



<input type="checkbox"/>	Kandidaatintutkielma
<input checked="" type="checkbox"/>	Pro gradu -tutkielma
<input type="checkbox"/>	Lisensiaatintutkielma
<input type="checkbox"/>	Väitöskirja

Oppiaine	Laskentatoimi ja rahoitus	Päivämäärä	5.11.2019
Tekijä(t)	Antti Bister	Matrikkelinumero	507228
		Sivumäärä	99 s. + liitteet
Otsikko	Big Datan ja Big Data-analytiikan vaikutukset tilintarkastajan kognitiiviseen työskentelyyn ja tilintarkastukseen		
Ohjaaja(t)	KTT, Erkki Vuorenmaa		

Tiivistelmä	
<p>Big Datalla tarkoitetaan valtavaa ja nopeasti lisääntyvää ihmisen luoman elektronisen ympäristön tuottamaa datamassaa. Tämän datamassan hyödyntämisen valjastamiseen myös nykypäivän liiketoiminta ja tämän perässä tilintarkastus ovat voimakkaasti suuntaamassa. Suurten tietomäärien tiedetään aiheuttavan kognitiivisella tasolla erinäisiä haasteita niitä käsittelevillä henkilöillä. Tämän johdosta tehokkaat analyttiset menetelmät ovat Big Datan käsittelyssä hyvin perusteltuja.</p> <p>Tämän tutkielman tavoitteena on selvittää miten Big Datan ominaisuudet ja Big Dataa prosessoiva analytiikka vaikuttavat tilintarkastajan työsuoritukseen hänen kognitiivisesta näkökulmastaan tarkasteltuna. Kiinnostuksen kohteena on samalla myös se, miten Big Datan vaikutukset tilintarkastajan kognition tasolla vaikuttavat lopulta koko tilintarkastustyöhön. Tutkielman tutkimusmenetelmäksi valikoitui laadullinen haastattelututkimus, joka toteutettiin kuutena erillisenä Big Data-analytiikan parissa säännöllisesti työskentelevään tilintarkastajaan kohdistuneena yksilöhaastatteluna.</p> <p>Tutkielman tulosten mukaan tehokkaista analyttisistä välineistä huolimatta Big Data haastaa tilintarkastajaa informaatioylikuorman, tiedon epäselkeyden ja epärelevantin, -mutta samalla myös relevantin tiedon runsauden kautta. Big Data-aikakauden tilintarkastajalta vaaditaan entistä data-analyttislähtöisempää ajattelutapaa ja kykyä erottaa olennainen tieto suuresta datamassasta. Lisäksi tilintarkastajan työtä luonnehtiva ammatillinen skeptisyys on muutoksessa Big Datan myötä. Tilintarkastaja on myös vaarassa ajautua kognitiivisiin vinoumiin ja laimennusvaikutuksen alaiseksi työstäessään mielessään Big Dataa. Ajastaan jäljessä oleviin tilintarkastusstandardeihin vastaaminen ei myöskään lainkaan helpota tilintarkastajan työskentelyä Big Datan parissa.</p> <p>Tutkielman keskeisimpänä johtopäätöksenä Big Data haastaa tilintarkastajaa kognitiivisesti uudennlaisin tavoin. Yllättävästi Big Data-analytiikkakaan ei suinkaan kokonaan poista tilintarkastajan kohtaamia kognitiivisia haasteita, vaan luo omansa. Big Datan myötä lisääntyvä kognitiivinen haastavuus ilmenee kuitenkin lopulta tilintarkastuksen laadun parantumisenä. Tämä tapahtuu tilintarkastajan perehtyessä asiakkaansa dataan ja liiketoiminnan mekanismeihin aiempaa syvällisemmin.</p>	
Asiasanat	tilintarkastus, Big Data, data-analytiikka, tilintarkastaja, kognitiivinen
Muita tietoja	







**TURUN  
YLIOPISTO**

Kauppakorkeakoulu

**BIG DATAN JA BIG DATA-ANALYTIIKAN  
VAIKUTUKSET TILINTARKASTAJAN  
KOGNITIIVISEEN TYÖSKENTELYYN JA  
TILINTARKASTUKSEEN**

Laskentatoimen ja rahoituksen  
pro gradu -tutkielma

Laatija:  
Antti Bister

Ohjaaja:  
KTT Erkki Vuorenmaa

5.11.2019  
Pori

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

## Sisällysluettelo

1	JOHDANTO .....	9
1.1	Johdatus aihepiiriin .....	9
1.2	Tutkielman tavoite, tutkimusongelman asettelu, perustelut ja rajaus .....	12
1.3	Tutkimusmetodologia ja tutkielman rakenne .....	13
2	BIG DATAN JA BIG DATA-ANALYTIIKAN VAIKUTUKSET TILINTARKASTUSTYÖHÖN .....	17
2.1	Big Datan ominaisuudet .....	17
2.2	Big Dataa prosessoivat analyttiset menetelmät tilintarkastuksessa .....	20
2.3	Koko populaation testaaminen .....	24
2.4	Jatkuva tilintarkastus .....	26
2.5	Tilintarkastuksen automatisaatio .....	30
2.6	Tilintarkastusstandardit ja Big Datan hyödyntäminen tilintarkastuksessa .....	31
3	SUURTEN DATAMÄÄRIEN PROSESSOINTIIN LIITTYVÄT TEKIJÄT BIG DATAA TYÖSTÄVÄLLÄ TILINTARKASTAJALLA .....	35
3.1	Tilintarkastajan kognitiivinen työskentely .....	36
3.2	Tilintarkastajan kohtaamat kognitiiviset vinoumat ja työperäinen stressi ...	38
3.3	Tilintarkastajan ammatillinen arviointikyky Big Data-aikakaudella .....	40
3.4	Data-analyttinen ajattelutapa .....	42
3.5	Big Datan aiheuttama tilintarkastusanalyttinen näkökulmamuutos .....	44
3.6	Tilintarkastajan kohtaamat yleiset haasteet ja ammatillinen skeptisyys Big Datan tulkinnassa .....	46
3.7	Big Datan tilintarkastajalle aiheuttama informaatioylikuorma .....	48
3.8	Olellisuuden hahmottamiseen liittyvät aspektit Big Dataa työstävällä tilintarkastajalla .....	51
3.9	Big Datan sisältämän tiedon relevanttius .....	53
3.10	Big Datan sisältämän tiedon selkeys .....	54
3.11	Tämän tutkielman kirjallisuuskatsaukseen perustuva tutkimusaukko ja teoriaosuuden yhteenveto .....	56
4	TUTKIELMAN EMPIIRINEN OSUUS .....	60
4.1	Empiirisen aineiston hankinnan toteutus ja tutkielmaa varten haastateltujen henkilöiden esittelemine .....	60

4.2	Big Datan määritelmään ja tilintarkastuksen automatisoitumiseen liittyvät tutkimustulokset .....	63
4.3	Olellisuuden hahmottamiseen, tilintarkastajan ajattelutavan muutokseen ja ammatilliseen skeptisyyteen liittyvät tutkimustulokset .....	66
4.4	Relevantin tiedon hahmottamiseen ja informaatioylikuormaan liittyvät tutkimustulokset .....	72
4.5	Big Datan ja Big Data-analytiikan vaikutukset tilintarkastajan kognitiiviseen työskentelyyn ja tilintarkastukseen – empiiristen tulosten pohdinta ja synteesi teoriaa vasten.....	77
5	TUTKIELMAN LOPUKSI.....	85
5.1	Johtopäätökset.....	85
5.2	Tutkielman reliabiliteetti, validiteetti ja yleistettävyys .....	87
5.3	Tutkielman luoma lisäarvo ja tutkielman rajoitteet .....	88
	LÄHDELUETTELO.....	90
	LIITTEET .....	100

## Kuvioluettelo

Kuvio 1: Tämän tutkielman teoriaosuuden rakenne .....	15
Kuvio 2: Teoreettinen esimerkki jatkuvan datapohjaisen tilintarkastusjärjestelmän toimintaperiaatteesta. (Mukaiillen Kogan ym. 2014, 226.).....	27
Kuvio 3: Kaksi tutkijoiden esittämää kuviota yhdistäen; perinteinen tilintarkastusprosessi (ylhäällä) ja audit by exception - tilintarkastusprosessi (alhaalla). (Mukaiillen Appelbaum ym. 2017, 11.)	29
Kuvio 4: Tilintarkastajan iteratiivinen ajatteluprosessi hänen analyyttisessä työskentelyssään. (Mukaiillen Koonce 1993, 59.) .....	37
Kuvio 5: Tilintarkastajan kokeman työperäisen stressin ja työsuorituksen välinen suhde. (Mukaiillen Choo 1986, 26.) .....	39
Kuvio 6: Tilintarkastajan ja muun laskenta-ammattilaisen analyysin näkökulmat. (Mukaiillen Richins ym. 2017, 65.).....	44
Kuvio 7: Tiedon prosessoinnin tehokkuuden ja tiedon määrän välinen suhde tietoa käsittelevällä henkilöllä. (Mukaiillen Hwang & Lin 1998, 214.).....	49
Kuvio 8: Tämän tutkielman tutkimusaukon viitekehys .....	57
Kuvio 9: Tämän tutkielman tulosten yhteenveto .....	77





# 1 JOHDANTO

## 1.1 Johdatus aihepiiriin

Kuvittele itsesi työstämässä niin suurta määrää tietoa, ettet edes kykenisi käsittämään tämän tietomäärän sisältämien datayksiköiden todellista lukumäärää. Vaikka sinulla olisikin kaikki tarvittava mahdollinen teknologia saatavilla tämän valtavan datamäärän työstämistä varten, työskentelevä mielesi kohtaisi väistämättä joitain haasteita datamäärän laajuuden vuoksi. Tämän kaltaisessa tilanteessa olet, mikäli käsittelet työssäsi nykypäivän tilintarkastajien tavoin Big Dataa ja teet sen kautta päätelmiä.

Miten Big Data -tyyppisen tiedon prosessoiminen vaikuttaa tätä käsittelevään henkilöön ja hänen työskentelyynsä? Kuten johdannosta pian ilmenee, Big Datan hyödyntäminen on laajasti liiketoimintaa sekä sitä myötä myös laskentatoimen ammattikuntaa tilintarkastus mukaan lukien ajankohtaisesti koskettava kysymys. Kuten tästä tutkielmasta kokonaisuudessaan ilmenee, tilintarkastajan työ sisältää runsaasti tiedon analysoimista. Tämä tieto voi nykypäivänä hyvinkin sisältää Big Dataa. Seuraavaksi on kuitenkin tarkoituksenmukaista jo heti tutkielman alkuun selventää, mitä Big Datalla ylipäätään tarkoitetaan.

Big Dataa voidaan pohjimmiltaan pitää ihmisen luoman kehittyneen ja jatkuvasti monimutkaistuvan nykyaikaisen teknologisen ympäristön datamääräisenä tuotteena. Tälle datamääräiselle tuotteelle on ominaista, että sitä syntyy erilaisissa muodoissa yhä kiihtyvämpään tahtiin tavattoman suuria määriä lisää. Ihmisen luomassa uudessa kehittyneessä teknologisessa ympäristössä, josta Big Data kumpuaa, lähes mikä tahansa signaali voidaan tallentaa, mitata ja kaapata digitaalisesti. (Cao, Chychyla & Stewart 2015, 423 – 424.) Nämä signaalit, joista Big Data koostuu voivat sisältää numeroita, tekstiä, kuvia, ääniä tai videoita (Byrnes & Pawlicki 2014, 7). Big Data on luonteeltaan erittäin totuudenmukaista ja puolueetonta, sillä se kattaa yleensä käytännössä lähes kaiken mahdollisen saatavilla olevan tiedon tarkastelun kohteena olevasta aihepiiristä. (Richards & King 2014, 432).

Big Datan ajankohtaisuus on kiistatonta nykypäivänä. Jo vuonna 2020 dataa ennustetaan muodostuvan 44 kertaa suurempi määrä kuin sitä muodostui vielä vuonna 2009 (Pansari 2016, 16). Päättävää 2010-luvun vuosikymmentä on täten hyvällä syyllä kutsuttu Big Datan vuosikymmeneksi (Ahmadi, Dileepa & Wheatley 2016, 289).

Big Data laajentaa maailman kuvaamme. Moninaisen ja lähes kaiken tiedon kattavan Big Datan tarkastelu luo meille laajemman tietoisuuden tason maailmasta. Tämä laajempi tietoisuuden taso auttaa meitä muun muassa entistä tarkempien ennustusten luomisessa ja moninaisten ongelmien ratkaisemisessa monilla tieteen osa-alueella. (Richards & King 2014, 405.)

Kuten edellä on ilmennyt, Big Data on monimuotoista. Big Datan monimuotoisuuden ansiosta se on myös käyttökelpoista hyvinkin moninaisiin tarkoituksiin. Big Dataa hyödynnetään nykypäivänä yhteiskunnassamme erityisesti liiketoiminnallisissa päätöksissä, mutta myös lukuisissa muissa eri yhteyksissä. Sitä hyödynnetään muun muassa urheilussa pelaajien arvioimisessa, kansallisen turvallisuuden arvioinnissa ja ihmiskuntaa uhkaavien epidemioiden ennustamisessa. (Byrnes & Pawlicki 2014, 7.)

Koyce (2017, 28 – 29) tiivistää Big Datan luonteen ja sen hyödyntämismahdollisuudet seuraavasti; ”*Big Data on äärimmäisen suuria datamääriä, joiden paljastamia ihmisen käyttäytymiseen ja vuorovaikutukseen kytkeytyviä kuvioita, trendejä ja syy-yhteyksiä on mahdollista tarkastella laskennallisilla menetelmillä.*”

Big Data koskettaa erityisesti yritysmaailmaa. Yritykset tuottavat ja tallentavat valtavia määriä dataa joka päivä. Esimerkiksi yrityksen taloutta kuvaava tieto virtaa johdon informaatiojärjestelmiin, kuten laskentatoimen järjestelmiin. Tällöin jokainen liiketoiminnan transaktio tallentuu. Tämä ilmiö on laajentunut viime vuosina eksponentiaalisesti, mikä on johtanut mittasuhteiltaan laajaan uudenlaisten tietojärjestelmien implementointiin ympäri maailman. Tuloksena ovat suuret määrät kerättyä ja talletettua dataa organisaatioiden tietovarastoissa. (Issa 2013, 90.)

Big dataa hyödyntävä analytiikka on tällä hetkellä vauhdikkaasti yleistymässä sen luomien etujen vuoksi. Useat yritykset ovatkin omaksumassa sitä omaan käyttöönsä sängen joutuisaan tahtiin. Big Dataan liittyvä analytiikka nähdään useissa organisaatioissa välineenä, jonka kautta on mahdollista parantaa näiden operationaalista ja strategista tehokkuutta, löytää uusia tulonlähteitä ja saada kilpailuetua markkinoilla. (Sivarajah, Kamal, Irani & Weerakkody 2017, 263.) Big datasta onkin tulossa nopeassa tahdissa tulevaisuuden valuutta. He ketkä ovat kykeneviä valjastamaan ja käsittelemään tätä dataa tulevat myös todennäköisimmin menestymään myös liiketoiminnassaan. (Koyce 2017, 37.)

Esimerkit Big Datan hyödyntämisestä liiketoiminnassa ovat moninaisia. Aivan yksinkertainen esimerkki Big Data -pohjaisen tiedon hyödyntämisestä on jokaiseen varaston tuotteeseen kiinnitettävä RFID-tunniste (engl. ”*Radio frequency identification*). RFID-tunnisteet helpottavat yrityksen varastonhallintaa mahdollistamalla esimerkiksi automaattisen ja tehokkaan varaston sijainnin, lukumäärän ja varastointiajan tarkastelun. (Kraheil & Titera 2015, 411 – 412.) Big Dataa ja Big Data-analytiikkaa hyödynnetään nykyisin myös esimerkiksi myös aineettoman pääoman arvioinnissa. Aineettoman pääoman arviointi paranee yrityksen kyetessä oikea-aikaisesti ja globaalisti mittaamaan aineettoman pääoman markkinahintaa esimerkiksi sosiaalisen median tarjoaman datan kautta. (Vasarlehyi, Kogan & Tuttle 2015, 386.)

Tilintarkastus seuraa muuta liike-elämää Big Datan hyödyntämisessä. Big Datan hyödyntämisen yhä yleistyessä tilintarkastuksen asiakkaiden liiketoiminnassa, tulee tilintarkastuksen ja tilintarkastajien ennen pitkää tasavertaisesti myös mukautua sen

vaatimuksiin ja käsittelyyn. Näin ollen voidaan todeta, että tilintarkastuksen siirtymistä Big Datan hyödyntämisen äärelle ajaa lähtökohtaisesti tilintarkastukseen nähden ulkosyntyinen tarve. (Alles 2015, 442.) Myös Earley (2015, 499) mukaan yritykset kaikilla teollisuuden aloilla tekevät parhaillaan kiihtyvällä tahdilla investointeja Big Dataan parantaakseen kilpailukykyään, ja samalla tilintarkastukselta odotetaan vastalauseena tähän tapahtumakulkuun voimakkaampaa samansuuntaista kehitystä.

Tilintarkastus on epäilemättä jo itsessään koko yhteiskunnan mittakaavassa tärkeä ammattikunta. Tilintarkastusarvioinnin tuotteena syntyvän tilintarkastuskertomuksen lopullisena tavoitteena on kasvattaa sijoittajien ja muiden ulkoisten sidosryhmien luottamusta tilintarkastuksen kohteena olevaa asiakasta kohtaan. (Setiawan 2017, 242.) Suomen tilintarkastuslaki 1141/2015 3:5 täsmentää lakisääteisen tilintarkastuksen yhdyntävän tilinpäätöksen allekirjoittavan tilintarkastajan antamaan tilintarkastuskertomuksen lausuntoon siitä, antavatko päättyneen tilikauden pohjalta luodut tilinpäätösdokumentit ”oikeat ja riittävät tiedot” kohteen ”toiminnan tuloksesta ja taloudellisesta asemasta”, sekä ”ovatko toimintakertomuksen ja tilinpäätöksen tiedot ristiriidattomia”. (Finlex 2015.)

Tilintarkastaja on haastavaa ajattelutyötä tekevä oman alansa asiantuntija. Tilintarkastajan työskentelyä ja hänen löydöksiään kuvaavaa on subjektiivisuus, joka perustuu hänen ajatteluunsa (Hogart 1991, 227). Päätös asiakasyrityksen tilintarkastuskertomuksen julkaisemisesta pohjaa täysin tilintarkastajan harkinnanvaraiseen arvioon tilintarkastusprosessin aikana kerätyn tilintarkastusevidenssin riittävydestä ja oikeellisuudesta tilintarkastuskertomuksen paikkansapitävyyden taustalle (Wedemeyer 2010, 326). Tilintarkastajan kognitiivinen työskentely perustuu lähtökohtaisesti hypoteesien luomiseen asiakkaan datasta. Näiden hypoteesien perusteella tilintarkastaja valitsee kulloiseenkin tehtävään soveltuvan tilintarkastustoimenpiteen (Koonce 1993, 59). Tilintarkastajan tulee kuitenkin aina suhtautua kriittisesti tilintarkastustoimenpiteidensä avulla löytämäänsä tilintarkastusevidenssiä kohtaan (Griffith, Hammersley, Kadous & Young 2015, 49). Big Datan voidaan olettaa suurella todennäköisyydellä luovan täysin uudenlaisia haasteita tilintarkastajan kognitiiviselle työskentelylle, kun ottaa huomioon Big Datan ominaisuudet moniulotteisena, suurena ja nopeana epämääräisenä massana.

Tilintarkastajan tulee kyetä hyödyntämään Big Dataa oikein. Totuus on, että Big Data luo arvoa tilintarkastukselle ainoastaan siinä määrin kuin tilintarkastaja on itse kykeneväinen hyödyntämään ja käsittelemään siitä kumpuavaa tilintarkastusevidenssiä tarkastustoimenpiteissään (Ramulkan 2015, 17). EY (2015b, 2) mukaisesti tilintarkastajan tulee omaksua aiempaan nähden entistä data-analyttisempi ajattelutapa, jotta hän kykenee vastaamaan oikein Big Datan luomiin vaatimuksiin. Zhang, Yang ja Appelbaum (2015, 470) mukaan tilintarkastusinstituutiota tällä hetkellä kaikista eniten pohdintaa herättävä asia on, kuinka kerätä oikeaa tietoa räjähdysmäisesti

laajenevista Big Data varastoista tilintarkastusta varten. Esimerkiksi jatkuva erilaisista lähteistä kumpuava ja juontava Big Data ei tarjoa lainkaan lisäarvoa, mikäli datan sisältämien tietojen lähteiden välille ei kyetä muodostamaan yhteyksiä.

Big Datan työstäminen ei kuitenkaan tapahdu ilman sen luomia haasteita, vaikka kehittynyt data-analytiikka tuokin siihen oman helpotuksensa. Big Datan hyödyntäminen tilintarkastustyössä johtaa tilintarkastajilla varsinkin kasvaneen informaatiomäärän käsittelyn kuormittavuuden aiheuttamiin haasteisiin. Tästä aiheutuu muun muassa hankaluuksia tunnistaa ja kerätä relevanttia tietoa. Tilintarkastajat joutuvat muodoiltaan vaihtelevaa Big Dataa työstäessään kohtaamaan tiedon epäselvyyteen liittyviä tekijöitä, joita varten heillä tulee täten olla parempi toleranssi. Kun tilintarkastukselle nykyisin tyypillisestä otannasta siirrytään tehokkaan data-analytiikan siivittämänä koko populaation testaamiseen, on ilmeistä, että Big Datan hyödyntäminen lisää tilintarkastajan kohtaamia haasteita suuren datamäärän sisältämän tiedon arvioinnissa. Tällainen haaste voi olla esimerkiksi suuresta datamäärästä löydettävä suuri löydettyjen poikkeusten lukumäärä. (Vasarlehyi ym. 2015, 391.) Ajastaan jäljessä olevat, pieniä tilintarkastusotoksia varten laaditut tilintarkastusstandardit eivät suinkaan helpota tilintarkastajan työskentelyä Big Datan parissa. (Whithouse 2014, 28 – 29.)

Tiivistäen, Big Data on nopeassa tahdissa lisääntymässä ympärillämme ja liike-elämän käyttötarkoituksissa. Tilintarkastusyhteisöt seuraavat liike-elämän kehityksen mukana ottamalla uudenlaista Big Dataa hyödyntävää analytiikkaa osaksi tilintarkastustoimeksiantoja. Big Data vaikuttaa tämän johdosta myös suoraan tilintarkastajiin ja heidän työskentelyynsä. Big Datan ollessa varsin tuore ja ajankohtainen ilmiö tilintarkastuksessa, sen parissa työskentelemisen vaikutuksia tilintarkastajan arkipäivän työhön on mielenkiintoista ja järkevää tutkia syvällisemmin.

## **1.2 Tutkielman tavoite, tutkimusongelman asettelu, perustelut ja rajaus**

Tällä tutkielmalla on selkeä tavoite. Tämän tutkielman tavoitteena on selvittää Big Datan ja Big Dataa prosessoivan analytiikan vaikutuksia tilintarkastajan kognitiiviseen työskentelyyn ja tätä kautta myös koko tilintarkastukseen. Tutkielman tutkimuskysymys koostuu yhdestä päätutkimuskysymyksestä ja yhdestä tähän nivoutuvasta alatutkimuskysymyksestä. Nämä ovat esiteltyinä seuraavana.

Tutkielman päätutkimuskysymys;

*Miten Big Datan ominaisuudet ja Big Data-analytiikka vaikuttavat tilintarkastajan käytännön työhön tilintarkastajan kognitiivisesta näkökulmasta tarkasteltuna?*

Tutkielman alatutkimuskysymys;

*Miten Big Datan ja Big Data-analytiikan vaikutukset tilintarkastajan kognitiivisessa työssä heijastuvat tilintarkastukseen kokonaisuutena?*

Tämän tutkielman luomiselle on kolme pätevää perustetta. Tutkielman aihe on ensinnäkin ajankohtainen, koska suuret tilintarkastusyhteisöt implementoivat tutkielman luontihetkellä 2010 -ja 2020-lukujen vaihteessa jatkuvasti lisää Big Dataa hyödyntävää data-analytiikkaa osaksi toimeksiantojaan (Deloitte LLP 2016; Fullerton 2016; Rapoport 2016; EY 2016). Toiseksi, että Big Datan hyödyntäminen edellyttää tilintarkastajalta uudenlaista näkökulmaa ja ajattelua, jonka tutkiminen on tärkeää. Big Datan prosessoiminen myös luo tilintarkastajalle uudentyyppisiä haasteita. Näiden haasteiden ymmärtäminen on tärkeää tilintarkastusyhteisöjen implementoidessa uudentyyppistä analytiikkaa. Kolmanneksi, Mc Kinney, Yoos ja Snead (2017, 63) mukaan tilintarkastajille tärkeisiin Big Datan prosessoinnissa ja analysoinnissa ilmeneviin kognitiivisiin tekijöihin liittyvää kirjallisuutta ja tutkimusta on olemassa vielä valitettavan vähän. Tämä tutkielma paikkaa osaltaan tätä aukkoa.

Tämän tutkielman aihepiiri on selkeästi rajattu. Tutkielman teema keskittyy perinteiseen lakisääteiseen tilintarkastukseen rajaten Big Datan ja Big Data-analytiikan mahdollistamat muut mahdolliset tilintarkastusyhtiöiden tarjoamat lisäpalvelumahdollisuudet aihepiirin ulkopuolelle. Tutkielma on rajattu käsittelemään Big Datan ja Big Data-analytiikan vaikutuksia tilintarkastajan kognitiiviseen työskentelyyn ja tätä kautta koko tilintarkastukseen. Täten muut Big Dataan ja Big Data-analytiikkaan liittyvät tilintarkastusta koskettavat aiheet ja tarkastelunäkökulmat ovat rajattuina pois.

### **1.3 Tutkimusmetodologia ja tutkielman rakenne**

Tämä tutkielma on toteutettu laadullisena haastattelututkimuksena. Laadullista tutkimusta kutsutaan myös kvalitatiiviseksi tutkimukseksi. Tutkimuksen kohteena on tilintarkastajan oma henkilökohtainen kokemuspohjainen perustuva näkemys ilmiöstä, jota ei ole mahdollista tutkia määrällisin, kvantitatiivisin tutkimusmenetelmin. Laadulliselle tutkimukselle ominaisia piirteitä ovat tutkittavan joukon pieni otoskoko ja tutkittavien henkilöiden korkea tietämys tutkittavan ilmiön osalta (Tuomi & Sarajarvi 2006, 87 – 88). Laadullista tutkimusta kuvaavaa on, ettei tutkijalla tarvitse tai edes saa olla ennakkolehtämuksia tutkimuskohdetta tai tutkimuksen tuloksia kohtaan. Laadullinen tutkimus

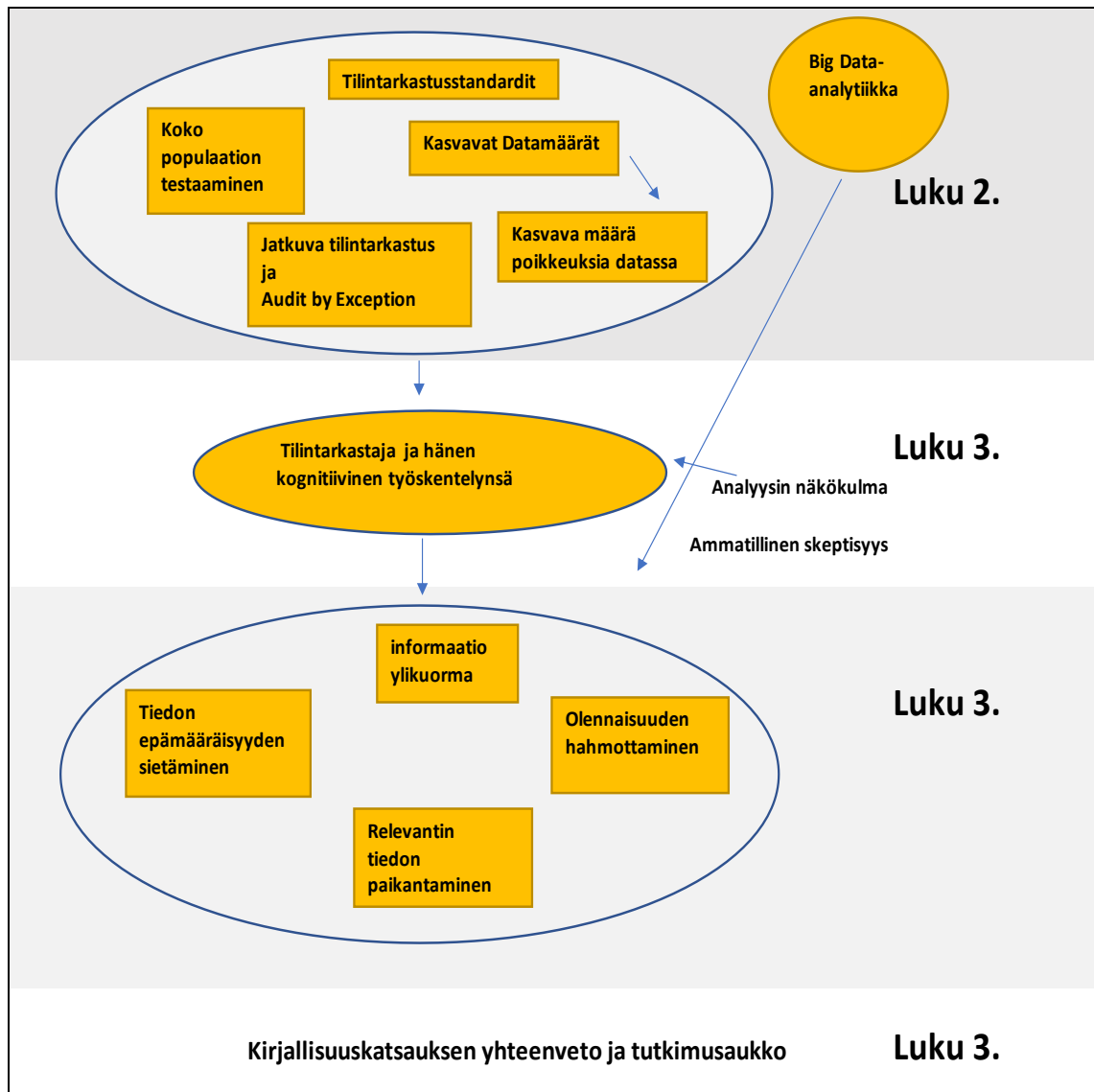
mahdollistaa sen, että aineiston analyysin kautta tutkimuksen tekijä voi oppia ja löytää täysin uusia näkökulmia tutkittavasta aiheesta tutkimusprosessin edetessä. Tällöin tutkija ei jää vain todentamaan ennakko-olettamuksiaan. (Eskola & Suoranta 2003, 19 –20.) Tutkimuksen kohteen suuntautuessa tilintarkastajien omakohtaisiin näkemyksiin, tarkoittaa tämä tutkielman tarkastelunäkökulman olevan subjektiivinen. Grönfors (1985,14) mainitsee subjektiivisuuden olevan erityisesti laadullista tutkimusta luonnehtiva piirre.

Tutkielman aineistonkeruumenetelmää ja aineistonkeruuprosessia tarkastellaan tarkemmin luvussa 4.1. Aineiston keruu menetelmäksi valikoitui Hirsijärvi ym. (1997, 203) mukainen puolistrukturoituitu haastattelumenetelmä, eli teemahaastattelumenetelmä. Tämänkaltainen menetelmä sopii hyvin laadullisen, subjektiivisen aiheen tarkasteluun.

Tämä tutkielma on rakenteeltaan johdonmukainen. Tämä tutkielma koostuu johdannosta, kahdesta teoreettisesta pääluvusta, empiirisestä osiosta ja johtopäätöksistä. Tutkimusongelman teoreettinen viitekehys rakentuu teemaa käsittelevien tutkimusten ja niitä tukevan muun kirjallisuuden perusteella. Teorian kautta vahvistuvalle tutkimusongelmalle pyritään tämän jälkeen löytämään uusia näkökulmia laadullisen empiirisen tutkimuksen avulla. Seuraavalla sivulla esiteltävässä kuviossa on hahmoteltuna tämän tutkielman teoriaosuuden rakenne. Tutkielman kaksi teorialukua (luvut 2 ja 3) käsittelevät samaa aihepiiriä hieman eri näkökulmista. Tutkielman teoreettinen viitekehys rakentuu tieteellisille artikkeleille, joissa pääpaino on annettu vuoden 2010 jälkeen tehdyille tutkimuksille. Vanhoja, edelleen tunnettuja käsitteitä ja teorioita on kuitenkin ammennettu melko vanhoistakin tutkimuksista tutkielman teoriataustaa vahvistamaan. Kummankin pääteorialuvun alussa esitellään ensin tiivistetysti niiden olennaisin sisältö alaluvuittain.

Teorialukuja on kaksi (luvut 2 ja 3), ja niillä omat tarkoituksensa. Ensimmäisessä teorialuvussa keskitytään Big Datan ja Big-Dataa hyödyntävän analytiikan vaikutuksiin tilintarkastustyöhön yleisellä tasolla. Ensimmäisen teorialuvun tarkoituksena on luoda yleisymmärrys siitä, mitä tilintarkastaja käytännössä tarkastelee, työstää ja prosessoi Big Datan ja tätä työstävän analytiikan parissa työskennellessään. Ensimmäisessä teorialuvussa käsiteltäviä teemoja ovat Big Dataa hyödyntävät analyttiset menetelmät ylipäätään, koko populaation testaaminen, jatkuva tilintarkastus, tilintarkastuksen automatisaatio ja ajankohtaisten tilintarkastusstandardien vaikutus tässä tutkielmassa käsiteltävään aiheeseen. Toisessa teorialuvussa peilataan Big Datan prosessointia tilintarkastajan kognitioon. Toisen teorialuvun tarkoituksena on tarkastella miten tilintarkastaja prosessoi mielessään Big Dataa tehdessään työtään. Tätä pohjustetaan teorialuvun alussa tarkastelemalla tilintarkastajan kognitiivista työskentelyä yleisesti. Big Datan kognitiiviseen työstämiseen liittyvät teemat koostuvat suuren datamäärän aiheuttamasta informaatioylikuormasta, olennaisen ja relevantin informaation

hahmottamisesta suuresta tietomäärästä, sekä tiedon epäselkeyden sietämisestä epämääräistä ja laajaa dataa käsiteltäessä. Tarkasteltavina ovat myös tilintarkastajan analyttinen näkökulman -ja ajattelutavan muutos Big Data-aikakaudella. Lisäksi toisessa teorialuvussa luodaan lyhyesti näkemys Big Datan vaikutuksista tilintarkastajan ammatilliseen skeptisyyteen. Toisen teorialuvun päättää kirjallisuuskatsauksen yhteenveto alaluvussa 3.11, jossa myös määritellään teorian pohjalta esiin tulevan tutkimusaukon myötä tarve empiirisen osuuden teettämiselle.



Kuvio 1: Tämän tutkielman teoriaosuuden rakenne

Tämän tutkielman empiirinen osuus sijoittuu päälukuun 4. Empiirisen osuuden alussa tehdään katsaus aineistonkeruumenetelmään ja aineistoa varten haastateltuihin tilintarkastajiin. Tämän jälkeen esitetään tutkielman tavoitteen kannalta olellaisin aineisto loogisessa järjestyksessä. Empiirisen luvun lopussa tehdään pohdinta

tutkimuksessa esiin tulleista tuloksista ja peilataan saatuja tuloksia tutkielman kirjallisuuskatsausosuuteen.

Tämän tutkielman empiriaa seuraa tutkielman yhteen vetävä luku 5. Tämä luku koostuu tämän tutkielman empiirisiin tuloksiin perustuvista johtopäätöksistä, sekä tämän tutkielman reliabiliteetin, validiteetin, yleistettävyyden, tutkielman luoman lisäarvon ja rajoitteiden tarkastelemisesta.



## 2 BIG DATAN JA BIG DATA-ANALYTIIKAN VAIKUTUKSET TILINTARKASTUSTYÖHÖN

Tutkielman teoriaosuuden ensimmäisessä pääluvussa keskitytään tarkastelemaan Big Datan ja Big Dataa prosessoivan analytiikan vaikutuksia tilintarkastustyöhön yleisellä tasolla. Big Datan ja Big Data-analytiikan kognitiiviset vaikutukset tilintarkastajan työskentelyyn jätetään tämän tutkielman myöhempään teorialukuun.

Tämä pääluke jakautuu kuuteen alalukuun. Tämän pääluvun ensimmäisen alaluvun alussa lähennetään johdantoa syventäen Big Datan ominaisuuksiin ja lähteisiin. Toisessa alaluvussa tarkastellaan Big Dataa prosessoivia analyttisiä menetelmiä, joita esiintyy yhä enenevässä määrin tilintarkastuksessa. Tämän pääluvun kolmannessa alaluvussa tarkastellaan koko populaation testaamista, neljännessä alaluvussa jatkuva tilintarkastusta ja viidennessä alaluvussa tilintarkastuksen automatisaatiota. Näihin teemoihin myös Big Dataa prosessoiva data-analytiikka perustuu. Tämän pääluvun kuudennessa alaluvussa tehdään katsaus ajankohtaisten tilintarkastusstandardien suhteesta Big Datan hyödyntämiseen.

### 2.1 Big Datan ominaisuudet

Tässä alaluvussa keskitytään kuvailemaan Big Datan ominaisuuksia johdantoa syventäen. Jo heti tämän aihepiirin käsittelemisen alussa on syytä mainita Big Datan olevan ilmiönä ja käsitteenä suhteellisen uusi (Cao ym. 2015, 423). Tämän johdosta täydellistä yhteisymmärrystä Big Datan tarkasta sisällöstä ei ole vielä olemassa (Labrinidis & Jagadish 2012, 2032). Tässä alaluvussa pyritäänkin seuraavaksi kuvailemaan näinkin epämääräisen olomuodon omaavaa ilmiötä karkeasti, jotta lukija saa riittävän ja olennaisen käsityksen Big Datan ominaisuuksista.

Ymmärrys Big Datan olemuksesta vaihtelee. Big Datan käsite on alkuaan syntynyt 1990-luvulla ja sen esiintyminen kirjallisuudessa on yleistynyt edelleen 2000-luvulla. Eri tutkijat ovat aikojen saatossa antaneet hieman toisistaan eroavia määritelmiä Big Datalle. Big Dataan poikkeuksetta liitettyjä ominaisuuksia ovat kuitenkin erittäin suuri volyyymi ja datamäärän eksponentiaalisen nopea lisääntyminen (Hasanat 2018, 581). Kaikista karkeimpiin Big Datan määritelmiin kuuluu Chen ym. (2013, 157) määritelmä, jota mukailleen Big Dataan sisältyvät kaikki tavanomaisten nykyaikaisten tietojärjestelmien varastointi- ja käsittelykapasiteetin ylittävät valtavat datamäärät. Big Datan olemusta voidaan tarkastella myös sen ominaisuuksien kautta. Big Dataan liitetyt perusominaisuuksiksi mainitaan volyyymi, variaatio, totuudenmukaisuus (engl. ”*veracity*”), nopeus (engl. ”*velocity*”), ja vaihtelevuus (engl. ”*variability*”) (EY 2015a).

Lähteestä riippuen Big Dataan liitettävien englannin kielisessä kirjallisuudessa V-kirjaimella alkavien ominaisuuksien määrä vaihtelee kuitenkin huomattavasti.

Kuten edellä kuvatusta Chen ym. (2013, 157) Big Datan määritelmästä ilmenee, Big Data on suurivolyymistä. Big DaBarnaghi, Sheth & Henson (2013, 10) kuvailevat Big Datan suuren volyymin tekevän sen määrittelystä, hakemisesta, prosessoimisesta ja integroimisesta muuhun tietoon ylipäättään haasteellista.

Big Dataan liittyy laaja variaatio (Grable ja Lyons 2018, 17 – 18). Big Datan variaatioon liittyen Labrinidis ja Jagadish (2012, 2032 – 2033) toteavat vaihtelevat, heterogeeniset datan muodot, lähteet, datan sisältämät sanomat ja laadun vaihtelut ovat samalla Big Dataan liittyvää pääomaa, mutta samalla myös koetinkivi sen prosessoimiselle, hallinnoimiselle ja hyödyntämiselle. Albeshri ja Thayananthan (2018, 582) mukaan laajan variaation vuoksi Big Datan hyödyntäminen vaatiikin tilastollisuuteen perustuvaa prosessointia taustalleen.

Big Dataa kuvaavaa on sen totuudenmukaisuus. Sivarajah ym. (2017, 273) käsittelevät Big Datan totuudenmukaisuutta, kuten suuren datamäärän aiheuttamia anonymiteettikysymyksiä ja epäjohdonmukaisuuksia hankalina taustatekijöinä sen ymmärtämisen ja hyödyntämisen suhteen. Koyce (2017, 29) mukaan totuudenmukaisuuden ominaisuus on suurin Big Dataan liittyvä haaste. Tämä johtuu siitä, että Big Datan ymmärtämistä hankaloittavat siihen liittyvät erilaiset häiriötekijät, harhat ja tavallista useammin esiintyvä epätavallinen informaatio.

Big Datan nopeudella tarkoitetaan datan synnyn huomattavaa yleisyyttä (Koyce, 2017, 29). Chen ym. (2013, 158) summaavat datan käsittelyn suhteen rakenteeltaan ei-homogeenisen Big Datan poikkeuksellisen suuren nopeuden olevan suuri haaste sen parissa työskenteleville henkilöille. Albeshri ja Thayananthan (2018, 582) mainitsevat esimerkiksi tunnin aikana tapahtuvien transaktioiden määrän nopeuden vaihtelevan lukemattomien syiden vuoksi. Oikeanlaisen data-analyttisten välineiden avulla on mahdollista analysoida teksti-, -ääni -ja videoperustaisen datan nopeuden syytä.

Big Datan vaihtelevuudella tarkoitetaan sen sisältämän sanoman nopeaa vaihtelevuutta ajan funktiona (Sivarajah ym. 2017, 273). Big Datan vaihtelevuus liittyy myös datan sisällön monitulkintaisuuteen, tässä hetkessä ja eri ajanhetkinä. Esimerkiksi samat sosiaalisessa mediassa ilmaistut sanat tai datapisteet ylipäättään voidaan tulkita hyvin monella tapaa eri ajanhetkinä. Big Datan hyödyntämisessä hankaluutena onkin se, että sen ymmärtämiseen hyödynnettävien data-analytiikan sisältämien algoritmien tulee olla kykeneviä ymmärtämään datan sisältämiä konteksteja. (Zhang ym. 2015b, 34.)

Taloudellinen, erityisesti tilintarkastajia ja muita taloushallinnon ammattilaisia kiinnostava Big Data voi olla jäsenneiltyä (eng. ”*structured*”), osittain jäsenneiltyä (engl. ”*semi-structured*”) tai ei-jäsenneiltyä (engl. ”*unstructured*”) (Chen ym. 2013, 158). Nämä kolme datatyyppeä lähteineen on syytä läpikäydä seuraavana.

Jäsennelty data on selkeämuotoista ja helposti johonkin tiettyyn kohteeseen kohdistettavaa. Jäsennelty data ei ole datatyypinä uusi Jäsenneltyä dataa on helppoa kerätä, rajata ja tulkita, sillä se on rakenteeltaan ja lähteiltään suhteellisen yhtenäistä. (Moffit & Vasarhelyi 2013, 9 – 10.) Puhuttaessa Big Datasta liiketoimintaympäristössä, Big Data -ekosysteemin keskiössä sijaitsee perinteinen taloudellinen data. Tämä data on tyypillisesti yrityksen sisäisistä organisaation laajuisista ERP-toiminnanohjausjärjestelmistä (*”Enterprise Resource Planning Systems”*) kerättyä dataa. (Moffit & Vasarhelyi 2013, 5.) ERP -järjestelmät keräävät ja koordinoivat tietoa ja toimintoja eri puolilta yritystä ja yrityksen ulkopuolelta samaan tietokantaan, joka toimii jaettuna datavarastona yrityksen omiin käyttötarkoituksiin (Appelbaum, Kogan, Vasarhelyi & Yan 2017, 31).

Viimeisen vuosikymmenen aikana teknologiset trendit ovat laajentaneet datan sisällön näkökulmaan myös ei-jäsennellyn, yrityksen rajapintojen ulkopuolella sijaitsevan datan. (Ramulkan 2015, 15.) Jäsennelty data muodostaa jopa 95 prosenttia aivan kaikesta nykypäivän datasta (Gandomi & Haider 2015, 143). Jäsennelty data on laajaa ja muodoiltaan vaihtelevaa. Jäsennelty data tulee erittäin moninaisista ulkoisista lähteistä, kuten äärimmillään sosiaalisesta mediasta, GPS-laitteista ja valvontakameroiden tallenteista. Se voi olla tekstimuodossa, GPS-muodossa, kuvina, numeroina, äänitallenteina tai videoituna (Moffit & Vasarhelyi 2013, 5; Cao ym. 2015, 425). ei-jäsennellyksi dataksi luokiteltavat video ja- audiodata ovat muuhun dataan nähden fyysisestä erillisyydestään huolimatta integroitumassa nopeassa tahdissa osaksi Big data ekosysteemiä. Videovalvonta ja älypuhelin videotallenteet edustavat valtavaa ja yhtenäistä datavirtaa. (Moffit & Vasarhelyi 2013, 9 – 10.)

Osittain jäsennelty data on ominaisuuksiltaan selkeästi jotain jäsennellyn ja ei-jäsennellyn datan väliltä. Osittain jäsennelty data voi esimerkiksi olla taloudellista dataa, joka merkitty esimerkiksi XML- tai XBRL- tunnisteella tehden siitä erityisen nopeaa (Moffit & Vasarhelyi 2013, 5; Vasarhelyi, Alles & Williams 2010, 10).

Big Data on jatkamassa laajenemistaan ja hakee uusia olomuotoja jäsennellyn ei-jäsennelty datan ulkopuolelta. Viimeisin ja jatkuvasti tuloaan tekevä Big Datan muoto on IoT-data (engl. *”Internet of Things”*). IoT-data voi olla yritykseen nähden sekä ulkoista tai sisäsyntyistä, mutta tyypillisintä sille on sen keskusteleva luonne yrityksen sisällä, rajapinnoilla ja ulkopuolella (Brown-Liburd & Vasarhelyi 2015, 3). IoT-data perustuu laajaan skaalaan internetiin kytkettäviin laitteita, jolloin ne voivat olla interaktiivisessa vuorovaikutuksessa toisten laitteiden ja ihmisten kanssa (Morgan 2014).

Kuten edellä on käynyt ilmi, Big Data voidaan määritellä hyvin monella tapaa. Myös Big Datan lähteet ja olomuodot vaihtelevat voimakkaasti. Seuraavassa alaluvussa käsitellään niitä menetelmiä, joilla Big Dataa on mahdollista käsitellä ja hyödyntää erityisesti tilintarkastuksen tarpeisiin.

## 2.2 Big Dataa prosessoivat analyttiset menetelmät tilintarkastuksessa

Datamäärän nopea ja jatkuva ympärillämme tapahtuva lisääntyminen luo tilintarkastajille mahdollisuuden tilintarkastuksen laadun parantamiseen. Tämä tapahtuu valjastamalla suuret datamäärät kehittyneen teknologian avulla aivan uudenkaltaiseksi tilintarkastusevidenssiksi (Wuerfel & Sikora 2017, 12). Big Data sisältää erityyppisiä informaatiolähteitä, jotka sallivat tilintarkastajan tutkia vapaasti asiakkaidensa transaktioiden ja liiketoimintaprosessien tilannetta. Big Data ylipäättään vähentää myös tilintarkastajien riippuvuutta asiakkaistaan datan saannin suhteen (Yun, Hoodguin & Zhang 2015, 346). Tämän johdosta tarkoituksenmukaisen Big Datan hyödyntäminen analytiikan avulla tuo runsaasti lisäarvoa tilintarkastajien työskentelylle (Sun & Vasarlehy 2017, 26).

Ensimmäiseksi tässä pääluvussa tehdään katsaus tilintarkastuksessa saatavilla olevan Big Data-analytiikan luonteeseen, ja myöhemmin käsitellään tarkemmin tilintarkastuksessa nykyisin saatavilla olevia erilaisia Big Data-analytiikan menetelmiä.

Big dataa analysoivat menetelmät ovat osa data-analytiikkaa ja data-analyysin konseptia (O'Donnell 2015, 24). Data-analyysin käsitteen alle sijoittuu hyvin moninainen määrä erilaisia teknologioita (Kwon, Lee & Shin 2014, 387 – 388). Data-analyysit voidaan karkeasti jaotella kuvaileviksi, diagnostisiksi, ennustaviksi ja ohjaileviksi analyyseiksi (Tschakert, Kokina, Kozlowski & Vasarlehy ym. 2016). Data-analytiikalla tarkoitetaan tilintarkastuksessa tietokoneavusteista tutkimusta, jonka avulla pyritään helpottamaan tilintarkastuksen toteutusta tehostamalla asiakkaaseen liittyvien yksityiskohtaisemman digitaalisen datan hyödyntämistä (Titera 2013, 325 – 326).

Vaihtelevaan Big Datan määritelmään nähden Big Data-analytiikka on selkeästi määriteltävissä. Big Data-analytiikkaa voidaan pitää oikea-aikaisena ja automatisoituna prosessina, jossa Big Datan tutkiminen, jaottelu, muokkaaminen, mallintaminen päätöksenteon kannalta hyödyllisiksi kuvioiksi ja johtopäätöksiksi seuraavat toisiaan (Cao ym. 2015, 424). Big Dataa prosessoivien analyttisten tekniikoiden luonne vaihtelee laskennallisesta matematiikasta, tilastollisiin malleihin ja koneoppimisen menetelmiin (Albeshri & Thayananthan 2018, 528). Cao ym. (2015, 425) mukaan Big Data-analytiikan avulla tyypillisesti yhdistellään dataa useista erinäisistä lähteistä. Tämä data voi olla siten sekä jäsenneiltyä että ei-jäsenneiltyä, tai sisältää kumpaakin näistä datatyypeistä.

Big Dataa hyödyntävät data-analyttiset työkalut koskettavat nykyisin myös tilintarkastusta. Tilintarkastajan tarkastustoimenpiteisiin liittyvä varsinainen problematiikka yhdentyy tilinpäätöksestä löytyvien virheiden diagnosoimiseen. Tarkastamattoman tilinpäätöksen sekä luontaisten riskitekijöiden alkukatsauksen jälkeen

tilintarkastaja luo perinteisesti sarjan hyväksyttäviä diagnooseja ja strategioita lisäevidenssin löytämiseksi, jota hyödynnetään tämän jälkeen tilinpäätöksen sisältämiä virheitä koskevien hypoteesien testaamisessa ja todentamisessa (Libby & Frederick 1990, 349). Cao ym. (2015, 424) summaavat tilintarkastajien työn pääasialliseksi tarkastuksen fokusalueeksi asiakkaidensa tilinpäätöstiedoissa ja niihin liittyvissä muissa asiakkaan informaatiokanavissa sijaitsevien transaktioiden, saldojen ja muiden taloudellisten tietojen kokonaisvaltaisen tutkimisen. De melo Mendes, Niyama ja Silva (2018, 363) toteavat myös asiakkaan omaisuuserien, kuten varaston ja käyttöomaisuuden arvostuksen arviointiin liittyvien kysymysten olevan myös eräs tilintarkastajien työnkuvaan olennaisesti liittyvä osa-alue. Big Data-analytiikalla on mahdollista tehostaa edellä kuvattua tilintarkastajan työnkuvaa.

Historiaan katsomalla on kuitenkin helppo todeta tilintarkastuksen omaksuneen kaikesta huolimatta suhteellisen hitaasti varsinaista liikemaailmaa mullistaneita teknologioita (Alles 2015, 447). Big Data-analytiikka ei ole tässä suhteessa poikkeus. Big Datan hyödyntäminen tilintarkastustyössä on kehittynyt ikään kuin jälkijunassa tilintarkastuksen asiakkaisiin verrattuna. Kysymykset siitä, miten Big Data ja sitä tukeva analytiikka tulevat vaikuttamaan tilintarkastukseen ovat nykyisin runsaan ja hedelmällisen tarkastelun alla. (Appelbaum, Kogan & Vasarlehyi 2017, 21 – 22.) Big Data-analytiikan tekniikoiden todetaan edustavan todellista arvoa luovaa mahdollisuutta tilintarkastukselle, jota ei kuitenkaan ole vielä hyödynnetty samoissa määrin kuin monilla muilla liiketoiminnan alueilla. On kuitenkin sanomattakin selvää, että perinteiset tilintarkastuksessa hyödynnetyt analyysimenetelmät ovat käyttökelvottomia suuren volyymin, vaihtelevuuden ja nopeuden kuvaaman Big Datan käsittelemisessä. (Gepp, Linnenluecke, O'Neill & Smith. 2018, 108.) On ylipäätään vain ajan kysymys, milloin myös tilintarkastuksessa omaksutaan samassa mittakaavassa osaksi päivittäisiä tarkastusrutiineja samankaltaisia Big Dataa prosessoivia analyttisiä työkaluja, kuin muualla liike-elämässä on sen edellä otettu käyttöön. (Cao ym. 2015, 428.)

Tilintarkastajat kokevat Big Dataa prosessoivan analytiikan tarpeelliseksi työssään. Appelbaum (2016, 17) toteaa tilintarkastajien olevan lähtökohtaisesti kiinnostuneita Big Datan hyödyntämisestä kahdesta syystä. Ensinnäkin sen vuoksi, että useimmat tilintarkastusasiakkaat hyödyntävät nykypäivänä Big Dataa omissa liiketoimintaprosesseissaan. Toiseksi, tilintarkastajat itse tahtovat päästä hyödyntämään Big Dataa tehostaakseen omaa työtään.

Tilintarkastus hyötyy Big Data-analytiikasta kokonaisvaltaisesti. Tämä ilmenee erityisesti asiakkaan riskien arvioinnissa, asiakkaan toimeksiannon jatkamiseen liittyvissä kannattavuusarvioinneissa, väärinkäytösten selvittämisessä, analyttisten menetelmien tehostumisena ja tehokkaampien johtopäätösten, sekä laajaulotteisemman ymmärryksen luomisena asiakkaastaan. (Cao ym. 2015, 424.) Tilintarkastukseen ja Big Data-analytiikkaan liittyen viime aikoina tehdyissä tutkimuksissa eräs tärkeä aihe on ollut

asiakkaan tuottamiin sisäisiin taloudellisiin lukuihin perustumattoman Big Datan soveltuvuus lakisääteiseen tilintarkastukseen (Wang & Cuthbertson 2015, 158.). Geat ja Xie (2017) havainnollistavat Big Dataa prosessoivan analytiikan luovan suuresta datamäärästä analyysyjä esimerkiksi kaavioiden muodossa, joita tilintarkastajat voivat hyödyntää tarkastuksessaan. Tämän myötä tilintarkastus tehostuu. Data-analyttiset menetelmät mahdollistavat myös koko populaation testaamisen, ja niiden avulla on mahdollista käsitellä taloudellisen datan lisäksi ei-taloudellista dataa.

Big Dataa hyödyntävä data-analytiikka hyödyttää tilintarkastusta kaikissa sen vaiheissa. Se auttaa jo tilintarkastuksen suunnitteluvaiheessa kullekin tilintarkastusasiakkaalle sopivimpien tarkastustoimenpiteiden suunnittelemisessa ja tarpeellisen tilintarkastusevidenssin hankinnan arvioimisessa. Tilintarkastuksen toteutusvaiheessa se auttaa tilintarkastajia tutkimaan asiakkaidensa tilejä aiempaa tarkemmin. (Capriotti 2014, 2 – 3.)

Myös ei-jäsenneltyä dataa on mahdollista hyödyntää tilintarkastuksessa data-analytiikan avulla. Tilintarkastajat kykenevät ainakin periaatteessa ulkoista Big Dataa hyödyntämällä varmistamaan asiakkaidensa transaktioita laskujen ja kuittien ohella myös jäsenneltymättömän datan tarjoamalla moniulotteisemmalla evidenssillä, kuten valokuvilla, videoilla, äänellä, tekstillä ja GPS-paikannuksilla (Moffit & Vasarlehyyi 2013, 5 ja 9). Esimerkiksi visuaalisten ja tekstuaalisten analyysien tekniikat helpottavat lähitulevaisuudessa tilintarkastajia yhä enemmissä määrin tarkastelemaan suurivolyymistä ei-jäsenneltyä dataa yhä luovemmin keinoin, sekä paikantamaan ei-jäsennellyssä datassa sijaitsevia piilotettuja väärinkäytöksen tai virheen jäljille johdattelevia johtolankoja. Tilintarkastus voi hyötyä suurissa määrin visuaalisista ja tekstuaalisen analyysin tekniikoista korkeariskisissä tilintarkastustoimeksiannoissa, johtuen näiden tekniikoiden tuottamasta korkeammasta tietoisuuden lisääntymisestä. (Aldhizer 2017, 30 – 31.) Ei-jäsenneltyä dataa edustavat audio-, audio valvonnan-, ja media audio virrat voidaan yhdistää tiettyyn tilintarkastusasiakkaaseen liittyväksi ja lopulta tallentaa automaattisesti tekstin muotoon. Sähköposteja ja ihmisten lisäämiä kommentteja mediassa on mahdollista kerätä ja yhdistää erilaisiin transaktioihin. Myös datan yhteen liitettävyyttä tukevat elementit, kuten tuotteisiin liitettävät paikannusta ja varastohallintaa helpottavat RFID-sirut ovat yhdistettävissä myös eri datavirtoihin. (Moffit & Vasarhelyi 2013, 9 – 10.)

Tällä ajanhetkellä on olemassa jo ilmeisen vahvoja todisteita siitä, että suuret tilintarkastusyhteisöt ja tilintarkastusala ylipäättään ovat tekemässä investointeja Big Data -analytiikkaan. Big Datan ja Big Dataa hyödyntävien analyttisten menetelmien hyödyntämisen väärinkäytösten tutkimisessa tiedetään lisääntyneen erityisesti ennustavissa tilintarkastuspalveluissa, joiden kysyntä on viime aikoina kasvanut. (Zabihollah & Wang 2017, 103 – 104.) Tilintarkastusyhteisö Deloitte on jo vuonna 2016 aloittanut yhteistyön Kira Systemsin kanssa kehittääkseen koneoppimiseen kykenevää

nopeaa tekstinlouhinnan ohjelmistoa. Tällä pyritään hyödyntämään liiketoimintadokumenttien sisältämää ei-jäsenneltyä dataa tilintarkastus -ja konsultointipalvelujen tukena. (Deloitte LLP 2016.) Tilintarkastusyhteisö EY on puolestaan aloittanut viime vuosina kartoittamaan Big Data-analytiikkaan kuuluvan niin sanotun lohkoketniikan (blockchain technology) hyödyntämismahdollisuuksia tilintarkastuksessa (EY 2016). Samanaikaisesti myös KPMG aloitti yhteistyön IBM:n kanssa hyödyntääkseen IBM Watsonia, syväajatteluteknologiaan (engl. ”deep learning technology”) perustuvaa kognitiivista koneoppimista tukeakseen tilintarkastajan työskentelyä suurten datamäärien parissa (Fullerton 2016). Edellisten tilintarkastusyhteisöjen ohella PwC on tehnyt suuria investointeja uuteen automatisaatioon liittyvään tilintarkastusteknologiaan (Raport 2016).

Esimerkkinä Big Data-analytiikan tekniikoista tilintarkastuksessa ovat varsinkin erilaiset automaattiset koneoppimisen (engl. ”machine learning”) tekniikat, joihin myös pian kuvattu syväajatteluteknologia kuuluu. Koneoppimistekniikoiden avulla on mahdollista havaita sellaisia odottamattomia vuorovaikutussuhteita suuren datamäärän sisällä, joiden paikantaminen olisi mahdotonta yksinkertaisemmilla tekniikoilla. Koneoppimistekniikoiden käänttöpuolena on niin sanottu ”mustan laatikon ilmiö”, joka tarkoittaa sitä, ettei tilintarkastaja tai muu tekniikkaa hyödyntävä osallinen saa mitenkään tarkasti selville kuinka kyseinen monimutkainen teknologia ja sen algoritmit ovat päässeet lopulliseen johtopäätökseensä. (Andrew 2017, 1.)

Syväajatteluteknologian kaltaisen koneoppimistekniikan hedelmällisin hyöty on sen käyttökelpoisuus hyvin erinäisistä ja epämääräisistä lähteistä kumpuavan datan analysoimisessa ja yhdistelemisessä. Tämä tekee siitä arvokkaan menetelmän osana Big Datan prosessointiin suuntautuvaa teknologiaa, sillä Big Data on hyvin pitkälti kategorisoimatonta (Najafabadi, Villanustre, Khoshgoftaar, Seliya, Wald & Muharemagic. 2015, 1). Tekoälyyn ja automaatioon perustuvan syväajatteluteknologian avulla on mahdollista parantaa tilintarkastajan työtä tilinpäätöstarkastuksessa, valvoa sisäisiä kontroleja ja vähentää rutiinitoimenpiteitä. Syväajatteluteknologian avulla tilinpäätökseen liittyvän evidenssin osiin voidaan automaattisesti liittää varsinaista evidenssiä tukevaa materiaalia videon pätkistä, sekä mobiilisesti, sähköpostitse tai sosiaalisessa mediassa käydyistä keskusteluista ja haastatteluista. Tämä ei- taloudellinen data voidaan tämän jälkeen yhdistellä asiakkaan taloudellisen datan kanssa. (Sun & Vasarlehy 2017, 25 – 28.)

Visuaalisia ja tekstuaalisia analyysin data-analyttisiä tekniikoita ovat myös esimerkiksi jäseneltyä ja jäseneltyä dataa moninaisista ja kompleksisista lähteistä yhdistelevä monilähdeanalyysi (Aldhizer 2017, 33). (engl. ”multiple- source analysis”), tilastollisiin riippumattomien ja riippuvien muuttujien välisiin algoritmeihin ja ryhmittelyyn kautta löytyviin eroavaisuuksiin perustuva klusterianalyysi (”cluster analysis”) (Thiprungsri & Vasarlehy 2011, 80 – 81), suuresta tekstimuotoisesta

datamassasta ”rivien välistä” löytyvien konseptien tunnistus (engl. ”*concept extraction*”) (Aldhizer 2017, 33), datan tekijöiden ja asiayhteyksien välille tekstidatasta linkkejä luova linkkianalyysi (Engl. ”*link analysis*”) (Ashford, 2013), kuvien perusteella mahdollisten virheiden ja väärinkäytösten maantieteellisiä sijainteja erotteleva paikkatietoanalyysi (engl. ”*geospatial analysis*”) (Aldhizer 2017, 32) sekä erityyppisiä ja väärinkäyttösriskin dataelementtejä eri värein ja väriastein esittelevät lämpökartat (engl. ”*heat map*”) (Aldhizer 2017, 32).

### 2.3 Koko populaation testaaminen

Big Datan tehokas hyödyntäminen edellyttää tilintarkastukselta tarkastusnäkökulman siirtämistä rajallisista otoksista koko populaation kattavalle tasolle. Tilintarkastuksen luonteen muutos otantaan perustuvasta näkökulmasta koko datajoukon tarkastelemiseen on kuitenkin suuri askel tilintarkastukselle. Tämä muutos ole tapahtumassa nopealla aikataululla (Whithouse 2014, 28). Koko populaation testaamisen ollessa Big Datan hyödyntämisen ydinpiirre, on perusteltua tarkastella koko populaation testaaminen heti data-analyyttisten menetelmien läpikäynnin jälkeen tässä kohtaa omassa alaluvussaan.

Big Data-analytiikan avulla on mahdollista analysoida kaikki tai lähes kaikki populaation sisältämä data. Koko populaation testaaminen mahdollistaa koko datamäärän kannalta olennaisimpiin datan sanoman sisältämiin suuntaviivoihin keskittymisen, mikä on koko populaation testaamisen etu. Otosperusteisessa tilintarkastuksessa epärelevantin datan eliminoimiseen tarvitaan merkittävä määrä huomiota, sillä näytteet ovat pieniä. Datan sisältämien poikkeusten vaikutusten ollessa suhteellisesti vähäisempi suuressa datamäärässä, korkeampi epärelevantin datan taso on hyväksyttävämpi tarkistettavien datamäärien ollessa suuria. (Cao ym. 2015, 426; Byrnes & Pawlicki 2014, 7.)

Koko populaation testaaminen luo tilintarkastukselle uusia mahdollisuuksia. Koko populaation tarkasteleminen luo aivan uudenlaisten väärinkäytösten ja virheiden tarkastelun mahdollisuuden tilintarkastajalle, nimittäin koko populaatioon levinneiden väärinkäytösten ja virheiden tarkastelun (Mc Kinney Jr ym. 2017, 76). Whithouse (2014, 28) mukaan koko populaation testaamisen etuna tilintarkastajalle ovat varsinkin epätavallisten tapahtumaketjujen ja paikkansapitämättömien sanomien erottaminen datasta, kun on mahdollista tarkastella samanaikaisesti asiakkaan koko ”jalanjälkeä” yhtenä kokonaisuutena. Geat ja Xie (2017) lisäävät koko populaation testaamisen kehittyneen analytiikan avulla vähentävän tilintarkastustoimenpiteisiin käytettävää aikaa huomattavasti. Tämän johdosta tilintarkastustiimi voi hyödyntää jäljelle jäävän asiakkaaseen budjetoidun ajan tutkien entistä laadukkaammin tilintarkastusta varten tehtyjen testausten tuloksia, ottamalla tarkemmin selvää asiakkaan dataan liittyvistä



trendeistä ja löytämällä siitä tilintarkastuskokonaisuuden kannalta lisää merkitseviä tekijöitä.

Tilintarkastukseen on perinteisesti liittynyt otanta riski. Otanta riskiksi kutsutaan tilintarkastajan tekemää otokseen perustuvaa väärää johtopäätöstä, jota hän pyrkii alentamaan otoskokoa kasvattamalla (Burtescu & Grigore 2011, 165). Tähän pohjaten, koko populaation testaamiseen siirtyminen tilintarkastuksessa tulee mitä todennäköisimmin jättämään otanta riskin konseptin hamaan historiaan, tai ainakin tämä muuttaa sen merkitystä.

Koko populaatiota testattaessa sekä datan määrä, että datasta löytyvien poikkeusten määrä kasvavat huomattavasti (Richins, Stapleton, Stratopoulos & Wong. 2017, 72). Näin ollen kokonaisuuden kannalta olennaisen tiedon huomiointi ja ylimääräisen tiedon jättäminen omaan arvoonsa ovat tilintarkastajalle entistäkin tärkeämpiä tekijöitä yhä datapitoisemmassa tilintarkastuksessa. Kuitenkin Mc Kinney Jr ym. (2017, 73) näkökulmaa mukaillen koko populaatiota analysoitaessa analysoitavan datan suuri volyyymi voi myös vaarallisesti harhauttaa tilintarkastajan ohittamaan muutoin pienemmässä otoskoossa merkittäviksi osoittautuvia poikkeamia, mikäli nämä eivät valtavassa datamäärässä näytäkään enää olennaisilta.

Erityisesti automatisoidun koko populaation kattavan analyysin haasteeksi nousevat niin sanotut ”*false positive*” -havainnot (Earley 2015, 497). ”*False positive*” -havainnot eivät ole virheen tai väärinkäytöksen aiheuttamia, mutta muistuttavat niitä suurissa määrin. Big Datasta mahdollisesti kumpuavat ”*false positive*”-havainnot johtuvat automatisoidun data-analytiikan tekemistä datan luokitteluvirheistä, jolloin analytiikka saattaa reagoida vääränlaisiin poikkeamiin relevanttina tilintarkastusevidenssinä. ”*False Positive*” -havainnot ovat vahingollisia tilintarkastajan huomioidessa ne, sillä ne voivat johdattaa tilintarkastajan työskentelyä väärään suuntaan. (Issa 2013, 99.) Krahel ja Titera (2015, 417) summaavat ”*false positive*” -havaintojen aiheuttavien kustannusten olevan tulevaisuudessa erittäin tärkeä keskustelun aihepiiri tilintarkastuksessa. Geat ja Xie (2017) peräänkuuluttavat ”*false positive*” -havaintoja koskevan ohjeistuksen tilintarkastajille olevan nykyisyydessään kuitenkin valitettavan niukka.

On perusteltua, että suuremmasta mahdollisten anomalioiden ja poikkeusten määrästä johtuen koko populaation testaaminen vaatii tilintarkastajalta itseltään huomattavasti syvällisempää tiedon prosessointia verrattuna perinteiseen otantaan. Tilintarkastajien tulee kehittää erittäin vahva ymmärrys siitä mitkä kaikki taustatekijät liittyvät kunkin asiakkaan laskentatoimen prosesseihin, asiakkaan liiketoiminnallinen näkökulma huomioon ottaen. Issa ja Kogan (2014, 223-226) tutkimusten tulosten mukaan eräs kehittyneen Big Data- analytiikan piirre piilee sen luomissa mahdollisuuksissa vähentää dramaattisesti ”*false positive*” -havaintoja tarkemman anomalioiden ja poikkeusten määrittelyn priorisoinnin avulla.

Edellä kuvattu populaation testaaminen on Big Data-analytiikan tärkeimpiä piirteitä. Seuraavassa alaluvussa käsitellään koko populaation testaamisen ohella Big Data-analytiikan hyödyntämiseen kiinteästi liittyvä jatkuvan tilintarkastuksen konsepti, joka reaaliaikaistaa tilintarkastusta.

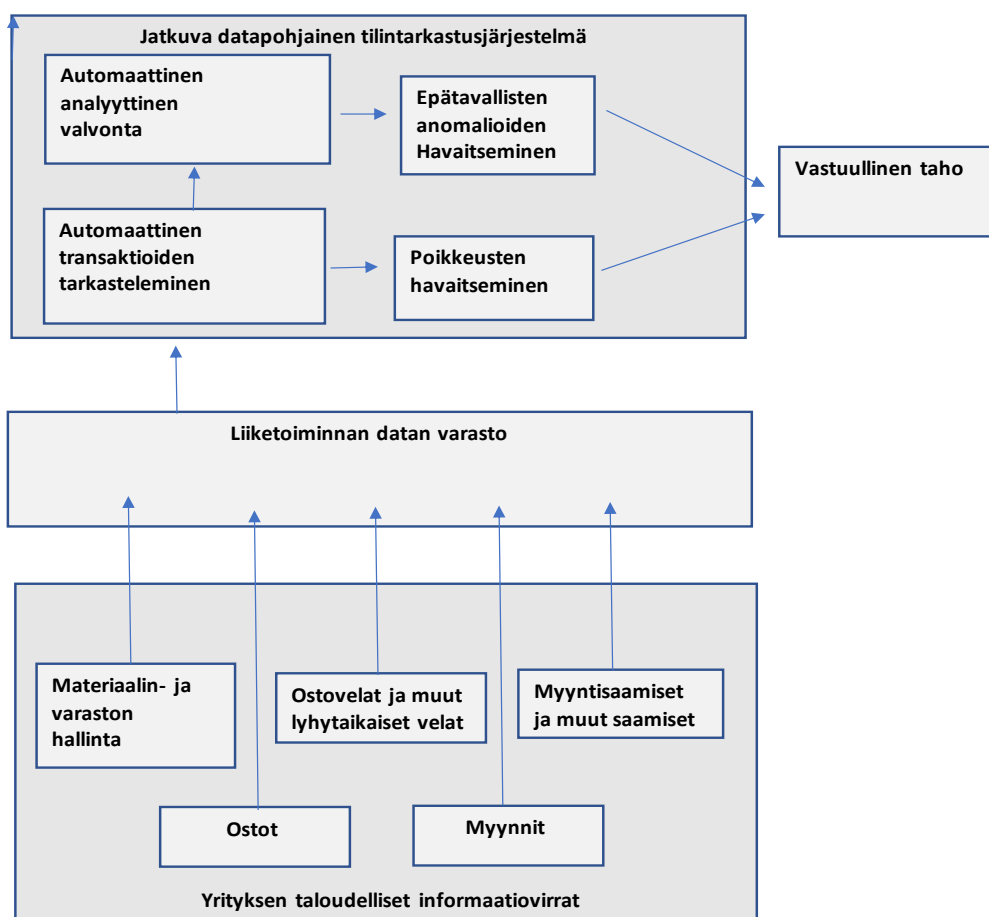
## 2.4 Jatkuva tilintarkastus

Koko populaation testaamisen ohella Big Dataa prosessoiva analytiikka mahdollista tilintarkastuksen luonteen muutoksen periodittaisesta jatkuvaksi. Big Data-aikana yleistyvä reaaliaikainen jatkuva tilintarkastus poikkeaa monin osin perinteisestä tilintarkastuksesta. Perinteinen tilintarkastus suoritetaan tyypillisesti kerran vuodessa, usein huomattavalla viiveellä tilinpäätöksen sisältämiin lukuihin vaikuttaneiden taloudellisten tapahtumien jälkeen. Tämä on toisinaan aivan liian myöhäistä taloudellisen menetyksen ehkäisemiseksi tarkastettavassa kohteessa. Jatkuva tilintarkastus poikkeaa perinteisestä yhdessä tai muutamassa kohtaa vuotta tehdystä tilintarkastuksesta siinä, että sen avulla tilintarkastajat pystyvät reaaliajassa tutkimaan asiakasyrityksen transaktioita heti niiden tapahtuessa. Tämän ansiosta väärinkäytökset ja havaitut virheet voidaan ratkaista välittömästi. Jatkuvan tilintarkastuksen eroavaisuudeksi perinteiseen tilintarkastukseen nähden voidaan pitää myös tilintarkastajan työpanoksen jakautumista tasaisemmin kalenterivuodelle. (Byrnes & Pawlicki 2014, 2 – 3, 7 – 8; Vasarlehyi & Haper 1991, 114.)

Jatkovaa tilintarkastusta voidaan kuvailla monin tavoin. Siinä missä perinteinen tilintarkastus nojaa otanta perusteiseen näkökulmaan, jatkuva tilintarkastus perustuu koko asiakkaan transaktioiden populaation testaamiselle (Issa & Kogan 2014, 213). Jatkovaa tilintarkastusta leimaavaa on koko populaation testaamisen lisäksi automaatio, jossa tietokonejärjestelmä poimii reaaliaikaisesti ja itsenäisesti suuresta datavirrasta huolellisempaa tarkastamista vaativia tapahtumia (Hunton & Rose 2010, 307–308). Läheisemmin tarkasteltuna jatkuvan tilintarkastuksen prosessissa tietojärjestelmä yhdistelee eri lähteistä kerätyt tiettyä objektia koskevat datavirrat yhteen. Jatkuvan tilintarkastusprosessin sisältämä datan yhdisteleminen ei vähennä datan ja siitä löytyvien poikkeusten määrää, mutta lisää prosessin lopputulemana tulevan tiedon relevanttiutta datavirtojen ollessa yhdisteltynä saman nimittäjän alle (Perlos & Murthy 2012, 36). Jatkuvaan tilintarkastukseen on mahdollista kytkeä toimintoja, jotka mahdollistavat poikkeusten löytämisen lisäksi anomalioiden, eli epätavallisten ja yleisten sääntöjen kautta hankalasti ennustettavien epäsäännöllisyyksien havaitsemisen suuresta datajoukosta (Kogan, Alles, Vasarlehyi & Wu 2014, 226.). Jatkuvan tilintarkastuksen nähdään koostuvan perinteisen tilintarkastuksen ohella kahdesta osasta; tilikauden

aikaista tarkastusta vastaavasta jatkuvasta sisäisten kontrollien valvonnasta ja tilinpäätöstarkastusta vastaavasta jatkuvasta varmistustoiminnasta. (Alles, Kogan & Vasarlehyi 2008, 200.) Jatkuvan tilintarkastuksen perustuessa jatkuvaan valvontaan ajoittain tapahtuvan perinteisen tilintarkastuksen sijaan, hämärtyvät myös tilintarkastuksen asiakkaan operationaalisen johdon ja tilintarkastajan välisen työnkuvan rajalinjat (Alles ym. 2008, 212 – 213).

Jatkuva tilintarkastus on yleistymässä, mutta vaatii samalla tilintarkastajilta uudenlaista asennoitumista tilintarkastusta kohtaan. Jatkuvaa tilintarkastusta on implementoitu viimeisen vuosikymmenen aikana useilla toimialoilla (Issa & Kogan 2013, 214). Vasarlehyi, Alles & Williams (2010, 10, 35) havainnoivat taloudellisen datan siirryttyä yhä enemmän erittäin nopeaan XBRL- muotoon olevan esimerkki jatkuvan varmistustoiminnan kasvaneesta tarpeesta. Kogan, Alles, Vasarlehyi & Wu. (2014, 221) esittävät jatkuvan tilintarkastuksen tuottavan vaikuttavia tuloksia ainoastaan tilintarkastajien muuttaessa toimenpiteitään ja asennoitumistaan tukemaan datan reaaliaikaisen saatavuuden reagointiin.



Kuvio 2: Teoreettinen esimerkki jatkuvan datapohjaisen tilintarkastusjärjestelmän toimintaperiaatteesta. (Mukaiillen Kogan ym. 2014, 226.)

Edellisellä sivulla olevassa kuviossa on havainnollistettu Kogan ym. (2014, 224 – 234) mukailten teoreettinen, tutkijoiden esittämään alkuperäiseen kuvioon nähden pelkistetympi esimerkki jatkuvan datapohjaisen tilintarkastusjärjestelmän toimintaperiaatteesta. Tässä teoreettisessa jatkuvan tilintarkastuksen toiminnan esimerkissä kaksi tärkeintä automatisoitua päätoimintoa ovat transaktioiden tarkasteleminen ja jatkuva kokonaisvaltainen analyttinen valvonta. Jatkuvassa tilintarkastuksessa otoksen teon sijaan yrityksen toiminnan koko informaatiovirta soljuu liiketoiminnan datavarastoon (engl. *”Business Data Warehouse”*), joka toimii reaaliaikaisen analyttisen prosessoinnin tukena. Tämä varasto voi olla integroituna asiakkaan ERP- järjestelmään (engl. *”Enterprise Resource Planning”*). Joskus ERP- järjestelmän sisältämä data on kuitenkin liian hajautettua ja reaaliaikaisen käytön suhteen hankalasti hyödynnettävissä. Tällöin erillinen liiketoiminnan datavarasto on järkevä ratkaisu jatkuvaa tilintarkastusta varten. Data siirtyy liiketoiminnan datavarastossa automaattiseen transaktioiden tarkasteluun, joka reagoi datavirrasta löytyviin poikkeamiin. Seuraavana data etenee automaattiseen analyttiseen valvontaan, joka luo datan välille tilastollisia korrelaatioita. Tästä seuraa mahdollisten datasta ilmenevien anomalioiden havaitseminen. Prosessin lopputulema siirtyy yrityksen tilintarkastuksesta vastuulliselle taholle. (Kogan ym. 2014, 224 – 234.)

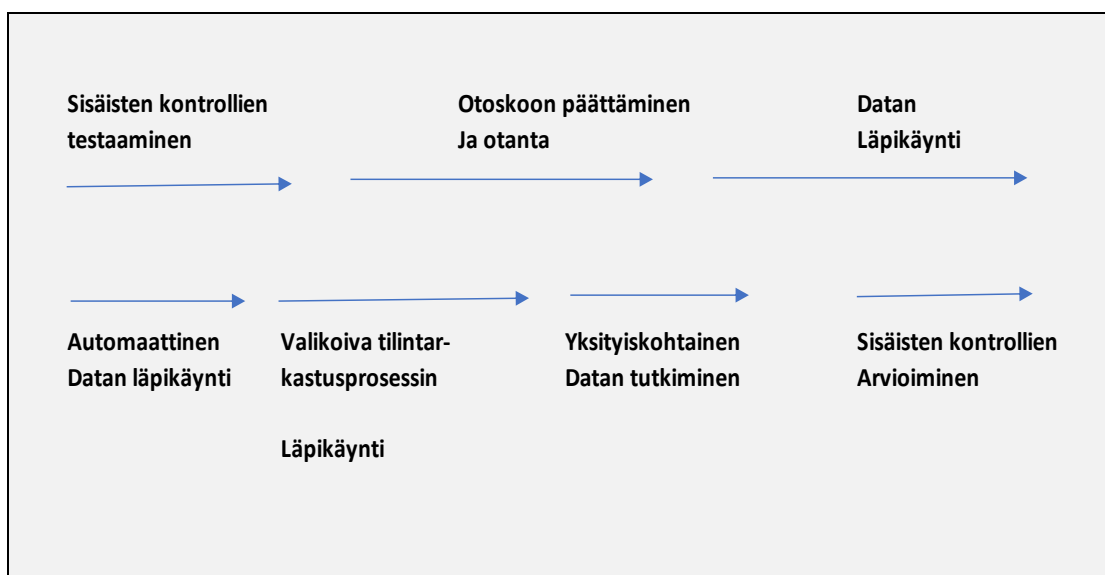
Datan yhdistelemisestä ja rakenteen yhdenmukaisuuden varmistamisesta tulee entistä tärkeämpi toimenpide jatkuvassa tilintarkastuksessa. Tämä korostuu, mikäli sen avulla ryhdytään käsittelemään yksinkertaistamista ja yhteenvetoa vaativaa ei-jäsenneltyä Big Dataa, joka kumpuaa hyvinkin erilaisista lähteistä. Toisaalta myös datan erottelemisestä tulee yhä tärkeämpää. Jäsennellyn datan, kuten tietyn myyntisumman kohdistaminen transaktiolle on helppoa. Kuitenkin ei-jäsennellyn datan, esimerkiksi tekstimuotoisen tarkoituserältään epäselvän datan kohdistaminen tietylle transaktiolle on haasteellisempaa. (Zhang, Yang & Appelbaum. 2015, 470 – 473.)

Big Datan prosessointi aiheuttaa myös muita haasteita jatkuvan tilintarkastuksen toimintaperiaatteille. Big Datan suuri volyyymi, nopeus ja variaatio vaikeuttavat datan yhdistelemistä. Nämä ominaisuudet tekevät Big Datasta epäyhtenäistä, epäjohdonmukaista ja vaikeasti muusta datasta eroteltavaa, sekä alentavat datan sisältävien tietojen luotettavuutta. Nämä Big Data -peräiset tekijät johtavat jatkuvan tilintarkastuksen kohtaamaan haasteita työstää ja integroida yhteen hyvinkin hankalissa muodoissa olevaa dataa, toisiinsa nähden erilaisia tunnisteita sisältävää dataa, toistensa kanssa konfliktissa olevaa dataa, tai muutoin epätäydellistä dataa. (Zhang ym. 2015, 471.)

Jatkuvaan tilintarkastukseen ja yhä datapitoisempaan ympäristöön siirtyminen johtavat niin sanotun *”audit by exception”* -tilintarkastusnäkökulman (ABE) ja tilintarkastusprosessin yleistymiseen tilintarkastustyössä (Appelbaum, Kogan & Vasarlehyy 2017, 11). ABE-tilintarkastusnäkökulmassa tilintarkastusprosessi aktivoituu

automaattisesti silloin kun tarkastetusta datavirrasta löytyy poikkeamia (Vasarhelyi & Haper 1991, 114, 117). ABE-lähestymistapa sopii sekä asiakkaan taloudellisen, että eitaloudellisen datavirran tarkasteluun (Brazel, Jones & Zimbelman 2009, 1161 – 1163). Jatkuvaan sisäiseen valvontaan perustuvana ABE-näkökulma kyseenalaistaa sisäisen tarkastajan ja tilintarkastajan työnkuvien välistä eroavaisuutta ja nykyisten tilintarkastusstandardien käyttökelpoisuutta tämänkaltaisessa tilintarkastuksessa (Appelbaum ym. 2017, 12).

Appelbaum ym. (2017, 11) julkaisussa havainnollistamien kuvioiden mukaisesti perinteinen tilintarkastusprosessi ja ABE-tilintarkastusprosessi eroavat toisistaan selvästi. Nämä kaksi erityyppistä tilintarkastusprosessia ovat toisiinsa nähden ikään kuin käänteisiä. Perinteinen otokseen perustuva tilintarkastussuunnitelma aloitetaan sisäisten kontrollien testaamisella, jonka perusteella määritellään perusjoukkoa tarpeellisesti edustava otoskoko. Otannan jälkeen otokseen sisältyvä data käsitellään. ABE-tilintarkastusprosessille taas on käänteisesti ominaista, että se käynnistyy valikoivasti datavirtaan ilmaantuneiden löydettyjen poikkeusten perusteella. Tämä johtaa yksityiskohtaisempaan datan tutkimiseen ja vasta lopulta tämän jälkeen sisäisten kontrollien arviointiin.



Kuvio 3: Kaksi tutkijoiden esittämää kuviota yhdistäen; perinteinen tilintarkastusprosessi (ylhäällä) ja audit by exception -tilintarkastusprosessi (alhaalla). (Mukaiillen Appelbaum ym. 2017, 11.)

ABE-tilintarkastusnäkökulmassa tilintarkastajat saattavat kuitenkin kohdata aivan uudenlaisia haasteita; kaikkien transaktioiden testaaminen voi tuottaa tuhansia poikkeuksia ja aiemmin tässä tutkielmassa käsiteltyjä ”false positive” -havaintoja mikäli

poikkeusten kategorisointiin ei anneta täsmällisiä määritelmiä (IIAE 2016, 18; Earley 2015, 497).

Jotta jatkuvaan tilintarkastukseen perustuvasta toimintamallista saadaan mahdollisimman suuri hyöty, on erittäin tärkeää rajoittaa poikkeusten määrä asiakaskohtaisesti järkevälle tasolle ennen niiden vientiä ihmismielen avulla työskentelevän tilintarkastajan päätöksentekoon (Issa & Kogan 2014, 215). On maininnan arvoista, että tilintarkastajan tulee hallita laaja-alaisesti data-analyttisiä työkaluja työskennelläkseen tehokkaasti ABE-tilintarkastusnäkökulmaa tukevassa ympäristössä (Appelbaum ym. 2017, 11). Käytännössä ABE-tilintarkastusnäkökulma on tilintarkastuksessa toimiva ainoastaan, mikäli analyysissä hyödynnettävän datan korkeasta laadusta pidetään huolta. Analyysin lopputulema voi nimittäin olla lopulta ainoastaan yhtä laadukas kuin analyysissä hyödynnettävä data on alun perin ollut. (Issa 2013, 87, 91.)

Yhteen vetäen, jatkuvan tilintarkastuksen voidaan todeta muuttavan tilintarkastuksen luonnetta kokonaisuudessaan. Jatkuvan tilintarkastuksen yleistymisen on kuitenkin vain yksi tekijä, jonka kautta Big Data-analytiikka vaikuttaa tilintarkastajan päivittäiseen työhön. Seuraavassa alaluvussa tarkastellaan laajemmin, miten Big Data-analytiikka muokkaa tilintarkastajan työn konseptia.

## 2.5 Tilintarkastuksen automatisaatio

Big Data-analytiikkaan ja kasvaneeseen analysoitavan datan määrään liittyy kiinteästi tilintarkastuksen automatisaatio. Käytännössä kaikki tässä tutkielmassa edellä kuvatut data-analytiikan tekniikat toimivat enemmän tai vähemmän automatisaatiolla. (Richins ym. 2017, 63; Frey & Osborne 2013, 62). Big Dataan kytketty data-analytiikka omaa potentiaalinsa automatisoida lukuisia tehtäviä, joita tilintarkastajat ovat perinteisesti tehneet. Data-analytiikan mahdollisuus korvata tilintarkastuksen rutiinitehtäviä onkin merkittävä kysymys tilintarkastuksen konseptissa (Richins ym. 2017, 63). Automaattisten prosessien avulla on mahdollista tarkastaa tehokkaammin koko populaatio epätavallisten poikkeusten löytämiseksi (Earley 2015, 496; Krahel & Titera 2015, 418).

Automatisaation vaikutusten laajuudesta tilintarkastuksessa tutkijat ovat kuitenkin erimielisiä. Frey ja Osborne (2013, 69) ennustavat tilintarkastuksen automatisoituvan 94 prosentin todennäköisyydellä kokonaan. Kuitenkin Richins ym. (2017, 72 – 74) mukaan tilintarkastajilla on edelleen tulevaisuudessakin paikkansa lopullisena arvioijana tilintarkastuksen informaation arvioinnissa. Tilintarkastuksen luonne siirtyisikin heidän mukaansa vain enemmän konsultointiin ja asiakkaan liiketoiminnan ymmärrykseen mekaanisesta lukujen täsmäytyksestä.

Tilintarkastajan työskentely vaatii lopultakin taustalleen korkeatasoista arviointia ja subjektiivisuutta, joten se on lopulta hankalasti kokonaan automatisoitavissa (Frey & Osborne 2013, 27 – 28). Richins ym. (2017, 72) muistuttavat, että automaattisesti tapahtuva koko populaation testaaminen ei vähennä tilintarkastajan tarvetta tilintarkastuksessa. Tilintarkastajaa tarvitaan edelleen tekemään kriittistä harkintaa sen suhteen, mitkä koko populaatiosta löytyvät poikkeukset ovat relevantteja tilinpäätöksen sisältämän tilintarkastusriskin arvioinnin kannalta.

Myös sosiaalista älykkyyttä vaativat tilintarkastajan tehtävät ovat vaikeasti automatisoitavissa. Esimerkkinä sosiaalista älykkyyttä vaativasta tilintarkastusjohtopäätösten arvioinnista on johdon väärinkäytökseen johtavien käyttäytymismallien ja aikomusten havainnointi. Väärinkäytösten havainnoimisen muuttaminen kokonaan automaattiseksi olisikin ratkaisuna toimimaton. Mikäli väärinkäytösten havainnointi muuttuisi kokonaan automatisoiduksi, se pohjautuisi väistämättä tietyille säännönmukaisuuksille. Tämä avaisi ovet kyseisistä säännöistä tietoisille väärinkäyttäjille, jotka kykenisivät mahdollisesti muuttamaan käyttäytymistään väärinkäytöksiä havainnoiville toimintoille huomaamattomaksi. Näin ollen ihmisajattelua vaativia toimia vaaditaan ainakin täydentämään automaattisia tekniikoita, mikäli tilintarkastus automatisoituisi niin pitkälle kuin mahdollista. (Richins ym. 2017, 72 – 73.)

## **2.6 Tilintarkastusstandardit ja Big Datan hyödyntäminen tilintarkastuksessa**

Tilintarkastustyö on riippuvainen hyvin pitkälti tilintarkastajaan nähden ulkoisista tekijöistä (Wedemeyer 2010, 332). Myös se asiantuntijarooli, joka tilintarkastajille määrittyy Big Data analytiikan parissa, riippuu varsinkin tilintarkastusstandardien kehityksestä (Richins ym. 2017, 75). Kuten pian teoriataustasta ilmenee, nykyisellään tilintarkastusstandardit eivät kuitenkaan juuri ota kantaa Big Datan olemassaoloon tilintarkastuksessa. Ne eivät tue tällä hetkellä riittävästi edes koko populaation testaamista. Tilintarkastusstandardit luovat viitekehyksen koko tilintarkastuksen toteutukselle. Tämän vuoksi aihe on erityisellä syyllä käsitellä täysin omana osionaan.

Nykyiset tilintarkastusstandardit vaativat tilintarkastajan työssä hyödynnettävältä tilintarkastusevidenssiltä erityisesti todennettavuutta, selkeyttä ja vakuuttavuutta (Wedemeyer 2010, 328). Tilintarkastusstandardien tulee kuitenkin ylipäätään tulevaisuudessa muuttua luonteeltaan samanaikaisesti tilintarkastajan työn muuttuessa faktojen todistamisesta yhä analyttisemmiksi ja keskustelelevammiksi (Kraheil & Titera 2015, 419). Tilintarkastusstandardit vaativat tilintarkastajien keräämältä

tilintarkastusevidenssiltä myös riittävyttä, relevanttiutta ja luotettavuutta. Kuitenkin keinot, joilla tilintarkastusevidenssin riittävyttä, relevanttiutta ja luotettavuutta standardien pohjalta arvioidaan, eivät ole riittäviä kohtaamaan Big Datan monimutkaista luonnetta. Tämän vuoksi on myös tärkeää, että uutta teknologiaa kyetään hyödyntämään oikein, jotta standardeja vastaavat vaatimukset täyttyvät. (Brown-Liburd & Vasarlehyi 2015, 6.)

Kansainväliset tilintarkastusstandardit eivät kuitenkaan määrittele tarpeeksi täsmällisesti millaista ja mihin tarkoitukseen soveltuvaa evidenssiä data-analytiikan ja Big Data analytiikan kautta on mahdollista kerätä. Standardit eivät myöskään huomioi riittävän tarpeeksi Big Data-analytiikan tuottaman tilintarkastusevidenssin hyödynnettävyyttä tilintarkastukseen liittyvissä arvioinneissa. Nykyhetken tilintarkastusstandardit eivät lainkaan huomioi Big Datan kaltaisten laajojen ja monimutkaisten lähteiden luotettavuuden varmistettavuuden monimutkaisuutta. (Ramulkan 2015, 17 – 19.)

Moffit ja Vasarlehyi (2013, 14 – 15) listaavat joukon seikkoja, kuinka tilintarkastusstandardien tulisi mukautua Big Datan ja Big Data-analytiikan aiheuttamiin muutoksiin. Yleisesti he ovat sitä mieltä, että informaation prosessoinnin ja informaatiota prosessoivan teknologian muutoksessa tilintarkastusprosessin tavoitteet tulee tilintarkastuksen näkökulmasta määrittää tilintarkastusstandardeissa uudelleen. Tilintarkastusstandardeissa tulisi luopua perinteisestä populaatioon kohdistuvan pienen näytteenoton konseptista, ymmärtää Big Datan vaikutukset tilintarkastuksessa ja huomioida kokonaisvaltaisemmin koko populaation tarkastaminen.

Tilintarkastusstandardit ovat aikaansa jäljessä. Whithouse (2014, 28 – 29) mukaan nykyisten tilintarkastusstandardien ongelmaksi muodostuu niiden perustuminen menneeseen aikaan, jolloin tilintarkastus on tukeutunut vahvasti otanta perusteisiin tekniikoihin. Tämän johdosta, vaikka tilintarkastajat kykenisivätkin jo tarkistamaan kaiken datan, tilintarkastusstandardit edellyttävät edelleen voimakkaasti perinteisiä otanta menetelmiä. Tilintarkastusstandardeja tulisi myös uudistaa kannustamaan tilintarkastajia hyödyntämään kehittyneempiä analyttisiä tekniikoita tilintarkastuksessa, jolloin perinteiset otanta tekniikat jäisivät lopulta tarpeettomiksi. Tilintarkastusstandardit edellyttävät myös nykyisin tilintarkastajalta jokaisen datalähteen paikkansapitävyyden varmistamista ja tarkkuutta, josta he hakevat tilintarkastusevidenssiä. Tämä on huomattavasti haasteellisempaa, mikäli tilintarkastajat hyödyntävät tilintarkastusevidenssin lähteenä myös asiakkaaseen nähden lukuisia ulkoisia kanavia. Ramulkan (2015, 19) täsmentää, etteivät tilintarkastusstandardit ota myöskään lainkaan kantaa siihen, kuinka tarkkaa tietoa uudennlaisilla analyttisillä menetelmillä tulee saada aikaan asiakkaasta.

Eräät merkittävät asiantuntijat ja tutkijat ovat ilmaisseet epävarmuutensa Big Datapohjaisen evidenssin roolista osana tilintarkastusevidenssin hierarkiaa.



Epävarmuutta herättää erityisesti, kykenevätkö tilintarkastusstandardit ollenkaan määrittelemään Big Data perustaista evidenssiä muuna kuin indikatiivisena tai tukevana evidenssinä perinteistä tilintarkastusevidenssiä täydentämässä. (Ramulkan 2015, 18 – 19.) Tämän johdosta niin kauan kuin tilintarkastusstandardit eivät ole täysin vielä mukautuneet määrittelemään Big Datan soveltuvuutta tilintarkastuksessa, myös Big Dataa hyödyntävien tekniikoiden lopullinen soveltuvuus korvata perinteisiä tilintarkastustekniikoita tilintarkastusevidenssin keräämisessä pysyy epäselvänä.

Jatkuvassa tilintarkastuksessa ja koko populaation testaamisessa suuren havaittujen poikkeusten lukumäärän tekee ongelmalliseksi se, että nykyiset, otantoja varten suunnatut tilintarkastusstandardit velvoittavat tilintarkastajaa tarkastamaan kaikki havaitut poikkeukset (PCAOB 2010c, Standardi No.2305). Tämä on ongelmallinen lähestymisnäkökulma, mikäli tarkastelun alaisena on samanaikaisesti koko populaatio. Byrnes ja Pawlicki (2014, 10) mukaan tilintarkastusstandardien asettajien tulisi mukauttaa olemassa olevia tilintarkastusstandardeja tukemaan populaation täydellistä testaamista. Standardit, jotka perustuvat tilintarkastajan näyteperusteiseen evidenssiin pohjautuvaan päätöksentekoon olisivat hyvä muuttaa tukemaan sataprosenttista transaktioiden testaamista.

Uusien ei-jäsennelyyn ja jäsennelyyn dataan soveltuvien standardien kehittäminen on tärkeää. Tilintarkastusstandardien tulee olla integroitavissa myös laajasti sekä aiemmin tässä tutkielmassa esitelyihin ongelmalähtöiseen- että aineistolähtöiseen analyysiin pohjautuviin tekniikoihin. Tämän avulla tilintarkastus ei jämähtäisi aikaan ennen Big Dataa. (Richins ym. 2017, 75.)

Datan muuntuessa yhä laajempaan määrällisempään muotoon, tulee myös tilintarkastusstandardien huomioida jatkuvasti laajentuva datamäärä. Tilintarkastus tulee myös tehostumaan nykytilanteeseen nähden tilintarkastusstandardien mukautuessa jatkuvaan tilintarkastukseen, koko populaation testaamiseen ja automaatioon. On nimittäin sängen ilmeistä, että väärinkäytösten paikantaminen tehostuu, kun koko populaatio kyetään testaamaan automaattisesti ja nopeasti, sekä vähemmällä määrällä ihmisillä erehdyksille altistavia testausmenetelmiä. (Krahel & Titera 2015, 416 – 417 & 420.)

Brown-Liburd ja Vasarlehyi (2015, 6) peilaavat Big Datan ominaisuuksia tilintarkastusstandardien luomiin vaatimukseen tilintarkastusevidenssiä kohtaan. He listaavat joukon vaatimuksia, joiden Big Datan tulee täyttää kohdatakseen asianmukaisen tilintarkastusevidenssin vaatimukset. Näitä tilintarkastusevidenssille olennaisia vaatimuksia ovat heidän mukaansa hankala muokattavuus, uskottavuus, täydellisyys, helppo hyödynnettävyys ja selkeys. Ensinnäkin evidenssin tulee olla vaikeasti muokattavissa, mutta ulkoisen Big Datan ongelmaksi tässä muodostuu, että se ei ole lainkaan asiakkaan liiketoiminnan kontrolloitavissa. Tilintarkastusstandardit vaativat tilintarkastusevidenssiltä myös uskottavuutta. Kuitenkin Big Datan sisältäessä

parhaimmillaan kaiken mahdollisen aiheeseen liittyvän datan, on uskottavuuden vaatimus osittain koetuksella. Big Data- pohjaisen evidenssin on myös hankala täyttää selkeyden vaatimusta Big Datan äärettömän luonteen vuoksi. Tilintarkastusevidenssin helppo hyödynnettävyys vaatii Big Datan tullessa kysymykseen sitä, että uudentyypisiä automaattisia metodeja kehitetään suurten datamäärien käsittelyä varten. Tilintarkastusevidenssin selkeyden vaatimukselle taas luo oman haasteensa erityisesti Big Datan määrällinen luonne.

Yhteenvedona voidaan todeta nykyaikaisten tilintarkastusstandardien olevan aikaansa jäljessä. Teorian ydinsanoma viittaa vahvasti siihen, että tilintarkastusstandardit tarvitsevat olennaista uudelleenmäärittelyä ollakseen ajan tasalla jo alkaneella Big Data- aikakaudella. Tilintarkastusstandardien tulisi kyetä vastaamaan varsinkin koko populaation testaamisen tarpeisiin, sekä Big Data -pohjaisen tilintarkastusevidenssin hyödynnettävyyteen liittyviin kysymyksiin.

### **3 SUURTEN DATAMÄÄRIEN PROSESSOINTIIN LIITTYVÄT TEKIJÄT BIG DATAA TYÖSTÄVÄLLÄ TILINTARKASTAJALLA**

Siinä missä tämän tutkielman ensimmäisen pääluvun tavoitteena on selvittää kuinka Big Data ja Big Data-analytiikka muuttavat tilintarkastustyötä suuressa mittakaavassa, toisen pääluvun tarkoituksena on luoda yksityiskohtaisempi näkemys tilintarkastajan työn muutoksesta hänen henkilökohtaisella kognitiivisen kokemuksen tasollaan. Tämän kautta on myös mahdollista selvittää miten Big Datan tilintarkastajassa aiheuttamat kognitiiviset tekijät vaikuttavat tilintarkastukseen laajemmassa mittakaavassa. Tämän avulla pääsemme lopulta luontevasti tämän tutkielman johdannossa läpikäytyjen tutkimuskysymysten ratkaisujen teoreettisille juurille.

Brown-Liburud ym. (2015, 460–462) toteavat tilintarkastajalta Big Data ympäristössä vaadittavien ominaisuuksien ja taitojen tutkimustarpeen olevan suuri esittäen listauksen aiheeseen liittyvistä relevanteista tutkimusaiheista. Big Dataa koskevien tilintarkastukseen liittyvien tutkimusten tulisi suuntautua yhä voimakkaammin käsittelemään tilintarkastajan työssä vaadittavia henkisiä toimintamalleja ja tilintarkastajan kykyä käsitellä suuria datamääriä.

Ihmisen kognitiivisen työskentelyn vaikuttavuutta on ylipäättään mahdollista arvioida nopeuden ja tarkkuuden dimensioiden kautta (Grant, Shen, Carswel & Santhanam 2010, 1360). Tässä pääluvussa jäljempänä ilmenevät Big Datan ominaisuuksiin liittyvät kognitiivista työskentelyä haastavat teemat kuten informaatioylikuorma, olennaisuuden hahmottaminen, tiedon epäselvyyden sietäminen ja relevantin tiedon havaitseminen peilautuvat vahvasti myös tämän kaksidimensionalisuuden kautta. Tämä pääluku on jaoteltuna yhteentoista alalukuun, jotka esitellään lyhyesti seuraavana.

Tämän pääluvun kaksi ensimmäistä alalukua käsittelevät yleisellä tasolla tilintarkastajan henkisiä toimintamalleja tilintarkastustyössä. Ensimmäisessä alaluvussa tehdään katsaus tilintarkastajassa itsessään työsuorituksen aikana tapahtuvaan kognitiiviseen ajatteluprosessiin. Lisäksi ensimmäisessä alaluvussa tarkastellaan lyhyesti tilintarkastajan asennoitumista ympäristön epävarmuustekijöihin. Tämän pääluvun toisessa alaluvussa tarkastellaan tilintarkastajassa mahdollisesti työsuoritusten aikana ilmeneviä kognitiivisia ilmiöitä kuten kognitiivisia vinoumia ja tilintarkastustyön aiheuttamaa stressiä.

Tämän pääluvun alaluvut 3, 4 ja 5 tarkastelevat tilintarkastajan ajattelutavan ja analyysin näkökulman muutosta Big Data-aikakaudella. Tähän liittyy erityisesti tilintarkastajan ajattelun muutos aiempaan nähden data-analyttisemmäksi.

Tämän pääluvun alaluvut 6, 7, 8, 9 ja 10 käsittelevät tilintarkastajan kohtaamia haasteita Big Datan käsittelemisessä. Tähän liittyvät ammatillisen skeptisyyden muutos, sekä informaatioylikuormaan, tiedon relevanttiuteen, tiedon selkeyteen ja olennaisuuden hahmottamiseen liittyvät kysymykset.

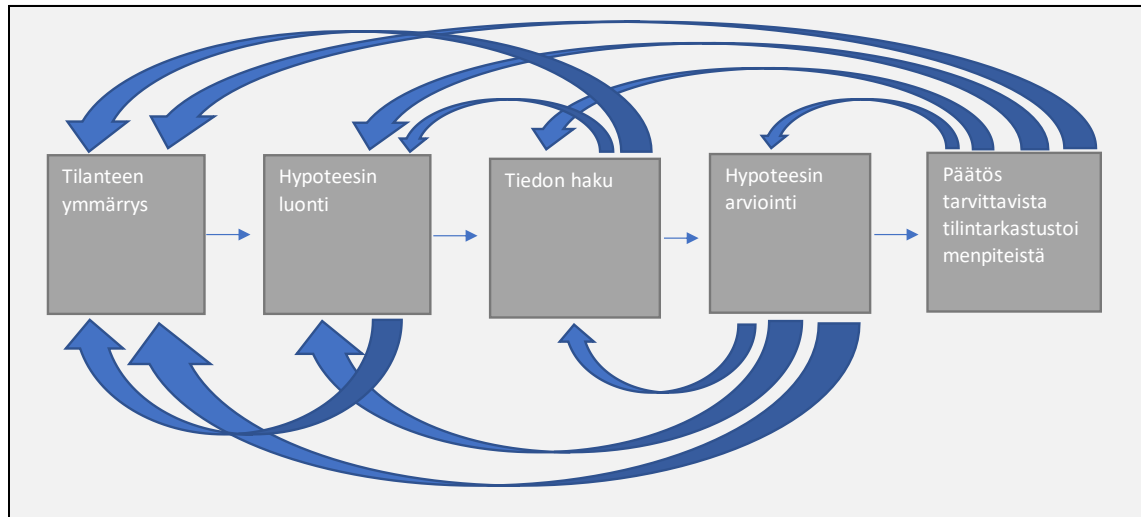
Tämän pääluvun yhdennessätoista alaluvussa esitellään koko tutkielman teoriaosuutta käsittelevä yhteenveto. Tämän yhteenvedon avulla pyritään perustelemaan syytä tehdä empiirinen tutkimus tutkimusaukosta, jota tällä tutkielmalla pyritään täydentämään.

### **3.1 Tilintarkastajan kognitiivinen työskentely**

Laskentatoimen työtä sängen kuvaavaa on sen lopputuleman olevan suurissa määrin lopultakin vain ihmisen kognitiivisiin ominaisuuksiin perustuvan ajattelutyön tuote. Tämä pätee, oli kyseessä sitten tilintarkastuskertomus, tilinpäätös, kirjanpidon kuukausiyhteenveto, kustannusten seuranta raportti tai muu laskentatoimen ja kirjanpidon prosessien aikaansaama dokumentti (Elbannan 2006, 484). Laskentatoimen työntekijän kognitiiviset ominaisuudet vaikuttavatkin jossain määrin hänen näkemykseensä siitä, mikä tieto on hänen työnsä kannalta olennaista (Dremer 1973, 511).

Tilintarkastajan työ on asiantuntijatyötä. Tilintarkastajan työhön liittyy mitä suuremmissa määrin asiakkaan datan sisältämien odottamattomien poikkeusten ja muiden tavallisesta poikkeavien taloudellisten seikkojen löytäminen ja määrittäminen (Koonce 1993, 57). Tilintarkastajan työskentelyä ja hänen löydöksiään erittäin kuvaavaa on myös verrattaen korkea subjektiivisuuden taso (Hogart 1991, 227).

Kooncen vuonna 1993 esittelemä kognitiiviseen psykologiaan ja päätöksentekopsykologiaan perustuva tilintarkastajan ajatteluprosessia käsittelevä malli tarjoaa lähtökohdan tilintarkastajan kognitiivisen työskentelyn teoreettiselle tarkastelulle (Wright 1993, 80). Koonce (1993, 58 – 60) malli edustaa tilintarkastajan iteratiivista ajatteluprosessia hänen tehdessään analyyttistä tilintarkastustyötä. Tämä prosessi lähtee liikkeelle tilanteen ymmärtämisestä ja päättyy päätökseen tarvittavista tilintarkastustoimenpiteistä. Näiden kahden vaiheen välillä tilintarkastaja luo alkuperäisen hypoteesin tehtävän lopputulemasta, hakee erinäisistä lähteistä tietoa hypoteesinsa tueksi ja lopulta arvioi aluksi luomansa hypoteesin oikeellisuutta. Tilintarkastaja saattaa palata missä tahansa kohtaa tätä prosessia johonkin edellisistä työstämistään ajatteluvaiheista, mikäli hän kokee tämän tarvittavien tilintarkastustoimenpiteiden valinnan kannalta tarpeelliseksi Kooncen 1993 esittelemä tilintarkastajan iteratiivinen ajatteluprosessi on havainnollistettu kuviona seuraavalla sivulla.



Kuvio 4: Tilintarkastajan iteratiivinen ajatteluprosessi hänen analyttisessä työskentelyssään. (Mukaillen Koonce 1993, 59.)

Siinä missä tilintarkastajan työskentelyssä tapahtuvaa ajatteluprosessia on tutkittu, myös tilintarkastajan kognitiivisiin ominaisuuksiin liittyviä työskentelytyylejä on niin ikään selvitetty tutkimusten kautta. Tilintarkastajien tiedetään ensinnäkin olevan konsultteihin verrattuna huomattavasti vähemmän innovatiivisia ongelmanratkaisutyyleiltään (Summers, Sweeney & Wolk 2000, 12).

Tilintarkastajan kognitiivisiin ominaisuuksiin perustuvat työskentelytyylit voidaan suuripiirteisesti jaotella intuitiiviseen ja analyttiseen ongelmanratkaisuun perustuvaan työskentelytyyliin. Näiden kahden työskentelytyylin välissä on erotettavissa kummankin edellisen työskentelytyylin piirteitä sisältävä työskentelytyyli. Tilintarkastuksessa työntekijälle ominainen kognitiivinen työskentelytyyli voi rajoittaa yleisiä kognitiivisia valmiuksia tilanteessa, jossa tilintarkastaja on ratkomassa hänen kognitiiviselle työskentelytyylilleen epäsoveliaa tehtävää. Yleisellä tasolla analyttiset tilintarkastajat suoriutuvat paremmin analyttisistä kuin intuitiivisista tehtävistä, ja intuitiiviset tilintarkastajat vuorostaan paremmin intuitiivisista tehtävistä kuin analyttisistä tehtävistä. (Fuller & Kaplan 2004, 131 – 133.)

Ammatillinen skeptisyys luonnehtii tilintarkastajan ajattelutyötä. Skeptinen tilintarkastaja on lähtökohtaisesti kykeneväinen muodostamaan kysymyksiä asiakkaan numeroiden taakse kätkeytyvien sanomien suhteen (Mc Kinney 2017, 68). Ammatillinen skeptisyys on yksi tilintarkastajan kognitiivisen työskentelyn taustalla piilevistä tärkeimmistä kulmakivistä. Nelson (2009, 1) määrittää ammatillisen skeptisyyden tilintarkastajalle tärkeäksi kyvyksi epäillä asiakasta koskevan alkuperäisen riskienarvioinnin täsmällisyyttä.

Tilintarkastusstandardien luonteella on myös vaikutusta tilintarkastajan tapaan ratkaista hänen työssään vastaantulevia haasteita. Laajojen pääperiaate -tyyppisten

standardien noudattamisen todetaan kasvattavan tilintarkastajan tietoteoreettista motivaatiota hakea uutta tietoa tilintarkastusevidenssin tueksi. Tätä ei ole samalla tapaa havaittavissa yksityiskohtaisempien sääntöperustaisten standardien noudattamisen osalta. (Peytcheva, Wright & Majoor 2014, 66 – 67.)

Tilintarkastajan kykyyn hyödyntää saatavilla olevaa uutta tietoa vaikuttaa hänen asennoitumisensa ympäröivässä informaatiota tuottavassa ympäristössä piileviin epävarmuustekijöihin. Tämä asennoituminen vaikuttaa siihen, millaisen tiedon tilintarkastaja kokee relevantiksi. Mikäli tilintarkastajan asennoituminen ympäristönsä epävarmuustekijöihin on heikko, asennoituu hän todennäköisesti epämääräistä tietoa ja uusia tilanteita kohtaan uhkina. Tämä johtaa tilintarkastuksen aikana uudenlaisen tiedon hyödyntämiseen johtavien tilanteiden välttelemiseen (Dremer 1973, 512 & 518). Epämääräisen, täysin uudenlaisen tiedon vältteleminen on ymmärrettävää, epädiagnostisen tiedon prosessointiin käytetyn ajan ollessa tilintarkastajan työn tehoa vähentävä tekijä (Peytcheva ym. 2014, 53).

### **3.2 Tilintarkastajan kohtaamat kognitiiviset vinoumat ja työperäinen stressi**

Työskentelytyyleistä ja asennoitumisesta tilintarkastusympäristöön on seuraavaksi luontevaa siirtyä tarkastelemaan tilintarkastajan kohtaamia kognitiivisia vinoumia, jotka ovat tiiviisti sidoksissa erityisesti edellä käsiteltyihin kognitiivisiin työskentelytyyleihin. Kognitiivisten vinoumien tarkastelun jälkeen tehdään katsaus tilintarkastajan kokeman stressin vaikutuksiin tilintarkastustyössä.

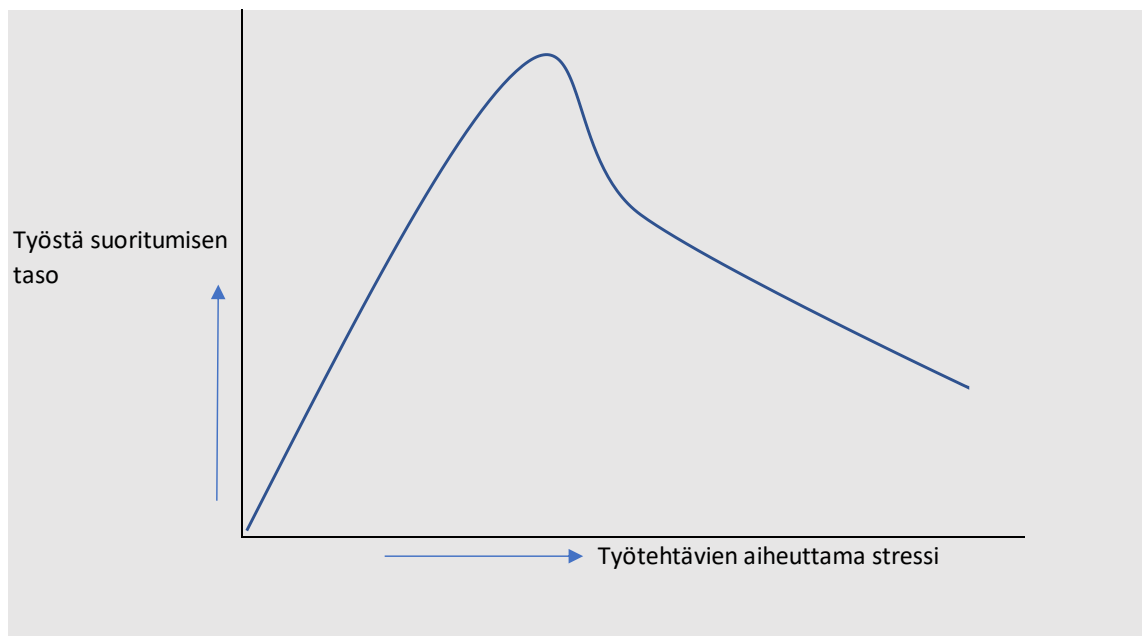
Jokainen tilintarkastaja altistuu ennemmin tai myöhemmin kognitiivisille vinoumille kuten kuka tahansa ihminen. Kognitiivinen vinouma tarkoittaa ihmiselle ominaista taipuvaisuutta painottaa ja hahmottaa tulkintoja henkilökohtaisesti ominaisella tyylillä ja perusteella. Tämä voi johtaa virhearvioihin, mikäli tehtävää olisi pitänyt lähestyä toisesta näkökulmasta käsin. Tilintarkastaja on erityisen altis kognitiivisille vinoumille etenkin kiirekautena, jolloin tarkastukseen tarvittava aika ja muut resurssit ovat rajalliset. Tällöin tilintarkastaja tukeutuu herkästi tottumuksellisiin päättelymalleihinsa uuden tiedon kyseenalaistamisen sijaan. (Jacka 2018, 67.)

Kognitiivisten vinoumien tiedetään ilmenevän tilintarkastajan työskentelyssä hyvin erilaisin tavoin. Ensinnäkin tilintarkastajan tulee usein arvioida asiakkaaseen liittyvien osapuolten ja sidosryhmien tekemiä päätöksiä, jotka ovat jo tällöin altistuneet kognitiivisille vinoumille. Toiseksi, tilintarkastusevidenssi voi aiheuttaa erityyppisiä kognitiivisia vinoumia kulkiessaan tilintarkastustiimin hierarkiassa assistenttien ja senioreiden kautta ylemmäksi kohti lopullisen tilintarkastuskertomuksen allekirjoittavaa partneria. Tällöin riskinä on evidenssiin liittyvien arvioiden muuttuminen epäedulliseen

suuntaan sen kulkiessa läpi koko tilintarkastustiimin. Kolmanneksi, tilintarkastajan työn kannalta tärkeimmät sidosryhmät kuten asiakkaan johto ja kirjanpito-osasto voivat olla tietoisia vinoutumista, joille tilintarkastaja voi tarjoamansa materiaalin kautta altistaa. Tämä tarjoaa asiakkaalle mahdollisuuden manipuloida tilintarkastajan työskentelyä asiakkaan kannalta edulliseen suuntaan. (Knapp & Knapp 2012, 41 – 42.)

Parker ja Fogarty (2012, 11) nostavat esille edelliseen liittyvän huomion, jonka mukaan tilintarkastajat, joilla on vahvat etukäteisolettamukset asiakkaastaan löytyvän evidenssin suhteen yleensä pitäytyvät sitkeästi ja suhteellisen kauan alkuperäisessä olettamuksessaan. Näin tapahtuu, vaikka evidenssin sanoma kertoisikin lähtökohtaisesti jotain muuta tilintarkastajan alkuperäiseen olettamukseen nähden.

Tilintarkastajien työssä ilmenee monen muun ammatin tapaan stressiä. Choo (1986, 25 – 26) esittää kuvaajan tilintarkastajan työssä kokeman stressin ja työstä suoriutumisen laadun suhteesta. Tämän perusteella tilintarkastajan työtehtävien aiheuttaman stressin lievä lisääntyminen aiheuttaa aluksi työssä suoriutumisen tason kohentumista. Kuitenkin työperäisen stressin lisääntyessä edelleen, tilintarkastajan työssä suoriutuminen laskee voimakkaasti. Tilintarkastajan kokeman stressin voidaan nähdä olevan jatkuvassa muutoksessa riippuen hänen tietyllä ajanhetkellä käynnissä olevien tarkastusten luonteista. Choo (1986, 26) esittämä kuvio tilintarkastajan kokeman työperäisen stressin ja työsuorituksen suhteesta on esitetty seuraavana. Tässä tutkielmassa esiteltävä kuvio on kuitenkin tutkijan havainnollistamaan kuvioon nähden pelkistetympi.



Kuvio 5: Tilintarkastajan kokeman työperäisen stressin ja työsuorituksen välinen suhde. (Mukaillen Choo 1986, 26.)

Stressillä on vaikutusta tilintarkastajan kognitiiviseen suoriutumiseen hänen työssään. Tilintarkastajien työssään kokeman stressin todetaan vaikuttavan negatiivisesti tilintarkastajien työssä suoriutumiseen työssä viihtymisen ja organisaatioonsa tyytyväisyyden kautta. Sinällään pelkästään stressaava työ ei kuitenkaan suoraan vaikuta tilintarkastajan työssä suoriutumiseen olennaisesti (Masihabadi, Rajaei, Koloukhi. & Parsian 2015, 312). Mikäli tilintarkastaja joutuu työskentelemään suuren aikapaineen alla valtavan työmäärän kanssa, on tilintarkastuksen laatu kuitenkin stressin kautta uhattuna. Tämä on todettavissa etenkin tilintarkastuksen kiirekausien aikana (Svanström 2016, 49). Myös, mikäli tilintarkastaja altistuu liialliselle työmäärälle, hänen työtyytyväisyytensä laskee työperäisen stressin lisääntymisen tahdissa (Pradana & Salehudin 2015, 119). Toisaalta, Summers ym. (2000, 1) toteavat tilintarkastajan ongelmanratkaisutyylin ollessa sopiva hänen funktionaaliseen rooliinsa, vähenee keskimäärin myös hänen työssään kokemansa stressi.

Seuraavissa alaluvuissa siirrytään tarkastelemaan erityisesti tilintarkastajan kykyä prosessoida suuria datamääriä. Lisäksi myöhemmin seuraavissa alaluvuissa selvennetään tilintarkastajan kohtaamia Big Datan prosessointiin haasteita luovia ilmiöitä.

### **3.3 Tilintarkastajan ammatillinen arviointikyky Big Data-aikakaudella**

Kuten jo aiemmin tässä tutkielmassa on tullut esille, tilintarkastajan työskentelyyn liittyvää perustavanaatuista taustavoimaa kuvaa erittäin korkea subjektiivisuuden taso, joka perustuu sekä kvantitatiivisille että kvalitatiivisille arvioille hänen asiakkaistaan. Laskentatoimeen liittyvän tietämyksen lisäksi tilintarkastajan tulee ymmärtää laajasti ja objektiivisesti jokaisen yksilöllisen asiakkaansa liiketoimintaa ja sisäistää tilintarkastusmetodologian kannalta kriittisenä pidettävä olennaisuuden konsepti, johon sisältyvät ammatillinen skeptisyys, riippumattomuus, ja rehellisyys. (Wedemeyer 2010, 329, 331.)

Tilintarkastaja hyödyntää aina ammatillista arviointikykyään valitessaan ja tehdessään tilintarkastustoimenpiteitä asiakkaastaan. Tämä ammatillinen arviointikyky syntyy yhdistelmästä relevanttia harjoitusta, tietämystä ja kokemusta. Lisäksi hänen ammatillisen arviointikykynsä taustalla vaikuttavat tilintarkastuksen eettiset ja ammatilliset standardit (International Auditing and Assurance Standards Board 2015, 84).

Tilintarkastajat tarvitsevat ammatillista arviointikykyä hyvin moninaisissa tilanteissa. Tilintarkastajat hyödyntävät ammatillista arviointikykyään etenkin arvioidessaan tilinpäätöksen sisältämää tilintarkastusriskiä, tilintarkastusriskiin vastattavia toimenpiteitä valikoidessaan, sekä tilintarkastusevidenssin laatua ja sanomaa tulkitessaan. Jotkin tilintarkastusraportointiin liittyvät arvioinnit sisältävät myös



subjektiivista asiakkaan tulevaisuuden tapahtumien läpikäyntiä. Esimerkkinä tämän kaltaisesta raportoinnista ovat taloudellisissa vaikeuksissa olevista asiakkaista laadittavat going concern-arviot, jotka perustuvat asiakkaan liiketoiminnan jatkumismahdollisuuksien arviointiin lähitulevaisuudessa. Olosuhteiden yllättävien muutosten johdosta nämä arvioinnit eivät suinkaan aina johda ennalta määriteltyyn lopputulokseen, vaikka niiden tekijöinä olisivatkin parhaimman kokemuksen omaavat tilintarkastajat sekä heidän käytettävissään kaikki mahdollinen relevantti, tarkoituksenmukainen ja ajankohtainen data ja siitä kumpuava evidenssi. (Wedemeyer 2010, 321 & 325.)

Miten Big Data ja Big Data-analytiikka vaikuttavat tilintarkastajan ammatilliseen arviointikykyyn ja ajattelumalleihin? Tähän kysymykseen haetaan tässä alaluvussa vastausta teorialähteisiin perustaen seuraavana.

Byrnes ja Pawlicki (2014, 7) luonnehtivat Big Data-analytiikkaa tieteeksi ja taiteeksi, jonka avulla on mahdollista syventää tilintarkastajan tietämystä analyysin kohteena olevasta asiakkaasta tunnistamalla ja analysoimalla toisiinsa kytkeytyviä kuvioita ja korrelaatioita tämän liiketoiminnan transaktiovirrasta. Tämä tarkoittaa olennaisesti tilintarkastajan ammatillista arviota asiakkaastaan. Richins ym. (2017, 65 & 72) mukaan nykypäivän tilintarkastajat soveltavat tavallisimmin data-analyttisiä taitojaan jäseneltyyn dataan, joka on kerätty asiakkaan laskentatoimen järjestelmistä. Nämä järjestelmät ovat kooltaan rajoitettuja ja niiden sisältämällä datalla on selkeä ominaisrakenteensa ja luonteensa. Tämän kaltaista dataa on suhteellisen suoraviivaista analysoida.

Uudet ja edelleen kehittyvät data-analytiikan menetelmät ovat viimeisimpinä vuosina parantaneet huomattavasti tilintarkastajien kykyä arvioida asiakkaidensa olennaisen virheellisyyden riskiä ja väärinkäytöksiä tilinpäätöksistä. Perinteisten pieneen relevanttiin näytteeseen perustuneiden otannan menetelmien todetaan johtavan koko populaation testaamiseen verrattuna johtavan epätodennäköisemmin olennaisen virheellisyyden tai väärinkäytösten arviointiin liittyvien johtolankojen paikantamiseen tilinpäätöksestä. Jäseneltyyn dataan sovelletut uudet menetelmät auttavat tilintarkastajia löytämään datasta vanhoja manuaalisia menetelmiä tehokkaammin ja alhaisemmin kustannuksin väärinkäytöksiin ja virheisiin liittyviä johtolankoja, esimerkiksi asiakkaan transaktioihin liittyviä tavanomaisesta poikkeavia tapahtumia. Ei-jäseneltyyn, vaikeasti ennustettavaan ja standardittomassa muodossa olevaan dataan sovellettavien koneoppimiseen perustuvien visuaalisen ja tekstianalyysin menetelmien odotetaan löytävän vielä ylimääräisiä johtolankoja olennaisten virheellisyyksien havaitsemiseksi. (Aldhizer 2017, 30 – 32.)

Tilintarkastaja hyötyy Big Datan luomasta tiedosta tilintarkastusprosessista, sillä se parantaa asiakkaan riskien arviointia ja luo lisää relevanttia evidenssiä kokonaisvaltaisen arvioinnin tueksi. Big Datan todetaan myös laajentavan tilintarkastajan tietämystä

asiakkaan liiketoiminnasta ja teollisuudenalasta. Näin tilintarkastajan ymmärrys ei keskity pelkästään asiakkaan tilinpäätöksen tasolle, vaan laajenee asiakkaan kokonaisvaltaiseen arviointiin. (Appelbaum 2016, 18.)

Big Dataa prosessoivan analytiikan hyödyntäminen johtaa väistämättä tilintarkastajan ajattelumallin siirtymiseen yksinkertaisista syy-seuraussuhteista monimutkaisempien korrelaatioiden arviointiin. Sen sijaan, että pyrittäisiin ymmärtämään yksittäisiä syitä vaikeaselkoiseen ja monimutkaiseen ilmiöön tilintarkastuksessa, on Big Dataa prosessoivan analytiikan avulla enemmissä määrin mahdollista ymmärtää kahden tai useamman muuttujan välisiä riippuvuussuhteita. (Cao ym. 2015, 426–427.)

Big Datan kaltaisten erittäin suurten tietomäärien hyödyntäminen edellyttää tilintarkastajan luomalta analyysiltä ja arviolta myös entistä määrällisempää tarkastelunäkökulmaa (Kraheil ja Titera 2015, 417). Määrällisesti lisääntynyt informaatio voi epäilemättä lisätä tiedon kokonaisarvoa, mutta se myös vähentää tilintarkastuksen näkökulmasta sen asymmetriaa eri osapuolten välillä. Big Datan totuudenmukaisuuden ja lävitseen luottavuuden aiheuttaman tiedon asymmetrian vähenemisen todetaan uhkaavan päämies agenttiongelman olemassaoloa tilintarkastuksessa. Päämies-agentti ongelma on tärkeä tilintarkastuksen olemassaolon kulmakivi, koska sen mukaisesti tilintarkastus on olemassa yrityksen ulkoisten sidosryhmien ja johdon ”välikätenä” varmistamassa kummankin osapuolen tietävän intressien kohteena olevasta yrityksestä aukoton totuus. (Appelbaum ym. 2017, 20 – 21.)

Big Datan käsittäessä laajimmillaan kaiken mahdollisen asiakkaaseen liittyvän datan, myös eettisen ajattelun merkitys asiakkaaseen liittyvää tietoa kohtaan korostuu. Haron, Ishak ja Ibrahim (2014, 49) kuvailevat tilintarkastusta ylipäätään ammatiksi, jossa työntekijältä vaaditaan työskentelynsä taustalle julkisen luottamuksen ylläpitämiseksi nimenomaan korkeatasoista eettistä ajattelu- ja päätöksentekokykyä. Nunan & Di Domenico (2013, 4 – 5) selventävät Big Datan tuovan oman haasteensa erityisesti asiakkaan yksityisyyden suojaa ja tietoturvaa koskeviin kysymyksiin. Kun Big Datan keruu tapahtuu automaattisten tekniikoiden kautta autonomisesti ilman ihmisen puuttumista prosessiin, herää eettinen huoli ihmisen kykenevyydestä vaikuttamaan käyttötarkoitukseensa nähden epäeettisen datan, kuten yksityisyyden suojaa rikkovan datan keräämisen kontrollointiin.

### 3.4 Data-analyttinen ajattelutapa

Suurien ja reaaliaikaisten datamäärien analysointi vaatii tilintarkastajalta uudenlaista ajattelutapaa verrattuna perinteiseen ajattelutapaan. Tilintarkastajan perinteinen ajattelutapa pyrkii toteuttamaan PCAOB:n määrittämiä tilintarkastustoimenpiteitä, kuten

näytteiden otantaa ja analyttisiä toimenpiteitä poikkeamien paikantamiseksi asiakkaan datasta muodostetuista otoksista (PCAOB, 2010).

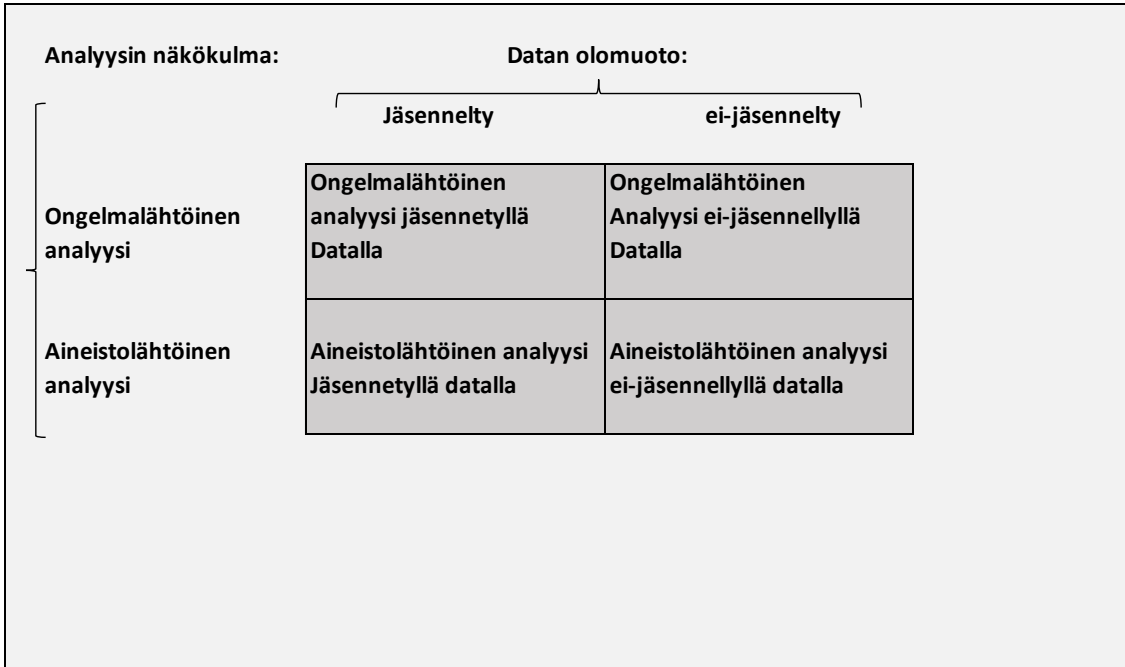
Kuten pian ilmenee, tilintarkastajan ajattelutavan todetaan olevan muutoksessa. Uudenlaisen data-analyttisen ajattelutavan voidaan nähdä olevan tilintarkastajan perinteisen ajattelutavan vastinpari Big Data-aikakaudella. EY (2015b) määrittelee tyhjentävästi sen, mitä data-analyttisellä ajattelutavalla tarkoitetaan tilintarkastuksen konseptissa; *“Data-analyttisellä ajattelutavalla tarkoitetaan tilintarkastajan kykyä suorittaa tilintarkastustoimenpiteitä sisällyttämällä niihin tavallista suurempia määriä erityyppisistä lähteistä peräisin olevaa tietoa, jotta tilintarkastuksen laajuus, ajoitus ja luonne olisivat riittäviä luomaan tarkoituksenmukaisen näkemyksen asiakkaasta”*. Tang ja Khondkar (2017, 36) mukaan tilintarkastajien tulee erityisesti kasvattaa tietämystään erilaisista ei-jäsenneilyn ja jäsenneilyn datan lähteistä pystyäkseen ymmärtämään ja palvelemaan Big Dataa tuottavia ja hyödyntäviä asiakkaitaan.

Jatkuvan tilintarkastuksen yleistyessä tilintarkastajalta edellytetään myös yhä voimakkaammin tulevaisuusorientoitunutta ajattelutapaa ongelmien ratkaisemiseksi perinteiseen tilintarkastukseen liittyneen menneisyysorientoituneen ajattelutavan sijaan. Tämä auttaa tilintarkastajaa reagoimaan tehokkaasti ongelmiin heti niiden sattuessa ja työskentelemään yhdessä asiakkaan johdon kanssa havaittujen ongelmien ratkaisemiseksi. (Vasarlehyi ym. 2010, 56 – 57.)

Ramulkan (2015, 17) mainitsee Big Datan luoman arvon tilintarkastukselle olevan olennaista ainoastaan, mikäli tilintarkastaja on itse kykeneväinen hyödyntämään data analytiikan luomaa evidenssiä ja luomaan sen avulla arvoa luovia johtopäätöksiä asiakkaastaan. Brown-Liburd ym. (2015, 455) mukaan tilintarkastuksen laatu jää puutteelliseksi ilman tilintarkastajan oikeanlaatuista kykyä liittää Big Dataa perinteisen tilintarkastusevidenssin osaksi

Big datan analysoiminen ei tosin ole todettavasti tilintarkastajien keskuudessa helppoa. Kaksikymmentäviisi prosenttia AICPAn vuoden 2014 kyselytutkimukseen osallistuneista tilintarkastajista kokivat Big Data-analyysit yhdeksi tilintarkastuksen suurimmista haasteista (AICPA 2014b, 8). Tähän liittyen, Rasso (2015, 44 – 45) tutkimustulokset viittaavat abstraktimmalla ja itsenäisemällä ajattelutavalla tilintarkastustoimenpiteiden työstämisen olevan todella tilintarkastajille haastavampaa ja aikavievämpää kuin tilintarkastusmateriaaleihin suoraviivaisen keskittymisen tarkkojen ohjeiden sanelemana. Tämä ajattelutapa kantaa hänen mukaansa kuitenkin lopulta hedelmää vaihtelevamman ja syvällisemmän tilintarkastusevidenssin muodossa.

### 3.5 Big Datan aiheuttama tilintarkastusanalyttinen näkökulmamuutos



Kuvio 6: Tilintarkastajan ja muun laskenta-ammattilaisen analyysin näkökulmat. (Mukaiillen Richins ym. 2017, 65.)

Richins ym. (2017, 65) mukaan Big Data-aikaan siirtyminen lisää tilintarkastajien soveltamia analyysinäkökulmia tilintarkastuksessa. Dataan kohdistuvan analyysin luonnetta tätä hyödyntävän henkilön jäsentelytavoissa voidaan tarkastella kahden dimension kautta; datan tyyppi (jäsennelty ja jäsenneltymätön) ja analyysin näkökulma. Analyysin näkökulmat sisältävät ongelmalähtöisen analyysin näkökulman (engl. ”problem based analysis”) ja aineistolähtöisen analyysin näkökulman (engl. ”exploratory analysis”). Näistä muodostuu nelikenttä, jonka kentistä kolme kenttää edustavat Big Datan tulkitsemiseen soveltuvia analyysityyppejä ja yksi perinteistä analyysityyppiä. Richins ym. (2017, 65) mukaisesti mukailtu kuvio tilintarkastajan ja muun laskenta-ammattilaisen analyysin näkökulmista on havainnollistettu edellisellä sivulla.

Ongelmalähtöinen analyysi on toinen analyysin näkökulmista. Ongelmalähtöistä analyysiä hyödynnettäessä ongelmat tunnistetaan, hypoteesit muodostetaan jo etukäteen mahdollisten syiden pohjalta ja ratkaisut kehitetään näiden hypoteesien kautta vastaamaan alussa määriteltyihin ongelmiin. Suurin osa perinteisestä tilintarkastuksesta on nykyisellään keskittynyt hyödyntämään lähinnä ongelmalähtöistä analyysinäkökulmaa jäsenneelyn datan analysointiin. (Richins ym. 2017, 65 & 72.) Byrnes ja Pawlicki (2014, 5 – 6) määrittelee aineistolähtöisen analyysin vastinpariksi

ongelmalähtöisen analyysin sijaan varmistavan analyysin (engl. ”confirmatory analysis”), joka vastaa määritelmällisesti Richins ym. (2017, 65 & 72) kuvaamaa ongelmalähtöistä analyysiä.

Ongelmalähtöisen analyysin vastaparina voidaan nähdä aineistolähtöisen analyysin. Aineistolähtöinen analyysi on tällä hetkellä varsin nopeasti yleistyvä analyysityyppi koko laskentatoimen alueella johtuen datamäärien nopeasta kasvusta. Aineistolähtöisessä analyysissä tehdään yhteenveto suurista datamääristä, jotta ymmärrettäisiin dataan liittyvien ominaisuuksien pääpiirteet ja poikkeamat. Nämä yhteenvedot ovat luonteiltaan itsenäisiä muodollisiin etukäteen asetettujen hypoteesien testaukseen verrattuna. Toisin sanoen analyysin tekijä ei ole tehnyt aineistolähtöisessä analyysissä etukäteen datasta oletuksia. Aineistolähtöinen analyysi on luonteeltaan korkeatasoiseen teknologiaan tukeutuvaa. (Richins ym. 2017, 65.)

Amerikkalainen, jo edesmennyt matemaatikko John. W Tukey loi urallaan aineistolähtöisen analyysin käsitteen. Tällä oli merkittäviä seurauksia fyysisille ja sosiaalisille tieteille, sekä myös myöhemmin data-analyysille. Hänen omaa alkuperäistä, hyvin perinpohjaista kuvaustaan aineistolähtöiselle analyysille on kuitenkin hankala määritellä ytimekkäästi. Tämä ei ole ihme, sillä John W Tukey vaikutti nauttivan suuresti useiden näkökulmien mahdollistavien hankalien konseptien parissa työskentelemisestä. (Klingemann 2011, 530.)

Big Data -aikakaudella ongelmalähtöinen analyysi jäsentämättömällä datalla on tulossa yhä tärkeämmäksi, kuten myös aineistolähtöisen analyysin kumpikin muoto. Tilintarkastajien ollessa jo alkujaan päteviä suorittamaan ongelmalähtöistä analyysiä jäsenneyllä datalla, tulee heidän aluksi integroida tätä taitoa jäsenneytyn datan ohella myös suurivolyymistä ei-jäsenneytä dataa kohtaan luodakseen Big Datan kautta lisäarvoa tilintarkastukselle. Tämä auttaa myös tilintarkastajia pysymään varmistuspalvelujen kilpailussa mukana. (Richins ym. 2017, 67, 72 & 74.)

Vaikka tilintarkastajan työssä hyödynnettävien analyysien näkökulmien tuleekin laajeta Big Datan yleistyessä tilintarkastuksessa, tärkeäksi kysymykseksi nousee se, missä määrin Big Datalla ja varsinkin asiakkaan ulkoisista lähteistä syntyvällä ei-jäsenneyllä datalla on mahdollista ylipäättään korvata perinteistä tilintarkastusevidenssiä tilintarkastajan työssä. Kyunghie, Hoogduin ja Li (2015, 436–437) mukaan Big Data -pohjainen, erityisesti ei-jäsenneyttilä tilintarkastusevidenssi on ominaisuuksiensa kautta vain arvokas lisä perinteiselle tilintarkastusevidenssille, sillä se vain jonkin verran vähentää tilintarkastajan riippuvuutta suoraan asiakkaalta saatua dataa kohtaan. Toisaalta Mc Kinney ym. (2017, 77) toteavat ei-jäsenneyllä Big Datan liittämisen perinteisen datan kanssa samaan analyysiin voivan helpottaa useiden tilinpäätöserien arviointeja. Tällöin vaikutus arvioituun asiakkaan sisältämään tilintarkastusrisktiin on ilmeinen. Vassarley ym. (2015, 388) kuitenkin täydentävät, että Big Data pohjaisen evidenssin integroiminen perinteisen tilintarkastusevidenssin kanssa ei ole helppoa. Tämä johtuu Big Datalle

yleisestä ei-jäsennellystä olomuodosta, jolloin se on rakenteeltaan epämääräistä. Jotta tilintarkastajat kykenisivät parhaiten liittämään ei-jäsennettyä Big Dataa tilintarkastukseen, heidän pitäisi löytää sopivia ”siltaratkaisuja” perinteisen ja uudentyyppisen datan linkittämiseksi keskenään osaksi laatimiaan analyysejään.

Tilintarkastuksen analyyttisen näkökulman voidaan todeta laajenevan ja monipuolistuvan Big Datan hyödyntämisen lisääntymisen myötä yhä useammaksi erilaiseksi lähestymistavaksi. Ongelmalähtöinen analyysi laajenee jäsenneilyn datan tarkastelusta myös ei-jäsennettyä Big Dataa kohtaan ja sen rinnalle nousevat aineistolähtöisen analyysin erilaiset näkökulmat. Seuraavassa alaluvussa tarkastellaan Big Datan prosessoimisen aiheuttamia ilmiöitä tilintarkastajan kognitiivisessa työskentelyssä.

### **3.6 Tilintarkastajan kohtaamat yleiset haasteet ja ammatillinen skeptisyys Big Datan tulkinnessa**

Kuten jo aiemmin tässä tutkielmassa on selvennetty, Big Datan käsittelyyn perustuva teknologia automatisoi ja nopeuttaa tilintarkastusevidenssin keräämistä. Tästä huolimatta olennaisimman datan määrittäminen ja data-analytiikan luomien tulosten arviointi jää vastedeskin tilintarkastajalle itselleen. Erityisesti olennaisen ja epäolennaisen tilintarkastusevidenssin erottamisen tärkeys korostuu evidenssin määrän lisääntyessä. (Wuerfel & Sikora 2017, 12.) Tyypillisesti tilintarkastajien tarkastelemat Big Data-analyysit sisältävät muun muassa erinäisiä graafisia taulukoita, joista on mahdollista havaita dataan liittyviä syy-yhteyksiä ja merkitseviä kuvioita (Jagadish, Gehrke, Labridinis, Papakonstantinou, Patel, Ramakrishnan & Sahabi 2014, 93).

Nopeasti kasvavien datamäärien luomat haasteet näkyvät jo tilintarkastajilta kaivatun osaamisportfolion muutoksena. PWC (2015, 13 – 16.) mukaan tilintarkastajien koulutustarpeet ovat Big Datan käsittelyn yleistymisen ja data-analyysien myötä muuttumassa. Erityisesti suurten datamäärien hallintaan liittyvät koulutukset ovat yhtä tarpeellisempia tilintarkastajiksi aikoville opiskelijoille. Tilintarkastajalle olisikin hyödyllistä sisällyttää korkeakoulututkintoonsa enemmän ohjelmoinnin kurseja, jäsenneilyjen ja ei-jäsenneilyjen tietokantojen hallintaa ja tilastollisten monimuuttujamenetelmien hallintaa.

Big Datan tulkitseminen ei ole ongelmaton. Big Datan ja Big Data-analytiikan tiedetään asettavan ainutlaatuisia haasteita data-analyysille ja niitä analysoiville henkilöille datan suuren määrän vuoksi. Erityisesti laajan datamäärän varastointi ja lukuisten erityyppisten datalähteiden keskinäinen yhdisteleminen ovat Big Datan

prosessointiin liittyviä avainhaasteita. Näiden haasteiden juurisyyhin kuuluvat datan muodon runsas vaihtelu, datan nopea virtaus, laajalle levinneet datan lähteet, runsas häiriöitä sisältävä ja huonolaatuinen data sekä seuraavana mainittava datan moniulotteisuus. (Najafabadi ym. 2015, 2.) Big Datan tilintarkastukselle tuottaman lisäarvon todetaan jäävän rajoittuneeksi, mikäli analytiikkaa hyödyntävät henkilöt ovat kykenemättömiä ymmärtämään ja työstämään Big Data-analyysin lopputulemaa. Tyypillisesti Big data analytiikan tuottaman analyysin käsittely sisältää erilaisten oletusten ja päätelmien muodostamista analyysin lopputulemasta ja yhteyksistä asiakkaan liiketoimintaan. (Labrinidis & Jagadish 2012, 2032.)

Big datan analysoimisen ja prosessoimisen päähaasteita ovat datan moniulotteisuus Big Datan hyödyntämiseen liittyvien suurten otoskokojen ohella. Big Datan moniulotteisuus aiheuttaa tarkastelupopulaatiossa täysin erilaisten datamuotojen olemassaolon aiheuttamaa alhaista homogeenisyyttä, sekä tuo esiin datan käyttötarkoitukseen nähden epärelevantteja korrelaatiota eri datapisteiden välillä. Big Dataan liittyvät suuret otoskoot kerätään sangen usein useammasta lähteestä ja eri ajanhetkiltä. Tämä luo dataan tilastollisia vääristymiä ja heterogeenisyyttä, jolloin dataa on hankalaa yhdistellä. (Fan, Han & Liu 2014, 2.)

Krumholz (2014, 1066) muistuttaa siitä vaarasta, että Big Dataan perustuva analyysi voi olla aina lopulta jossain määrin väärä. Tämä johtuu Big Datan monimutkaisesta ja moniulotteisesta luonteesta. Tämän vuoksi Big Data-analyysin tuloksia ei tulisi lähtökohtaisesti luokitella valideiksi tai epävalideiksi, vaan pikemminkin tarkoituksenmukaisiksi tai epätarkoituksenmukaisiksi. Myös Mc Kinney Jr ym. (2017, 68) toteavat ettei Big Datan suuri volyymi saisi sokaista Big Dataa hyödyntävää tahoja liiaksi luottamaan Big Data-analyysin lopputuleman totuudellisuuteen, sillä valtavan datamäärän taakse kätkeytyy myös aina suuri lukumäärä erilaisia selittäviä ja ei-selittäviä tekijöitä.

Edellä läpikäytyt tutkimustulokset viittaavat Big Dataa prosessoivalta tilintarkastajalta vaadittavan myös aivan uudentyyppistä ammatillista skeptisyyttä toisaalta aiemmin ilmenneen Big Datan tulkitsemisen moninaisuuden ja toisaalta Big Datan laajuuden vuoksi. Tilintarkastajan ajattelutyötä leimaava ammatillinen skeptisyys on määritelty aiemmin tässä tutkielmassa alaluvussa määritelty alaluvussa 3.1.1. Tilintarkastajilta vaadittava ammatillinen skeptisyys todettavasti suhteellisesti vähenee heidän työskennellessään erittäin suurten tietomäärien parissa. Ammatillinen skeptisyys auttaa tilintarkastajia muun muassa paikantamaan asiakkaastaan enemmän väärinkäytöksiä, virheitä ja ristiriitaisuuksia. (Hurt, Brown-Liburd, Earley & Krishnamoorthy 2013, 59 & 72.)

### 3.7 Big Datan tilintarkastajalle aiheuttama informaatioylikuorma

Vaikka informaatioteknologian kehittyminen ja datamäärien jatkuva laajentuminen ovat vahvoja merkkejä yhteiskuntamme kehittymisestä yhä edistyksellisempään suuntaan, niillä on myös kääntöpuolensa. Näitä kääntöpuolia ovat muun muassa tätä suurta tietomäärää prosessoivan henkilön keskittymisen hajaantuminen, sekä erityisesti hänen kohtaamansa informaatioylikuorman lisääntyminen (Dean & Webb 2011, 80). Tässä alaluvussa keskitytään edellä esitetyistä kahdesta tekijästä jälkimmäiseen tilintarkastuksen näkökulmasta.

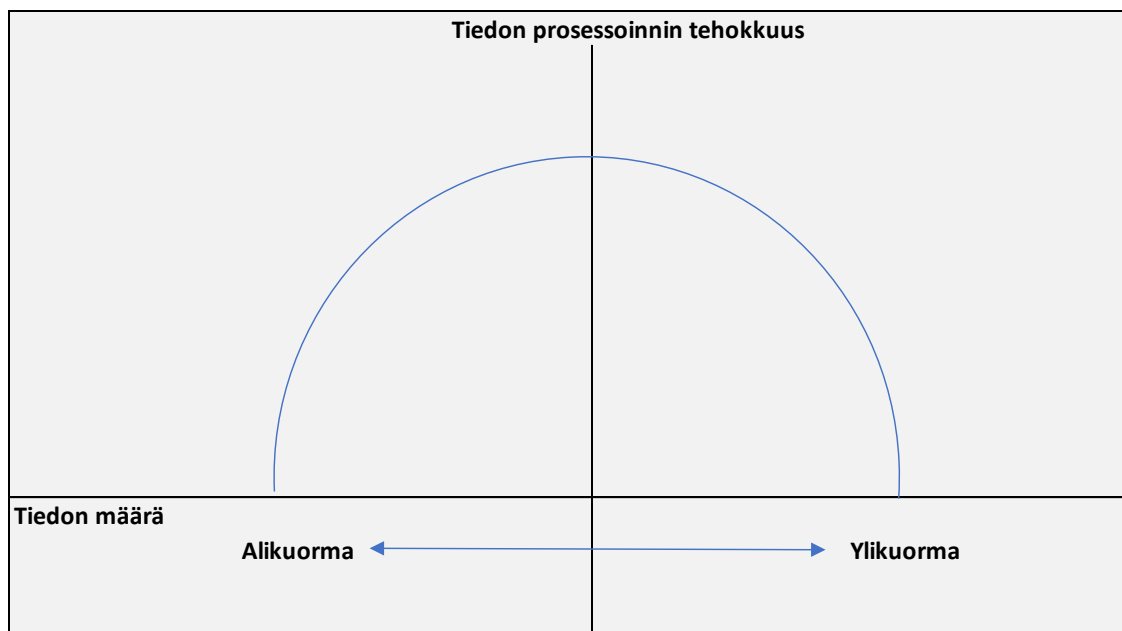
Henkilön kohtaamat informaatioylikuorman haitat ovat moninaiset. Tutkimusten valossa työntekijöiden kokema informaatioylikuorma aiheuttaa työsuoritusten kokonaislaadun kärsimistä (Casey 1980, 36), työsuorituksiin käytetyn ajan pitenemistä (Jacoby 1984, 435), kognitiivisiin työsuorituksiin liittyvien johtopäätösten epäjohtonmukaistumista ja lopullisen yhteisymmärryksen kärsimistä työn lopputulemaa kohtaan (Chewning & Harrell 1990, 527). Krahel ja Titera (2015, 417) mukaan datamäärien yhä laajetessa myös tilintarkastajan kokeman informaatioylikuorman vaikutukset ja kustannukset ovat nousemassa yhä ajankohtaisemmiksi aiheiksi tilintarkastuksessa. Lopez & Peters (2012, 141 – 142 & 162) toteavat tilintarkastajan kokeman liiallisen tietokuorman laskevan tilintarkastuksen laatua. Mitä informaatioylikuormalla syvällisemmin tarkoitetaan kognitiivisesta näkökulmasta käsin? Tähän on seuraavaksi syytä hakea vastausta aiheesta käsittelevästä tutkimuksesta.

Informaatioylikuormalla (engl. ”*information overload*”) viitataan yksinkertaisesti tiedon vastaanottajan olemassa oleviin resursseihin nähden liian suuren tietomäärän vastaanottamiseen. Kuluneina vuosikymmeninä Informaatioylikuormaan liittyen on tehty lukuisia tutkimuksia. Näiden tutkimusten avulla on pyritty ymmärtämään paremmin työn tehokkuuden ja tiedon prosessointikyvyn vaihtelua yksilön kohtaaman tietomäärän muuttuessa (Eppler & Mengis 2004, 326). Vaikka informaatioylikuormasta on keskusteltu jo melko kauan, sitä koskeva keskustelu on voimistunut huomattavasti viime vuosina tapahtuneen datan määrän eksponentiaalisen lisääntymisen myötä. Tätä ilmentävät jo Brandel (2008, 19, 23) vuosikymmen sitten esittämät toteamukset tiedon määrän nopeasta lisääntymisestä. Informaatioylikuorma onkin nykyisin tärkeä Big Data-perustaiseen työskentelyyn liittyvä, itse työntekijää läheisesti koskettava ilmiö (Saxena & Lamest 2018, 290, 295).

Informaatioylikuorma hankaloittaa tietoa prosessoivan henkilön kognitiivista työskentelyä (O'Reilly 1980, 684). Ihmisellä on vain rajallinen kapasiteetti sisällyttää tietoa tulkintoihinsa. Aluksi lisääntynyt tiedon määrä tekee kuitenkin ihmisen tiedon prosessoinnista tehokkaampaa. Tietomäärän ylitettyä tietyn pisteen tiedon prosessointi muuttuu kuitenkin selvästi tehottomammaksi. Tämä ilmiö muistuttaa visualisoituessaan käänteistä U-käyrää, jossa X-akselilla on tietomäärä ja Y-akselilla tiedon prosessoinnin



tehokkuus. Tiedon prosessoinnin tehokkuus on huipussaan tämän käänteisen U-käyrän lakipisteessä. Tämän jälkeen tietomäärän yhä kasvaessa tiedon prosessoinnin tehokkuus laskee dramaattisesti (Hwang & Lin 1998, 214). Myös tarve yhdistellä tietoa yhä useammista tietolähteistä vaikuttaa vastaavalla tavalla tiedon prosessointikyvyn tehokkuuteen, kuin kasvanut tietomäärä (Iselin 1988, 147). Hwang & Lin (1998, 214) mukaillen havainnollistettu kuvio tietoa käsittelevän henkilön tiedon prosessoinnin tehokkuuden ja tiedon määrän välisestä suhteesta esiteltynä seuraavana.



Kuvio 7: Tiedon prosessoinnin tehokkuuden ja tiedon määrän välinen suhde tietoa käsittelevällä henkilöllä. (Mukaillen Hwang & Lin 1998, 214.)

Big Data on potentiaalinen informaatioylikuorman aiheuttaja sen suuren määrän vuoksi. Uutta dataa siirtyy internetiin joka sekunti enemmän kuin olimme kokonaisuudessaan tallettaneet sinne 20-vuotta sitten (McAfee ja Brynjolfsson 2012, 62). Saxena ja Lamest (2018, 290 – 291 & 295) listaavat erilaisia tekijöitä sen suhteen, miten Big Data aiheuttaa ominaisuuksiensa vuoksi sitä työstävälle henkilölle informaatioylikuormaa. He toteavat Big Datan suuren volyymin hidastavan sitä prosessoivan henkilön työskentelyä. Tämä yhdistettynä Big Datalle mahdolliseen kvalitatiiviseen luonteeseen ja odotuksiin nähden epätyypillisiin datan lähteisiin tekevät sen prosessoimisesta melkoisen raskasta.

Myös tilintarkastajat altistuvat tutkimusten mukaan lisääntyneelle informaatioylikuormalle Big datan parissa työskentelyn kautta. Kun otannoista siirrytään koko populaation testaamiseen ja jatkuvaan tilintarkastukseen, on tilintarkastajaa uhkaavan informaatioylikuorman riski ilmeinen huomattavasti lisääntyneen datamäärän

ja sen sisältämien lukuisten poikkeusten myötä (Brown-Liburd ym. 2015, 458; Krahel & Titera 2015, 417 – 418).

Vaikka esimerkiksi koko populaation testaaminen onkin lävitseen luotaavuutensa johdosta nerokas keksintö tilintarkastuksessa, koko populaation testaamisesta johtuva informaatioylikuorma johtaa väistämättä tilintarkastustiimin sivuuttamaan useita tilintarkastustoimeksiannon yhteydessä ilmeneviä poikkeuksia. (Krahel & Titera 2015, 418.)

Tutkimusten mukaan myös jatkuvaa tilintarkastusta luonnehtiva ABE-näkökulma altistaa tilintarkastajan informaatioylikuormalle. Esimerkiksi Perlos & Murthy (2012, 35) mainitsevat jatkuvan tilintarkastuksen potentiaalista luoda kestäättömän suuri määrä poikkeuksia ihmismielen prosessoitavaksi.

Vaikka Big Dataa prosessoivan data-analytiikan sovellukset mahdollistavatkin tilintarkastuksessa suurten datamäärien nopean ja tehokkaan tarkastelun, analyttisten työkalujen tuottaman datan analyysi ja tulkinta voivat silti olla toisinaan hankalia tehtäviä tilintarkastajille datan suurivolyymisyyden vuoksi (Issa & Kogan 2014, 209). Vaikka tilintarkastaja olisi tottunut sisällyttämään suuria määriä eri lähteistä kumpuavaa tietoa analyysiinsä, analyysin luonne ja lopputulema voivat silti olla vähemmän laadukkaat tilintarkastajan kokemasta informaatioylikuormasta johtuen. (Brown-Liburd ym. 2015, 455).

Olemassa oleva informaatioylikuormaa käsittelevä kirjallisuus ja tutkimus eivät ainoastaan käsittele tiedon ylikuormituksen pääasiallisia syitä ja seurauksia, vaan myös ehdottavat vastatoimenpiteitä sitä vastaan. Tutkimuksissa esitetyt vastatoimenpiteet vaihtelevat yleisistä tiedon käsittelijän toimintaan liittyvistä muutoksista monimutkaisten analyttisten työkalujen hyödyntämiseen. (Eppler & Mengis 2004, 334.)

Tietoa käsittelevien henkilöiden toimintaan liittyen Eppler & Mengis (2004, 335) listaavat henkilöiden vastatoimenpiteiksi informaatioylikuormaa vastaan muun muassa ajankäytönhallinnan parantamisen, tiedon läpikäynnin tehostamisen, tiedon järjestelmällisen luokittelun ja organisoinnin, systemaattisen prioriteettien asettamisen työskentelylle ja keskittymisen ainoastaan työtehtävän kannalta arvoa luovaan tietoon. Chewning ja Harrell (1990, 539) taas ehdottavat lähestymistavaksi informaatioylikuormaa vastaan kenelle tahansa tiedon käsittelijälle ennalta määritellyjä henkisiä toimintamalleja, jotka ovat suunnattu juuri tiettyihin tilanteisiin. Tämä organisoitu toiminta voi vähentää suuren tietomäärän prosessoinnin hankaluutta.

Myös data-analyttisten työkalujen todetaan kirjallisuudessa helpottavan suurten datamäärien työstämisen parissa selviytymistä. Redman (2009, 61) mukaan suuren tietomäärän summaaminen sitä havainnollistaviksi kuvaajiksi ja kaavioiksi helpottaa tilintarkastajaa ymmärtämään ja prosessoimaan tehokkaammin suuren datamäärän kehitystrendejä, vähentäen hänen kokemaansa informaatioylikuormaa.

### 3.8 Olennaisuuden hahmottamiseen liittyvät aspektit Big Dataa työstävällä tilintarkastajalla

Big Datan suureen volyymiin sisältyy mahdollisuus paikantaa populaation sisältämiä yhteneväisyyksiä ja eroavaisuuksia, joiden tarkastelu ei olisi mahdollisia pienemmillä datamäärillä (Fan ym. 2014, 1). Kuten pian teorian tarjoamien esimerkkien kautta havainnollistetaan, Big Datan hyödyntämisen suhteen on kuitenkin erittäin ratkaisevaa myös tilintarkastajan oma kykeneväisyys hahmottaa tarkastustoimenpiteen kannalta ratkaisevia kuvioita suuresta datamäärästä.

Kun uusi tilanne tai ongelma tulee vastaan, useimmat ihmiset kehittävät aluksi mielessään itselleen heidän ajatteluaan ohjaavan kokonaiskuvan ongelmasta. Psykologian oppien mukaisesti tämä kokonaiskuva vaikuttaa tämän jälkeen tapahtuvaan henkilön suorittamaan kognitiiviseen työskentelyyn ja ongelmanratkaisuun (Bierstaker, Berard & Biggs 1999, 19). Tätä kokonaiskuvan luomisprosessia kutsutaan tässä tutkielmassa olennaisuuden -tai olennaisten kuvioden hahmottamiseksi (engl. ”*pattern recognition*”). Olennaisuuden hahmottamista käsitellään enimmäkseen psykologiaan liittyvissä tutkimuksissa, kuten tätä aihetta rönsyilevästi käsittelevässä Igushi ym. (2009, 49-60) julkaisussa, sekä useissa tekoälyä käsittelevissä tutkimuksissa (esim. Coakley & Brown 1993, 19 – 39; Libano, Rech, Tambara & Tonfrat 2018, 288 – 295).

Olennaisuuden hahmottaminen liitetään myös kiinteäksi osaksi tilintarkastajan työnkuvaa. Tyypillisesti olennaisuuden hahmottaminen painottuu tilintarkastustoimenpiteen alkuun, jolloin tehdään ensin analyttinen katsaus tarkastettavasta erästä tai populaatiosta. Esimerkiksi asiakkaan kahdenkertaisessa kirjanpidossa säännöllisesti toistuvien epätavallisten tapahtumien ja tilintarkastusriskin syyn erottaminen tilintarkastusasiakkaiden transaktiovirroissa vaativat ennen tarkempaan tutkimiseen ryhtymistä selkeän kokonaiskuvan hahmottamista tilanteesta. (Coakley & Brown 1993, 19 – 20.) Jotta tilintarkastajat hahmottaisivat asiakkaan datan sisältämät olennaiset kuviot, tulee heidän usein erottaa olennainen ja toisiinsa korreloituva informaatio liittyen useisiin toisiinsa linkittyviin kirjanpidon eriin. Yksinkertaistettu esimerkki tästä on myynnin lisääntymisen aikaan saama myyntisaamisten lisääntyminen. (O’Donnell & Perkins 2011, 274.)

Tilintarkastustyön alussa tapahtuvan olennaisten kuvioden hahmottamisen taidon todetaan vaikuttavan oleellisesti tilintarkastustoimeksiannon aikana ilmenevien ongelmien ratkaisuun. Heikko olennaisten kuvioden hahmottamisen taito johtaa ilmiöön, jossa tilintarkastaja ajautuu tarkistamaan toistuvasti ongelman alkutilanteessa ehdotettuja alkuperäisiä hypoteeseja, kykenemättä luomaan enää uusia, vaikka tilanne tätä edellyttäisikin. Edellä kuvattua tilannetta kutsutaan nimellä ”*wheel spinning*”-ilmiö. Mikäli alkuperäinen hypoteesi kuitenkin todetaan lopulta vääräksi, tilintarkastaja, jolla

on harjaantumaton olennaisten kuvioiden hahmottamisen taito, epäonnistuu helposti uuden hypoteesin luomisessa. Tilintarkastaja voi olla jopa kokonaan kykenemätön irtaantumaan alun perin ongelmaan asetetusta hypoteesista, vaikka hypoteesin vastaista evidenssiä ilmaantuisikin. Tämän seurauksena alhaisen olennaisten kuvioiden hahmottamiskyvyn omaavat tilintarkastajat ajautuvat helposti ”*wheel spinning*”- ilmiöön takertumiseen. Sen sijaan tilintarkastajat, joiden todetaan kykenevän kehittämään täydellinen ymmärrys asiakkaan datavirtojen välisistä yhteyksistä, ovat myös kykeneväisempiä luomaan uusia hypoteeseja datasta ilmeneviin ongelma-kohtiin. (Bierstaker ym. 1999, 18 – 19, 32.)

Bierstaker ym. (1999, 34.) ehdottavat olennaisten kuvioiden hahmottamisen laadun parantamiseksi ja ”*wheel spinning*”-ilmiön ehkäisemiseksi strategiaa, jossa alun perin asetettua hypoteesia paneudutaan aluksi pohtimaan tavallista pidemmäksi ajaksi. Tällöin tilintarkastajan näkemys ongelman luonteesta kokonaisvaltaistuu, ja myöhemmin tapahtuvilta umpikujilta ja jatkuvalta ”*wheel spinning*”- ilmiöön takertumiselta on mahdollista välttyä.

Tilintarkastajan tulee kyetä Big Datan myötä tunnistamaan diagnostisia kuvioita myös laajasta määrästä jäsenneilyä ja ei-jäsenneilyä Big Dataa. Big Datan suuren volyymin avulla kyetään paljastamaan koko populaatiota koskevia säännönmukaisuuksia. Esimerkiksi tekijä, joka saattaa olla tiettyssä pienessä satunnaisessa otoksessa muodostavan selkeän säännönmukaisuuden, ei välttämättä olekaan enää koko populaatiota tarkasteltaessa relevantti. (Fan ym. 2014, 8.)

Rose ym. (2017) tutkimuksen mukaan tilintarkastajan kyky hahmottaa olennaisia kuvioita Big Datasta on rajoittunut, mikäli hän tarkastelee Big Data-analytiikan tuottamia tuloksia ennen perinteiseen evidenssiin perehtymistä. Suurissa tilintarkastusyhtiöissä Big Dataan liittyvä taustaoletus kuitenkin yleensä on, että tilintarkastajat tarkastelevat suuresta datamäärästä löydettäviä tilastollisia kuvioita jo ennen perinteiseen evidenssiin perehtymistä. Tällöin oletuksena on, että Big Data voisi paljastaa hyödyllisiä kuvioita ja arvokasta tietoa tarkastuksen tueksi, mikäli tilintarkastajalla on vielä niin sanotusti ”puhtaat paperit” tapauksen suhteen. Tämä tarkoittaa, että hän ei ole vielä ehtinyt muodostaa perinteisen tilintarkastusevidenssin kautta hypoteeseja ja johtopäätöksiä, jotka voisivat vaikuttaa Big Data-analyysin tulkinnessa. (Rose ym. 2017, 82.)

Viimeaikaiset teknologiset trendit tarjoavat tietoa siitä, että tilintarkastajat voivat päästä data-analytiikan avulla yli joistain olennaisuuden hahmottamiseen liittyvistä ongelmista. Varsinkin suurten datamäärien ollessa kyseessä. Edellisessä pääluvussa mainittu jatkuvuusyhtälömenetelmä ja useat visuaalisen ja tekstuaalisen kontekstianalyysin tekniikat ovat tapoja muodostaa data-analytiikan avulla olennaisia kuvioita suurista datamääristä (Kogan ym. 2014, 233 – 234; Aldhizer 2017, 31 – 33). Thiprungsri ja Vasarlehvi (2011, 80 – 81) mukaan vakuutusosalalla hyödynnetty, valittujen parametrien kautta ryhmittelyä tuottava klusterianalyysi on osoittautunut hyödylliseksi

kuvioiden ja anomalioiden hahmottamista avustavaksi teknologiaksi suuria datamääriä tarkasteltaessa. Klusterianalyysin tulokset tarjoavat tarkempaa tietoa esimerkiksi transaktioiden luonteesta asettamalla samankaltaiset transaktiot eri ryhmiin, ja nostamalla esiin samankaltaisten transaktioiden heterogeenisyyden. Klusterianalyysin kaltaisen data-analytiikan hyödyntäminen auttaisi vakuutusalan ohella yhtäläisesti myös tilintarkastusta.

Olennaisten kuvioiden hahmottaminen on tilintarkastajalle tärkeä ominaisuus, ja siinä epäonnistuminen vaikeuttaa tässä tutkielmassa hyödynnetyn kirjallisuuden mukaisesti tilintarkastajan kognitiivista työskentelyä. Olennaisten kuvioiden hahmottamisen tärkeys ei kirjallisuuden mukaan vähene tilintarkastajilla Big Datan ja Big Data-analytiikan myötä. Kehittynyt analytiikka tarjoaa tähän kuitenkin apuvälineitä.

### 3.9 Big Datan sisältämän tiedon relevanttius

Tilintarkastus on ja tulee jatkossakin olemaan prosessi, jossa kerätään ja arvioidaan asiakkaalta saatavaa tilintarkastusevidenssiä. Tämän kautta asiakkaan taloudesta voidaan varmentaa oikeat ja riittävät tiedot. Relevanttia tilintarkastusevidenssiä voidaan pitää laadukkaana tilintarkastustyön perustana (Parker & Fogarty 2012, 11). Täten, mitä relevanttimman informaation parissa tilintarkastaja työskentelee, sitä enemmän ja laadukkaampaa työtä hän saa lähtökohtaisesti aikaan (Streufert 1973, 227).

Tilintarkastaja kohtaa paljon sekä relevanttia että irrelevanttia tietoa jokaisen yksittäisen tilintarkastustoimeksiannon yhteydessä (Shelton 1999, 217). Ideaalitilanteessa tilintarkastaja kykenee kuitenkin hyödyntämään kaiken relevantin informaation täsmällisen tilintarkastusarvioinnin perusteeksi (Blay 2005, 766).

Tämän päivän polttava puheenaihe Big Data-analytiikalle on, millaisia ovat relevantille analytiikan tuottamalle datalle asetellut kriteerit (Najafabadi ym. 2015, 18). Liiallisen irrelevantin informaation läsnäolon tarkastelupopulaatioissa tiedetään nimittäin heikentävän tilintarkastajan kognitiivista työssä suoriutumista (Shelton 1999, 2017). Kasvanut tiedon määrä muutenkin vaikeuttaa relevanttien vihjeiden paikantamista (O'Reilly 1980, 686). Edellisiin päätelmiin pohjaten on selvää, että Big Datan suuri volyyymi ja lukemattomat lähteet tuovat lähtökohtaisesti lisää haasteita relevantin informaation hankinnalle ja hyödyntämiselle Big Dataa hyödynnettäessä tilintarkastuksessa.

Big Datan suuri volyyymi ja datan tyyppien suuri lukumäärä hankaloittavat relevantin datan tunnistamista datamassasta (Zhang ym. 2015, 472). Aikaisemmin tässä pääluvussa käsitellyn informaatioylikuorman lopputulemana aiheutuu lopulta kykenemättömyys erotella osaa irrelevanttia ja relevanttia tietoa toisistaan. Suuret määrät irrelevanttia tietoa todistettavasti alentavat tätä tietoa käsittelevän henkilön kykyä paikantaa relevanttia

tietoa. Tämä laskee työnteon laatua. (Streufert 1973, 225.) Tätä ilmiötä kutsutaan laimennusvaikutukseksi (engl. ”dilution effect”) (Waller & Zimelman 2003, 254).

Laimennusvaikutus on havaittavissa myös tilintarkastustyössä. Tavallisesti tilintarkastajat ovat tottuneet käsittelemään rutiininomaisesti sekä relevanttia että epärelevanttia tietoa. Kuitenkin vain näistä ensimmäisen tulisi vaikuttaa heidän luomiinsa johtopäätöksiinsä. Tilintarkastajien kykeneväisyys riskien arviointiin asiakkaastaan alentuu selvästi, mikäli heille annetaan työstettäväksi poikkeuksellisen suuren määrän irrelevanttia tietoa sisältävä tapaus. Tilintarkastajien on kuitenkin mahdollista lieventää tiedon prosessointiin liittyvää laimennusvaikutusta. Tilintarkastusyhteisöt voivat esimerkiksi tehdä tilintarkastajat tietoisiksi ongelmasta, toivoen heidän etsivän itse tätä ongelmaa torjuvia työskentelystrategioita. Tilintarkastajat ovat myös lähtökohtaisesti kykeneväisiä erottamaan relevantit ja epärelevantit väärinkäytöksiin viittaavat indikaattorit toisistaan saatuaan tähän ensin koulutusta. (Hoffman & Patton 1997, 227 – 228 & 236.) Laimennusvaikutusta on mahdollista lieventää myös muutoin. Shelton (1999, 223) mukaan jo pelkästään tilintarkastajalle kertynyt kokemus on omiaan vähentämään laimennusvaikutusta.

Big Data-analytiikan etuna voidaan nähdä sen, että koko populaatiota prosessoiva Big Data-analytiikka ei jätä lainkaan relevanttia tietoa tarkastelun ulkopuolelle (Aldhizer 2017, 30). Big Datan luonne hankaloittaa kuitenkin relevantin tiedon hahmottamista kokonaisuudesta. Relevantin jäsentyneen datan erottaminen ja sille merkityksen kohdentaminen on kuitenkin melko helppoa. Mikäli tarkastelun alla oleva Big Data kuitenkin sisältää myös jäsentymätöntä dataa, relevantin datan tunnistus ja tietojen yhdisteleminen toisiinsa nähden hankaloituu selkeästi data-analyysin lopputulemasta. (Zhang ym. 2015, 472)

Tilintarkastajat kohtaavat tehokkaiden ja enemmän yksityiskohtia paljastavien data-analyttisten menetelmien ja koko asiakkaan datamäärien tutkimisen myötä myös yhä enemmän erilaisia pienehköjä virheellisyyksiä asiakkaiden datasta. Tällöin tilintarkastajat joutuvat vaikean päätöksen eteen arvioidessaan yksittäisten poikkeamien tutkimisen tärkeyttä koko tilintarkastuksen kannalta. (Whithouse 2014, 29.)

### **3.10 Big Datan sisältämän tiedon selkeys**

Kuten jo aiemmin tässä tutkielmassa on ilmennyt, Big Dataa kiinteästi leimaava piirre on selkeän ja jäsentyneen olomuodon lisäksi myös jäsentymättömyys ja lukuisat eri olomuodot. Brown-Liburud ym. (2015, 457) tarkentavat jälkimmäisten ominaisuuksien lisäävän datassa tiedon epäselkeyttä (engl. ”ambiguity”).

Dataa prosessoivan henkilön suhtautumista tiedon epäselvyyteen ja epäselviin tilanteisiin yleisellä tasolla on mahdollista mitata esimerkiksi Budnerin asteikolla.

Ihmisen psykologiaa 1960-luvulla tutkineen Stanley Budnerin tutkimusten mukaan tiedon epäselvyyttä huonosti sietävät henkilöt kokevat epäselkeät tilanteet uhkaavina (Budner 1962a, 29 – 50). Stanley Budner on kehittänyt testin, jossa kyselyyn vastaaja valitsee seitsemänportaisen asteikon skaalan sisältävältä täysin eri mieltä-täysin samaa mieltä -asteikolta itseään kuvaavimmat vaihtoehdot kompleksisen ongelman kohtaamista ja sen ratkomista varten. Lisäksi samalla kyselyllä selvitetään uudentyyppisissä tilanteissa toimimiseen liittyviä kysymyksiä. Mitä korkeamman pistemäärän vastaaja saa, sitä heikommin hän sietää tiedon epäselkeyttä. (Budner 1962b.) Lowe ja Reckers (1997, 53) toteavat tietoa prosessoivan henkilön toleranssin asteen tiedon epäselvyyttä kohtaan olevan tärkeä tekijä arvioiden ja päätelmien luomisessa jostain tietystä kohteesta.

Tiedon epäselvyyden sietokyky koskettaa myös tilintarkastajan työskentelyä. Lowe ja Reckers (1997, 53) mukaan tiedon epäselvyyttä hyvin sietävät tilintarkastajat ovat tavallista kykeneväisempiä liittämään päätelmiensä tueksi ylimääräistä tietoa. Tämä johtaa usein enemmän lisäarvoa tuoviin tilintarkastustoimenpiteisiin kuin tiedon epäselvyyttä heikommin sietävillä tilintarkastajilla.

Tiedon epäselvyyden todetaan vaikuttavan konservatiivisesti tilintarkastajan työskentelyyn. Tiedon epäselvyyden tason ollessa kohtalaisen korkea, tilintarkastajat liittävät tilintarkastusdokumentteihin herkästi huomioita mahdollisista epävarmuustekijöistä. Tiedon epäselvyyden tason ollessa alhainen, tilintarkastajat raportoivat vastaavasti määrällisesti vähemmän epävarmuustekijöistä. (Nelson & Kinney 1997, 269.)

Big Datan hyödyntäminen tilintarkastuksessa tuo myös oman ulottuvuutensa tilintarkastustyötä leimaavalle tiedon epäselvyydessä toimimiselle. Big Datan onnistunut hyödyntäminen tilintarkastuksessa riippuu ensisijaisesti tilintarkastajan asennoitumisesta tiedon epäselvyyteen. On todennäköistä, että tiedon epäselvyyttä heikosti sietävät tilintarkastajat kokevat olevansa mukavuusalueensa ulkopuolella käsitellessään varsinkin ei-jäsenneltyä Big Dataa. Tällöin he myös saattavat vältellä sen käsittelyä mahdollisimman paljon, jolloin Big Data-analytiikan tuottamasta tiedosta ei saada lainkaan lisäarvoa tilintarkastukselle. (Brown-Liburd ym. 2015, 458.)

Heterogeenisestä Big Datasta puuttuu tyypillisesti selkeä yhteinen nimittäjä, jonka avulla sen kautta on mahdollista erottaa dataan linkittyvät yhteiset reaali maailman objektit. Jo pelkästään useissa yritysten datankeruujärjestelmissä kuten SAP:issa ja Oraclessa jäsennellyn datan rakenne, määrittely ja muoto vaihtelevat radikaalisti toisiinsa nähden. Näin ollen eri lähteistä kerätyn datan informaation laatu on riippuvainen jo siitä kuinka eri lähteet ovat yhdisteltävissä keskenään. (Issa 2013, 91 – 92 & 130.) Erityisesti kuitenkin ei-jäsenneltyyn dataan liittyy sen vaikea kohdistettavuus jollekin tietylle objektille, mikä hankaloittaa selkeän yleiskuvan luontia koko datapopulaatiosta (Mc Kinney ym. 2017, 70).

Big Datan luomaan tiedon epäselvyyteen liittyy myös dataan liittyvät häiriöt (engl. “noise”), jotka Jagadish ym. (2014, 90) mukaan ovat tunnusomaisia erityisesti ei-taloudelliselle Big Datalle. Ei-taloudellisen Big Datan sisältämällä häiriöillä tarkoitetaan tällaisen datan sisältämää suurta määrää merkityksetöntä tietoa. Vaikka ei-taloudellinen Big Data sisältääkin usein häiriöitä, on häiriöitä sisältävä koko populaatio kuitenkin informatiiviselta arvoltaan pientä vastaavasta datasta luotua vähähäiriöisempää otosta arvokkaampi. Earley (2015, 498) mukaan tilintarkastajan tulee tietää, millä tasolla häiriöitä sisältävä ei-taloudellinen data on soveltuvaa tilintarkastusevidenssiksi. Tilintarkastajan tulee myös tietää, milloin tietynlaisia häiriöitä sisältävä data on kelpaamatonta tähän tarkoitukseen.

Häiriöiden lisäksi Big Dataan liittyy myös datan helposti sekoitettavuus toiseen vastaavaan kaltaiseen dataan. Big Datan lähteiden lisääntyttyä myös datan väliset päällekkäisyydet ja samankaltaisuudet ovat lisääntyneet (Zhang ym. 2015, 471 – 472, 474). Jotkin ulkoiset datan lähteet voivat myös olla toisiinsa nähden korreloituneita. Tällöin useista lähteistä kerätyn evidenssin yhteisarvo jää rajoittuneeksi, eri lähteiden ollessa informaatioarvoiltaan ja sisällöiltään voimakkaasti yhteneviä. (Kyunghee ym. 2015, 435)

Sillä Big Dataa tulee sekä asiakkaaseen nähden ulkoisista että sisäistä kanavista, tulee tilintarkastajan olla myös varma datan tulevan suojatusta lähteestä (Earley 2015, 497). Tilintarkastusevidenssin laatu vaikuttaa tilintarkastusevidenssin luotettavuuteen (Appelbaum 2016, 20). Korruptoitunut tai manipuloitunut data voi johtaa tilintarkastajan väärään johtopäätökseen (Geat & Xie 2017). Asiakkaaseen nähden ulkoiseen Big Dataan liittyy ongelma myös sen lähteiden varmistamisen suhteen, jota ei ilmene perinteisen evidenssin kohdalla. Perinteisen datan lähteen varmistamisen perusteeksi on tyypillisesti löydettävissä myös paperista dokumentaatiota, mutta asiakkaaseen nähden ulkoisesta ei-jäsennellystä evidenssistä tämä tyypillisesti puuttuu. (Appelbaum 2016, 20).

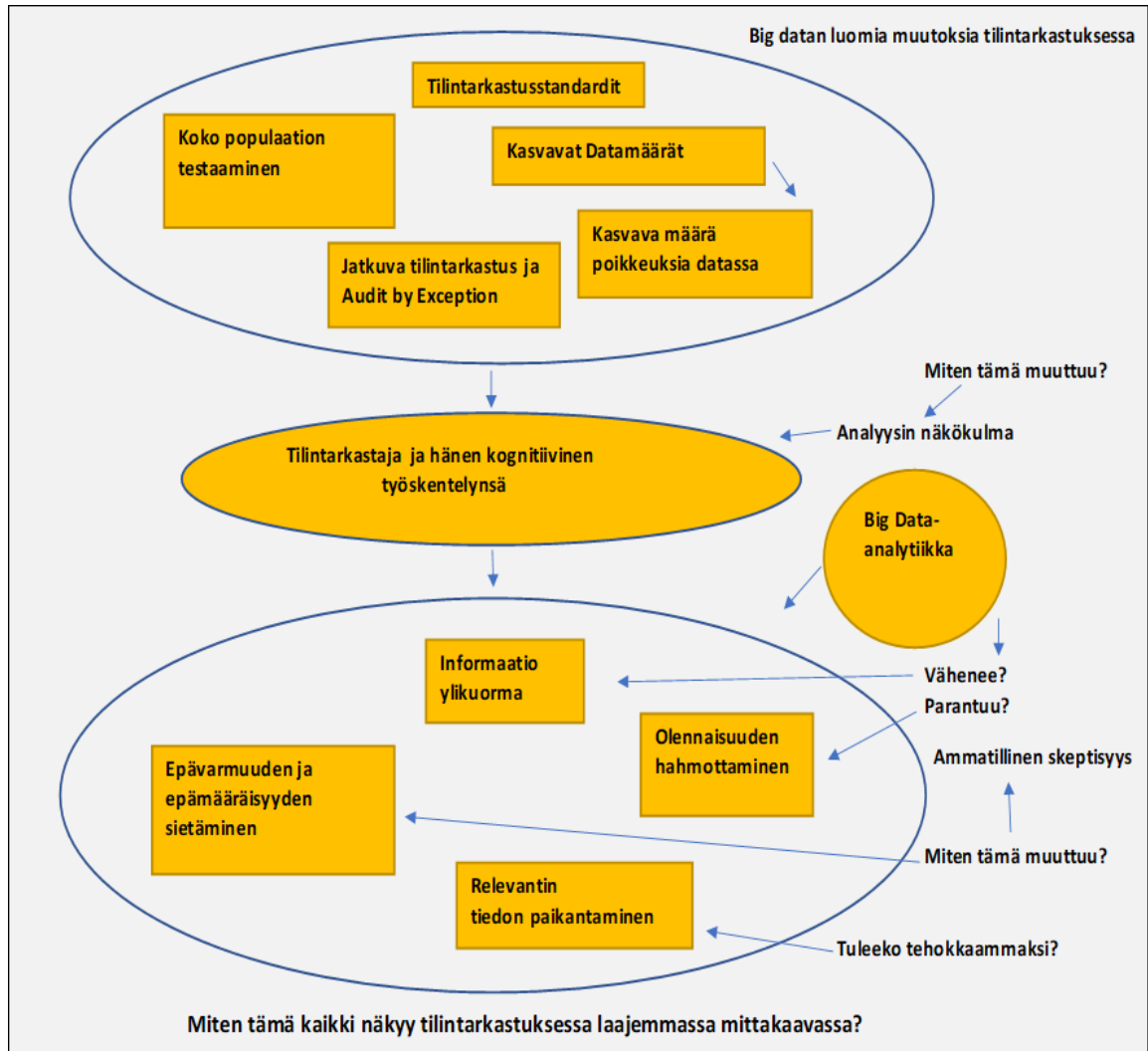
### **3.11 Tämän tutkielman kirjallisuuskatsauksen perustuva tutkimusaukko ja teoriaosuuden yhteenveto**

Tämän tutkielman teoriaosuuden päättävän yhteenveto-osuuden tehtävänä on selventää ja perustella tämän tutkielman empiirisen osuuden tarkoitusta ja olemassaoloa tässä tutkielmassa läpikäydyn kirjallisuuden pohjalta. Tätä kuvaillaan selväpiirteisesti seuraavana esitellyn, tämän tutkielman tutkimusaukon viitekehystä kuvaavan kuvion kautta.

Kuvion ylimmässä aiheryhmässä havainnollistetaan tämän tutkielman ensimmäisessä teorialuvussa käsiteltäviä teemoja, joiden voidaan todeta luovan pohjan tämän tutkielman



varsinaiselle tutkimusaukon käsittelemiselle myöhemmin teorian aikana. Ylemmän aihe ryhmän tekijät vaikuttavat nimittäin suoraan tilintarkastajan kognitiiviseen työskentelyyn, johon liittyviä tekijöitä on kuvattu teoriaosuuden toisessa pääluvussa. Alemmassa aihe ryhmässä esitelläänkin toisen teorialuvun pääteemoja, joille tutkielman empiria enimmäkseen suoraan pohjaa.



Kuvio 8: Tämän tutkielman tutkimusaukon viitekehys

Tutkielman teoriaosuus herättelee esiin edellisellä sivulla havainnollistetun kuvion mukaisesti monia kysymyksiä, joille on syytä etsiä vastauksia empiirisen tutkimuksen kautta. Ensinnäkin; miten Big Data ja Big Data-analytiikka vaikuttavat tilintarkastajan kokemaan informaatioylikuormaan, ammatilliseen skeptisyyteen, olellisuuden hahmottamiseen, epämääräisen tiedon sietämiseen tai relevantin tiedon paikantamiseen? Mikä siis on tilintarkastajan käytännön kokemus hänen kognitiivisesta työsuorituksestaan

hänen työskennellessään Big Datan parissa? Entä miten tämä kaikki edellä läpikäyty vaikuttaa tilintarkastukseen kokonaisuutena?

Jotta edellä mainittujen kysymysten luoma tutkimusaukko voidaan perustella, tulee tehdä yhteenveto tässä tutkielmassa käsitellystä teoriasta. Tämän yhteenvedon tarkoituksena on luoda selkeä viitekehys sille, mitkä tekijät vaikuttavat tilintarkastajan kokemukseen tilintarkastustyötä kohtaan juuri nyt heräilevällä Big Data -aikakaudella.

Kuten tässä tutkielmassa käsitellystä teoriasta ilmenee, tilintarkastus on Big Datan ja tätä hyödyntävän analytiikan ilmaantumisen myötä kokonaisvaltaisessa muutoksessa. Olennaisimpia muutoksia lienevät mahdollisuus koko populaation testaamiselle pelkän otannan sijaan sekä jatkuva tilintarkastus (Issa 2013, 51). Samalla kun Big Dataa on mahdollista kerätä reaaliaikaisesti, aivan uudenlainen, reaaliaikaisesti olennaisiin poikkeamiin reagoiva ”*audit by exception*” -tilintarkastusnäkökulma on kehittymässä vaihtoehdoksi perinteiselle periodeittain tehtävälle tilintarkastukselle (Appelbaum ym. 2017, 11). Vaikka tilintarkastus päivittyykin Big Data-aikakaudelle, olemassa olevat tilintarkastusstandardit ovat kuitenkin valitettavasti yhä aikaansa jäljessä (Whithouse 2014, 28 – 29; Moffit & Vasarlehyi 2013, 14 – 15).

Samalla kun tilintarkastuksen konsepti on muutoksessa, joutuvat tilintarkastajat tämän muutoksen johdosta aivan uudenlaisten haasteiden eteen käytännön työssään Big Datan johdosta. Tilintarkastajat joutuvat ensinnäkin siirtymään etukäteen luotuja hypoteesejä testaavasta ongelmälähtöisen analyysin näkökulmasta ennakoimattomia ratkaisuja hakevaan aineistolähtöisen analyysin näkökulmaan (Richins ym. 2017, 65). Big Datan suuri määrä ja sen lävitseen luotaava totuudenmukaisuuden luonne haastavat myös tilintarkastajan kykyä ylläpitää tilintarkastusalaa leimaavaa ammatillista skeptisyyttä (Mc Kinney Jr ym. 2017, 68).

Tilintarkastustyötä kuvaavat varsinkin kiirekauden aikaan nopeatempoisuus, stressi, suuri työmäärä ja aikapaineessa suoritettava aivo työskentely. Tilintarkastaja altistuu herkästi kognitiivisille vinoumille ja erittäin suurien kiireellisten työmäärien aiheuttamille sivuvaikutuksille työsuorituksissaan (Jacka 2018, 67; Pradana & Salehudin 2015, 119). Myös tilintarkastajan kokemalla stressillä on vaikutusta hänen työssään suoriutumiseensa (Choo 1986, 25 – 26). Mielenkiintoinen kysymys onkin, miten Big Data ja Big Data-analytiikka vaikuttavat tilintarkastajan kokemaan työn stressaavuuteen?

Tutkimuskirjallisuuden mukaan suurten datamäärien käsittelyn aiheuttamia sivuvaikutuksia tilintarkastajassa ovat muun muassa informaatioylikuorma ja relevantin tiedon havaitsemisen hankaloituminen. Näiden tekijöiden tiedetään heikentävän työsuorituksen laatua (Saxena & Lamest 2018, 290–291, 295; Brown-Liburd ym. 2015, 455; Zhang ym. 2015, 472). Varsinkin informaatioylikuorman tiedetään jo ennestään hankaloittavan tietoa prosessoivan henkilön kognitiivista työskentelyä (O’Reilly 1980, 684). Edelliseen perustaen, Big Dataa leimaavan suuren volyymin voidaan olettaa kasvattavan tilintarkastajan kokemaan informaatioylikuormaa ja lisäävän sen aiheuttamia

sivuvaikutuksia tilintarkastukselle. Tilintarkastajan kokemaa informaatioylikuormaa ja siihen liittyviä tekijöitä onkin teoriakirjallisuuteen perustuen suotavaa tutkia Big Datan ja Big Data-analytiikan näkökulmista empiirisen tutkimuksen kautta lisää.

Olennaisten kuvioiden hahmottamisen kyvyn merkitys kasvaa tarkastettavien tietomassojen yhä laajetessa (Fan ym. 2014, 8). Kuitenkin kykenemättömyys erotella osaa irrelevanttia ja relevanttia tietoa toisistaan irrelevantin tiedon osuuden ja määrän lisääntyessä otoksessa johtaa niin sanottuun laimennusvaikutukseen (Waller & Zimelman 2003, 254; Streufert 1973, 225), joka on havaittavissa myös tilintarkastuksessa. Big Datan epäselvä luonne todistettavasti hankaloittaa relevantin tiedon havaitsemista datasta (Zhang ym. 2015, 472).

Edellä läpikäydyn kirjallisuuskatsauksen yhteenvedon perusteella on erittäin selvää, että tilintarkastajan kognitiivista suoriutumista käsittelevälle tutkimukselle on lisääntyntä tarvetta Big Datan työstämisen näkökulmasta tarkasteltuna. Tämän tutkielman kautta on tarkoituksena esittää täydentäviä näkökulmia tälle tutkimusaukolle.

## 4 TUTKIELMAN EMPIIRINEN OSUUS

Tämän tutkielman empiirinen osuus pohjaa teoriaosuuden päätteeksi luvussa 3.11 selvennettävään tutkimusaukkoon, jolle pyritään löytämään empirian kautta vastauksia. Tämän tutkielman empiirisen osuuden ensimmäisessä alaluvussa tarkastellaan ensin, miten tässä tutkielmassa hyödynnetty empiirinen aineisto on hankittu. Tämän jälkeen luodaan yleiskatsaus empiirisen aineiston lähteinä olleisiin haastateltuihin henkilöihin. Tämän pääluvun toisessa alaluvussa esitellään empiirisen aineiston esiintuomat tulokset. Pääluvun viimeisessä alaluvussa tehdään syvällisempi pohdinta havaituista empiirisistä tuloksista ja peilataan saatuja tuloksia tämän tutkielman teoreettiseen aineistoon. Tällä pyritään täydentämään tämän tutkielman luvussa 3.11 käsiteltävää tutkimusaukkoa.

### 4.1 Empiirisen aineiston hankinnan toteutus ja tutkielmaa varten haastateltujen henkilöiden esitteleminen

Kuten jo tämän tutkielman johdannosta ilmenee, tämän tutkielman empiirinen osuus rakentuu laadulliselle haastattelututkimukselle. Nämä haastattelut ovat teetetty Hirsijärvi ym. (1997, 203) mukaisina puolistrukturoituina haastatteluina, joiden tarkoituksena tuottaa syvällistä tietoa juuri tietystä ilmiöstä. Puolistrukturoidussa haastattelumenetelmässä kysymykset ovat haastattelijan valitsemat, mutta haastateltavalle annetaan vapaus vastata kysymyksiin täysin oman näkökantansa mukaisesti (Koskinen, Alasuutari & Peltonen 2005, 104). Tämän tutkielman empiirinen aineisto on kerätty haastattelemalla yhteensä kuutta tilintarkastustyötä tekevää henkilöä elokuun 2018 ja marraskuun 2018 välisenä aikana. Suurin osa haastatteluista ajoittuu kuitenkin syyskuun 2018 lopulle. Aineiston keruun perustaksi valikoitui haastattelumenetelmä, sillä tutkittavaa aihetta on sen kokemuseräisyytensä vuoksi otollisinta lähestyä henkilökohtaisten yksilöhaastattelujen kautta. Haastateltavia valikoitui kuusi ja heidän kattavuutensa on neljä suurta tilintarkastusyhteisöä Suomessa.

Haastateltavista viidellä kuudesta on ollut KHT-tilintarkastajan pätevyys haastatteluhetkellä. Haastateltaviksi valikoitui henkilöitä, jotka tekevät töitä muiden työtehtävien ohessa tilintarkastuksessa hyödynnettävän data-analytiikan parissa. Haastateltavia lähestyttiin LinkedIn-verkostoitumissivuston kautta, tai haastattelija otti heihin henkilökohtaisesti yhteyttä. Haastattelurunko lähetettiin haastateltaville jo etukäteen. Haastattelurungon sisältävä haastattelulomake löytyy tämän tutkielman lopun liitteistä. Samaa haastattelurunkoa on hyödynnetty kaikissa kuudessa haastattelussa. Haastattelut toteutettiin anonymieinä. Haastateltavista esitellään tässä tutkielmassa vain heidän tilintarkastuskokemuksensa ajallinen pituus, koulutus, työtehtävien pääasiallinen

sisältö ja työnimike. Heihin viitataan tässä tutkielmassa termein ”haastateltava 1, 2, 3” jne.

Tämän tutkielman tutkimusmenetelmäksi valikoitunutta puolistrukturoitua haastattelumenetelmää kuvaavaa on joustavuus (Horton, Macve & Struyven 2004, 340). Puolistrukturoidun haastattelumenetelmän joustavuus ilmenee tässä tutkielmassa haastattelujen vapaamuotoisuudessa. Toteutetut haastattelut rönsyivät ajoittain monenlaisia pohtivia näkökulmia kohti ja haastattelija esitti vapaasti ennalta valikoimattomia lisäkysymyksiä kunkin haastattelutilanteen edetessä. Tämä mahdollisti myös haastattelujen aikana uudelleenlaisille teema-alueille jalkautumisen.

Seuraavaksi haastateltavat esitellään ajallisessa haastattelujärjestyksessä vanhimmasta uusimpaan. Jokaista haastateltavaa yhdistäviä piirteitä ovat kauppatieteiden maisterin koulutus laskentatoimen pääaineen alta, sekä useamman vuoden kokemus data-analytiikasta suuressa tilintarkastusyhteisössä. Heidän roolinsa data-analytiikan parissa vaihtelevat jonkin verran, mikä tuo lisäarvoa tutkimusotokselle.

*Haastateltava 1.* on ollut haastatteluhetkellä associate -ranking nimikkeellä tilintarkastustyötä tekevä henkilö. Hän on ollut nykyisessään työpaikassaan haastatteluhetkellä kaksi vuotta. Hänellä on kuitenkin jo ennen tilintarkastuspuolelle siirtymistä kokemusta data-analytiikan parissa työskentelystä samalla työnantajalla. Ensimmäinen haastateltavista valikoitui haastateltavien otokseen kiinnostuksen kohteidensa ja käytännön tason työkokemuksensa kautta.

Haastateltava 1. kuvailee työnkuvaansa seuraavanlaisesti; *”Työtäni ovat tilintarkastus ja tilintarkastukseen liittyvien data-analyysien tekeminen ja sisäisissä hankkeissa data-analyyseihin ja muuhun dataan liittyvään työhön liittyvissä projekteissa myös mukana, ja myös henkilökunnan koulutusta esimerkiksi Excelin tiimoilta. Data-analyysit... no ne on ihan mun rutiinityötä, eli tota vuoden aikana aina tiettyihin yrityksiin tehään tiettyjä analyyseja ja osassa niistä mä oon ollu mukana useamman vuoden, tai siis osaa teen toista vuotta ja osaa nyt sitten ensimmäistä vuotta”.*

*Haastateltava 2.* on tilintarkastusrankingiltaan senior manager. Hän kuvailee työrooliaan data-analytiikan parissa seuraavanlaisesti; *”Mä oon ollu kaheksan vuotta tilintarkastajana. Käytännössä nyt koko työurani jossain määrin oon ainakin data-analyyseissä mukana. Tää big data on varmasti tossa viime aikoina tullu enemmän ja enemmän, että nyt sinänsä aika läheistä”.*

*Haastateltava 3.,* haastatteluhetkellä senior -rankingin tilintarkastaja, kuvailee työksensä yksinkertaisesti ”tilien tarkastamisen”. Hän on päässyt hyödyntämään eri toimeksiannoissa data-analytiikkaa ja kohdannut usein Big Dataa arkipäivän työssään.

Hän kertoo hakeutuneensa tehtäviin, joissa hän on päässyt työskentelemään data-analytiikan parissa.

*Haastateltava 4.* edustaa otoksen manager-rankingia. Hänet on valittu haastateltavien otokseen tilintarkastajalle tavallista spesifimmän työnkuvansa perusteella, jossa hän työskentelee ikään kuin edelläkävijänä uudenlaisten tilintarkastusteknologioiden implementointien parissa. Hän kuvailee omaa työnkuvaansa seuraavanlaisesti; ”*Toimin managerina täällä meidän tilintarkastuksessa ja pääasiassa, no perustilintarkastusteen ja vedän tiimejä ja sitten toinen rooli on myös meillä olla tällain innovaatiomanagerina.*”

*Haastateltava 5.* on edellisen tapaan työnimikkeeltään manager, kuvailee rooliaan data-analytiikan parissa seuraavanlaisesti; ”*Toimin neljättä vuotta tilintarkastustehtävissä ja toimin managerina tällä hetkellä. Oikeestaan päivittäisten töiden lisäksi oon mukana meidän analytiikkaryhmässä ja sen myötä saanu kosketusta analytiikkaan ja miten se vaikuttaa tilintarkastustyöhön.*”

*Haastateltava 6,* manager, selventää vuorostaan omaa työtehtäväkenttäänsä seuraavassa tyhjentevästi; ”*Oikeestaan päivittäisten töiden lisäksi oon mukana meidän analytiikkaryhmässä ja sen myötä saanu kosketusta analytiikkaan ja miten vaikuttaa tilintarkastustyöhön*”. Hänen kuvailunsa antaa hyvät perustelut sille, miksi hän on valikoitunut osaksi tämän tutkielman tutkimusotosta.

Ennen haastatteluja laadittiin kahdenkymmenen kysymyksen lista, joka etenee johdattelevista kysymyksistä yhä syvällisempiin kysymyksiin. Kysymyslista haastattelukysymyksistä löytyy tämän tutkielman lopun liitteistä. Samaa haastattelukysymyslistaa hyödynnetään kaikissa tämän tutkielman kuudessa haastattelussa, mutta kahdessa viimeisimmässä haastattelussa kysymyslista on läpikäyty nurinkurisessa järjestyksessä. Tämä on toteutettu sen riskin poissulkemiseksi, että haastateltavat vastaisivat kattavammin vain kysymyslistan alkupään kysymyksiin. Haastattelut on toteutettu muutoin haastateltavien henkilöiden organisaatiossa neuvotteluhuoneessa, paitsi yksi suoritettiin julkisella paikalla työajan ulkopuolella. Haastattelutyökaluna on hyödynnetty Turun yliopistolta lainattua litteraa sekä haastattelijan oman matkapuhelimen ääninauhuria. Äänitettyjen haastattelujen kesto vaihtelee reilun puolen tunnin ja kahden tunnin välillä.

Haastattelut on muokattu kaikille yhteiseen MP3 muotoon. Tätä on seurannut aineiston litterointi tekstimuotoon ennen varsinaisen empiirisen osion kirjoittamista. Kukin haastattelu on johdettu omiksi erillisiksi Microsoft Word -tiedostoiksi. Tämän jälkeen on siirretty tarkastelemaan haastattelujen tuloksia, joista lisää seuraavissa alaluvuissa.

Haastatteluaineistosta on hyödynnetty ainoastaan tämän tutkielman tavoitteen kannalta relevantit vastaukset ja kommentit, joten vain murto-osa kaikista haastatteluissa käydyistä keskusteluista ja kommenteista ilmenee tässä tutkielmassa.

Seuraavissa alaluvuissa esitellään tutkielman empiriaa varten tehdyissä haastatteluissa ilmennneet merkitsevät tutkimustulokset. Tutkimustulokset jaotellaan tarkempien alalukujen avulla tämän tutkielman teoriaan perustuviin teemoihin. Teemat ovat järjestyksessä; *Big Datan määritelmään ja tilintarkastuksen automatisoitumiseen liittyvät tutkimustulokset, Olennaisuuden hahmottamiseen, tilintarkastajan ajattelutavan muutokseen ja tilintarkastajan ammatilliseen skeptisyyteen liittyvät tutkimustulokset, sekä relevantin tiedon hahmottamiseen ja informaatioylikuormaan liittyvät tutkimustulokset.*

## **4.2 Big Datan määritelmään ja tilintarkastuksen automatisoitumiseen liittyvät tutkimustulokset**

Haastateltavien myöhempien vastausten ymmärryksen parantamiseksi on ensimmäiseksi hyvä tehdä kartoitus heidän määritelmistään koskien Big Dataa. Big Datan määritelmä on nimittäin kirjallisuuden valossa yhä vakiintumaton, kuten muun muassa Labrinidis ja Jagadish (2012, 2032) julkaisussaan esittävät. Haastateltavan 1. määritelmä mukailee hyvin pitkälle haastateltavan 3. määritelmää, joten tätä ei erikseen esitetä seuraavissa tuloksissa.

*”No Big Dataa on kaikki tieto ja mitä yrityksestä ja toimintaympäristöstä on tarjolla ja mitä saadaan muokattua tommoseen datamuotoon.”*

(Haastateltava 3, senior)

*”Se (Big Datan määritelmä) on vähän vaikee kysymys aina, kun tavallaan se riippuu ihan siitä käyttötarkotuksesta. Toistaseks suurimmaks osaks tilintarkastuksii me käytetään vaan sitä pääkirjadataa, eli kaikki viennit sieltä. Mutta sit isommissa tarkastuksissa meillä myös ne ostotilaus -ja myyntitilaustiedot ja voi myös olla et me käytetään pankkien viitesiertolistoja, joista saadaan tietää datana se tieto mitä on tullu rahana yhtiöön sisään”*

(Haastateltava 4, manager)

*”No varmaan yleisiä määrittelyjä voi olla hyvinkin paljon erilaisia, mutta ehkä tilintarkastuksen kannalta niin siihen vois liittyä sekä sitten tällaista taloudellista informaatiota ja tai dataa ja sitten jotain relevanttia ei-taloudellista dataa. ”*

(Haastateltava 6, manager)

*”Niin toihan on aika laaja käsite ja ite mä nyt monesti aina peilaan sitä tilintarkastuksen kannalta, mut Big Data niin ehkä se on kaikki tieto mikä on saatavilla vaikka järjestelmien kautta, ja mitä pystytään sitten hyödyntämään eri tarkoituksissa. Ehkä mä näen sitä voi tulla sosiaalisen median kautta, sitä voi tulla ihan totta kai asiakkaiden toiminnanohjausjärjestelmien kautta, siellä on todella paljon eri lähteitä. Ehkä semmosta hyvin selkeätä määritelmää on vaikee sanoa että nyky maailmassa se on mun mielestä hyvin laaja käsite.”*

(Haastateltava 4, manager)

*”Big data on, mä ehkä lähtisin hakee sitä tämmösestä isosta datamassasta, joka yleensä on jossain pilvessä niin sanotusti. Oikeestaan se yks määrittely on, se on isoo, se ei välttämättä oo kauheen järjestelmällistä, sitä on niin sanotusti saatavilla helposti ja enemmän semmonen tietovarasto, joka ei oo sillai niin aidattu. Tää nyt on, miten mä ehkä sitä käsittelen omassa työssäni.”*

(Haastateltava 2, senior manager)

Vastausten perusteella on havaittavissa myös tätä tutkielmaa varten haastateltujen tilintarkastajien käsittävän tällä ajanhetkellä Big Datan tilintarkastuksen konseptissa hyvin erilaisin tavoin. Toisille Big Data edustaa tilintarkastuksen konseptissa lähinnä asiakkaan pääkirjadataa, joka koostuu asiakkaan kaikista kirjanpitoivienneistä. Toisille haastateltaville Big Datan määritelmä ei ole tilintarkastuksenkaan konseptissa niin tarkasti rajattu, sen käsittäessä sekä taloudellista, että ei-taloudellista dataa. Tämä lähentelee tämän tutkielman ensimmäisessä teorialuvussa käsiteltyä Chen ym. (2013, 157) melko karkealuonteista Big Datan määritelmää.

Big Datan määritelmän ohella on tärkeää ymmärtää ennen syvällisempiin kysymyksiin siirtymistä millaiseen suuntaan haastateltavat tilintarkastajat kokevat tilintarkastuksen olevan parhaillaan yleisesti kehitymässä. Olemassa oleva kirjallisuus esittää lakisääteisen tilintarkastuksen joko automatisoituvan erittäin varmasti kokonaan (esim. Frey & Osborne 2013, 69), tai vähintäänkin muuttavan tilintarkastajan roolin enemmänkin konsultoivaksi liiketoiminnan ymmärtäjäksi (esim. Richins ym. 2017, 72 – 74).

Seuraavana esitellään miten tämän tutkielman empiiriseen aineistoon haastatellut tilintarkastajat ovat vakuuttavasti sitä mieltä, ettei tällä hetkellä käynnissä oleva teknologinen kehitys tilintarkastuksessa ole johtamassa tilintarkastuksen täydelliseen automatisoitumiseen. Haastattelujen perusteella voidaan tulkita kirjanpidon kirjauksissa ilmenevän inhimillisiä virheitä niin kauan kuin kirjanpidon tekee koneen sijasta ihminen. Näitä virheitä eivät koneet kykene selvittämään. Tällöin tarvitaan toista ihmismieltä tilintarkastajan roolissa ”ottamaan koppia” kirjanpidossa tapahtuneista inhimillisistä



virheellisyyksistä. Haastatteluissa korostuu täten se tosiaasia, että tilinpäätöksen luvut ovat lopultakin vain Elbannan (2006, 484) määritelmän mukaisia inhimillisen ajattelutyön tuotteita, joihin vaikuttavat tilinpäätösten laatijoiden ja tilinpäätösten sidosryhmien tarpeet ja tavoitteet. Haastatteluaineistosta tulee myös ilmi automaattisen, jatkuvan tilintarkastuksen purkavan tilintarkastustyön kausiluontoisuutta. Tämä parantaa tilintarkastuksen ”tehokkuutta”, eli toisin sanoen tilintarkastuksen laatua. Seuraavana esiteltyinä olennaisimmat tulokset.

*”No ei se koskaan täysin automatisoidu, että aina siinä on sitä tulkinnanvaraa ja jonkun pitää sen takia kuitenkin käydä se data läpi. Varmasti tällöinen tekninen työ tulee vähentymään, kuten kaikilla muillakin aloilla, sit se menee tähän enemmän että analyttiset ominaisuudet yms korostuu ja sit se, että ymmärtää sitä liiketoimintaa ja ymmärtää niitä kirjauksia yms. Totta kai myös se kausiluonteisuus tulee vähentymään, kun pystytään aikasemmin reagoimaan, totta kai se piikki on edelleen siellä alkuvuodessa kun tilikaudet on päättyneet mutta esim tapahtumia pystytään kattamaan jo kauden aikana hyvin ja tehokkaasti.”*

(Haastateltava 3, senior)

*”Ei, ei siis täysin (ole automatisoitumassa) tietenkään, jonkunhan se nimi pitää sinne paperiin vetää lopulta. Siinä vaiheessa siinä pitää olla kova luotto siihen, että se tarkastus on mennyt oikein ja kyllähän siinä aina jonkun verran tulee väkisinkin semmosia tulkinnanvaraisia tilanteita. Automaattisia tehtäviä pystyy kyllä automatisoimaan hyvinkin pitkälti. Niin, ja ne laskennat pystyy toteuttaan koneellisesti joo. Mut kyllä siinä pitää ihmisen tietää, et mitä se kone on laskenut ja totta kai sieltä tulee aina välillä jotain havaintoja, niinku inhimillisiä virheitä ja tällöisiä, niin kyllähän ne jää ihmisille selvitettäväks, että mitä minkä takia jotkut kirjaukset on tietyllä tavalla kirjattu.”*

*”..... Ja ne rutiinityötkin on meillä siis loppupeleissä, ne on semmosii, et ku maailma muuttuu koko ajan, myös asiakkaan ja liiketoimintaympäristö plus niiden järjestelmät muuttuu myös, et ku toi sykli on yksi vuosi, niin voi olla, että yhtenä vuotena asiakkaan järjestelmät on yhenlaiset, ja seuraavana vuonna ne on ihan täysin uudet, erilaiset. Että mitä se rutiinityö niinku on, et ehkä siinä jää tilintarkastajalle se uusien mallien kehittäminen ja analyysien laatiminen ja järjestelmien tunteminen ja tuntemus ja ymmärtäminen.”*

(Haastateltava 1, associate)

*”Ei sitä tulla ikinä kokonaan automatisoimaan. Siis se on ihan mahotonta, koska se, miten sä näät kaikki riskit siinä asiakkaas, koska joka erälle määritellään oma riski, onko se korkeempi riski vai matalampi riski ja se taas liittyy siihen asiakkaiden, ihmisten*

*tuntemiseen, onks he pätevii vai eiks he oo pätevii. Mitä pätevämpi on, sitä vähemmän virheitä on todennäköisesti... Ja sit sun täytyy tuntea sen yhtiön kontrollit, et onks siellä ollu virheitä tai poikkeamia aiempina vuosina ja sit just, et onks prosesseissa ollu muutoksia, jos on niin tarkastetaan enemmän ja on korkeempi riski, niin jo tästä riskien mieltimisestä ja suunnittelusta lähtien siihen vaaditaan ihan hirveesti tilintarkastajien omaa ajattelua ja panosta ja se ei oo mikään sellanen, et mikään kone sitä pystyis tekee. Sitten tilintarkastajia tarvitaan erityisesti siihen, näihin johdon harkintaa vaativien erien tarkastukseen, et me voidaan tehdä jotain arvonalentumistestauksia, tai siel käytetään eri näkösiä diskonttokorkoja, tulevaisuuden rahavirtoja. Kaikkii tällasii analyysei, mis pitää ottaa kantaa, et onks ne järkevii vai ei.”*

(Haastateltava 4, manager)

*”Asiakkaathan on siirtymäs kans robotteihin. Mutta sit lähinnä se, ku se tilinpäätös ei periaattees oo pelkästää numeroita, että kyl me käydää yleensä asiakkaan kanssa, että miten niillä menee ja onks jotain sellasia asioita, mitä pitäis tilinpäätöksessä näkyä. Yks ainoo tapa on se keskustelu. Ei se oo poistumassa eikä välttämättä ees vähene. Se automatisointi liittyy tähän, että meidän, no parempaa sanaa keksimättä, ni hanttihommat vähenee. Me voidaa käyttää se säästö, mitä niistä saadaa, näihin mielenkiintoisempiin asioihin. Nää keskustelut, jotta me ymmärretää paremmin, että mitä siellä liikkuu ja mihin yritys tähtää ja onks siellä jotain semmosia erilaisia yrityksen sisäisiä tai ulkopuolisia tapahtumia, mitä ei välttämättä datasta nää, mitkä sitten pitäis huomioida jossain määrin tilinpäätöksessä.”*

(Haastateltava 2, senior manager)

### **4.3 Olennaisuuden hahmottamiseen, tilintarkastajan ajattelutavan muutokseen ja ammatilliseen skeptisyyteen liittyvät tutkimustulokset**

Haastatteluista viestittyä olennaisuuden hahmottamisen ja aineistolähtöisen analyysin (esim. Richins ym.2017, 65) näkökulman tärkeyden korostuminen Big Dataa työstettäessä. Tämä viittaa niin ikään EY:n (2015b) määrittelemän Data-analyttisen ajattelutavan nousevan tärkeäksi tilintarkastajan ominaisuudeksi. Haastatteluista ilmenee ylipäätään Big Datan haastavan tilintarkastajaa ymmärtämään entistä syvällisemmin asiakkaan toimintaa ja koko tarkastelupopulaation luonnetta.

Tulosten valossa on havaittavissa, että datan monimutkaistuuessa ja analyysityökalujen monipuolistuuessa myös tilintarkastajalta vaaditaan yhä enemmän heittäytymistä mukavuusalueensa yli uudenlaisiin ajattelutapoihin uudentyyppistä dataa

prosessoidessaan. Tästä voidaan vetää johtopäätös, että tilintarkastajalla tulee olla vahva tiedon epäselvyyden sietokyky hänen prosessoidessaan Big Dataa.

*”No ehkä varmaan helpoin ois lähteä siitä, tavallaan ajattelutavan muutoksesta, että päästään pois jostain satunnaisotantoihin perustuvista testausmenetelmistä sinne, että meillä onkin koko populaatio käytössä ja pystytään muodostamaan kokonaiskuvia. Ennen vanhaan on riittänyt se, että tarkastaa X määrä myyntilaskuja, nyt se ei välttämättä riitäkään tai tavallaan sun pitää silti ymmärtää koko populaatio ja yhtiön toiminnan luonne ja se haastaa. Se haastaa tilintarkastajaa ja mun mielestä se ei oo ollenkaan huono juttu. Meidän ei tarvii tällöin tietyillä testausmäärillä perustella sitä, että on tehty tarvittavat toimenpiteet, vaan pystytään keskittymään mielenkiintoisimpiin asioihin. Se kuitenkin vaatii jotain, että saat odotukset ja johtopäätökset työpaperille kirjoitettua. Ei päästä kyllä helpolla.”*

(Haastateltava 6, manager)

*”Eryteisesti tarvitaan monimutkaisten järjestelmien ymmärrystä. kyllä ne asiakasyritykset enimmäkseen lopulta kasvaa suuremmiks ja suuremmiks kuitenkin, et siinä mielessä järjestelmäkin muuttuu monimutkaisemmaksi, et kyl se vaatii nykyisin enemmän monimutkaisten asioiden ymmärtämistä isoimmissa yrityksissä ainakin.”*

(Haastateltava 1, associate)

*”Enemmän pitäis ymmärtää sitä kokonaisuutta, että siihen toi kyl ajaa. Jos se on aikasemmin ollu vaan sitä että tehdään joku otos ja sitten pyydetään otoksesta materiaalit ja käydään se materiaali läpi, niin nykyään enemmän pitää kiinnittää harkintaa siihen otoksen tekemiseen että mitkä me sieltä valitaan sieltä populaatiosta ja miten me perustellaan se että mitkä me ollaan valittu. se tarkoittaa myös sitä että sillä tilintarkastajalle ketä tekee sitä substanssityötä nyt sen Big Datan avulla niin on kyllä hyvä olla hyvä tietämys siitä asiakkaasta ja totta kai myös, esimerkiksi jos käydään vaikka kirjausketjua läpi vaikka sieltä pääkirjalta niin sit pitää olla kyllä tietämys niistä asiakkaan prosesseista ja miten ne kirjaukset sieltä kulkee että muuten sieltä on vaikea tulkita sitä, että kyllä tää (Big Data) muuttaa sitä.”*

(Haastateltava 5, manager)

*”Se on se, kyl se datan ymmärtäminen on, mitä dataa siellä on, mitä kaikkea sillä voi ylipäätään tehdä ja työkalujenkin käyttö, että millä niitä analysoidaan. Kyllä se on varmasti se, mitä nyt vaaditaan.”*

(Haastateltava 2, senior manager)

Seuraava, haastateltavan 1. kommentti antaa viitteitä tilintarkastajan analyysin näkökulman olevan muutoksessa aiempaa aineistolähtöisemmäksi uusien datan muotojen lisääntyessä. Aineistolähtöisessä analyysissä pyritään Richins ym. (2017, 65) mukaan olematta luomatta ennakkohypoteeseja uudessa tilanteessa, vaan hypoteesit kehittyvät vasta tilanteen mukana asianosaisen arvioidessa ilman ennakkokäsityksiä uudenlaista tietoa.

*”Tulis pitää silmät auki, et koko ajan tulee kuitenkin uudenlaisia järjestelmiä ja järjestelmät muuttuu asiakkailta, niin siellä saattaa yhtäkkiä tulla uusia mahdollisuuksia saada uudenlaista dataa sieltä ulos. Siinä just se, tommosista asioista pitäis tilintarkastajan olla hereillä mun mielestä, et jos yhtäkkiä onkin mahdollisuus testata se koko populaatio, niin sitten edetä siihen suuntaan. Tilintarkastajan tulis myös olla kiinnostunut siitä datasta mitä sieltä on mahdollista saada, jos siitä mahdollisesti pystyis vääntämään uutta analyysia, kenties säästää jopa vähän aikaa siinä samalla.”*

(Haastateltava 1, associate)

Seuraavaksi esiteltävän, haastateltavan 4. kommentin perusteella on vahvasti tulkittavissa tilintarkastajan altistuvan Big Dataa käsitellessään Knapp ja Knapp (2012, 41 – 42) tutkimuksen mukaisille kognitiivisille vinoumille, mikäli hän ei ole valmis muuttamaan asennettaan ja ymmärtämään suuren datamäärän taakse kätkeytyviä syy-seuraussuhteita sekä datan sisältämää sanomaa ylipäättään. Tämän johdosta Big Datan parissa työskentelevällä tilintarkastajalla tulee olla uutta tietoa aktiivisesti vastaanottavainen ajattelumaailma ja uusiin tilanteisiin mukautuva asenne.

*”Et pystyy just hahmottaan niit kokonaisuuksii, se on tosi tärkeä. On se ollu kyl tähän astikin tärkeä. Mut ehkä sen lisäksi kaikist eniten myös se asennepuoli, et sä haluat ymmärtää sen koko datan ja ne kaikki syy-seuraussuhteet siellä, sen sijaan et sä katot vaan yksittäisii tapahtumii, niin ehkä se koko ajattelumaailma, mikä tällä alalla on ollu, niin sen täytyy muuttuu.”*

(Haastateltava 4, manager)

Haastateltava 6. tuo seuraavana esiin mielenkiintoisen näkökulman, jonka mukaan Big Datan luoma haasteellisempi ajattelumalli tilintarkastuksen lähtökohtana parantaa lopulta tilintarkastuksen laatua. Tämä johtuu hänen aiemmin kuvaamastaan tilintarkastajaan kohdistuvasta haastavuudesta, joka pakottaa hänet ymmärtämään syvällisemmin asiakkaan koko datapopulaatiota ja liiketoiminnan luonnetta.

*”No nyt kun se (Big Data) haastaa tilintarkastajaa niin siinä ehkä tällaisena kolikon positiivisena puolena on myös se, että laatu parantuu.”* (Haastateltava 6, manager)

Eräät haastateltavat toteavat tilintarkastusstandardien ja tilintarkastusmetodologian määrittävän tilintarkastajalle lähtökohtaisesti sen mihin kysymyksiin hänen tulee tarkastuksen kautta saada vastaus. Teorian mukaan tilintarkastusstandardien voidaan todeta olevan kuitenkin puutteellisia vastaamaan Big Datan luomiin aivan uudenlaisiin kysymyksiin (Ramulkan 2015, 19). Vanhanaikaisten tilintarkastusstandardien voidaan täten tulkita rajoittavan tällä hetkellä data-analyttisen ajattelutavan kehittymistä tilintarkastuksessa.

*”Enemmän se on ajatustavan muuttumista. Että pystytään käyttää sitä dataa lähtökohtaisesti siihen mistä ne standardit lähtee, niin pystytään käyttää sitä evidenssiä siten että se vastaa niiden kysymyksiin. ISA-standardit määrittää aika tarkasti mitä tilintarkastajan tulee tehdä, että vastataan niihin samoihin asioihin dataa hyödyntämällä, niin se on ehkä tää tärkein ajatustapa siellä taustalla.”*

(Haastateltava 3. senior)

*”Voisin sanoa, että tää (Big Data) on jo muokannut Big Four-yhteisöjen tilintarkastusmetodologiaa, että metodologia menee koko aika siihen suuntaan, että ne huomioivat jatkuvasti enemmän sitä, että tätä Big Dataa hyödynnetään tarkastuksissa. Elikkä me ollaan siihen suuntaan menossa, eli ollaan vastattu siihen. Mutta sitten nää ulkoset regulaattorit niin siellä ei oo tullu vielä selkeitä muutoksia siihen, - no ehkä puhetta on ollu, esimerkiksi Suomessa sen näkee siten että PRH, tilintarkastusvalvonta puhuu jo tästä asiasta että mitkä ovat heidän mielipiteitensä siitä, siellä on jo keskustelua. Mutta sitten ihan tuolla tilintarkastusstandardeissa se ei vielä suoranaisesti näy.”*

(Haastateltava 5, manager)

Edellisiä tuloksia yhteen vetäen, tilintarkastajalta itseltään vaaditaan ajattelutavan muutosta analyttisempään ja olennaisuuksia hahmottavampaan suuntaan Big Datan parissa työskennellessään.

Kuten seuraavista kommentteista ilmenee, kehittynyt data-analyttinen teknologia on kuitenkin tilintarkastajan puolella auttamassa olennaisten tekijöiden hahmottamisessa suuresta datamäärästä. Haastateltavat painottavat Big Data-aikakauden tilintarkastajalta lisääntyntä kiinnostusta uudenlaisten teknologioiden hyödyntämistä ja niiden avulla kehittyneempien analyysien tuottamista kohtaan.

*”Mutta se ymmärrys ja kokonaiskuvan käsitys paranee sitä kautta, kun me oikeesti nähdään koko populaatio ja saadaan siitä sellainen yleiskuva. Nythän, jos puhutaan tällasesta perinteisestä aineistotarkastuksesta mikä perustuu johonkin tietynlaiseen*

*satunnaisotantaan, niin välttämättä se otanta ei sitten kuitenkaan kerro meille ihan kaikkea siitä populaatiosta, plus -että se on sitten huomattavasti tehottomampaa. Tavallaan vaikka meillä olisi niin kuin testatut määrät olisi isojakin niin populaatioon saattaa silti jäädä sellaisia ikään kuin riskisiä tapahtumia mitä me ei pystytä tunnistamaan ja mitkä ei noudata jotain tiettyä kaavaa, mutta sitten analytiikkaa käyttäen meidän on helppo bongata nää tapahtumat.”*

(Haastateltava 6, manager)

*”No ei siinä ajattelutavassa oikeastaan ehkä niin suurta eroa. Ehkä kuitenkin, että ehkä pitää olla kiinnostuneempi varsinaisesti tämmösestä analyysipyörittelystä, mutta totta kai sitä analyttisyyttä on myöskin vaadittu myös aikasemmin kun noita työkaluja ei sinällään oo ollu käytössä. Mutta ehkä teknologialähtösempää on sitten tuo jatkossa tuo homma, että totta kai on pitänyt aikasemminkin ymmärtää eri asioiden väliset yhteydet.”*

(Haastateltava 3. senior)

Haastateltava 5. mainitsee Big Data-analytiikan kapasiteetin tulevan kuitenkin lopulta vastaan. Äärimmäisessä tapauksessa tilintarkastajan tuleekin tietää juuri ne olennaisimmat osat Big Dataa, jotka analytiikkaan on tarkoitus syöttää. Haasteena on lisäksi se, että tilintarkastajan tulee kyetä myös tulkitsemaan näitä valitsemallaan datalla luotuja analyysijään.

*” Se on tän volyymin ongelmakin, kun sitä on, niin sitä haluis jossain määrin käyttääki. Ehkä tässä on nyt sit ohjaavana tekijänä kuitenkin vaikka just nää työkalut. Sinne ei välttämättä ihan kaikkea pysty laittaakaa. Ehkä seki vähän auttaa tai ohjaa sitä tarkastajaaki menee oikeeseen suntaa Mutta sitä tietoo toki tulee, niin sitte pitäis niin sanotusti tehdä jotain johtopäätöksiäki sillä tiedolla. Niin kyllä se varmasti haastaa enemmän, että mitä näillä kaikilla tiedoilla tekee ja mikä on se pihvi. Mutta tää on, jos puhutaan taas tästä kokemuksesta.”*

(Haastateltava 5, manager)

Seuraavana havainnollistetaan, miten Big Datan vaikutukset tilintarkastajan ammatilliseen skeptisyyteen jakaa haastatteluiden perusteella osittain tilintarkastajien mielipiteitä. Haastateltavat asennoituvat ammatillisella skeptisyydellä myös Big Dataan, mutta tuloksista on tulkittavissa tilintarkastajan ammatillisen skeptisyyden olevan muutoksessa.

Erytyisesti Big Datan sisältämän sanoman verifiointin tärkeys ja mahdollinen taloudellisesta kirjanpituodatasta puuttuva tieto nousevat haastatteluiden perusteella relevanteiksi kysymyksiksi koskien tilintarkastajan ammatillista skeptisyyttä Big Data-aikakaudella. Haastateltava 3. toteaa tosin ammatillisen skeptisyyden olevan Big Data-

aikakaudellakin samankaltaista kuin aiemmin, sillä tilintarkastajan ammatillinen perusolemus perustuu ylipäättään olemassa oleviin tilintarkastusstandardeihin. Haastateltava 1. arvioi ammatillisen skeptisyyden olevan jopa vähentymässä Big Dataa käsiteltäessä tämän sisältämän kaiken kattavan tiedon läpinäkyvyyden vuoksi. Tämä mukailee Mc Kinneyn (2017, 68) kuvailemaa riskiä, jonka mukaan tilintarkastaja saattaa alkaa luottamaan aivan liian voimakkaasti Big Data-analyysin lopputuleman oletettuun totuudenmukaisuuteen. Kaikki haastateltavat toteavat kuitenkin yksimielisesti, ettei ammatillinen skeptisyys voi kuitenkaan koskaan kokonaan kadota tilintarkastajan työstä. Haastateltavan 2. mukaan tilintarkastajan ammatillinen skeptisyys on jopa lisääntymässä Big Datan myötä, kun yhä suurempi määrä dataa tulee verifioida.

*”Se skeptisyyden kohde on vaihtumassa, että nyt meidän täytyy olla hyvin tarkkoja siinä, että mitä se, ensinnäkin mitä se data meille kertoo. Kertooko se oikeita asioita. Onko se luotettavaa se data. Ja sitten toinen asia on se, että vaikkapa nyt pääkirjan kohdalla, niin aina nähdään vain se tarina mitä pääkirjaan on merkattu. Meidän täytyy sitten miettiä, että puuttuuko sieltä datasta jotain mahdollisesti.”*

(Haastateltava 6, manager)

*”Ei se poista sitä, ei tietenkään. Kyllä se on niin perusominaisuus, et se tulee kyllä aina olemaan. Kyllä tilintarkastajilla pitää kuitenkin aina lopulta olla vähän pelisilmää ja olla hereillä ja kattoo, tai ettei nyt usko ihan kaikkeen kuitenkaan. Mut ite mä lähtökohtaisesti luottaisin enemmän siihen Big Dataan tietyllä tavalla. Tietysti yksittäisiä laskuja on jotenkin helpompi ujuttaa jonnekin isoon massaan sekaan, ja jos aattelee sitten kuitenkin otospohjaista tarkastusta, niin se ei todennäköisesti tuu sieltä esille. Big dataa, tai isoja populaatioita -totta kai sinnekin on helppo ujuttaa jotain. Mutta sitä dataa on ehkä vaikeempi vääristellä. Ja et kun ne kaikki kuitenkin aina täsmäytetään johonkin lukuun.”*

(Haastateltava 1, associate)

*”Kyllä mä uskon, että se skeptismi ei ainakaa vähene. Kyllä se enemmänkin lisääntyy. Dataa mitä saadaa pitää edelleen verifioida. Toki se haaste tulee myös siinä, kun sitä dataa on enemmän, niin ehkä niitä näkökulmiakin on, että miten niitä verifioidaa. Kyl se on just tää datan määrä. Jos niitä lähteitä on varsinki useampia, ni kyl se periaattes pitäis myös kaikki jossai määrin ymmärtää. Meillä se on ehkä vähän sisäänrakennettua, että ollaa skeptisiä siihen tietoon, että mitä me saadaan. Yks mahollisesti tän haasteen taklaamiseksi, käytännössä pitää vähän sitten käyttää enemmän niitä lähteitä, mitkä on tunnustettuja tai muuten hyväksi koettuja historiaan perustuen, että tää on osunu vaikka tuolta tuleva tieto totuudenmukaseks. Tää on hyvin kokemusperäistäkin välillä.”*

(Haastateltava 2, senior manager)

*”No tää tulee aika standardipohjaisesti tää ammatillinen skeptisyys. Näähän on vaan toimenpiteitä mitä tehdään riittävän varmuuden saavuttamiseksi nää data-analytiikkahommat, että samalla tavalla näitä lukuja on aiemminkin analysoitu, ehkä työkalut yms on vähän erilaisia. periaatteessa kun datan luotettavuuteen on aina kiinnitetty huomiota että ei se sinällään tää big data muuta sitä että sitä on pitänyt ammatillisesti skeptisesti arvioida niitä lukuja yms mitä siellä nyt vaikutuksia voinu olla ja että onks siellä taustalla jotain pyrkimyksiä tai tämmösiä. Ei se perusajatus siellä sinällään muutu taustalla.”*

(Haastateltava 3, senior)

#### **4.4 Relevantin tiedon hahmottamiseen ja informaatioylikuormaan liittyvät tutkimustulokset**

Kirjallisuudessa todetaan kasvaneen tiedon määrän vaikeuttavan relevantin informaation löytämistä suuresta tietomäärästä (O'Reilly 1980, 686). Ei ole siis mitenkään yllättävää, että myös haastatteluista nousee esiin tilintarkastajien huoli relevanteimman tiedon paikantamisen hankaluudesta yhä laajenevista datamääristä. Tilintarkastajat kokevat huolta kuitenkin myös sellaista tilannetta kohtaan, jossa pelkästään relevanttia tietoa on jo liikaa. Tämä viittaa vahvasti siihen, että tilintarkastaja altistuu Waller ja Zimbelman (2003, 254) mukaiselle laimennusvaikutukselle käytännössä kahdella tavalla Big Datan parissa työskennellessään.

Haastatteluun osallistuneet tilintarkastajat ehdottavat tilintarkastajan ajattelutavan keskittämistä olennaisiin tekijöihin Big Dataa työstettäessä. Tilintarkastajan tulisi pitää kirkkaana mielessä se, mitä käsiteltävällä datalla pyritään ratkaisemaan kussakin tarkastustilanteessa. Tällöin päästään tehokkaasti käsiksi kaikista relevanteimpaan tietoon. Datan sisältämän irrelevantin tiedon suhteellisella osuudella on kuitenkin rajansa. Haastateltava 6. mainitsee nimittäin, ettei liian epätarkalla datalla tee juurikaan mitään.

*”Totta kai ensimmäisenä tulee mieleen se, että mitä jos on liian paljon aineistoa, niin sit ei osaa löytää sieltä mikä niistä on se relevantti tieto. Mut sit ku toisaalta yleensä se lähtee tarpeesta se ymmärrys, että ei semmosta epärelevanttiuteen liittyvää riskiä oikeesti välttämättä oo ihan niin isosti, et joku myyntireskontra on nyt aina vaan se myyntireskontra. Jos tarkastetaan myyntisaamisia ja myyntisaamisten ikääntymistä, niin sit me tarvitaan vaan se myyntireskontra siihen. Et kun tietää mitä haluaa, niin sitte. Vaatii tietysti ammattitaitoa siinä mielessä, et jos asiakas tyrkyttää vähän kaikenlaista aineistoa, niin sit osaa sanoa mikä on oikeaa ja mikä ei oo oikeaa aineistoo.”*

(Haastateltava 1, associate)



*”Seki on tieteenki yks, mikä pitää monesti miettiä, että vaikka data on relevanttia, mutta jos sitä on törkeän paljon, mikä vie paljon aikaa ja resursseja siitä, että sitä pystys analysoimaa, ni seki vois siinä mielessä haaste, että ehkä sitä jää jopa käyttämättäkin.”*

(Haastateltava 2, senior manager)

*”Varmaan yks ihan lähtökohta kaikessa analytiikassa pitää olla se, että sen datan pitää olla tarkkaa. Ja sitten toinen on fokus kirkkaana mielessä, eli meidän pitää oikeestaan tietää, että minkälaista dataa me tarvitaan. Mutta jos dataa ei oo tarkkaa, niin sille ei juurikaan tee mitään.”*

(Haastateltava 6, manager)

Seuraavaksi ilmenee, miten käytännön tasolla haastateltava 1. ehdottaa suurten ja suhteellisesti paljon irrelevanttia tietoa sisältävien datamäärien parissa työskentelevää tilintarkastajaa pilkkomaan tehtävät pienempiin osiin. Tämä auttaa häntä sivuuttamaan mahdollisen suuren irrelevantin datamäärän takia haasteellisen työtehtävän aiheuttaman stressaavuuden. Tällöin datan linkittäminen erillisinä osina osaksi suurta kokonaiskuvaa helpottuu.

*”No ihan vaan pilkkomalla sen palasiksi ja ymmärtää sitten pala kerrallaan, totta kai se on tosi stressaava tilanne, mutta sit se vaan vaatii semmosta tiettyä taitoa, kykyä hahmottaa isoja kokonaisuuksia varmasti.”*

(Haastateltava 1, associate)

Tilintarkastajalle on myös olennaista se, että hänellä on stressaavasta tilanteesta huolimatta kirkkaana mielessään, millaista tietoa suuresta datamäärästä ollaan kulloinkin etsimässä. Tällöin irrelevantti data on helpompaa sivuuttaa suuresta tietomäärästä. Tässä tilanteessa tilintarkastaja ei ”huku” valtavaan datamäärään tutkiessaan tarjolla olevan Big Datan ilmentämää poikkeamien valikoimaa.

*”Ainahan sitä pyritään löytää se relevantti data, että on se sitten big data tai mitä muuta tahansa dataa niin pyritään sitten käyttää mikä on tilintarkastuksen kannalta olennaista. Menee aika paljon case-by-case että missä muodossa se tieto on. Voi olla helpompaa ja totta kai siellä nyt on enemmän mahdollisuuksia löytää ne relevantit asiat sieltä on big datan avulla. Toisaalta sit kun ääretön määrä tietoo niin sit sen olennaisen tunnistaminen voi olla haastavampaakin.”*

(Haastateltava 3, senior)

*”Joku vois ajatella, että tieto lisää tuskaa, mutta se on mun mielestä yksi sellanen peruspilari mikä pitää olla kunnossa eli tiedetään mitä me ollaan tekemässä ja*

*minkälaisista dataa siihen tarvitaan, jos erehdytään käymään läpi ihan meille irrelevanttia dataa niin kyllähän se kuormittaa tosi paljon. Eli se fokus täytyy olla kirkkaana mielessä, että mitä mitä ollaan tekemässä.”*

(Haastateltava 6, manager)

Tutkimuskirjallisuuden valossa tilintarkastajien todetaan kokevan informaatioylikuormaa heidän työskennellessään suurten tietomäärien parissa (Kraheil & Titera 2015, 417). Haastateltava 6. mainitsee tämän johtuvan vakiintuneiden toimintatapojen puuttumisesta kohdata poikkeuksellisen suuri määrä tietoa.

*”Joo. Onhan se riskinä (informaatioylikuorma). Kyllä varmaan jonkun verran sorrutaan siihen, että yritetään käydä läpi sellaista dataa mitä ei oikeesti tarvitsisi. Se varmaan johtuu tällä hetkellä vielä siitä, että tää on aika uusi juttu ja ei oo olemassa mitään vakiintuneita toimintatapoja...”*

(Haastateltava 6, manager)

Tätä tutkielmaa varten haastatellut haastateltavat kuitenkin mainitsevat tehokkaan data-analytiikan olevan suora vastaus suurten tietomäärien aiheuttamiin haasteisiin tilintarkastustyössä. Kehittynyt data-analytiikka helpottaa tilintarkastajaa suodattamaan hyvinkin suuresta datamassasta olennaisimman tiedon.

*”Meillä vastauksena tähän on se, että on kehitetty semmoset tehokkaat työkalut mitkä tukee sen tiedon käsittelyä. Eli mun mielestä se on se vastaus tähän, että pitää olla oikeesti järkevät työkalut ja tekniikka, että millä sitä dataa käsitellään että sieltä pystytään oikeesti havaitsemaan ne tarpeelliset poikkeamat ja tekemään ne riittävät analyysit siihen. Ja sitten kun nää työkalut on hyvät niin silloin siellä ei oo enää mitään väliä että kuinka iso se datamäärä oikeesti on, että onks siellä miljoona riviä vai 15 miljoonaa riviä, sillä ei oo siinä vaiheessa merkitystä. Mutta korostan vielä, että se on tärkeä että ne käsittelyyn tehdyt työkalut on sitten tehokkaita ja järkeviä, muuten se homma ei pelaa.”*

(Haastateltava 5, manager)

*”Tavallaan vaikka meillä olisi niin kuin testatut määrät olisi isojakin niin populaatioon saattaa silti jäädä sellaisia ikään kuin riskisiä tapahtumia mitä me ei pystytä tunnistamaan mitkä ei noudata jotain tiettyä kaavaa, mutta sitten analytiikkaa käyttäen meidän on helppo bongata suuresta datamassasta nää tärkeimmät tapahtumat ja sitten riskiarvioon perustuen niin keskittyä niihin ja jättää tää iso osa jotain tiettyä kaavaa noudattava populaatio sitten vähän vähemmälle tarkastelulle.”*

(Haastateltava 6, manager)

Vaikka kehittyneet data-analyttiset menetelmät auttavat suurten datamäärien käsittelyä, haastatteluaineiston perusteella on havaittavissa tilintarkastajien olevan hyvin yksimielisiä Big Datan olevan tästäkin huolimatta informaatioylikuorman lähde. Aineistosta ilmenee, että sinällään datan lukumäärä ei ole ainoa tekijä, joka suoraan aiheuttaa informaatioylikuormaa tilintarkastajalle. Informaatioylikuormaa aiheutuu tilintarkastajalle haastateltavan 1. mukaisesti Big Data analytiikan tuotoksina tulevien suurten tietomäärien yhteenvedojen kautta. Tilintarkastajan tulee tehdä nopeasti päätelmiä suurten datamäärien taakse piiloutuvista taustatekijöistä tarkastellessaan data-analytiikan tuottamia tuloksia. Haastateltavan 1. mukaan data-analytiikka tekee analyyseistä ketteriä ja ajallisesti tehokkaita, jolloin tilintarkastaja saattaa käsitellä useampia valtavia analyyssi-yhteenvedoja päivittäin. Tämä altistaa tilintarkastajaa informaatioylikuormalle.

*”Ehdottomasti tätä ilmenee (Tilintarkastajan altistumista informaatioylikuormalle suuria tietomääriä käsitellessään) Tehokkuuden ja kaiken muun kautta ihmisillä on yhtäkkiä yhä enemmän mahdollisuuksia toteuttaa lukumääräisesti suurempi määrä erilaisia tehtäviä työpäivän aikana, niin ensinnäkin tilintarkastuksessa se, et hyppää päivän aikana, jos puhutaan automaattisista analyyseistä, niin se, et pystyy siirtämään sen ajatusmaailman yhen asiakkaan jostain analyysistä siihen seuraavan asiakkaan toiseen analyysiin, ja hyppiin päivän aikana useammassa eri tämmöisessä jutuissa, et mä uskon et se tulee oleen aivoille esimerkiksi raskasta tietyllä tavalla. Uppoutua erilaisiin maailmoihin periaatteessa päivän aikana tietyllä tavalla, ihan erilaiseen yritykseen, ihan erilaiseen analyysiin, ja sit tietysti se, et jos ne järjestelmät ei toimikaan, niin sittenhän se koko juttu hajoaa sit täysin käsiin tietysti, et siinä on se riski. Koko populaation testaaminen ei sinällään lähtökohtaisesti taas välttämättä aiheuta tommosta. Ei siis siinä mielessä, kuitenkin toimitaan niin olennaisuusrajojen sisällä, että jos siellä on paljon pieniä virheitä, niin todennäköisesti niille saattaa olla useammalle sama selitys, jota pystyy käyttämään, tai sitten jos ne on alle olennaisuuden rajojen, jotain tosi tosi pieniä virheellisyyksiä, sittenhän niitä ei tarvi huomioida siis ollenkaan. Et se, et sieltä poimii ne isoimmat.”*

(Haastateltava 1, associate)

Edellä käydyn mukaisesti haastateltavat tilintarkastajat kertovat data-analytiikan olevan vastaus olennaisen tiedon löytämiseen suuresta datamäärästä. Seuraavasta vastauksesta kuitenkin ilmenee, miten erittäin suurten tietomäärien kohtaaminen aiheuttaa pahimmillaan tilintarkastajassa turhautuneisuutta ja mahdollisesti myös lamaantuneisuutta. Raskas informaatioylikuorma voi erään haastateltavan mukaan johtaa nimittäin jopa siihen, ettei data-analytiikkaakaan pystytä hyödyntämään suuren tietomäärän analysoinnissa parhaalla mahdollisella tavalla. Tämä on ilmeistä varsinkin silloin, kun tilintarkastajan oma osaaminen tämän analytiikan hyödyntämisen suhteen ei

ole tarpeeksi riittävää. Tästä voidaan tulkita Big Data-analytiikan olevan tilintarkastajan kannalta stressaava tekijä, mikäli hänen oma osaamisensa ei ole riittävää tarkastustoimenpiteessä vaaditun analyttisen työkalun hyödyntämiseen. Tällöin tilintarkastajan kokema stressi on kohonnut Choo (1986, 25) tutkimusta mukaillen niin korkealle tasolle, että hänen työsuorituksensa data-analytiikan mukaan ottamisesta hyötyvässä työtehtävässä laskee olennaisesti data-analytiikan jäädessä pois toimenpiteestä. Tämän voidaan tulkita johtavan tilintarkastustoimenpiteen laadun alenemiseen.

*”Kyllähän se (kuormittavan suuri tietomäärä) saattaa aiheuttaa turhautuneisuutta ja loppupeleissä sitten sitä, että niitä hyötyjä mitä analytiikalla voidaan saavuttaa, niin niitä ei osata tuoda esiin tai siinä vaiheessa, jos se, vaikka joku työkalu, niin jos me ei osata sitä käyttää, niin hyvin helposti saattaa analytiikka jäädä meidän dokumentaation ulkopuolelle ja sinne pöytälaatikkoon, kun niitä toimintatapoja ei ole.”*

(Haastateltava 6, manager)

Kuten jo tämän tutkielman teoreettisesta osiosta ilmenee, tilintarkastusstandardit eivät ole juuri pysyneet dataympäristön muutoksessa mukana (Kraheil & Titera 2015, 416 – 417). Tämän ilmiön vaikutuksista tilintarkastajan työkuormaan ilmaisee huolensa myös haastateltava 3. Myös haastateltava 1. korostaa tilintarkastusstandardeihin vastaamisen olevan tärkeää siitäkin huolimatta, että ne ovat aikaansa jäljessä. Mikäli standardit ovat luotu pienempiä otoksia silmällä pitäen, voi samoihin standardeihin vastaaminen suuremman datamäärän kanssa olla hyvinkin työlästä. Tämän voidaan todeta aiheuttavan informaatioylikuormaa.

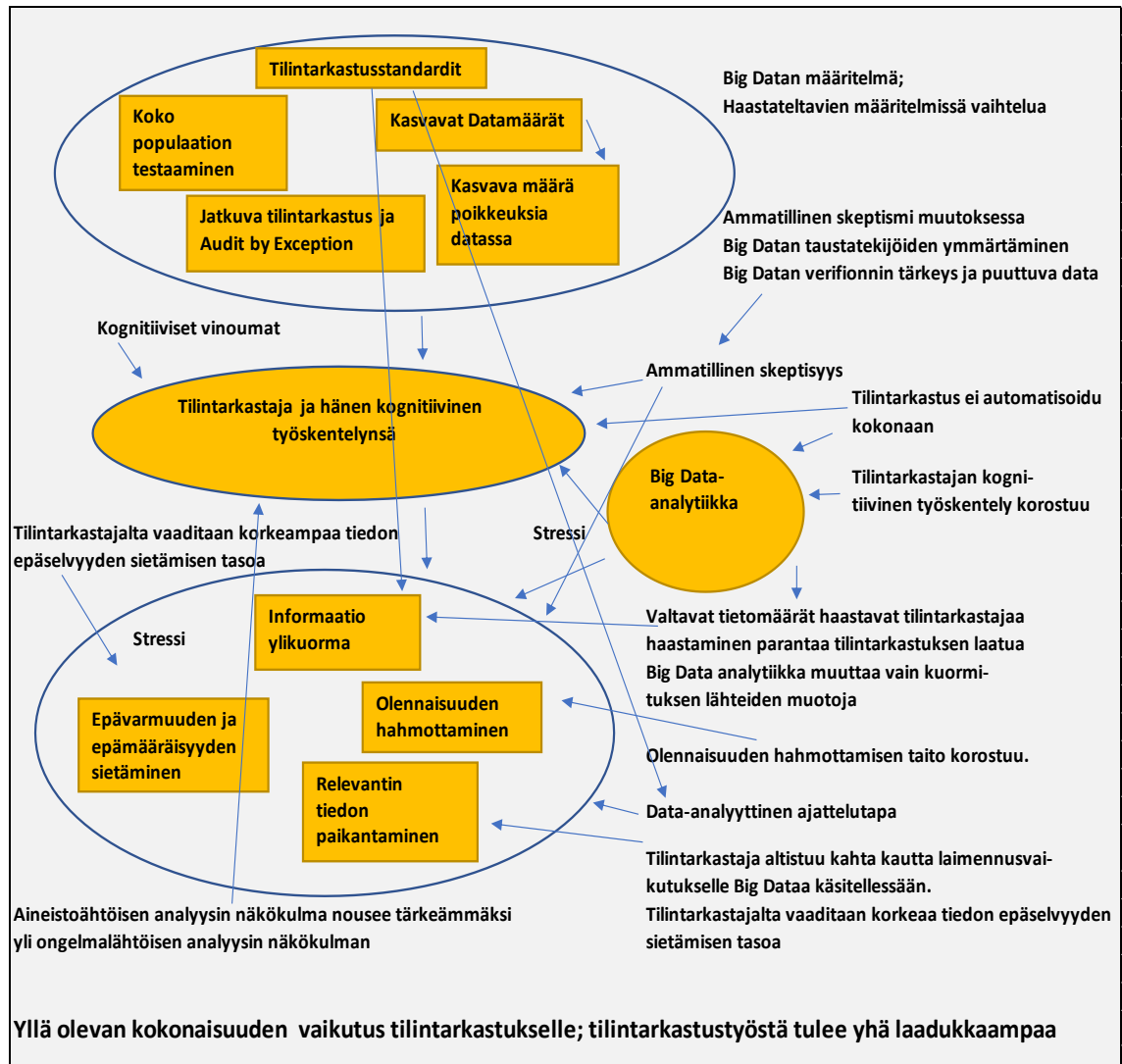
*”No periaatteessahan tässä, kun tehokkuutta tulee lisää niin totta kai se standardivaatimuksesta riippuu enemmän se työn kuormitus kuin siitä että mitä työvälinettä käytetään että missä määrin voidaan luottaa näihin ja mikä sitten katsotaan riittäväksi.”*

(Haastateltava 3, senior)

*”Tällä alalla on semmonen vallitseva trendi. Kaikki; tilinpäätös, tilintarkastusstandardit ja kaikki tämmöset on laadittu ennen Big Data -aikaa. Ne standardit pitää silti, ne kaikki standardit pitää edelleen tyydyttää ja niitä noudattaa.”*

(Haastateltava 1, associate)

#### 4.5 Big Datan ja Big Data-analytiikan vaikutukset tilintarkastajan kognitiiviseen työskentelyyn ja tilintarkastukseen – empiiristen tulosten pohdinta ja synteesi teoriaa vasten



Kuvio 9: Tämän tutkielman tulosten yhteenveto

Tässä alaluvussa tehdään pohdinnallinen yhteenveto ja tulkinta teoriaa vasten tämän tutkielman empiirisistä tuloksista, joiden tarkoituksena on tuoda lisäarvoa luvussa 3.11 esitellylle tämän tutkielman tutkimusaukolle. Palatkaamme siis luvussa 3.11 esiteltyyn tämän tutkielman tutkimusongelman viitekehykseen, mutta tällä kertaa luvun 4 tarjoamien empiiristen tulosten kanssa yllä olevan, lievästi kaoottisen näköisen kuvion johdattamana.

Tässä alaluvussa tehty pohdinta perustuu edellisessä alaluvussa esitettyihin empiriaa varten haastateltujen haastateltavien kommentille, joita myös hyödynnetään paikoin

suoraan tämän alaluvun pohdintoja tukemassa. Haastatteluaineistosta otettuja lainauksia ei ole tässä alaluvussa kuitenkaan kohdennettu enää erikseen kullekin haastateltavalle, sillä nämä kohdennukset ovat jo tehty edellisissä alaluvuissa, joissa varsinaiset empiiriset tulokset on käyty yksilöidysti läpi haastateltavittain. Tämän vuoksi tässä alaluvussa haastateltavista käytetään esimerkiksi nimityksiä ”eräs haastateltava” tai ”muuan haastateltava”. Tässä alaluvussa tärkeintä onkin tutkielman tekijän suorittama haastattelujen luomien tulosten pohdinta ja niiden peilaaminen teoreettista viitekehystä vasten, kuin haastatteluaineiston tulosten yksilöinti haastateltavittain.

Big Datalla ei ole olemassa olevan kirjallisuuden valossa yksiselitteistä määritelmää (Labrinidis & Jagadish 2012, 2032). Tämän tutkielman kirjallisuuskatsauksesta ilmenee Big Datan saavan useita määritelmiä, joista tyypillisimmät ovat erinäisen lukumäärän V-kirjaimia sisältävät Big Datan määritelmät. Tällainen määritelmä on esimerkiksi EY (2015a) määritelmä, jonka mukaan Big Datan ominaisuudet ovat; volyymi, variaatio, totuudenmukaisuus (engl. ”*veracity*”), nopeus (engl. ”*velocity*”), ja vaihtelevuus (engl. ”*variability*”). Tämän tutkielman empiirinen aineisto täydentää kirjallisuuskatsauksessa ilmenevää vakiintumatonta Big Datan määritelmää tilintarkastuksen konseptissa. Eräs haastateltavista rajasi tilintarkastuksessa hyödynnettävän Big Datan ainoastaan lähinnä asiakkaan pääkirjadataan. Osalle haastateltavista Big Data tarkoittaa kuitenkin tilintarkastuksen konseptissa hyvinkin laajaa datan määritelmää, johon lukeutuvat myös asiakkaan kirjanpidon ulkopuoliset datan lähteet. Tämä laajempi Big Datan määritelmä mukailee Chen ym. (2013, 157) määritelmää, jonka mukaan Big Dataan sisältyvät kaikki tavanomaisten nykyaikaisten tietojärjestelmien varastointi- ja käsittelykapasiteetin ylittävät datamäärät. Erään haastateltavan vastaus mukailee hyvin pitkälle edellä kuvattua Chen ym. (2013, 157) laajaa Big Datan määritelmää; *”Oikeestaan se yks määrittely on, se on isoo, se ei välttämättä oo kauheen järjestelmällistä, sitä on niin sanotusti saatavilla helposti ja enemmän semmonen tietovarasto, joka ei oo sillai niin aidattu”*.

Vaikka haastateltavat vaikuttavat olevan pääsääntöisesti erittäin selvillä Big Datan eri lähteistä, huomion arvoista kuitenkin on, ettei yhdessäkään haastattelussa tuotu kuitenkaan esille IoT-dataa. IoT-data on Brown-Liburd & Vasarhelyi (2015, 3) mukaan hyvin uusi ja vielä suuntaansa hakeva Big Datan muoto.

Tämän tutkielman empiirinen aineisto tukee Richins ym. (2017, 72) käsitystä siitä, ettei tilintarkastajan työ ole katoamassa automatisaation myötä, vaan sen voidaan todeta olevan enemmänkin jatkuvassa muutoksessa kognitiivisesti haastavampaan suuntaan. Kuten Elbannan (2006, 484) kuvailee laskentatoimen raporttien olevan loppujen lopuksi ihmisten ajattelutyön tuotteita, myös tämän tutkielman empiiriset tulokset viittaavat tilintarkastajaa tarvittavan vielä pitkään tarkastamaan inhimillisille virheille alttiita laskentatoimen ja kirjanpidon raportteja. Myös esimerkiksi muuan haastateltavan vastaus kiteyttää, miten koneet eivät kykene korvaamaan Richins ym. (2017, 72 – 73) mukaisesti tilintarkastajalle kuuluvia sosiaalista älykkyyttä vaativia tehtäviä; *”Ei sitä tulla ikinä*

*kokonaan automatisoimaan. Siis se on ihan mahotonta, koska se, miten sä näät kaikki riskit siinä asiakkaas, koska joka erälle määritellään oma riski, onko se korkeempi riski vai matalampi riski ja se taas liittyy siihen asiakkaiden, ihmisten tuntemiseen, onks he päteviä vai eiks he oo päteviä. Mitä pätevämpi on, sitä vähemmän virheitä on todennäköisesti”.* Edellisestä vastauksesta on tulkittavissa tilintarkastajan kognitiivisten ja henkisten valmiuksien nostavan tärkeyttään Big Data-aikakauden tilintarkastustyössä.

Empiirinen aineisto täydentää kirjallisuuskatsauksessa läpi käytyä teoriaa sen osalta, että tilintarkastajat todella kokevat tällä ajanhetkellä EY (2015b) määrittelemän data-analyyttisen ajattelutavan nostavan tärkeyttään tilintarkastuksessa Big Datan myötä. Tämän tutkielman empiirisen aineiston perusteella on kuitenkin havaittavissa vanhakantaisiin, pienempiin datamääriin suunnattujen tilintarkastusstandardeihin vastaamisen rajoittavan data-analyyttisen ajattelutavan kehittymistä tilintarkastuksessa. Muuan haastateltavista mainitsee tilintarkastusta säätelevien ISA-standardien määrittävän melko tarkasti mitä tilintarkastajan tulee työssään tehdä. Näin ISA-standardit luovat samalla osaltaan raamit tilintarkastajan ajatustavalle. Tämä on uudentyypinen, ja tämän tutkielman tarkoituksen kannalta merkittävä tulkinta olemassa olevaa tilintarkastajan data-analyyttistä ajattelutapaa käsittelevää teoriaa täydentämään. Whithouse (2014, 28 – 29) mukaisesti empiirinen aineisto kuitenkin tukee edellisellä tavalla teorian käsitystä olemassa olevien tilintarkastusstandardien vanhanaikaisuudesta Big Datan työstäminen silmällä pitäen.

Big Dataa työssään käsittelevän tilintarkastajan tulisi tämän tutkielman empiirisen aineiston perusteella olla teknisesti osaava ja kiinnostunut siitä, mitä kaikkea suurella, epäselkeällä ja vaihtelevalla datamäärällä voidaan saada aikaan. Hänen tulisi sietää suurtakin tiedon epäselkeyden aiheuttamaa epävarmuutta (kuten. Brown-Liburd ym. 2015, 458), sillä Big Data sisältää relevantin tiedon ohella myös erittäin paljon irrelevanttia tietoa. Tätä kuvaa erään haastateltavan kommentista viestittyvä epävarmuus ja tilintarkastajalta selvästi vaadittava epävarmuuden sietokyky Big Datan sisältämän tiedon epärelevanttiutta kohtaan; *”Totta kai ensimmäisenä tulee mieleen se, että mitä jos on liian paljon aineistoa, niin sit ei osaa löytää sieltä mikä niistä on se relevantti tieto”*, -sekä *”Et kun tietää mitä haluaa, niin sitte. Vaatii tietysti ammattitaitoa siinä mielessä, et jos asiakas tyrkyttää vähän kaikenlaista aineistoa, niin sit osaa sanoa mikä on oikeaa ja mikä ei oo oikeaa aineistoo”.* Haastatteluista ilmenee myös Big Datalta toivottavan kuitenkin riittävää tarkkuutta. Liian epätarkalla datalla ei erään haastateltavan mukaan tee juurikaan mitään.

Tämän tutkielman tutkimusaukkoa käsittelevässä alaluvussa 3.11 todetaan tilintarkastajan kokema informaatioylikuormaa ja siihen liittyviä tekijöitä olevan syytä tutkia lisää empiirisen tutkimuksen kautta Big Datan ja Big Data-analytiikan näkökulmista. Seuraavaksi esitetään tätä täydentämään kaksi mielenkiintoista tämän tutkielman empiriasta ilmenevää tilintarkastajan kokemaan informaatioylikuormaan

liittyvää tutkimustulosta, joista toinen liittyy olemassa oleviin tilintarkastusstandardeihin vastaamiseen Big Datan avulla ja toinen Big Dataa prosessoivan analytiikan hyödyntämiseen.

Olemassa olevat vanhanaikaiset tilintarkastusstandardit ovat tämän tutkielman empiiristen tulosten valossa informaatioylikuorman lähde, mikäli niihin pyritään vastaamaan Big Datan avulla. Empiirinen aineisto tuo tässä kysymyksessä lisäarvoa olemassa olevalle teorialle (esim. Whithouse 2014, 28 – 29) sen suhteen, että pieniin otoskokokoihin suunnattujen tilintarkastusstandardien hyödyntäminen näyttää ilmenevän Big Dataa työstävän tilintarkastajan kognitiivisessa työskentelyssä pahimmillaan informaatioylikuorman lähteenä. Eräs haastateltavista toteaaakin työn kuormituksen olevan riippuvainen enemmän standardivaatimuksesta kuin siitä mitä työvälinettä tilintarkastuksessa hyödynnetään. Nykyaikaisten standardien on edellisen mukaisesti suoraviivaista tulkita täten lisäävän tilintarkastajan kokemaa Big Datan aiheuttamaa kuormitusta. Nykyaikaiset tilintarkastusstandardit eivät huomioi millään tavalla Big Datan laajuutta.

Teorian pohjalta on mahdollista olettaa Big Dataa prosessoivan data-analytiikan helpottavan kaikkia Big Datan luomia tilintarkastajan kognitiivisen työskentelyn haasteita. Esimerkiksi Appelbaum (2016, 17) mukaan tilintarkastajat ovat kiinnostuneita suuria tietomääriä prosessoivasta data-analytiikasta, sillä tämän oletetaan tehostavan tilintarkastajan omaa työtä. Kuitenkin empiirisen aineiston perusteella Big Data-analytiikan luomat tehokkaat datayhteenvedot pikemminkin kuormittavat tilintarkastajaa kognitiivisesti. Tämä tapahtuu erityisesti hänen perehtyessä useampiin suuria datamassoja käsitteleviin yhteenvetoihin ja niiden taustatekijöihin monta kertaa kerran lyhyen aikaperiodin sisällä. Tämä viittaa Big Dataa prosessoivan data-analytiikan luovan tilintarkastajalle aivan uudentyypin informaatioylikuorman lähteen. Tulos on jokseenkin yhtenevä myös Issa & Kogan (2014, 209) pohdinnan kanssa, jonka mukaisesti analyttisten työkalujen tuottaman datan analyysi ja tulkinta voivat olla toisinaan hankalia tehtäviä tilintarkastajille Big Datan suurivolyymisyyden vuoksi.

Edellä mainittu tulos kuitenkin myös jossain määrin kyseenalaistaa Appelbaum (2016, 17) esittämän väitteen tilintarkastajien halusta tehostaa omaa työtään data-analyttisten menetelmien kautta, mikäli tilintarkastaja tulee niiden hyödyntämisestä hyvin kuormittuneeksi informaatioylikuorman kautta. Mahdollisesti Big dataa prosessoiva analytiikka ei siis kuitenkaan tee tilintarkastajaa täysin autuaaksi hänen yrittäessään tämän avulla helpottaa tiedon analysointia. Tästä huolimatta empiiristä aineistoa varten haastatellut tilintarkastajat korostavat myös teorian (esim. Cao ym. 2015, 424) kanssa yhteneväisesti data-analyttisten menetelmien tarjoamaa suurten datamäärien työstämistä helpottavaa vaikutusta Big Dataan perustuvassa tilintarkastustyössä.

Tämän tutkielman empiiristen tulosten perusteella Big Datan käsittely aiheuttaa tilintarkastajilla informaatioylikuorman kautta turhautumista ja mahdollisesti jopa



lamaantuneisuutta. Vastaavasti teorian valossa informaatioylikuorma aiheuttaa työsuoritusten kokonaislaadun kärsimistä (Casey 1980, 36; Lopez & Peters 2012, 141 – 142 & 162). Empiiristen tulosten valossa informaatioylikuorma saattaa tilintarkastajan pahimmillaan niinkin kuormittuneeksi ja turhautuneeksi, ettei hän kykene enää hyödyntämään edes kehittyntä analytiikkaakaan suuren datamäärän tarkastamisessa. Tämä on ilmeistä varsinkin tilanteessa, jossa hänen taitonsa hyödyntää tätä analytiikkaa on vähäänkään puutteellinen. Tämä tulos tukee teoriaa, jonka mukaisesti informaatioylikuorma aiheuttaa työsuorituksiin käytetyn ajan pitenemistä (Jacoby, 1984, 435) ja kognitiivisiin työsuorituksiin liittyvien johtopäätösten epäjohtonmukaistumista ja lopullisen yhteisymmärryksen kärsimistä työn lopputulemaa kohtaan (Chewning & Harrell 1990, 527).

Edellä kuvatussa tilanteessa data-analytiikka on myös tilintarkastajalle lähinnä ylimääräisen stressin lähde. Liika stressaavuus saa pahimmillaan tilintarkastajan jättämään analytiikan hyödyntämisen kokonaan pois tarkastustoimenpiteistään. Tämän voidaan tulkita johtavan työsuorituksen laadun alenemiseen, mikäli tilintarkastaja kuitenkin suorittaa annetun tehtävän tavalla tai toisella ilman toimenpidettä tehostavaa analytiikkaa. Edellä kuvailtu tapahtumaketju on yhtenevä teorian kanssa; Choo (1986, 25–26) tutkimuksen mukaisesti tilintarkastajan työsuorituksen laatu laskee, mikäli hänen kokemansa työperäinen stressi kasvaa liian suureksi. Edellinen tulos vastaa myös osaltaan tämän tutkielman tutkimusaukkoa käsittelevässä alaluvussa 3.11 esitettyyn kysymykseen siitä, miten Big Data ja Big Dataa prosessoiva analytiikka vaikuttavat tilintarkastajan kokemaan työn stressaavuuteen.

Teoriakirjallisuuden mukaisesti esimerkiksi O'Donnell & Perkins (2011, 274) toteavat tilintarkastajan kykeneväisyyden erottaa olennaisia kuvioita tarkastelupopulaatiosta tärkeäksi tilintarkastajan ominaisuudeksi. Empiirisen aineiston perusteella tämä on erityisen tärkeä ominaisuus myös tilintarkastajan tarkastellessa suuria tietomääriä. Eräs haastateltavista nimittäin toteaa suuren määrän irrelevanttia dataa sisältävän työtehtävän kohtaamisen voivan olevan stressaava tilanne tilintarkastajalle, jolloin tilintarkastajalla tulisi olla kirkkaana mielessä mihin kyseisellä datalla pyritään antamaan vastaus. Tämä tulos viittaa siihen, että tilintarkastajan kehittynyt olennaisuuden hahmottamisen kyky auttaa häntä selviytymään kohtaamastaan Big Datan luomasta informaatioylikuormasta. Haastateltavan toteama tilintarkastajan kokema stressi suuren irrelevantin datamäärän parissa myös täydentää Svanström (2016, 49) tutkimusta, jonka mukaan informaatioylikuorma lisää tilintarkastajan kokemaa stressiä.

Muuan haastateltava mainitsee hyväksi keinoksi suuresta irrelevantista datamäärästä selviytymiseen tämän pilkkomisen pienempiin osiin. Tämä mukailee Eppler & Mengis (2004, 335) näkemystä, jonka mukaan suuren datamäärän luokittelu on toimiva tapa selvittää suuren datamäärän aiheuttamasta informaatioylikuormasta.

Huomionarvoinen, ja tämän tutkielman tarkoituksen kannalta merkittävä tulos on erään haastateltavan kommentti tilintarkastuksen laadun paranemisesta suurten datamäärien analysoimisen aiheuttaman kognitiivisen haastavuuden kautta. Tämä on samassa linjassa Rasso (2015, 44 – 45) tutkimustulosten kanssa, jonka perusteella haasteellisempi ja itsenäisempi ajattelutapa on lopulta tilintarkastuksen lopputulemalle hedelmällinen. Sivuhuomiona edelliseen, mikäli tämä haastateltavan mainitsema kognitiivinen haastavuus ilmenee lievänä stressinä, täydentää tämä tulos Choo (1986, 25 – 26) tutkimusta, jonka mukaisesti lievä työstressi parantaa tilintarkastajan työssä suoriutumista.

Empiiristen tulosten mukaan tilintarkastuksen laatu paranee muutoinkin data-analytiikan myötä. Eräs haastateltava korostaa jatkuvan tilintarkastuksen parantavan tilintarkastuksen laatua, sillä poikkeamiin kyetään tällöin reagoimaan Vasarlehyi ja Haper (1991, 114) tutkimusta mukaillen oikea-aikaisesti.

Empiirinen aineisto ja teoria ovat samassa linjassa sen suhteen, että tilintarkastajalla on selvä riski altistua Knapp ja Knapp (2012, 41 – 42) ja Jacka (2018, 67) kuvailemille kognitiivisille vinoumille hänen työstäessään Big Dataa. Erään haastateltavan mukaan tilintarkastajalla tulee olla vastaanottavainen asenne ja halu ymmärtää kaikki datan sisältämät syy-seuraussuhteet, jotta suuresta datamäärästä saataisiin kaikki tarkastuksen kannalta tarvittava evidenssi oikealla tavalla hyödynnetyksi. Tällä voidaan tulkita olevan kognitiivisia vinoumia vaimentava vaikutus. Eräs haastateltavista kritisoikin ”rivien välistä” tilintarkastajien ajattelutapojen jämähtäneisyyttä alkavalla Big Data-aikakaudella; *”et sä haluut ymmärtää sen koko datan ja ne kaikki syy-seuraussuhteet siellä, sen sijaan et sä katot vaan yksittäisiä tapahtumia, niin ehkä se koko ajattelumaailma, mikä tällä alalla on ollu, niin sen täytyy muuttuu.”* Edellinen muuan haastateltavan kommentti viittaa vahvasti tilintarkastajalta vaadittavan mukautuvaisempaa ajattelumaailmaa Big Datan parissa työskennellessään. Jacka (2018, 67) mukaan kognitiiviset vinoumat johtuvat nimenomaan yksilölle tyypillisestä taipuvaisuudesta painottaa ja hahmottaa tulkintoja henkilökohtaisesti ominaisella tyylillä ja perusteella. Tämä voi johtaa virhearvioihin, mikäli tehtävää olisi pitänyt lähestyä toisenlaisesta näkökulmasta käsin

Tiedollisesti ennalta arvaamattomampaa ja laajaa Big Dataa käsittelevän tilintarkastajan analyysin näkökulman tulisi myös empiiristen tulosten perusteella laajeta vahvasti kohti Richins ym. (2017, 65) kuvailemaa ennakkohypoteeseja asettamatonta aineistolähtöisen analyysin näkökulmaa. Tämä ilmenee varsinkin seuraavasta muuan haastateltavan johtopäätöksestä, jonka mukaan datan muodot saavat uusia muotoja nykypäivän laajenevassa dataympäristössä nopeaan tahtiin ja tilintarkastajan tulisi olla kykeneväinen kohtaamaan tämän datan; *”Tulis pitää silmät auki, et koko ajan tulee kuitenkin uudenlaisia järjestelmiä ja järjestelmät muuttuu asiakkailta, niin siellä saattaa yhtäkkiä tulla uusia mahdollisuuksia saada uudenlaista dataa sieltä ulos”*. Edellisen

kommentin mukaiseen tilanteeseen vastaaminen vaatii väistämättä tilintarkastusnäkökulman muuttumista hyvinkin aineistoperusteiseksi. Tällöin tilintarkastaja vastaanottaa kiinnostuksella ilman ennakkokäsityksiä ennalta arvaamattomankin datapopulaation tarjoaman kaiken mahdollisen tiedon.

Empiiristen tulosten perusteella tilintarkastajan ammatillinen skeptisyys on muutoksessa Big Data-aikakaudella. Tätä ilmentävät haastateltavien ilmaiset vaihtelevat näkemykset siitä, onko ammatillinen skeptisyys vähenemässä vaiko lisääntymässä Big Datan myötä. Empiirinen aineisto täydentää ensinnäkin teoriaa sen suhteen, että Big Datan suuri määrä ja sen lävitseen luotaava totuudenmukaisuuden luonne todistetusti haastavat Mc Kinney Jr ym. (2017, 68) ja Hurtt ym. (2013, 59 ja 72) mukaisesti tilintarkastajan kykyä ylläpitää tilintarkastustyötä leimaavaa ammatillista skeptisyyttä. Tätä ammatillista skeptisyyttä vähentävää luottamusta kuvaa seuraava erään haastateltavan kommentti; ”*Mut ite mä lähtökohtaisesti luottaisin enemmän siihen Big Dataan tietyllä tavalla. Tietysti yksittäisiä laskuja on jotenkin helpompi ujuttaa jonnekin isoon massaan sekaan, ja jos aattelee sitten kuitenkin otospohjaista tarkastusta, niin se ei todennäköisesti tuu sieltä esille. Big dataa, tai isoja populaatioita -totta kai sinnekin on helppo ujuttaa jotain. Mutta sitä dataa on ehkä vaikeempi vääristellä*”. Vastakohtana edelliseen, muuan toinen haastateltava kuvaa erityisesti datan verifiointin tärkeyteen liittyvien kysymysten pikemminkin lisääntyvän puhuttaessa tilintarkastajan ammatillisesta skeptisyydestä Big Data-aikakaudella; ”*Kyllä mä uskon, että se skeptismi ei ainakaa vähene. Kyllä se enemmänkin lisääntyy. Dataa mitä saadaa pitää edelleen verifioida. Toki se haaste tulee myös siinä, kun sitä dataa on enemmän, niin ehkä niitä näkökulmiakin on, että miten niitä verifioidaa*”. Kolmas haastateltava ei totea ammatillisen skeptisyyden vähenevän eikä lisääntyvän Big Datan myötä, vaan olevan pikemminkin muutoksessa kohti kokonaisvaltaisempaa datamassakokonaisuuksia koskevaa skeptisyyttä; ”*Se skeptisyyden kohde on vaihtumassa, että nyt meidän täytyy olla hyvin tarkkoja siinä, että mitä se, ensinnäkin mitä se data meille kertoo. Kertooko se oikeita asioita*”. Edelliseen skeptisyyden kohteen muutokseen liittyy samaisen haastateltavan mukaan myös lisääntynyt skeptisyys siitä, puuttuuko täydelliseksi oletetusta Big Datasta vielä jotain olennaista tietoa.

Tiedon epäselvyyden aiheuttama laimennusvaikutus näyttää tämän tutkielman empiirisen aineiston perusteella ilmenevän tilintarkastajassa Big Dataa työstettäessä kahdella tavalla. Ensinnäkin laimennusvaikutus ilmenee Waller ja Zimbelman (2003, 254) mukaisesti relevantin tiedon löytämisen hankaluutena suuresta tietomäärästä. Tämän tutkielman empiiristen tulosten valossa Big Datan luoma aivan uudenlainen näkökulma tiedon epäselvyyden aiheuttamaan laimennusvaikutukseen piilee kuitenkin siinä, että vastaavanlaista laimennusvaikutusta ilmenee myös Big Datan sisältämän relevantin tiedon runsauden vuoksi. Tätä tilannetta kuvaa seuraava erään haastateltavan kommentti; ”*Seki on tiettenki yks, mikä pitää monesti mieltä, että vaikka data on*

*relevanttia, mutta jos sitä on törkeän paljon, mikä vie paljon aikaa ja resursseja siitä, että sitä pystys analysoimaa, ni seki vois siinä mielessä haaste, että ehkä sitä jää jopa käyttämättäkin”*. Edellä kuvatussa tilanteessa tilintarkastajan pohdittavaksi jääkin, mikä tieto jo relevantin tiedon sisällä on kaikista relevanteinta meneillään olevan tarkastustoimenpiteen kannalta. Mikäli tilintarkastaja onnistuu tässä, on hänellä mahdollisesti hallussaan kaikista relevantein tieto tarkastustoimenpiteensä suorittamista varten.

## 5 TUTKIELMAN LOPUKSI

### 5.1 Johtopäätökset

Johtopäätösosion tarkoituksena on vastata tämän tutkielman alaluvussa 1.2 esitettyihin tutkimuskysymyksiin, joista ensimmäinen edustaa tämän tutkielman päätutkimuskysymystä ja jälkimmäinen tämän tutkielman alatutkimuskysymystä. Tämän tutkielman päätutkimuskysymys on; ”*Miten Big Datan ominaisuudet ja Big Data-analytiikka vaikuttavat tilintarkastajan käytännön työhön tilintarkastajan kognitiivisesta näkökulmasta tarkasteltuna?*” Tämän tutkielman alatutkimuskysymys on; ”*Miten Big Datan ja Big Data-analytiikan vaikutukset tilintarkastajan kognitiivisessa työssä heijastuvat tilintarkastukseen kokonaisuutena?*” Seuraavaksi esitetään näihin kysymyksiin pohjustavat vastaukset ja jäljempänä todetaan suorat johtopäätökset tämän tutkielman pää -ja alatutkimuskysymyksiin.

Tilintarkastajan työnkuva on kehittyneemmän analytiikan seurauksena muuttumassa Richins ym. (2017, 72 – 74) tutkimustulosten kanssa yhteneväisesti mekaanisten työtehtävien suorittajasta yhä voimakkaammin kognitiivisesti haastavaa työtä tekevää asiakkaansa ja asiakkaansa liiketoiminnan ymmärtäjän roolia kohti, joka on kykenevä näkemään koko suureen datapopulaatioon liittyvät olennaiset asiat. Eräs haastateltavista toteaa tämän lisääntyneen haastavuuden parantavan tilintarkastuksen laatua, mikä on tämän tutkielman kannalta merkitsevä tutkimustulos. Tilintarkastajan kognitiivisen työskentelyn merkitys onnistuneessa tilintarkastustyössä korostumassa, eikä tilintarkastus suinkaan ole automatisoitumassa kuten Frey ja Osborne (2013, 69) esittävät.

Tilintarkastajalta vaaditaan lisääntyntä kiinnostusta datan syvällistä ymmärrystä ja data-analyttisiä työkaluja kohtaan. Tämä viittaa EY (2015b) kuvaileman data-analyttisen ajattelutavan tärkeyden korostumiseen Big Data-aikakauden tilintarkastustyössä. Vanhakantaisiin tilintarkastusstandardeihin vastaaminen kuitenkin rajoittaa tällä hetkellä data-analyttisen ajattelutavan kehittymistä tilintarkastuksessa. Tilintarkastajalla tulee olla halu ymmärtää koko data ja kaikki sen sisältämät olennaisimmat syy-seuraussuhteet, mikä viittaa Richins ym. (2017, 65) kuvaileman aineistolähtöisen analyysinäkökulman korostumiseen Big Dataa työstävällä tilintarkastajalla. Tämä vaatii tilintarkastajalta myös korkeampaa epäselvän tiedon sietämisen taitoa, sillä Big Data on muodoltaan usein vaihteleva elementti. Big Datan vaihtelevaluonteisuus viestittyy jo haastateltavien antamista Big Datan määritelmistä tilintarkastuksen konseptissa, jotka eivät ole yhteneviä toisiinsa nähden.

Tilintarkastaja altistuu Big Dataa käsitellessään Jacka (2018, 67) sekä Knapp ja Knapp (2012, 41 – 42) kuvailemille kognitiivisille vinoumille, mikäli hän ei ole valmis

muuttamaan asennettaan uudentyyppiselle datalle vastaanottavaisemmaksi. Hänen tulee myös ymmärtää suuren datamäärän taakse kätkeytyviä syy-seuraussuhteita ja datan sisältämää sanoma.

Tilintarkastajan ammatillinen skeptisyys on muutoksessa Big Datan myötä. Ammatillinen skeptisyys on muuttamassa Big Datan ja Big Data-analytiikan johdosta muotoaan siten, että Big Datan sanoman verifioinnin tärkeys ja mahdollinen Big Datasta puuttuva tieto ovat nousemassa ensisijaisiksi skeptisyyden kohteiksi. Big Datan verifioinnin tärkeys lisää ammatillista skeptisyyttä. Toisaalta empiiristen tulosten perusteella Big Dataa kohtaan ilmenee myös skeptisyyttä vähentävää luottamusta tämän lävitseen luotaavan totuudenmukaisuutensa vuoksi Hurtt ym. (2013, 59 ja 72) tutkimustulosten mukaisesti.

Tilintarkastaja altistuu tiedon laimennusvaikutukselle Big Datan ja Big Data analytiikan kautta kahta reittiä; ensinnäkin Waller ja Zimelman (2003, 254) mukaillen relevantin informaation paikantamisen hankaluudesta suuresta tietomäärästä. Toiseksi sitä aiheutuu myös suuren relevantin tietomäärän ylijäämästä tilanteesta, jossa pelkästään relevanttia tietoa on jo tarjolla liikaa.

Tutkielman empiirisen aineiston tarjoamat tulokset ovat kiistatta yhteneviä Saxena ja Lamest (2018, 290 – 291, 295) tutkimuksen kanssa siitä, että Big Data on informaatioylikuorman lähde tätä prosessoivalle tilintarkastajalle. Kuitenkin tämän tutkielman empiiristen tulosten perusteella myös Big Dataa hyödyntävä analytiikka muodostaa oman informaatioylikuorman lähteensä. Big Dataa hyödyntävän analytiikan tehokkuuden seurauksena informaatioylikuorman juurisyy vain vaihtuu. Mikäli tilintarkastaja siirtää huomionsa useaan kertaan saman työpäivän aikana eri asiakkaiden koko populaatioita käsitteleviin analyysihin, rasittuvat hänen aivonsa tästä tiedon paljoudesta hänen joutuessaan jokaisella kerralla syventymään kunkin analyysin sisältämiin liiketoiminnallisiin taustatekijöihin.

Vanhanaikaisiin tilintarkastusstandardeihin vastaamiseen pyrkiminen nähdään vartenotettavana informaatioylikuormaa lisäävänä tekijänä Big Dataa hyödynnettäessä tilintarkastustoimenpiteissä. Kuten esimerkiksi Whithouse (2014, 28 – 29) esittää, olemassa olevat tilintarkastusstandardit ovat lähtökohtaisesti suunnattu Big Dataa huomattavasti pienemmille tarkastettaville datamäärille.

Empiiristen tulosten mukaan Big Datan aiheuttama informaatioylikuorma on myös Svanström (2016, 49) tutkimustulosten kanssa yhteneväisesti stressiä aiheuttava tekijä tilintarkastajilla. Varsinkin Big Datan aiheuttamasta informaatioylikuormasta ylitsepääsemiseksi tilintarkastajalla tulee olla kirkkaana mielessä, millaiseen ongelmaan datalla pyritään löytämään vastaus. Hyvän olennaisuuden hahmottamisen kyvyn myötä myös Big Datan mahdollisesti sisältämän irrelevantin datan ja toisaalta myös osan relevantista datasta sivuuttaminen helpottuu.

Big Dataa prosessoiva data-analytiikkakaan ei kuitenkaan auta juuri lainkaan tiedon paljoudesta ylikuormittunutta tilintarkastajaa hänen työssään, mikäli hänen kykynsä hyödyntää tai tulkita tätä analytiikkaa on puutteellinen. Tällöin vaatimukset data-analytiikan hyödyntämisestä tietyssä tarkastustoimenpiteessä aiheuttaa tilintarkastajalle erittäin paljon ylimääräistä kognitiivista stressiä, jolloin tarkastukselle lisäarvoa tuonut analytiikka jää helposti pois koko tarkastustoimenpiteestä. Tässä tilanteessa tilintarkastajan työsuorituksen laatu laskee hänen kokemansa stressin vuoksi. Tämä on ilmiönä yhtenevä Choo (1986, 25–26) tutkimustulosten kanssa.

Johtopäätöksenä tämän tutkielman päättökysymykseen voidaan sekä tämän tutkielman teoriaan että empiiriseen aineistoon vedoten todeta Big Datan ja Big Data-analytiikan haastavan tilintarkastajaa kognitiivisesti uusin tavoin. Big Datan analysoinnin voidaan todeta muokkaavan tilintarkastajassa ilmeneviä kognitiivisia ilmiöitä, kuten ammatillista skeptisyyttä ja tiedon laimennusvaikutusta uudenslaisiin uomiin. Lisäksi Big Data muokkaa tilintarkastajan kokeman informaatioylikuorman lähteitä. Big Data-analytiikka ei suinkaan täysin poista tilintarkastajan kohtaamia kognitiivisia haasteita, vaan jopa luo osaltaan omat kognitiiviset haasteensa informaatioylikuorman ja toisinaan myös ylimääräisen stressin lähteenä. Big Datan analysointi haastaa tilintarkastajaa kognitiivisesti myös pieniin otoksiin laadittujen vanhanaikaisten tilintarkastusstandardien kautta. Tulokset viestittävät Big Datan parissa työskentelevältä tilintarkastajalta vaadittavan uudentyyppistä laajaa-alaisempaa data-analyttistä ajattelutapaa, sekä laajempaa, aineistolähtöisempää analyysin näkökulmaa. Tilintarkastajan tulee olla kykeneväinen hahmottamaan Big Datasta oleelliset tekijät ja sietää voimakkaasti vaihtelevan ja moninaisen datan sisältämää tiedon epäselkeyttä. Hänen tulee olla kykeneväinen sivuuttamaan Big Datasta irrelevantti tieto, mutta toisinaan myös osa relevantista tiedosta Big Datan sisältämän tiedon runsauden vuoksi. Edellisessä onnistuessaan tilintarkastajalla on hallussaan kaikista relevantein tieto.

Johtopäätöksenä tämän tutkielman alatutkimuskysymykseen tilintarkastuksen laatu kuitenkin kokonaisuutena paranee Big Datan prosessoinnin aiheuttaman lisääntyneen kognitiivisen haastavuuden kautta. Tämä tapahtuu tilintarkastajan perehtyessä asiakkaansa dataan ja liiketoiminnan mekanismeihin entistä syvemmillä tasolla Big Dataa prosessoivia data-analyttisiä menetelmiä hyödyntäessään.

## **5.2 Tutkielman reliabiliteetti, validiteetti ja yleistettävyys**

Laadullista tutkimusta, jota myös tämä tutkielma edustaa, voidaan arvioida sen reliabiliteetin ja validiteetin kautta. Reliaabeli tutkimus on toistettavassa uudelleen samankaltaisin tuloksin. Tutkimuksen reliabiliteetti tarkoittaa siis, että tehdyn tutkimuksen tulokset eivät ole sattumanvaraisia. Validiteetilla sen sijaan viitataan

tutkimusmenetelmän oikeaan valintaan huomioon ottaen tutkittava kohde. (Hirsijärvi, Remes & Sajavaara 2007, 226.)

Tämän tutkielman voidaan todeta olevan varsin reliabeeli. Haastattelujen jakautuminen eri tilintarkastusyhteisöihin ja eri ranking-tason tilintarkastajille voidaan nähdä reliabiliteettia lisäävänä tekijänä aineiston ollessa hajautettua ja laadullisesti monipuolista. Mikäli tämän tutkielman sisältämä empiirinen tutkimus toistettaisiin uudelleen samoille henkilöille, vastausten lopullinen sanoma tuskin eroaisi merkittävästi tässä tutkielmassa ilmenneestä sanomasta. Tulokset olisivat todennäköisesti myös enimmäkseen samankaltaisia, mikäli haastateltaisiin eri henkilöitä samoista työtehtävistä. Mikäli tämä tutkimus toistettaisiin muutaman vuoden kuluttua, saattaisivat tulokset olla edelleen samankaltaisia mutta entistä tarkempia.

Tämän tutkielman validiteetti on hyvä. Tämä tutkielma perustuu lähtökohtaisesti haastateltavien henkilökohtaisille kokemuksille työstä, ja validein menetelmä tämän mittaukseen on haastattelututkimus. Tämän tutkielman validiteettia lisäävää tässä tutkielmassa on tutkijan valmistautuminen haastatteluun monipuolisilla kysymyksin, joiden avulla haastatteluista on saatu mahdollisimman perusteellisia. Lisäksi haastateltavan on annettu kertoa vapaasti mielipiteitään haastattelukysymyksiin, haastattelijan mukautuessa haastateltavan johtamaan keskusteluun. Haastattelurunkoa rakentaessa on lisäksi kiinnitetty huomiota siihen, ettei se varsinaisesti johdattelisi haastateltavien vastauksia tiettyyn suuntaan.

Laadullinen tutkimus ei pohjautu tilastolliseen yleistettävyyteen (Lukka, 1991, 170). Tästä huolimatta tämän tutkielman tavoitteena on kuitenkin paikantaa tutkittavista ilmiöistä elementtejä, joita olisi mahdollista havaita edelleen, mikäli samankaltainen tutkimus toistettaisiin uudelleen. Tämän johdosta tätä tutkielmaa varten on haastateltu vain sellaisia henkilöitä, joilta löytyy aitoa asiantuntijuutta tutkielman aihepiiriä kohtaan.

### **5.3 Tutkielman luoma lisäarvo ja tutkielman rajoitteet**

Mikä on tämän tutkielman tiedollinen lisäarvo Big Dataa, Big Data-analyysejä sekä tilintarkastusta käsittelevälle tutkimukselle? Entä millaisia rajoitteita tämän tutkielman tuloksiin liittyy? Näitä kahta asiaa on tärkeää tarkastella tämän tutkielman päätteeksi.

Tämän tutkielman empiiristen tulosten voidaan todeta tuottavan lisäarvoa uudenlaisten näkemystensä kautta tutkielman tutkimuskysymyksiin. Empiirisistä tuloksista viestittävät varsinkin Big Datan ja Big Data-analytiikan tilintarkastajalle aiheuttamien uudenlaisten informaatioylikuorman, kognitiivisen stressaavuuden sekä laimennusvaikutuksen, tilintarkastajalta vaaditun ajattelutavan ja ammatillisen skeptisyyden lähteiden muutokset. Nämä ovat Big Datasta puhuttaessa ajankohtaisia ja uusia ilmiöitä. Tutkielman tulokset tukevat myös uusilla tavoilla kognitiivisen näkökulman kautta



olemassa olevia lukuisia toteamuksia tilintarkastusstandardien vanhanaikaisuudesta vastata Big Datan työstämisen luomiin vaatimuksiin. Tiedollisesti merkittävä tulos on myös erään haastateltavan toteama Big Data-analytiikan aiheuttaman kognitiivisen haastavuuden luoma tilintarkastuksen laadun paraneminen. Tutkielman tuottama tieto on arvokasta esimerkiksi tilintarkastusyhteisöille uutta analyttistä teknologiaa implementoitaessa osaksi tilintarkastusta.

Tämän tutkielman tuloksia on kuitenkin tarpeellista lähestyä kriittisesti empiriaa varten haastateltujen tilintarkastajien otoksen pienen koon vuoksi. Tämän tutkielman lisäarvona voidaan lähinnä nähdä syntynyt tarve laajemman tutkimuksen teettämiselle samasta aihepiiristä. Otoksen pienuuden aiheuttamia rajoitteita tämän tutkielman tuloksiin tarkastellaan seuraavana vielä tarkemmin tämän tutkielman rajoitteiden tarkastelun yhteydessä.

Tämän tutkielman rajoitteista olennaisimmat ovat haastateltavien otoksen pienuus ja haastatteluihin varattu aika. Tutkielman rajoitteita on syytä tarkastella, jotta tiedetään paremmin mitkä rajoittavat taustatekijät ovat mahdollisesti vaikuttaneet tutkielman empiirisen tutkimuksen lopputulemaan. Haastateltavia on valikoitunut tämän tutkielman empiiriseen aineistoon vain kuusi, mikä on tämän tutkielman tulosten tulkitsemisen kannalta rajoittava tekijä. Haastateltavien antamat vastaukset ovat toisiinsa nähden osittain erilaisia, joten tutkielman empiria voisi olla kattavampi esimerkiksi yli viidentoista haastateltavan aineistolla. Tällöin haastatteluiden perusteella pystyisi mahdollisesti havaitsemaan jo tutkimustuloksissa saturaatiota, eli toistoa. Tämän kautta empiriasta olisi mahdollista kyetä luomaan kattavampia yleistyksiä. Kuuden haastateltavan populaatiolla on mahdollista saada aikaan kuitenkin jo vaihtelevuutta ja erityyppisiä näkökulmia vastauksina tutkimuskysymyksiin. Mahdolliset tämän tutkielman teemaa jatkavat tutkimukset olisivatkin syytä toteuttaa suuremmilla populaatioilla. Haastatteluihin varattu aika on toinen tämän tutkielman tuloksiin vaikuttava rajoittava tekijä. Useimmat haastatteluihin osallistuneet tilintarkastajat olivat sangen kiireisiä ja haastatteluihin saatiin tavallisesti varattua aikaa puolesta tunnista tuntiin. Mikäli haastateltavilla olisi ollut enemmän aikaa pohtia tutkimuskysymyksiä, olisivat vastaukset saattaneet olla syvällisempiä. Mahdollisissa tämän tutkielman teemaa jatkavissa tutkimuksissa olisi haastateltavien tiukat aikataulut huomioidakseen hyvä keskittyä vain laadukkaisiin haastattelukysymyksiin, jolloin alun johdatteleville kysymyksille jätettäisiin mahdollisesti vähemmän sijaa.

## LÄHDELUETTELO

- Ahmadi, M.; Dileepa, P. & Wheatley, K. K. (2016). *A SWOT analysis of Big Data*. Journal of Education for Business, Vol. 91 (5), 289 – 294.
- AICPA. (2014b). *The 2014 AICPA Survey on international trends- in forensic and valuation services*. <<https://www.aicpa.org/InterestAreas/ForensicAndValuation/Resources/PracticeManagement/DownloadableDocuments/2014-fvs-trend-survey-results.pdf>> Haettu 18.06.2018
- Albeshri, A. & Thayananthan, V. (2018). *Analytical Techniques for decision making on information security for Big Data breaches*. International Journal of Information Technology & Decision making. Vol. 17 (2), 527 – 545.
- Aldhizer, G. (2017). *Visual and text analytics: The Next Step in Forensic auditing and accounting*. CPA Journal, Vol. 87 (6), 30 – 33.
- Alles, M. G. (2015). *Drivers of the use and facilitators and obstacles of the evolution of big data*. Accounting Horizons, Vol. 29 (2), 439 – 449.
- Alles, M. G.; Kogan, A.; Vasarlehyyi, M.A. (2008). *Putting Continuous Auditing Theory into Practice: Lessons from Two Pilot Implementations*. Journal of information systems. Vol. 22 (2), 195 – 214.
- Andrew, C. (2017). *Machine Learning in the \*Big Data Age\**. CIO Insight 4/19/2017, 1-1.
- Appelbaum, D. (2016). *Securing Big Data Provenance for auditors: The Big Data Provenance Black Box as Reliable Evidence*. Journal of emerging technologies in accounting, Vol. 13 (1), 17 – 36.
- Appelbaum, D.; Kogan, A. & Vasarlehyyi, M.A. (2017). *Big data and analytics in the modern audit engagement: Research needs*. Auditing: A Journal of practice & theory. Vol. 36 (4), 1 – 27.
- Appelbaum, D.; Kogan, A.; Vasarlehyyi, M.A. & Yan, Z. (2017). *Impact of Business Analytics and Enterprise Systems on Managerial Accounting*. International Journal of Accounting Information Systems, Vol. 25, 29 – 44.
- Ashford, W. (2013). *Ernst & Young email keyword analysis identifies fraudsters*. <<https://www.computerweekly.com/news/2240175638/Ernst-Young-email-keyword-analysis-identifies-ID-fraudsters>> haettu 22.07.2018.
- Barnaghi, P.; Sheth, A. & Henson, C. (2013). *From data to actionable knowledge, big data challenges in the web of things*. IEEE Intelligent Systems, Vol. 28 (6), 6 – 11.
- Bierstaker, J. L.; Bedard, J. C. & Biggs, S. F. (1999). *The Role of Problem Representation Shifts in Auditor Decision Processes in Analytical Procedures*. Auditing: A Journal of Practice & Theory. Vol. 18 (1), 18 – 36.

- Blay, A. D. (2005). *Independence treats, litigation risk, and the auditor's decision process*. Contemporary Accounting Research, Vol 22 (4), 759 – 789.
- Brandel, M. (2008). *Information OVERLOAD: is it time to go on a data diet? (cover story)*. Computerworld, Vol. 42 (34), 18 – 23.
- Brazel, J.; Jones, K. & Zimbelman, M. (2009). *Using nonfinancial measures to assess fraud risk*. Journal of accounting research. Vol. 47 (5), 1135 –1166.
- Brown-Liburud, H.; Vasarhelyi, M.A. (2015). *Big data and audit evidence*. Journal of emerging technologies in accounting. Vol. 12, 1 – 16
- Brown-Liburud, H.; Issa, H. & Lombardi, D. (2015). *Behavioral implications of Big data's impact on audit judgement and decision making and future research directions*. Accounting horizons. Vol. 29 (2), 451 – 468
- Budner, S. (1962a). *Intolerance of ambiguity as a personality variable*. Journal of Personality, Vol. 30, issue 1, 29 – 50
- Budner, S. (1962b). *Tolerance of Ambiguity*  
<http://faculty.www.edu/dunnc3/rprnt.toleranceofambiguityscale.pdf>  
 haettu 06.08.2018
- Burtescu, C. & Grigore, G. (2011). *Sampling in audit and professional reasoning*. Agricultural management. Vol. 13 (3), 165 – 172.
- Byrnes, P. E & Pawlicki, A. R (2014). *Reimagining auditing in a wired world*. AICPA, 1 – 11.  
 <[https://pdfs.semanticscholar.org/814c/67cb3365f4e1fad1a9aa1df1a8bc55046c9.pdf?\\_ga=2.20799736.1926382291.1572173120-1933717458.1572173120](https://pdfs.semanticscholar.org/814c/67cb3365f4e1fad1a9aa1df1a8bc55046c9.pdf?_ga=2.20799736.1926382291.1572173120-1933717458.1572173120)> haettu 29.06.2018.
- Capriotti, R. J. (2014). *Big Data Bringing Big Changes to Accounting*. Pennsylvania CPA Journal. Vol. 85 (2), 1 – 3
- Casey, C. J. Jr. (1980). *Variation in accounting information load: The effect on loan officers' prediction of bankruptcy*. The Accounting Review (Tammikuu): 36–49.
- Cao, M.; Chychyla, R. & Stewart, T. (2015). *Big Data Analytics in Financial Statement Audits*. Accounting Horizons. Vol. 29 (2), 432 – 429.
- Chen, J.; Chen, Y.; Du, X.; Li, C.; Lu, J.; Zhao, S. & Zhou, X. (2013). *Big Data challenge: a data management perspective*. Frontiers of Computer Science, Vol. 7 (2), 157–164.
- Chewning, E. G. Jr. & Harrell, A.M. (1990). *The effect of information load on decision makers' cue utilization levels and decision quality in a financial distress decision task*. Accounting, Organizations and Society, 15 (6) 527 – 542.
- Choo, F. (1986). *Job stress, job performance and auditor personality characteristics*. Auditing: A Journal of practicing and theory. Vol. 5 (2), 17 – 34

- Coakley, J. R. & Brown, C. (1993). *Artificial Neural Networks Applied to Ratio Analysis in the Analytical Review Process*. International Journal of Intelligent Systems in Accounting, Finance & Management. Vol. 2 (1), 19 – 39.
- Dean, D & Webb, C. (2011). *Recovering from information overload*. McKinsey quarterly. (1), 80 – 88.
- Deloitte LLP. (2016). *Deloitte forms Alliance with Kira Systems*. Deloitte US Press Release. < <https://info.kirasystems.com/news/deloitte-forms-alliance-with-kira-systems-to-drive-the-adoption-of-artificial-intelligence-in-the-workplace>> haettu 25.06.2018.
- De melo Mendes, P. C.; Katsumi Niyama, J; Tibrucio Silva. C. A. (2018). *The perception of Auditors in the measurement of instruments Financial institutions at Fair Value in Financial institutions*. Brazilian Business Review (English Edition), Vol. 15 (4), 363 – 381.
- Dremer, J. (1973). *Cognitive Characteristics and the Perceived Importance of Information*. Accounting Review, Vol. 48 (3), 511 – 519.
- Earley, C. E. (2015). *Data analytics in auditing: Opportunities and challenges*. Business Horizons, Vol. 58 (5), 493 – 500.
- Elbannan, M. A. (2006). *The Effects of Contexts and Cognitive Characteristics on Accounting Information Processing: A Review of Behavioral Accounting and Cognitive Psychology Research*. Emerging Trends and Challenges in Information Technology Management. Vol. 1 & 2, Idea Group Inc. 482 – 485.
- Ernst & Young LLP (EY). (2016). *Building Blocks of the Future*. EY Reporting. <<https://www.ey.com/Publication/vwLUAssets/EY-building-blocks-of-the-future/%24FILE/EY-building-blocks-of-the-future.pdf>> haettu 25.06.2018.
- Ernst & Young LLP (EY). (2015a). *Big Data analytics in the audit process: mitigating risk and unlocking value*. EY Center of Board matters.
- Ernst & Young LLP (EY). (2015b). *Defining the audit of the future: An analytics mindset holds the key*. EY Academic resource center (EYARC).
- Eskola, J. & Suoranta, J. (2003). *Johdatus laadulliseen tutkimukseen*. 6. painos. Jyväskylä
- Eppler, M. & Mengis, J. (2004). *The concept of information overload: A review of literature from organization science, accounting, marketing, MIS, and related disciplines*. The Information Society, Vol. 20 (5), 325 – 344.
- Fan, J.; Han, f. & Liu, H. (2014). *Challenges of Big Data Analysis*. Natural Science Review, Vol. 1 (2), 293 – 314. <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4236847>> haettu 29. 07. 2018 .

- Finlex. (2015). *Tilintarkastuslaki 1141/2015*.  
<<https://www.finlex.fi/fi/laki/alkup/2015/20151141>> Haettu 11.08.2018.
- Frey, C.B. & Osborne, M.A. (2013). *The future of employment: how susceptible are jobs to computerization?*. Oxford Martin School. <[https://www.oxfordmartin.ox.ac.uk/downloads/academic/The\\_Future\\_of\\_Employment.pdf](https://www.oxfordmartin.ox.ac.uk/downloads/academic/The_Future_of_Employment.pdf)> haettu 23.6.2018
- Fuller, L. R. & Kaplan, S. E. (2004). *A Note about the Effect of Auditor Cognitive Style on Task Performance*. Behavioral Research In Accounting. Vol. 16, 131 – 143.
- Fullerton, L (2016). *KPMG collaborates with IBM Watson to usher in era of cognitive computing*. <<https://www.thedrum.com/news/2016/03/08/kpmg-collaborates-ibm-watson-usher-era-cognitive-computing>> haettu 25.06.2018
- Gandomi, A. & Haider, M. (2015). *Beyond the hype: Big Data concepts, methods & analytics*. International Journal of Information management, Vol. 35 (2), 137 – 144.
- Geat, K. W. & Xie, Z. (2017). *Big Data analytics a Boon for Audits*. ISCA Journal <[https://journal.isca.org.sg/2017/06/15/data-analytics-a-boon-for-auditors/pugpig\\_index.html](https://journal.isca.org.sg/2017/06/15/data-analytics-a-boon-for-auditors/pugpig_index.html)> haettu 07.08.2018.
- Gepp, A; Linnenluecke, M. K; O'Neill, T. J. & Smith, T. (2018). *Big Data techniques in auditing research and practice: Current trends and future opportunities*. Journal of Accounting Literature, Vol. 40, 102 – 115.
- Grable, J. E. & Lyons, A. C. (2018). *An Introduction to Big Data*. Journal of Financial Service Professionals. Vol. 72 (5), 17 – 20.
- Grant, R.; Shen, M.; Carswell, M. & Santhanam, R. (2010). *The Effect of Ambiguity Tolerance and User Discretion on Spatial Task performance and Display Choices*. Conference: Pacific Asia Conference on Information Systems, PACIS 2010, 9 – 12. <<https://dblp.uni-trier.de/db/conf/pacis/pacis2010.html>> Haettu 16.7.2019
- Griffith, E. E; Hammersley, J. S; Kadous, K. & Young, D. (2015). *Auditor Mindsets and Audits of Complex Estimates*. Journal of Accounting Research, Vol. 53 (1), 49 – 77.
- Grönfors, M. (1985). *Kvalitatiiviset kenttätutkimusmenetelmät*. WSOY, Juva.
- Haron, H.; Ishak, I. & Ibrahim, D. N. (2014) *Factors Influencing Ethical Judgement of Auditors in Malaysia*. Malaysian accounting review, Vol 13, (2), 47-86.
- Hasanat, B. (2018). *Big Data: An institutional Perspective on Opportunities and Challenges*. Journal of economic issues (Taylor and Francis Ltd). Vol. 52 (2), 580 – 588.

- Hirsijärvi, S.; Remes, S. & Sajavaara, P. (1997). *Tutki ja kirjoita*. 2. painos. Helsinki: Kirjayhtymä
- Hirsijärvi, S.; Remes, S. & Sajavaara, P. (2007) *Tutki ja kirjoita*. 13. painos. Keuruu: Otavan kirjapaino
- Hogart, R. M. (1991) *A Perspective on cognitive research in accounting*. The Accounting Review, Vol. 66 (2), 277 – 290
- Hoffman, V.B. & Patton, J.M. (1997). *Accountability, the Dilution Effect and Conservatism in Auditor's Fraud Judgements*. Journal of Accounting Research, Vol. 35 (2), 227 – 237.
- Horton, J.; Macve, R. & Struyven, G. (2004). Qualitative Research: Experiences in Using Semi-Structured Interviews. Teos: *The Real Life Guide to Accounting Research: A Behind-the-Scenes view of Using Qualitative Research Methods* 339 – 357. Elsevier Ltd.
- Hunton, J.E. & Rose, J.M. (2010). *21<sup>st</sup> Century Auditing: Advancing Decision Support Systems to Achieve Continuous Auditing*. Accounting horizons, Vol. 24 (2), 297 – 312.
- Hurtt, K. R.; Brown-Libur, H.; Earley, C. E & Krishnamoorthy, G. (2013). *Research on Auditor Professional Skepticism: Literature Synthesis and Opportunities for future Research*. Auditing: A Journal of Practice & Theory, Vol. 32, Supplement 1, 45 – 97.
- Hwang, M. I. & Lin, J.W. (1999). *Information dimension, information overload and decision quality*. Journal of Information science, Vol. 25 (3), 213 – 218.
- Igushi, H.; Abe, K.; Misawa, T.; Kimura, H. & Daido, Y. (2009). *Recognition of grouping patterns in trademarks based on the Gestalt psychology*. Electronics & Communications in Japan, Vol. 92 (10), 49 – 60.
- International Auditing and Assurance Standards Board. (2015). *Handbook of International Quality control, Auditing, Review, Other Assurance, and Related Services Pronouncements*. (2015 Edition, Volume 1). New York: International Federation of Accountants.
- Iselin, E. R. (1988). *The effects of information load and information diversity on decision quality in a structured decision task*. Accounting, Organizations and Society, Vol. 13 (2), 147 – 164.
- Issa, H. & Kogan, I. (2014). *A Predictive Ordered Logistics Regression Model as a Tool for Quality Review of Control Risk Assessment*. Journal of information systems Vol. 28 (2), 209 – 215.
- Jacka, M. (2018). *I WAS RIGHT ALL ALONG*. Internal auditor. vol. 75 (4), 67– 67.
- Jacoby, J. (1984). *Perspectives on Information Overload*. Journal of Consumer Research, Vol. 10 (4), 432 – 435.

- Jagadish, H.V.; Gehrke, J.; Labridinis, A.; Papakonstantinou, Y.; Patel, J. M.; Ramakrishnan, R. & Sahabi, C. (2014). *Big Data and its technical challenges*. Communications of the ACM. Vol. 57 (7), 86 – 94.
- Klingemann, H. D (2011). *Exploratory Data analysis* . US Berkley Statistics <<https://www.stat.berkeley.edu/~brill/Papers/EDASage.pdf>> haettu 11.08.2018.
- Knapp, M. C & Knapp, C.A. (2012). *Cognitive biases in audit engagements*. CPA Journal, Vol. 82 (6), 40 – 45.
- Kogan, A.; Alles, M. G.; Vasarlehyyi, M. A. & Jia, W. (2014). *Design and Evaluation of a Continuous Data Level Auditing System*. Auditing: A Journal of Practice & Theory. Vol 33 (4), 221 – 245
- Koonce, L. (1991). *A Cognitive Characterization of Audit analytical review*. Auditing: A Journal of Practice & Theory, Vol. 12 (2), 57 – 76.
- Koskinen, I.; Alasuutari, P. & Peltonen, T. (2005). *Laadulliset menetelmät kauppateieteissä*. Gummerrus, Jyväskylä.
- Krahel, J. & Titera, W.R. (2015). *Consequences of Big Data and formalization on accounting and auditing standards*. Accounting Horizons, Vol. 29 (2), 409 – 422.
- Krumholz, H. M. (2014). *Big Data and new knowledge in medicine: Thinking, Training and tools needed for a learning health system*. Health Affairs, Vol. 33 (7), 1163 – 1170
- Koyce, K. (2017). *the Challenges of using Big Data effectively*. 6<sup>th</sup> Annual International Conference on Enterprise Marketing and Globalization - EMG 2017, 28 – 37
- Kwon, O.; Lee, N. & Shin, B. (2014). *Data quality management, data usage experience and acquisition intention of big data analytics*. International Journal of Information management, Vol. 34 (3), 387 – 394.
- Kyunghee, Y. & Hoogduin, L. & Li, Z. (2015). *Big data as complementary audit evidence*. Accounting horizons. Vol. 29 (2), 431 – 438.
- Labrinidis, A. & Jagadish, H.V. (2012). *Challenges and opportunities with Big Data*. Proceedings of the WLDB endowment, 5 (12), Vol. 2032 – 2033.
- Libano, F.; Rech, P.; Tambara, L. & Tonfrat, J. & Kastensmidt, F. (2018). *On the Reliability of Linear Regression and Pattern Recognition Feedforward artificial Neural Networks in FPGAs*. IEEE Transactions on Nuclear Science. Vol. 65 (1), 288 – 295.
- Libby, R. & Frederick, D. M. (1990). *Experience and the Ability to Explain Audit Findings*. Journal of Accounting Research (Wiley-Blackwell), Vol. 28 (2), 348 – 367.

- Lopez, D.M. & Peters, G.F. (2012). *The effect of workload compression on audit quality?* Auditing: A Journal of Practice & Theory. Vol. 31 (4), 139 – 165.
- Lowe, D. J. & Reckers, M. J. (1997). *The Influence of Outcome Effects, Decision Aid Usage and Intolerance of Ambiguity on Evaluations of Professional Audit Judgement.* International Journal of Accounting, Vol. 1 (1), 43 – 59.
- Lukka, K. (1991). *Laskentatoimen tutkimuksen epistemologiset perusteet.* Liiketalouden aikakauskirja. Vol. 40 (2), 162 – 184.
- Masihabadi, A; Rajaei, A; Koloukhi, A.S. & Parsian, H. (2015). *Effects of stress on auditors' organizational commitment, job satisfaction, and job performance?* International journal of organizational leadership. Vol. 4 (2015), 303 – 314
- McAfee, A. & Brynjolfsson, E. (2012). *Big Data: The management revolution. (cover story).* Harvard Business Review, Vol. 90 (10), 60– 68.
- McKinney Jr, E.; Yoos, C, J. & Snead, K. (2017). *The need for 'skeptical' accountants in the era of Big Data.* Journal of Accounting education. Vol. 38, 63 – 80
- Moffit, K.C. & Vasarlehyyi, M.A. (2013). *AIS in an age of Big data.* Journal of information systems. Vol. 27 (2), 1– 19.
- Morgan, J. (2014). *A Simple Explanation of 'The internet of things'.*  
 <<https://www.forbes.com/sites/jacobmorgan/2014/05/13/simple-explanation-internet-things-that-anyone-can-understand/#29566b071d09>> haettu 04.08.2018
- Najafabadi, M. M.; Villanustre, F.; Khoshgoftaar, T.; Seliya, N.; Wald, R. & Muharemagic, E. (2015). *Deep Learning Applications and challenges in Big Data analytics.* Journal of Big Data. Vol. 2 (1), 1 – 21.
- Nelson, M. W. (2009). *A Model and Literature review of professional skepticism in auditing.* Auditing: A Journal of practice & theory, Vol 28 (2), 1 –34.
- Nelson, M. W. & Kinney Jr., W. R. (1997). *The effect of ambiguity on loss contingency reporting judgement.* Accounting Review, Vol. 72 (2), 257 – 274.
- Nunan, D. & Di Domenico, ML. (2013). *Market research and the ethics of big data.* International Journal of Market Research. Vol. 55 (4), 2 – 13.
- O'Donnell, R. (2015). *Data, analytics and your audit: what financial executives need to know.* Vol. 31 (3 & 4), 24 – 29.
- O'Donnell, E. & Perkins, J. D. (2011). *Assessing Risk with Analytical Procedures: Do Systems-Thinking Tools Help Auditors Focus on Diagnostic Patterns?* Auditing: A Journal of Practice & theory, Vol. 30 (4), 273 – 283.
- O'Reilly, C. A. (1980). *Individual and information overload in organizations: Is more Necessarily better?* Academy of management journal. Vol 23 (4), 684 –696.
- Parker, L. M. & Fogarty, T. J. (2012). *Seeing what you want to see: Perceptual biases of auditors.* Journal of management Policy & Practice. Vol. 13 (2), 11 – 25.
- Pansari, N. (2016). *Analytics-Driven audits.* Internal auditor, Vol. 73 (4), 16 – 17.



- Perlos, J. L. & Murthy, U. S. (2012). *Information Fusion in Continuous Assurance*. Journal of Information Systems, Vol 26 (2), 35 – 52.
- Peytcheva, M.; Wright, A. M. & Majoor, B. (2014). *The Impact of principles-based versus rules-based standards on auditors' motivations and evidence demands*. Behavioral Research in Accounting, Vol. 26 (2), 51 – 72.
- Pradana, A.; Salehudin, I. (2015). *Work overload and turnover intention of junior auditors in greater Jakarta, Indonesia?* The South East Asian Journal of Management. Vol. 9 (2), 108 – 124.
- Public Company Accounting Oversight Board. (2010). *Auditing Standard NO.15: Audit evidence*.  
[https://pcaobus.org/Standards/Auditing/pages/auditing\\_standard\\_15.aspx](https://pcaobus.org/Standards/Auditing/pages/auditing_standard_15.aspx)  
haettu 07.06.2018
- Public Company Accounting Oversight Board (PCAOB) 2010c. (2010). AS 2035: *Substantive Analytical Procedures*.  
<https://pcaobus.org/Standards/Auditing/Pages/AS2305.aspx> haettu 18.7.2018
- PwC. (2015). *Data driven: What students need to succeed in a rapidly changing business world*. Lontoo: PricewaterhouseCoopers LLC. <https://cpb-us-w2.wpmucdn.com/sites.gsu.edu/dist/1/1670/files/2015/08/pwc-data-driven-paper-1wdb00u.pdf> .haettu 5.7.2018
- Ramulkan, R. (2015). *How big data and analytics are transforming the audit*. FEY daily. <<http://daily.financialexecutives.org/how-big-data-and-analytics-are-transforming-the-audit/>> haettu 24.06.2018.
- Rapoport, M. (2016). *Auditing firms count on technology for backup*. Wall Street Journal (Maaliskuu 8) <<https://www.wsj.com/articles/auditing-firms-count-on-technology-for-backup-1457398380>> haettu 25.6.2018
- Rasso, J. T. (2015). *Construal instructions and professional skepticism in evaluating complex estimates*. Accounting, Organizations & Society. Vol. 46, 44 – 55.
- Redman, K. S. (2009). *A first step in audit committee effectiveness*. Director & Boards, Vol. 34, (1), 60 – 61.
- Richards, N. M. & King, J. H. (2014). *Big Data ethics*. Wake Forest Law Review, Vol 49 (2), 393 – 432.
- Richins, G; Stapleton, A.; Stratopoulos, T. C & Wong, C. (2017). *Big Data Analytics: Opportunity or Threat for the Accounting Profession?*. Journal of Information systems. Vol. 31 (3), 63 – 79.
- Rose, A. M.; Rose, J. M; Sandersson, K-A & Thibodeau, J. C. (2017). *When should Audit Firms Introduce Analyses of Big Data into the Audit Process?* Journal of Information Systems. Vol. 31 (3), 81 – 99.

- Saxena, D. & Lamest, M. (2018). *Information overload and coping strategies in the big data context: Evidence from the hospitality sector*. Journal of Information Science, Vol. 44(3), 287 – 297.
- Setiawan, W. Y. (2017) *Need for Cognition: Does it influence Professional Judgement* . Review of Integrative Business and Economics Research, Vol. 6 (1), 240 – 248.
- Sivarajah, U.; Kamal, M. M.; Irani, Z. & Weerakkody, V. (2017). *Critical analysis of Big Data challenges and analytical methods*. Journal of Business Research, Vol. 70, 263 – 286.
- Shelton, S. W. (1999). *Effect of xperience on the use of irrelevant evidence in auditor judgement*. The Accounting review, Vol. 74 (2), 217 – 224.
- Streufert, S. C. (1973). *Effects of information relevance on decision making complex environments*. Memory & Cognition, Vol. 1 (3), 224 – 228.
- Summers, S. L.; Sweeney, J. T & Wolk, C. M (2000). *Problem solving style and fit in Consulting and Auditing*. Journal of information systems, Vol. 14 (1), 1 – 15.
- Sun, T. & Vasarlehyyi, M. (2017). *Deep learning and the future of auditing: How an evolving technology could transform analysis and improve judgement* . CPA Journal. Vol. 87 (6), 24 – 29.
- Svanström, S. E. (2016). *Time pressure, Training activities and dysfunctional audtiour behaviour: Evidence from small audit firms*. International Journal of Auditing. Vol. 20 (1), 42 – 51.
- Tang, J. & Khondkar, K. (2017). *BIG DATA in Business Analytics: Implications for the Audit Profession*. Vol. 87 (6), 34 – 39.
- Tschakert, N.; Kokina, J.; Kozlowski, S. & Vasarlehyyi, M. (2016). *The next frontier n data analytics*. Journal of accountancy, Vol. 222 (2), 58 – 63.
- Thiprungsri, S. & Vasarlehyyi, M. A. (2011). *Cluster Analysis for Anomaly Detection in Accounting Data: An Audit Approach*. The International Journal of Digital Accounting Research. Vol. 11 (17), 69 – 84.
- Titera, W. R. (2013). *Updating Audit Standard-Enabling Audit Data Analysis*. Journal of Information Systems, Vol. 27 (1), 325 – 331.
- Tuomi, J. & Sarajärvi, A. (2006). *Laadullinen tutkimus ja sisällön analyysi*. Jyväskylä. Tammi.
- Vasarlehyyi, M. A; Alles, M. & Williams, K. T. (2010). *Continuous Assurance for the Now Economy*.  
<<https://pdfs.semanticscholar.org/d63e/6cb0b6df5392f2b2530e4b448674369094.pdf>> Haettu 23.07.2018.
- Vasarlehyyi, M. A. & Haper, F. B. (1991). *The Continuous Audit of online systems*. Auditing: A Journal of Practicing & Theory, Vol. 10 (1), 110 – 125.

- Vasarlehyi, M.; Kogan, A. & Tuttle, B.M (2015). *Big Data in accounting: An Overview*. Accounting Horizons. Vol. 29 (2), 381 – 396
- Waller, W. S. & Zimbelman, M. F. (2003). *A Cognitive footprint in archival data: Generalizing the dilution effect from laboratory to field settings*. Organizational behavior & human Decision Processes, Vol 91 (2), 254 – 268.
- Wang, T. & Cuthbertson, R. (2015). *Eight Issues on Audit Data Analytics We Would like Researched*. Journal of Information Systems. Vol. 29 (1), 155 – 162.
- Wedemeyer, P. (2010). *A Discussion of Auditor judgement as the critical component of audit quality- A practitioner's perspective*. International Journal of Disclosure and Governance, Vol. 7 (4), 320 – 333.
- Whithouse, T. (2014). *Auditing in the Era of Big Data*. Compliance Week, Vol. 11 (126), 28 – 67.
- Wright, W. F. (1993). *Discussion of a cognitive characterization of audit analytical review*, Auditing: A Journal of Practice and Theory. Vol. 12 (2), 79 – 81.
- Wuerfel, A. & Sikora, A. (2017). *Cognitive Technologies and the impact on audits*. New Jersey CPA. marras 2017/joulu 2017 (66), 12.
- Zabihollah, R.; Wang, J. (2017). *Relevance of Big Data to Forensic Accounting Practice and Education: Insight from China*. Annual International Conference on Accounting and Finance 2017, 103 – 109.
- Zhang, J.; Yang, X.; Appelbaum, D. (2015a) *Toward Effective Big Data Analysis in Continuous Auditing*. Accounting Horizons, Vol. 29 (2), 469 – 476.
- Zhang, X.; Hu, Y.; Xie, K.; Zhang, W.; Su, L. & Liu, M. (2015b). *An evolutionary trend reversion model for stock trading rule discovery*. Knowledge-Based Systems, Vol. 79 (toukokuu 2015), 27 – 35.

## LIITTEET

Haastattelujen pohjana käytetty haastattelulomake;

**Haastattelija;**

**Antti Bister, Turun kauppakorkeakoulu, Porin yksikkö**

**Haastattelurunko**

- Ammattinimike ja työtehtävien sisältö pääpiirteittäin
- Koulutustausta, miten kauan töissä alalla?
- Miten Big Data ja data-analyysit liittyvät työnkuvaan tai koskettavat sinua?

**Haastattelukysymys sarja 1. – Miten Big Data ja Big Data Analytiikka, sekä niiden tuottama informaatio vaikuttavat tilintarkastukseen?**

1. Miten määrittelet Big Datan?
2. Miten Big Dataa ja Big Data-analytiikkaa hyödynnetään tällä hetkellä tilintarkastuksessa? Mikä ajaa tilintarkastusta Big Datan hyödyntämisen ja Big Data analytiikan pariin?
3. Millaisia Big Data analytiikan menetelmiä on tällä hetkellä otettu käyttöön tilintarkastuksessa? Miksi juuri nämä menetelmät? Millaisia haasteita ja mahdollisuuksia näiden menetelmien hyödyntäminen luo tilintarkastajan työhön?
4. Big Datan ominaisuudet; Volyyymi, Variaatio, Veracity (totuudenmukaisuus), Velocity (nopeus) ja vaihtelevuus (Variability) luovat haasteita ja mahdollisuuksia tällaiset ominaisuudet sisältävän datan parissa työskenteleville henkilöille. Millaisia?
5. Miten Big Datan hyödyntäminen tilintarkastuksessa helpottaa väärinkäytösten havaitsemista ja asiakkaan riskien arviointia?
6. Big Data analytiikkaan ja siihen liittyvään koko populaation testaamiseen liittyy enemmän tai vähemmän tilintarkastuksen automatisaatio. Voiko tilintarkastusta tai tilintarkastajan päätöksentekoa automatisoida täysin? Miltä osin voi ja miltä osin ei voi?
7. Miten koko populaation testaamiseen siirtyminen vaikuttaa tilintarkastajan työnkuvaan? Millaisia etuja koko populaation testaamiseen liittyy? Mitä haasteita koko populaation testaamiseen liittyy? Miten koko populaation testaamiseen liittyvät False- positive – havainnot mahdollisesti vaikuttavat tilintarkastajan päätöksentekoon ja työhön, sekä miten tilintarkastaja voi toimia niiden parissa?
8. Miten jatkuva tilintarkastus muuttaa tilintarkastajan työtä? Miten jatkuva tilintarkastus vaikuttaa tilintarkastusprosessiin? Mikäli tilintarkastuksessa siirrytään enemmän kohti audit by exception-näkökulmaa, millaisia haasteita ja hyötyjä tämä toisi tilintarkastukselle? Millaisia kysymyksiä audit-by exception tyyppinen tilintarkastus ylipäätään herättää?
9. Millä tavoin Big Dataa voidaan hyödyntää nyt ja tulevaisuudessa tilintarkastusevidenssinä? Millaista lisäarvoa Big Data tuo perinteiselle tilintarkastusevidenssille, ja missä määrin perinteistä evidenssiä on mahdollista korvata Big Datalla? Miten Big Dataa on mahdollista integroida perinteisen evidenssin kanssa?
10. Miten tilintarkastaja voi varmistua asiakkaaseen nähden ulkoisen evidenssin totuudenmukaisuudesta ja relevanttiudesta?

11. Big Data analytiikkaan siirtymiseen tilintarkastuksessa liittyy myös tilintarkastuksen siirtyminen syy-seuraussuhteista korrelaatioihin ja runsaampien tilastollisten menetelmien hyödyntämiseen, sillä Big Data pohjainen evidenssi on perinteistä evidenssiä stokastisempaa ja todennäköisyyspohjaan perustuvampaa. Missä määrin tilintarkastajan päätöksenteko ja johtopäätökset voivat olla todennäköisyyspohjaisia tai määrällisiä?
12. Missä määrin tilintarkastusstandardit tukevat tällä hetkellä Big Datan ja Big Data analytiikan hyödyntämistä tilintarkastuksessa? Millä tavoin standardien tulisi muuttua nykyisestä?

#### Kysymyssarja 2 – Tilintarkastajan ajattelu, toiminta ja Big Datan prosessointi

1. Tulisiko yliopistojen ja ammattikorkeakoulujen liiketalouden ja kauppatieteiden opintosuunnitelmia muokata nykyisestä enemmän Data analytiikan ja Big Datan parissa työskentelyn vaatimukset silmälläpitäen? Millaisia kursseja tai osaamista tulisi painottaa nykyisestä?
2. Millaisia uusia ominaisuuksia tai työskentelytapoja tulevaisuuden tilintarkastajalla tulisi olla, kun perinteisestä tilintarkastuksesta siirrytään jatkuvaan tilintarkastukseen ja Big Datan hyödyntämiseen? Tulisiko tilintarkastajan ajattelutavan jollain tapaa mukautaa uudella tavalla Big Data aikakaudelle?
3. Mitä vaikutuksia Big Datalla ja Big Data analytiikan hyödyntämisellä on tai tulisi olla tilintarkastajan ammatilliselle skeptisyydelle? Mitkä piirteet Big Datassa erityisesti vaikuttavat ammatilliseen skeptisyyteen?
4. Miten Big Datan suuri volyymi vaikuttaa tilintarkastajan päätöksentekoon? Ovatko tilintarkastajat entistä alttiimpia tiedon ylikuormitukselle (information overload) Big Datan johdosta? Kuinka tilintarkastaja voi selviytyä suuren datamäärän parissa työskennellessään? Helpottaako Big Data analytiikka mahdollisesti tätä tilannetta?
5. Kuvioiden hahmottaminen (Pattern recognition) liitetään tilintarkastajalle erityisen tärkeäksi ominaisuudeksi. Miten Big Data ja Big Data analytiikka haastavat tai luovat mahdollisuuksia tilintarkastajan kuvioiden hahmottamiselle?
6. Miten tilintarkastajan päätöksenteko ja työskentely ylipäättään mahtavat muuttua mikäli hänelle annetaan työstettäväkseen poikkeuksellisen suuren määrän irrelevanttia tietoa (kuten erinäistä Big Dataa) sisältävä tapaus?
7. Vaikka jäsentymättömän (unstructured) datan kuten sosiaalisesta mediasta louhittavan datan helppo saatavuus on jossain määrin Big Data analytiikan etu, kuitenkin suuri kysymys on edelleen millaisia ovat relevantille datalle asetetut kriteerit. Vaikeutuuko/helpottuuko relevantin tiedon kerääminen ja hyödyntäminen Big Data analytiikan avulla?
8. Jäsentymätöntä Big Dataa leimaa sen epämääräinen, epäselvä, heterogeeninen niin ominaisuuksiltaan kuin lähteiltään oleva luonne ja alkuperä. Miten tämä voi vaikuttaa tilintarkastajan työskentelyyn ja päätöksentekoon?
9. Miltä tilanne tilintarkastuksessa ja Big Data analytiikan hyödyntämisessä tilintarkastuksessa näyttää 10-vuoden kuluttua?