ilmoittamaan asianmukaisesti, selkeästi ja yksiselitteisesti, jotta ei jää epäselväksi, mikä on lainattua ja mikä omaa tekstiä.

Tutkimuksen luotettavuutta pohdittiin myös haastateltavien valinnassa. Haastateltaviksi valittiin Suomessa toimivien pankkien edustajia, jotka toimivat rahanpesun ja terrorismin rahoittamisen estämisen sekä data-analytiikan parissa. Tällä päätöksellä, tutkimusaineistosta saatiin luotettavaa, haastateltavien ollessa alan asiantuntijoita. Jokaisella haastateltavalla oli lisäksi jo ehtinyt kertymään useamman vuoden kokemus alalla toimisesta, joka antaa luottamusta siihen, että aihe on riittävän tuttu heille, jotta pystyvät antamaan syvällisiä ja laadukkaita vastauksia. Haastateltavien yhteystietoja tiedusteltiin puhelimitse Finanssiala Ry:ltä. Haastateltaviin otettiin yhteyttä sähköpostitse, jossa kerrottiin, mistä hei-

dän sähköpostinsa on saatu ja miksi heihin ollaan yhteydessä. Sähköpostitse tiedustelemisella ja haastattelusta sopimisella jää kaikille tutkimusosapuolille dokumentaatiota siitä, mitä on sovittu ja miksi, vaikkei sähköposti ehkä itsellään ole kaikkein luotettavin valinta, sillä keskustelun osapuolet jäävät sähköpostia käyttämällä hieman etäisiksi ja tuntemattomiksi.

Luotettavuutta ja eettisiä näkökohtia oli pohdittava myös haastatteluiden toteutuksessa ja paikan valinnassa. Covid-19 aiheuttaman tilanteen takia, haastateltaville haluttiin antaa mahdollisuus esittää toive haastattelupaikan valinnasta. Yhteisymmärryksessä haastattelujen kanssa haastattelut toteutettiin virtuaalisesti Microsoft Teams -alustan välityksellä yhtä haastattelua lukuun ottamatta. Tällä tavoin haastateltavat saivat olla itselle mieluisassa paikassa, osa kotonaan ja osa toimistolla, jolloin haastattelu oli mahdollista toteuttaa siihen soveltuvassa tilassa ilman häiriötekijöitä ja ennen kaikkea turvallisesti. Virtuaalisesti toteutetusta haastattelusta huolimatta, haastattelutilanne oli neutraali ja osallistujien välillä saatiin aikaan luottamus.

Tutkimuseettisen neuvottelukunnan vuoden 2019 ohjeen mukaan, tutkijan tulee toteuttaa tutkimus siten, että siitä ei aiheudu tutkittaville, tässä tapauksessa haastateltaville ja heidän edustamille yrityksilleen, merkittäviä riskejä, vahinkoja tai haittoja. Yksi tähän liittyvistä tutkimuksessa pohdittavista asioista on anonymisyyden taso. Empiirisessä tutkimuksessa tulee aina pitää huolta etiikan näkökulmasta siitä, että yksittäistä henkilöä ei tunnista (Tuomi & Sarajärvi 2018, 27). Tässä opinnäytetyössä haastateltaville annettiin mahdollisuus anonymisointiin. Työssä haastateltavat pankit nimettiin sattumanvaraisessa järjestyksessä pankki A:ksi, B:ksi, C:ksi ja D:ksi. Tämän lisäksi tutkimuksessa ei käytetty haastateltavien nimiä. Nämä huomioon ottaen, haastateltavien anonymisyyden taso, jonka he halusivat, toteutui. Toinen huomioon otettava asia on, että tutkimuksesta tulisi informoida haastateltavaa, ja haastateltavalla pitäisi olla kaikissa vaiheissa mahdollisuus vetäytyä tai kieltäytyä tutkimuksesta. Tuomi ja Sarajärven kirjan mukaan eettisesti perusteltua on kertoa haastateltaville, mitä aihetta haastattelu koskee (Tuomi & Sarajärvi 2018, 86). Haastattelujen mahdollisuutta selvittäessä, mahdollisille haastateltaville kerrottiin, mistä haastattelussa on kysymys ja mihin tarkoitukseen sitä käytetään. Haastateltavilta myös tiedusteltiin nauhoittamisen mahdollisuutta etukäteen, johon haastateltavat antoivat suostumuksensa. Haastateltavilla oli mahdollisuus vetäytyä tai kieltäytyä tutkimuksesta sen kaikissa vaiheissa. Lisäksi tutkimuksen tulokset lähetettiin haastateltaville luettavaksi ja kommentoitavaksi. (Tutkimuseettinen neuvottelukunta 2019, 7).

Työn luotettavuudesta ja eettisyydestä kertoo myös tutkimuksen raportointi. Tutkimuksen luotettavuutta ja eettisyyttä arvioitaessa on arvioitava, miten tutkimusaineisto on koottu ja

analysoitu (Tuomi & Sarajärvi 2018, 164). Tutkimuksen toteutus avattiin perusteellisesti vaihe vaiheelta ja muun muassa tavoitteiden, rajausten ja menetelmien valinta perusteltiin johdonmukaisesti. Tämä vahvistaa myös tutkimuksen toistettavuutta. Tutkimus olisi mahdollista toistaa samanlaisena ja todennäköisesti tutkimuksen vastaukset olisivat samankaltaisia tämän tutkimuksen kanssa. Tämä siitä syystä, että pankkisektori on hyvin homogeeninen, mitä rahanpesun estämisen menetelmien käyttöön tulee, ja vastaajat edustivat joukkoa melko hyvin. Vastauksissa myös ilmeni selkeitä yhteyksiä ja painopisteitä myös kannattavuuden kokemusten osalta. Opinnäytetyön tutkimusaineisto ja tulokset esitettiin mahdollisimman selkeästi ja totuudenmukaisesti ilmaisten, vääristelemättä tai muuntelematta tuloksia.

6.4 Opinnäytetyöprosessin ja oman oppimisen arviointi

Opinnäytetyöprosessi alkoi suunnittelulla syksyllä 2020, josta prosessi eteni jouhevasti työn kirjoittamiseen. Ennen varsinaista opinnäytetyöprosessia olin jo koittanut miettiä itseleni mielenkiintoista, mutta samalla ajankohtaista aihetta. Sitä ei tuntunut aluksi löytyvän millään, kunnes silmiini osui teksti, joka käsitteli data-analytiikan mahdollisuuksia talousrikollisuuden torjunnassa. Se kiinnosti heti itseäni ja oli samalla myös erittäin ajankohtainen aihe. Lopulta aihe muotoutui sen viimeisimpään muotoonsa eli: Data-analytiikka rahanpesun estämisessä. Koin aiheen tukevan myös mahdollisuuksiani toimia tällä itseäni kiinnostavalla alalla valmistumisen jälkeen. Vaikka valitessani kyseisen aiheen, tuli minun varautua siihen, että joudun vaihtamaan aihetta, mikäli en saisikaan järjestettyä yhtäkään haastattelua, halusin sinnikkäästi toteuttaa opinnäytetyötä juuri tästä aiheesta.

Haasteeksi prosessissa tuntui kuitenkin muodostuvan tutkimustavoitteiden ja rajausten valinta. Tuntui, että aihe oli jo itsessään kovin rajattu, mutta koska aiheesta on olemassa laajasti paljon tietoa, oli raja pakko vetää johonkin. Tutkimustavoitteet ja rajaukset muotoutuivat kuitenkin pitkän pohdinnan sekä useiden kokeilujen jälkeen ja olin niihin myös lopulta tyytyväinen. Tavoitteet, rajaukset ja näkökulma tukivat omaa kiinnostustani, mutta myös sitä hyötyä, jonka alalla toimijat opinnäytetyöstä saivat. Työ olisi ollut mahdollista rajata myös koskemaan yhtä tiettyä data-analytiikan aihealuetta, esimerkiksi koneoppimista, mutta koin, että tämä olisi ollut liian rajattu, etenkin kun pankeissa käytetään yhtäaikaaisesti useita menetelmiä rahanpesun estämiseksi.

Aiheen ajankohtaisuuden vuoksi alkutietämykseni oli melko vähäistä, joka hankaloitti alkuun kirjoittamisprosessia. Data-analytiikan ja rahoituksen kurssit sekä työkokemus henkivakuutusyhtiössä kuitenkin antoivat pohjaa opinnäytetyön tekemiselle aiheesta. Mutta koska aihe oli mielestäni erittäin mielenkiintoinen, lähdin innolla tutustumaan aiheeseen

tarkemmin. Huomasin opinnäytetyöprosessin aikana sisäistäväni asioita helpommin ja kirjoittaneeni selkeämpää tekstiä sitä mukaa kun tietämystä kertyi.

Aiheen ajankohtaisuuden vuoksi tuoretta aineistoa oli kuitenkin onnekseni melko runsaasti saatavilla. Käytössäni oli pääasiallisesti kansainvälisiä lähteitä, sillä aihetta on toistaiseksi enemmän tutkittu muualla maailmalla. Data-analytiikan hyödyntäminen rahanpesun estämisessä on globaali aihe ja se vaikuttaa ympäri maailmaa. Tärkeää lähteiden, esimerkiksi artikkeleiden ja tutkimusten valinnassa, oli se, että kirjoittajalla on asiantuntemusta aiheeseen esimerkiksi ammatin puolesta. Pyrin löytämään ja käyttämään samaa aihetta käsitteleviä lähteitä, joka tukisi lähteiden luotettavuutta. Tärkeässä roolissa opinnäytetyön lähteissä olivat myös laki rahanpesun ja terrorismin rahoittamisen estämisestä sekä rikoslaki, koska ala on hyvin säänneltyä. Onnistuin siis mielestäni käyttämään työssäni runsaasti ja monipuolisesti lähteitä, jotka olivat korkealaatuisia ja ajantasaisia. Useiden lähteiden tutkiminen tuki omaa substanssiosaamista, sekä koin kansainvälisten lähteiden lukemisen parantavan myös kielellistä osaamistani. Pyrin merkitsemään lähteet heti niihin viitatessani. Koin tämän helpottavan huomattavasti opinnäytetyöprosessia, kun niiden viimeistelyyn ei tarvinnut käyttää lähes lainkaan aikaa, kun ne olivat alusta lähtien asiallisesti merkattuna.

Tutkimusmenetelmäksi valitsin haastattelun, joka mielestäni oli kaikkein luontevin valinta tutkimustavoitteiden kannalta. Koska tavoitteena oli tutkia pankkien kokemuksia ja tapoja, oli haastattelu perusteltu valinta. Tutkimuksen tekemiseen tukea sain henkilökohtaisen ohjaajani lisäksi menetelmäohjaajalta, jolta tiedustelin haastattelurungon pätevyyttä ja kysymysten yhteyttä tavoitteisiin. Menetelmäohjaajan tuen avulla sain mielestäni luotua hyvän yhteyden alatavoitteiden ja teoreettisen viitekehyksen, haastattelun kysymysten ja tutkimustulosten välillä. Tämä on havaittavissa myös johdannossa olevasta peittomatriisista. (Liite 1; taulukko 1.)

Haastatteluja tehdessä huomasin kuitenkin, että osa kysymyksistä saattoi olla liian laajoja, joka vaikeutti osaltaan haastateltavien vastaamista. Myös esimerkiksi kysymys, joka käsiteli data-analytiikan hyödyntämisen vaikutusta henkilötyöntuntemuksiin asiakasta kohden, osoittautui huonoksi. Haastateltavat eivät joko voineet vastata kysymykseen kovin tarkasti tai eivät olleet sellaista laskeneetkaan. Kysymyksiin saadut vastaukset olivat joiltakin osin hyvin yleisellä tasolla, johon osaltaan syynä myös aiheen sensitiivisyys. Kysymysten laajuuteen ja epärelevanttiuteen syynä voi olla se, että halusin sopia haastattelut mahdollisimman varhaisessa vaiheessa, koska aihe on kovin sensitiivinen ja riskinä oli, ettei haastatteluja saada lainkaan, enkä täten ollut ehtinyt perehtyä itse aiheeseen kovin syvästi. Mikäli siis saisin tehdä opinnäytetyöprosessissa jotain toisin, tutustuisin aiheeseen syvällisemmin ennen haastattelurungon luomista ja toteuttaisin haastattelut vasta sitten,

kun vaadittava syvälinen tietämys olisi saavutettu. Kysymyksistä olisin tehnyt tarkempia ja yksityiskohtaisempia, jotta kysymyksiin olisi ollut mahdollista saada rajatumpia vastauksia. Itse haastattelut sujuivat kuitenkin jouhevasti ja keskustelua saatiin aikaiseksi. Haastattelut opettivat minulle paljon, niin kysymysten asettamisesta kuin itse haastattelutilanteesta-kin. Haastattelut toivat rohkeutta ja itsevarmuutta haastattelemiseen sekä opettivat paljon itse alasta. Viimeisen haastattelun sain kuitenkin järjestettyä vasta aivan työn loppuhetkellä, joka meinasi hidastaa työn valmistumista, mutta otin tavoitteekseni kiritä työssä laa-
tua heikentämättä, ja koen onnistuneeni tässä hyvin. (Liite 1.)

Opinnäytetyöprosessi eteni kokonaisuudessaan sujuvasti ja se saatiin päätökseen mar-
raskuussa 2020 suunnitelmien mukaisesti. Tutkimuksen perusteella saadut tulokset vasta-
sivat hyvin työn päätavoitteeseen ja olivat merkittäviä niin tutkimuksen kohteille kuin yh-
teiskunnallekin, ja sitä kautta koen, että opinnäytetyö oli onnistunut. Tämän lisäksi onnis-
tumisesta kertoo, että pysyin tavoitteessani saada opinnäytetyö valmiiksi kahden periodin
aikana suunnitelmien mukaisesti, vaikka tähän haastetta toivat haastatteluiden viivästymi-
sen lisäksi samanaikaisesti työskentely ja pari muuta kurssia. Opinnäytetyöprosessi opetti
ajanhallintaa, itseohjautuvuutta ja tavoitteisiin pyrkimistä. Opinnäytetyö oli kaiken kaikki-
aan erittäin opettavainen ja kehityin niin kirjoittajana kuin data-analytiikan ja rahanpesun
estämisen osaaajanakin. Opinnäytetyö kehitti myös analysointitaitojani, josta on myös hyö-
tyä tällä alalla toimimiseen. Opinnäytetyö vahvisti entisestään kiinnostustani data-analyti-
kan ja rahanpesun estämisen parissa toimimiseen.

Lähteet

Ambientia s.a. Tekoäly – Lopeta yrityksesi tiedon päällä istuminen, hyödynnä sitä. Luettavissa: <https://www.ambientia.fi/palvelut/tekoaly/>. Luettu: 22.11.2020.

Breslow, S., Hagstroem, M., Mikkelsen, D. & Robu, K. 7.11.2017. The new frontier in anti-money laundering. McKinsey & Company. Luettavissa: <https://www.mckinsey.com/business-functions/risk/our-insights/the-new-frontier-in-anti-money-laundering>. Luettu: 20.10.2020.

Bughin, J., McCarthy, B. & Chui, M. 2017. A Survey of 3,000 Executives Reveals How Businesses Succeed with AI. Harvard Business Review. Luettavissa: <https://hbr.org/2017/08/a-survey-of-3000-executives-reveals-how-businesses-succeed-with-ai>. Luettu: 14.10.2020.

Chattopadhyay, R. 2016. Effective Business Solutions with Big Data Analytics: Key for Business Growth. Globsyn Management Journal, 10, 1/2, 87-96.

Chen, Z., Van Khoa, L.D., Teoh, E.N., Nazir, A., Karupiah, E.K. & Lam, K.S. 2018. Machine learning techniques for anti-money laundering (AML) solutions in suspicious transaction detection: a review. Knowledge and Information Systems. 57, 2, 245–285.

Delle-Case, A., Bailey, N., Carr, B. & Ekberg, M. 2018. Machine Learning in Anti-Money Laundering – Summary Report. Institute of International Finance. Washington, D.C. Luettavissa: https://www.iif.com/portals/0/Files/private/32370132_iif_machine_learning_in_aml_-_public_summary_report.pdf. Luettu: 21.10.2020.

Empirica s.a. Kehittynyt analytiikka. Luettavissa: <https://www.empirica.fi/data-analytiikka/>. Luettu: 22.11.2020.

Eskola, J. & Suoranta, J. 1998. Johdatus laadulliseen tutkimukseen. Vastapaino. Tampere.

Euroopan parlamentin ja neuvoston asetus luonnollisten henkilötietojen käsittelyssä sekä näiden tietojen vapaasta liikkuvuudesta ja direktiivin 96/46/EY kumoamisesta (yleinen tietosuoja-asetus) (EU) 2016/679. Annettu 27.4.2016.

EY Global 2019. How data analytics is leading the fight against financial crime. Luettavissa: https://www.ey.com/en_gl/consulting/how-data-analytics-is-leading-the-fight-against-financial-crime. Luettu: 4.10.2020.

FATF 2019. Anti-money laundering and counter-terrorist financing measures, Finland, Mutual Evaluation Report. FATF. Pariisi. Luettavissa: <http://www.fatf-gafi.org/media/fatf/documents/reports/mer4/MER-Finland-2019.pdf>. Luettu: 18.10.2020.

FATF s.a. Money Laundering. Luettavissa: <https://www.fatf-gafi.org/faq/moneylaundering/>. Luettu: 18.10.2020.

Finanssiala ry 2019. Viranomaisten sähköiset tilikyselyt tehostavat rahanpesun torjuntaa. Luettavissa: <https://www.finanssiala.fi/uutismajakka/Sivut/sahkoiset-tilikyselyt-tehostavat-rahampesun-torjuntaa.aspx>. Luettu: 15.10.2020.

Finanssivalvonta 16.4.2019. Finanssivalvonta vahvistaa rahanpesun estämisen valvontaa – Myös eurooppalaista valvontaa tehostetaan. Lehdistötiedote. Luettavissa: <https://www.finanssivalvonta.fi/tiedotteet-ja-julkaisut/lehdistotiedotteet/2019/finanssivalvonta-vahvistaa-rahampesun-estamisen-valvontaa--myos-eurooppalaista-valvontaa-tehostetaan/>. Luettu: 21.10.2020.

Finanssivalvonta s.a. Rahanpesun ja terrorismin rahoittamisen estäminen. Luettavissa: <https://www.finanssivalvonta.fi/rahampesun-estaminen/>. Luettu: 2.10.2020.

Garg, A., Grande, D., Miranda, G, M-L., Sporleder, C. & Windhagen, E. 2017. Analytics in banking: Time to realize the value. McKinsey & Company. Luettavissa: <https://www.mckinsey.com/industries/financial-services/our-insights/analytics-in-banking-time-to-realize-the-value>. Luettu: 14.10.2020.

Griboval, T. 12.1.2020. Combining KYC, AML and data-privacy regulations. Aion Consulting. Luettavissa: <https://www.aionconsulting.eu/post/combining-kyc-aml-and-data-privacy-regulations>. Luettu: 21.10.2020.

Gupta, B. 2016. Interview Questions in Business Analytics. Bangalore. Luettavissa: <https://learning.oreilly.com/library/view/interview-questions-in/9781484205990/ACover-HTML.html>. Luettu: 13.10.2020.

Hyvärinen, H. 28.4.2020. Data-analytiikka, tekoäly ja koneoppiminen – trenditermit suomeksi. Teknologiayrityksen blogi. Luettavissa: <https://www.vincit.fi/fi/data-analytiikka-tekoaly-ja-koneoppiminen-trenditermit-suomeksi/>. Luettu: 22.11.2020.

Hyvärinen, M., Nikander, P. & Ruusuvuori, J. 2017. Tutkimushaastattelun käsikirja. Vastapaino. Tampere.

Kalyvas, J. & Overly, M. 2014. Big Data. Auerbach Publications. Luettavissa: <https://learning.oreilly.com/library/view/big-data/9781466592377/?ar>. Luettu: 21.10.2020.

Kananen, H. & Puolitaival, H. 2019. Tekoäly: bisneksen uudet työkalut. Alma Talent. Helsinki. Luettavissa: [https://bisneskirjasto-almatalent-fi.ezproxy.haaga-helia.fi/teos/BAX-BBXATCBIED#/kohta:TEKO\(\(c4\)LY\(\(20\)-\(\(20\)Bisneksen\(\(20\)uudet\(\(20\)tyokalut/piste:tqe](https://bisneskirjasto-almatalent-fi.ezproxy.haaga-helia.fi/teos/BAX-BBXATCBIED#/kohta:TEKO((c4)LY((20)-((20)Bisneksen((20)uudet((20)tyokalut/piste:tqe). Luettu: 7.10.2020.

Kananen, J. 2017. Laadullinen tutkimus pro graduna ja opinnäytetyönä. Jyväskylän ammattikorkeakoulu. Jyväskylä. Luettavissa: <https://www.booky-fi.ezproxy.haaga-helia.fi/lainaa/1168>. Luettu: 3.11.2020.

Kohonen, I., Kuula-Luumi, A. & Spoof, S-K. 2019. Ihmiseen kohdistuvan tutkimuksen eettiset periaatteet ja ihmistieteiden eettinen ennakkoarviointi Suomessa. Tutkimuseettinen neuvottelukunta. Helsinki. Luettavissa: https://www.tenk.fi/sites/tenk.fi/files/Ihmistieteiden_eettisen_ennakkoarvioinnin_ohje_2019.pdf. Luettu: 19.11.2020.

Laki pankki- ja maksutilien valvontajärjestelmästä 571/2019.

Laki rahanpesun ja terrorismin rahoittamisen estämisestä 28.6.2017/444.

Lehto, T. 27.5.2019. ”Ohjelma ei leimaa ketään huijariksi” – Yritykset etsivät rahanpesun paljastamiseen apua analytiikasta. Talouselämä. Luettavissa: <https://www.talouselama.fi/uutiset/ohjelma-ei-leimaa-ketaan-huijariksi-yritykset-etsivat-rahampesun-paljastamiseen-apua-analytiikasta/1bdae826-56f2-4d67-9895-1ce678b54335>. Luettu: 18.10.2020.

Marr, B. 2018. How Much Data Do We Create Every Day? The Mind-Blowing Stats Everyone Should Read. Luettavissa: <https://www.forbes.com/sites/bernardmarr/2018/05/21/how-much-data-do-we-create-every-day-the-mind-blowing-stats-everyone-should-read/#2cad04e460ba>. Luettu: 14.10.2020.

Martinsuo, M. & Kärri, T. 2017. Teollinen internet uudistaa palveluliiketoimintaa ja kunnossapitoa. Kunnossapitoyhdistys Promaint ry. Helsinki. Luettavissa: https://tutcris.tut.fi/portal/files/10991976/teollinen_internet_kirja_digi.pdf. Luettu: 27.9.2020.

Merilehto, A. 2018. Tekoäly: matkaopas johtajalle. Alma Talent. Helsinki. Luettavissa: [https://bisneskirjasto-almatalent-fi.ezproxy.haaga-h-lia.fi/teos/GADBDXDTEB#kohta:TEKO\(\(c4\)LY\(\(20\)\)](https://bisneskirjasto-almatalent-fi.ezproxy.haaga-h-lia.fi/teos/GADBDXDTEB#kohta:TEKO((c4)LY((20))). Luettu: 7.10.2020.

Nordea s.a. Rahanpesun estäminen. Luettavissa: <https://www.nordea.com/fi/tietoa-nordeasta/nordea-yhteiskunnassa/Rahanpesun-estaaminen/>. Luettu 22.9.2020.

Poliisi s.a. Rahanpesun torjunta. Luettavissa: <https://www.poliisi.fi/rahanpesu>. Luettu: 18.10.2020.

Reese, B. 24.5.2018. GDPR and EU AML Directives – A Regulatory Tug-of-War? Konsulttiyrityksen blogi. Luettavissa: <https://blog.protiviti.com/2018/05/24/gdpr-eu-aml-directives-regulatory-tug-war/>. Luettu: 18.10.2020.

Rikoslaki 19.12.1889/39.

Rouse, M. 2020. Data Analytics (DA). Tech Target. Luettavissa: <https://searchdata-management.techtarget.com/definition/data-analytics>. Luettu: 27.9.2020.

Salo, I. 2013. Big Data – Tiedon vallankumous. Saarijärven Offset Oy. Saarijärvi.
SAS. Analytics - What it is and why it matters. Luettavissa: https://www.sas.com/en_us/insights/analytics/what-is-analytics.html. Luettu: 6.10.2020.

Shainski, R. 20.11.2019. For Banks, Data Privacy and Anti-Money Laundering Don't Have to Be Incompatible. CPO Magazine. Luettavissa: <https://www.cpomagazine.com/data-privacy/for-banks-data-privacy-and-anti-money-laundering-dont-have-to-be-incompatible/>. Luettu: 21.10.2020.

Taanila, A. 26.8.2020. Data-analytiikka antaa vastauksia kysymyksiin. Menetelmäopettajan blogi. Luettavissa: <https://tilastoapu.wordpress.com/tag/kuvaileva-analytiikka/>. Luettu: 17.10.2020.

Tietosuojalaki 5.12.2018/1050.

Tuomi, J. & Sarajärvi, A. 2018. Laadullinen tutkimus ja sisällönanalyysi. Kustannusosakeyhtiö Tammi. Helsinki.

United Nations Office on Drugs and Crime. Money-Laundering and Globalization s.a. Luettavissa: <https://www.unodc.org/unodc/en/money-laundering/globalization.html>. Luettu: 2.10.2020.

Valtiovarainministeriö 2018. Rahanpesun torjunnan lainsäädäntöä tiukennetaan. Luettavissa: <https://vm.fi/-/rahanpesun-torjunnan-lainsaadantoa-tiukennetaan>. Luettu: 15.10.2020.

Varantola, K., Launis, V., Helin, M., Spoof, S-K. & Jäppinen, S. 2012. Hyvä tieteellinen käytäntö ja sen loukkausepäilyjen käsitteleminen Suomessa. Tutkimuseettinen neuvottelukunta. Helsinki. Luettavissa: https://www.tenk.fi/sites/tenk.fi/files/HTK_ohje_2012.pdf. Luettu: 19.11.2020.

Weeks-Brown, R. 2018. Cleaning Up. Finance & Development, 55, 4.

Liitteet

Liite 1. Haastattelurunko

Millainen on työnkuvasi?

Kuinka pitkään olet toiminut alalla, entä kyseisessä työyhteisössä?

Millaisia menetelmiä teillä on rahanpesun estämiseen liittyen?

Miten hyödynnätte data-analytiikkaa rahanpesun estämisessä?

Miten paljon rahanpesuun liittyvää dataa teillä on saatavissa?

Miten dataa pystytään hyödyntämään rahanpesun estämiseksi?

Onko asiakkaasta saatavissa tiedoissa käyttörajoituksia?

Miten paljon data-analytiikan hyödyntäminen vaatii henkilötyötunteja asiakasta kohden?

Millaisia kustannuksia data-analytiikan hyödyntämisestä aiheutuu?

Mitä lisäarvoa data-analytiikan hyödyntäminen tuo rahanpesun estämiseen?

Miten kannattavana koette data-analytiikan hyödyntämisen rahanpesun estämiseksi?

Mitä mahdollisuuksia data-analytiikan hyödyntämisellä voisi vielä olla?