

Revisjonsbevis generert fra Big Data og Big Data Analytics

*Hvordan sikrer revisor påliteligheten av revisjonsbevis når Big Data og
Big Data Analytics anvendes i revisjon?*

Malin Nordhus & Vårin H. Bjørnø

Veileder: Carmen Olsen

Masteroppgave i Regnskap og Revisjon

Handelshøgskolen ved UiS



Universitetet
i Stavanger

HANDELSHØGSKOLEN VED UIS
MASTEROPPGAVE

STUDIEPROGRAM:

MRRMAS, Master i Regnskap og Revisjon

ER OPPGAVEN KONFIDENSIELL? Nei

(NB! Bruk rødt skjema ved konfidensiell oppgave)

TITTEL:

Revisjonsbevis generert fra Big Data og Big Data Analytics

ENGELSK TITTEL:

Audit evidence generated from Big Data and Big Data Analytics

FORFATTER(E)

Kandidatnummer:

6011

.....

6003

.....

Navn:

Vårin Haaland Bjørnø

.....

Malin Nordhus

.....

VEILEDER:

Carmen Olsen

Forord

Denne oppgaven inngår som en del av mastergradstudiet i regnskap og revisjon ved Handelshøgskolen ved Universitetet i Stavanger. Masteroppgaven vår markerer slutten på en flott studietid for oss begge. Vi har fulgt hverandre i fem år og utviklet oss akademisk gjennom et tett samarbeid men også knyttet et nært vennskap.

Bakgrunnen for det valgte temaet har sitt utspring i vår interesse for digitalisering i revisjon. Til tross for at digitalisering i revisjonsfaget nesten har vært utelukket fra pensum ser vi at revisjonsbransjen er i utvikling mot en mer digitalisert revisjonsmetodikk. Temaet er relativt nytt, men har et potensiale til å endre bransjen betydelig. Vår interesse og ønske om å gi et nyttig bidrag til et lite utforsket område i revisjon har dermed vært et viktig utgangspunkt for arbeidet.

Vi ønsker å takke vår veileder Carmen Olsen for all støtten under arbeidet med oppgaven. Vi ønsker også å rette vår takk til våre medstudenter og “kontorgjengen” som har bidratt med gode innspill og fungert som diskusjonspartnere underveis.

Til slutt ønsker vi å takke alle informantene som har tatt seg tid til å bli intervjuet i en hektisk periode. Denne studien ville ikke vært mulig å gjennomføre uten deres bidrag.

Stavanger, 15 juni 2018

Malin Nordhus & Vårin H. Bjørnø

Sammendrag

Ved inntoget av Big Data og Big Data Analytics i revisjonsmetodikken endres arten av revisjonsbevis og nye utfordringer rundt påliteligheten av revisjonsbevis oppstår (Appelbaum, Kogan & Miklos, 2017, s.8). Formålet med vår oppgave er å klargjør Big Data, Big Data Analytics og pålitelighetskravet i henhold til ISA 500. Ved utgangspunkt i Appelbaums (2016) teoretiske rammeverk, med egenskaper som bygger opp under pålitelighetskravet til revisjonsbevis, har vi også undersøkt hvordan revisor tar stilling til de aktuelle egenskapene og dermed pålitelighetskravet når revisjonsbevis genereres fra Big Data og Big Data Analytics.

Vi har gjennomført et kvalitativt flercasestudie med intervju av fem informanter fra fire store revisjonsselskap. For å klargjøre Big Data og Big Data Analytics for revisjonsformål har vi gjennom intervjudataene utformet to revisjonsrettede definisjoner. Vi presenterer også et utvidet rammeverk for hvordan digitale revisjonsbevis sikres i dag og forslag for hvordan digitale revisjonsbevis kan sikres i fremtiden. For eksempel har vi funnet at revisor anvender koder eller skripter til datasett for å gjenskape digitale revisjonsbevis og som i sin tur skal sikre egenskapen klarhet i rammeverket.

Nøkkelord: Big Data; Big Data Analytics; revisjonsbevis; pålitelighetskrav

Innholdsfortegnelse

Figur liste	V
Tabell liste.....	V
Forkortelser	V
1. Introduksjon.....	1
2. Teori.....	4
2.1 Big Data og Big Data Analytics i revisjon	4
2.1.1 Big Data	4
2.1.2 Big Data Analytics.....	6
2.2 Revisjonsbevis i henhold til revisjonsstandard	7
2.2.1 Definisjon av revisjonsbevis	9
2.2.2 Relevans og pålitelighet	9
2.3 Rammeverk for egenskaper til revisjonsbevis	10
2.3.1 Risiko for endring	12
2.3.2 Troverdighet	12
2.3.3 Fullstendighet	13
2.3.4 Bekreftelse på attestasjoner	13
2.3.5 Brukervennlighet	13
2.3.6 Klarhet.....	14
3. Metode	15
3.1 Intervju.....	15
3.1.1 Pilottesting	16
3.1.2 Analyse av intervju og koding	17
4. Resultater.....	18

4.1 Definisjon av Big Data og Big Data Analytics i revisjon, (F1).....	18
4.1.1 Big Data	18
4.1.2 Big Data Analytics.....	20
4.2 Pålitelighetskrav til revisjonsbevis, (F2).....	22
4.2.1 Risiko for endring.....	22
4.2.2 Troverdighet.....	23
4.2.3 Fullstendighet.....	24
4.2.4 Bekreftelse på attestasjon.....	26
4.2.5 Brukervennlighet:.....	26
4.2.6 Klarhet.....	28
5. Diskusjon.....	29
6. Konklusjon	34
Litteraturliste.....	36
Vedlegg: Big Data og Big Data Analytics i Revisjon	40

Figur liste

Figur 1 Illustrasjon av revisjonsrisiko	8
---	---

Tabell liste

Tabell 1 Rammeverk for egenskaper til revisjonsbevis.....	11
Tabell 2 Oversikt over informanter	15
Tabell 3 Illustrasjon over Big Data Analytics som revisjonshandling	21
Tabell 4 Utvidet rammeverk med digitale revisjonsbevis	31
Tabell 5 Oppsummering av digitale revisjonsbevis	33

Forkortelser

ERP-system: Enterprise Resource Planning - System

IAASB: International Auditing and Assurance Standards Board

IBM: International Business Machines Corporation

ISA: International Standards on Auditing

RSS feeds: Rich Site Summary/Really Simple Syndication

XML: eXtensible Markup Language

1. Introduksjon

Innføringen av Big Data og Big Data Analytics i revisjonsbransjen byr på utfordringer som det må tas stilling til. Dagens revisjonsstandard er utdatert og mangler veiledning angående Big Data og Big Data Analytics (Cao, Chychyla, Stewart, 2015, s. 5). International Audit and Assurance Standard Board (IAASB) har utarbeidet et høringsutkast som adresserer noen av de mest utbredte utfordringene, i tillegg er standardsetterne i gang med å utarbeide en veiledning rundt temaet (IAASB, 2016). En omtalt utfordring omhandler synet på revisjonsbevis og hvordan det endres i omgivelsene med Big Data og Big Data Analytics (Brown-Liburud and Vasarhelyi, 2015). Dagens komplekse IT omgivelser og fremveksten av Big Data i revisjon har blant annet endret arten av revisjonsbevis (Appelbaum et al., 2017, s.8). Det kan tenkes at de tradisjonelle karakteristikkene som definerer revisjonsbevis ikke lenger er holdbare (Brown-Liburud and Vasarhelyi, 2015). Med Big Data er ikke mengden av data et problem, men kvaliteten av dataen kan by på utfordringer (Appelbaum et al., 2017, s.8). Som følge av utfordringene som råder i revisjonsbransjen er forskning rundt kvaliteten av dataen som revisjonsbeviset bygger på etterspurt (Appelbaum et al., 2017, s.8). Formålet med denne oppgaven er dermed å undersøke problemstillingen: *Hvordan sikrer revisor påliteligheten av revisjonsbevis når Big Data og Big Data Analytics anvendes i revisjon?*

Utforming av oppgaven er todelt hvor vi har utarbeidet to forskningsspørsmål som skal belyse sentrale aspekt og bygge opp under problemstillingen. Det første forskningsspørsmålet er: *Hvordan defineres Big Data og Big Data Analytics i revisjon?* (F1). Uten en klar definisjon av Big Data og Big Data Analytics for revisjonsformål kan det være vanskelig å ha en felles forståelse for hva som ligger i begrepene. Med utgangspunkt i generelle definisjoner har vi dermed utarbeidet forslag til definisjoner av Big Data og Big Data Analytics for revisjonsformål. Dette kan i sin tur være til hjelp for å møte utfordringene ved implementeringen av Big Data og Big Data Analytics i revisjonsmetodikken. Det andre forskningsspørsmålet er: *Hvordan sikres egenskaper til revisjonsbevis ved bruk av Big Data og Big Data Analytics?* (F2). Sett opp mot problemstillingen skal vi undersøke egenskaper som oppfyller pålitelighetskravet i henhold til ISA 500.A31 i revisjonsstandarden. Egenskapene er hentet fra et teoretiske rammeverk som tar utgangspunkt i attri-

butter fra revisjonsstandarden (Appelbaum, 2016, s.30) og skal sees i lys av revisjonsbevis generert fra Big Data og Big Data Analytics. Ved å undersøke hvilke handlinger revisor utfører for å oppfylle egenskapene i rammeverket kan vi få svar på hvordan pålitelighetskravet til revisjonsbevis blir møtt når revisor anvender Big Data og Big Data Analytics.

I oppgaven har vi anvendt et kvalitativt casedesign med intervju for å undersøke problemstillingen. Vi har intervjuet fem informanter fra fire forskjellige revisjonsselskap. Til tross for stor variasjon i stilling og erfaring hadde alle informantene kjennskap til revisjonsbransjen gjennom stilling som revisor. I tillegg var alle informantene involvert i en satsning rettet mot digitalisering og Big Data Analytics i sitt respektive revisjonsselskap. Revisjonsselskapene er kategorisert som store revisjonsselskap med betydelige markedsandeler. Selskapene er valgt ut på grunnlag av størrelse og relevans for oppgaven. I tillegg ga intervju med informanter fra forskjellige revisjonsselskap mulighet til å undersøke likheter og ulikheter på tvers av bransjen. Intervjuene ble gjennomført på forskjellige plattformer som følge av tilrettelegging for informantene. Både personlig oppmøte, telefon, Skype og andre kommunikasjonsverktøy via internett ble anvendt. Det ble også benyttet en intervjuguide med åpne spørsmål for å gi informanten mulighet til å få frem komplekse og nyanserte svar.

Først fant vi at Big Data i revisjonssammenheng kan defineres som *informasjon kjennetegnet ved høyt volum som krever nye metoder for prosessering og tilrettelegging. Dataen har som oftest liten grad av variasjon hvor det hovedsakelig brukes strukturert data som transaksjonsdata fra ERP-system. I tillegg må dataen inneha en høy grad av troverdighet for å fungere som grunnlag for bedre beslutninger, økt innsikt og optimalisering av prosesser.* Den foreslåtte definisjonen viser at volum, på lik linje med den generelle definisjonen, fortsatt er hoveddriveren. På en annen side baserer den revisjonsrettede definisjonen seg på liten grad av variasjon, men høy grad av troverdighet. Et interessant funn er at troverdighet og variasjon kan gå på kompromiss av hverandre i en Big Data sammenheng. Med andre ord kan ustrukturert data innebærer større utfordringer med å sikre troverdighet i revisjon.

Sett opp mot Big Data Analytics viser våre funn mindre forskjeller mellom den generelle og den revisjonsrettede definisjonen. Derimot er det interessant å se antydninger til at Big Data Analytics som revisjonshandling kan erstatte test av kontroller gjennom detaljtesting av hele populasjoner. Våre funn indikerer også at Big Data Analytics kan klassifiseres som alle de tre revisjonshandlingene: test av detalj, analytisk substanshandling og test av kontroll.

Våre funn har også gjort oss i stand til å utvide Appelbaums (2016) teoretiske rammeverk til å omfatte revisjonsbevis generert fra Big Data og Big Data Analytics. Funnene viser revisjonshandlinger som er ment å dekke egenskapene i rammeverket. For eksempel må revisjonsbevis kunne gjenskapes for at en annen revisor med lignende kompetanse skal være i stand til å trekke samme konklusjon og dermed sikre egenskapen klarhet. I følge våre funn kan instruksjonssett som koder eller skripter gjøre revisor i stand til å regenerere digitale revisjonsbevis. I tillegg kan det tenkes at teknologien bak blockchain og digitale signaturer vil bidra til å sikre pålitelig revisjonsbevis mer tilstrekkelig i fremtiden. I sum viser våre funn at egenskapene per i dag sikres i varierende grad, hvor majoritetene av handlingene fortsatt er basert på tradisjonell revisjonsmetodikk. Til tross for dette er verktøyene, som brukes for å gjennomføre handlingene, tilpasset et større volum av data og kan derfor anses egnet til å dekke egenskapene. Undersøkelsens utforskende natur har derimot resultert i flere kartleggende funn. Det vil i sin tur gjøre det mer utfordrende å garantere at handlingene i tilstrekkelig grad dekker egenskapene i rammeverket og dermed pålitelighetskravet.

Opgaven starter med teori sett opp mot oppgavens problemstilling og forskningsspørsmål. Dette innebærer blant annet Big Data, Big Data Analytics, revisjonsbevis og det teoretiske rammeverket tilpasset og oversatt fra Appelbaum (2016). Fremgangsmåte for metode og hvordan analysen er gjennomført blir forklart etterfulgt av en gjennomgående presentasjon av resultatene. Deretter diskuteres funnene for å få frem ulike perspektiv og meninger. Til slutt presenteres en konklusjon på oppgavens problemstilling.

2. Teori

Her blir teori som er aktuell for vår problemstilling presentert. Kapittelet starter med teori om Big Data og Big Data Analytics og kobler dette opp mot forskningsspørsmål F1. For å sette forskningsspørsmål F2 i en større sammenheng presenteres generell teori om revisjonsbevis sett opp mot revisjonsstandard. Deretter presenteres en definisjon av revisjonsbevis etterfulgt av en presisering av hva som ligger i revisjonsbevisets relevans og pålitelighet. Videre vil vi ta for oss det teoretiske rammeverket til Appelbaum (2016) hvor egenskapene blir beskrevet inngående for å få en forståelse av hva som skal til for at digitale revisjonsbevis skal ansees som pålitelig i henhold til rammeverket.

2.1 Big Data og Big Data Analytics i revisjon

Fremveksten av Big Data og Big Data Analytics har fått økende oppmerksomhet i revisjonsbransjen de siste årene (Appelbaum et al., 2017, s.4). Til tross for fremveksten ser vi at flere revisjonsselskap ikke anvender en revisjonsrettet definisjon av begrepene Big Data og Big Data Analytics, kun definisjoner utarbeidet på generelt grunnlag (Andersen & Bakkeli, 2015, s.2; Lillebekk & Lyngstad, 2015, s.2). Det er mangelen på en slik definisjon som er motivasjonen bak forskningsspørsmål F1: *Hvordan defineres Big Data og Big Data Analytics i revisjon?* Vi mener det er viktig å ha en spesifisert definisjon for å kunne plassere begrepene i revisjonsmetodikken. For å komme nærmere en slik definisjon er det først og fremst hensiktsmessig å forstå hva som menes med de generelle begrepene til Big Data og Big Data Analytics. Nedenfor har vi derfor greid ut om noen aspekter ved de generelle definisjoner av Big Data og Big Data Analytics.

2.1.1 Big Data

Big Data har blitt definert på ulikt vis opp gjennom tidene. Mest sannsynlig har definisjonen inneholdt en forklaring på ordet data, i tillegg til at dataen har foreligget i et stort volum, derav Big. Hva som ansees som “Big” har derimot variert opp gjennom tidene etter som at teknologi for datalagring har blitt stadig kraftigere og mengden lagrede data er blitt større (Brown-Liburd, H., Issa, H., & Lombardi, D., 2015, s. 452). Big Data kan bestå av datasett som er for store og komplekse til å manipulere, eller for store til å bli behandlet med standard metoder eller verktøy

(Cao et al., 2015, s. 423). I en annen kjent definisjon av Doug Laney defineres Big Data ut i fra de tre V`ene - volume, variety and velocity (Laney, 2001). Nedenfor skal vi forklare nærmere hva som ligger i Laney`s definisjon av Big Data.

Den første egenskapen er størrelsen, altså volum (Laney, 2001, s. 1). Dette er følgelig den egenskapen som preger Big Data i størst grad, og som nevnt over er størrelsen den egenskapen som mange definerer Big Data ut i fra. Selv om Big Data ofte måles ut i fra en måleenhet som terabytes kan Big Data kvantifiseres på andre måter, som for eksempel ved antall transaksjoner (Russom, 2011, s. 6).

Den andre egenskapen er variasjon (Laney, 2001, s. 2). Variasjon refererer til de mange ulike kildene som Big Data kan komme fra (Cao et al., 2015, s. 423). Noen nye typer datakilder er webkilder som logging, clickstreams og sosiale medier. Egenskapen variasjon indikerer at dataen kan være både strukturert, ustrukturert, semistrukturert eller en blanding. Tradisjonelt har data stort sett bestått av strukturert data mens Big Data oftere inneholder en komponent av ustrukturert data eller semistrukturert data, eksempelvis XML¹ og RSS² feeds. Andre typer data fra audio, video og andre enheter kan være vanskelig å kategorisere som enten strukturert, ustrukturert eller semistrukturert, men er ikke desto mindre en del av Big Data (Russom, 2011, s. 8).

Velocity er den siste egenskapen i Laney sin definisjon av Big Data. På norsk kan det oversettes til hastighet eller raskt endrede. Hastighet betyr med andre ord hyppigheten av datagenerering eller frekvensen av data som blir levert (Cao et al., 2015, s. 423). Et eksempel er strømmen av data som kommer fra en type enhet eller sensor, eksempelvis robot-produksjonsmaskiner, temperaturmålere, mikrofoner som lytter etter bevegelse i et sikkert område eller videokameraer som skanner etter et spesifikt ansikt i en folkemengde (Russom, 2011, s.7). Alle disse eksemplene generer data som blir oppdatert til stadighet. Dermed gir dataen brukeren mulighet til å være oppdatert i nåtid og dermed til å ta relevante beslutninger basert på fersk informasjon.

¹ XML er et verktøy for deling av strukturert data mellom informasjonssystemer, særlig over internett (Quin, 2016).

² RSS feeds brukes for å videreformidle utdrag av innhold fra en nettside som oppdateres jevnlig (ofte en weblogg) (Lacoma. 2017).

En siste karakteristikk av Big Data som er kommet i senere tid er veracity. IBM definerer veracity som den fjerde dimensjonen av Big Data og refererer til “the level of reliability associated with certain types of data, including truthfulness, accuracy or precision, correctness” (Schroeck, Shockley, Smart, Romero-Morales, & Tufano, 2012, s.5). Veracity kan oversettes til troverdig på norsk. Denne karakteristikken ble ikke involvert i Laney`s definisjon, men er blitt en akseptert del av Big Data i senere tid. Veracity kan omfatte blant annet nøyaktighet, fullstendighet, sannferdighet og opprinnelse av data (Schroeck et al., 2012). Til tross for utbredelsen av Big Data er opprinnelsen og behandlingen av datasettene stort sett ukjent (Taylor, Haggerty, Gresty & Hegarty, 2010, s. 306). For revisor er muligheten til å verifisere nøyaktigheten av data avgjørende for å kunne bruke informasjonen (Liao and Squicciarini, 2015). Til tross for at veracity er en sentral karakteristikk av Big Data er det gjort få forsøk på å undersøke veracity som et teoretisk fenomen, hva som ligger i fenomenet eller hvordan man kan måle det (Lukoianova, & Rubin, 2014, s. 5).

2.1.2 Big Data Analytics

På lik linje med Big Data ønsker vi også å utarbeide en definisjon av Big Data Analytics for revisjonsformål. En revisjonsrettet definisjon av Big Data Analytics kan gi en klarere formening av hva som ligger i begrepet for revisjonsformål, og dermed gjøre det lettere for revisor å anvende det i revisjonen. Nedenfor vil vi først presentere en generell definisjon av Big Data Analytics. Deretter vil vi vise til noen punkter i høringsutkastet som adresserer utfordringer knyttet opp mot Big Data Analytics i revisjonsmetodikken.

Cao et al. (2015) definerer Big Data Analytics som prosessen hvor man inspiserer, vasker, transformerer og modellerer Big Data for å oppdage og formidle nyttig informasjon og mønstre, foreslå konklusjoner og støtte opp under beslutningstaking (Cao et al 2015). Ut i fra definisjonen ser vi at Big Data Analytics ikke er et hvilket som helst verktøy, men verktøy som er tilpasset store, multi-terabyte datasett med minimal data forberedelse (Russom, 2011, s. 6). Formålet er større innsikt for å ta bedre beslutninger (Russom, 2011, s.15). Big Data Analytics kan innebære ulike teknikker og verktøy som for eksempel prediktiv analyse, datamining, statistisk analyse, data visualisering, kunstig intelligens og språkteknologi (Russom, 2011, s.6).

I høringsutkastet (2016) punkt 6 er data analytics definert som “the science and art of discovering and analyzing patterns, deviations and inconsistencies, and extracting other useful information in the data underlying or related to the subject matter of an audit through analysis, modeling and visualization for the purpose of planning or performing the audit”. Til tross for at høringsutkastet presenterer en definisjon av data analytics gir definisjonen liten indikasjon på hvor i revisjonsmetodikken det kan plasseres. Dagens standard skiller mellom analytisk substanshandling, test av detalj og test av kontroller (ISA 330.4). Hvilken type revisjonsbevis kan man da anse at Big Data Analytics genererer (IAASB, 2016, p.11). Kan Big Data Analytics generere revisjonsbevis som substanshandling og vill dette revisjonsbeviset klassifiseres i revisjonsstandarden som test av detalj eller som analytisk substanshandling (IAASB, 2016, p.11B). Det kan for eksempel tenkes at revisjonsbevis generert fra Big Data Analytics kan klassifiseres som alle overnevnte typene av revisjonsbevis (IAASB, 2016, p.19e).

2.2 Revisjonsbevis i henhold til revisjonsstandard

Revisors overordnede mål med revisjonen er å skaffe betryggende sikkerhet for at det totalt sett ikke er vesentlig feilinformasjon i regnskapet, verken som følge av misligheter eller feil (ISA 200.11). For å oppnå det overordnede målet skal revisor bruke de relevante ISA-er. Revisor må dermed vurdere om det er innhentet tilstrekkelig og hensiktsmessig revisjonsbevis i henhold til ISA 200 punkt 21.

I følge ISA 200 punkt 13 c er revisjonsrisiko “[r]isikoen for at revisor gir uttrykk for en uriktig mening i revisjonsberetningen når regnskapet inneholder vesentlig feilinformasjon”. Denne risikoen kan uttrykkes ved funksjonen;

Revisjonsrisiko = iboende risiko*kontrollrisiko* oppdagelsesrisiko (Figur 1).

Figur 1

Illustrasjon av revisjonsrisiko
Inspirert av Appelbaum (2016, s.20)



Det mest sentrale aspektet ved funksjonen i vårt studie er oppdagelsesrisiko, altså risikoen for at eksisterende feilinformasjon ikke blir avdekket av revisor gjennom revisjonshandlingene (ISA 200.13e). Det er viktig med hensiktsmessig og tilstrekkelig revisjonsbevis for å redusere revisjonsrisikoen til et akseptabelt nivå. Hensiktsmessighet er målet på revisjonsbevisets kvalitet. Det vil si dets relevans og pålitelighet når det gjelder å underbygge konklusjonene som revisors mening bygger på. Hvis beviset er relevant, pålitelighet og bearbeidet fra troverdige kilder kan revisor bruke det som substanshandling eller andre analytiske handlinger. Tradisjonelt sett har mye av bevisene vært papirbasert, basert på observasjoner, henvendelser og fysiske inspeksjoner. Som Figur 1 illustrerer er bevisets hensiktsmessighet sentralt sett opp mot revisors fastsettelse av oppdagelsesrisiko (Appelbaum, 2016, s.19).

2.2.1 Definisjon av revisjonsbevis

I henhold til ISA 500 punkt 5 (c) er revisjonsbevis “informasjon brukt av revisor for å komme frem til konklusjonen som revisors mening bygger på”. Videre heter det at revisjonsbevis kan være informasjon som kommer frem i regnskapsmaterialet men også informasjon innhentet fra andre kilder (ISA 500.5c).

Målet til revisor skal være å utforme og utføre revisjonshandlinger som gir tilstrekkelig og hensiktsmessig revisjonsbevis for å kunne trekke rimelige konklusjoner som grunnlag for revisors mening (ISA 500.6). I standarden defineres hensiktsmessighet med begrepene relevans og pålitelighet. Med andre ord skal det være relevant for revisjonen og være innhentet fra pålitelige kilder. Dette belyses nærmere i neste delkapittel.

2.2.2 Relevans og pålitelighet

I standardens veiledning er det en mer utdypende forklaring på hva som ligger i pålitelig og relevant revisjonsbevis. Relevans omhandler den logiske tilknytning til formålet med revisjonshandlingen og påstanden som vurderes (ISA 500.A27). Påliteligheten av informasjonen som skal brukes som revisjonsbevis påvirkes av kilden, type og omstendighetene rundt innhenting av informasjonen (Aurstad, 2017, s.26). Det er dermed vanskelig å generalisere påliteligheten av forskjellige typer revisjonsbevis. En generalisering i standarden er “informasjon innhentet fra en uavhengig ekstern kilde kan for eksempel være upålitelig dersom kilden ikke er kompetent, eller dersom en ekspert brukt av ledelsen ikke er objektiv” (ISA 500. A31). I standarden er det samtlige generaliseringer, men det påpekes at det kan forekomme unntak. Et av punktene vi mener er av særlig interesse er generaliseringen: “Revisjonsbevis er mer pålitelig når det innhentes fra uavhengige kilder utenfor enheten” (ISA 500. A31). En slik generalisering er ikke lenger like holdbar da Big Data potensielt kan frembringe den motsatte situasjonen. På grunn av den mangelfulle sporbarheten av kilden og troverdigheten kan det medføre at det blir mindre pålitelig bevis for revisor. Dette kommer av at Big Data i stor grad ikke er internt generert hos kunden og mest sannsynlig prosessert ekstern (Appelbaum, 2016, s.18). I tillegg har dagens kompliserte IT omgivelser og Big Data endret arten av revisjonsbevis. Hver fase av en transaksjon genereres og loggføres digitalt, noe som kan føre til at det også må verifiseres digitalt. Datagrunnlaget rundt

transaksjonene er loggført digitalt og kan være utsatt mulige endringer³. Sikring av opprinnelsen til kilden er derfor elementær og utfordrer generaliseringen om at eksterne bevis er mer pålitelige (Appelbaum, 2016, s.19).

2.3 Rammeverk for egenskaper til revisjonsbevis

Appelbaum (2016) foreslår et teoretisk rammeverk for egenskaper som revisjonsbevis bør bestå av, hvor alle egenskapene er etterspurt forskning (Appelbaum et al., 2017, s.53-54). Egenskapene er med å bygger opp under påliteligheten til revisjonsbevis og har tatt utgangspunkt i attributter fra revisjonsstandarden (Appelbaum, 2016, s.30). Spørsmålet er om revisjonsbevis generert fra Big Data og Big Data Analytics dekker egenskapene og i hvilken grad? Per i dag viser rammeverket egenskaper til tradisjonelle bevis, som papirbasert- og elektronisk⁴ revisjonsbevis, og adresserer ulikheter mellom dem. I tillegg påpekes hva som er styrker og svakheter med henholdsvis papirbasert og elektronisk revisjonsbevis. Det har vist seg at styrker ved papirbasert revisjonsbevis kan være svakheter ved elektronisk revisjonsbevis (Appelbaum, 2016, s.19). Vi ønsker å utvide Appelbaum (2016) sitt rammeverk slik at det også omfatter revisjonsbevis generert fra Big Data og Big Data Analytics, altså digitale bevis⁵. Ved å utvide rammeverket blir vi i stand til å svare på forskningsspørsmål F2; *Hvordan sikres egenskaper til revisjonsbevis ved bruk av Big Data og Big Data Analytics?*

³ Et aktuelt eksempel på betydningen av sikker lagring av dataens opprinnelse er Tidal-saken, hvor avspillings-data fra en musikkstrømmetjeneste er påstått manipulert. Denne saken viser hvor sentralt det er med relevant og pålitelig data for at næringslivet skal fungere rettferdig. I tillegg understreker Advokat Morten Andreassen hvor vanskelig det kan være for en alminnelig revisor å revidere en slik strømmetjeneste ved uttalelsen “Dessverre er dette området så bransjespesifikt og teknisk krevende, at det å be en norsk revisor revidere Tidal dessverre er som å be en analfabet om å skrive dikt” (Eckblad, Husby, Sæter & Tobiassen, 2018).

⁴ Elektronisk bevis er en generell beskrivelse som omfatter flere teknikker. Det kan være elektroniske dokumenter som skannede bilag, avtaler etc. Skanner du en signatur og sender denne på e-post, så er det en elektronisk signatur (Aurstad, 2017, s.26).

⁵ Digitale bevis er å benytte datatekniske metoder og verktøy for å erstatte eller effektivisere manuelle eller fysiske oppgaver (Aurstad, 2017, s.26).

Tabell 1

Rammeverk for egenskaper til revisjonsbevis
 Modifisert og oversatt fra Appelbaum, 2016, s. 21

Egenskaper til revisjonsbevis	Papirbasert bevis	Elektronisk bevis
<i>Risiko for endring</i> Bevis som kan endres fra sin opprinnelige form mangler troverdighet; revisjonsbevis bør være vanskelig å endre.	Det er vanskelig å endre papirbasert dokumentasjon uten av det blir oppdaget	Endring av elektronisk bevis kan være utfordrende å oppdage uten å gjennomføre nærmere bestemte tester
<i>Troverdighet</i> Dagens revisjonsstandard hevder at eksterne kilder øker troverdigheten når det innhentes uavhengig av kunden og kan bekreftes av revisor	Eksterne kilder av papirbasert bevis innhentet direkte av revisor øker troverdigheten	Internkontroller øker troverdigheten til elektronisk dokument. Ekstern elektronisk dokumentasjon mangler sikkerheten internkontrollene medbringer; at beviset ikke er falskt eller endret.
<i>Fullstendighet</i> Alle essensielle sider av en transaksjon skal være vertifiserbare	Vanligvis er de viktigste vilkårene inkludert i tekst format	Et elektronisk system kan erstatte tekstformat med koder og kryssreferanser til annen data som ikke er tilgjengelig for revisor
<i>Bekreftelse på attestasjon</i> Viktige aspekter ved internkontrollen skal være transparent og lett å vertifisere	Papirdokumentasjon som inkluderer attestasjoner øker dokumentasjonens fullstendighet	Ved elektronisk attestasjon kan det være behov for ytterligere vertifisering
<i>Brukervennlighet</i> Brukervennlige programmer og tilganger støtter etterlevelse (compliance)	Papirbasert bevis kan vanligvis evalueres uten bruk av tilleggsverktøy og/eller ferdigheter	Elektroniske bevismaterialet kan muligens kreve bistand fra en ekspert for å innhente data
<i>Klarhet</i> Et bevis anses som kompetent når en annen revisor kan gjenskape beviset og trekke samme konklusjon	Papirbasert dokumentasjon er klar og dermed intuitiv for en annen revisor	Elektronisk bevis er ikke alltid like klar for en annen revisor, spesielt ved fravær av passende kontroller

Digitale bevis, bevis generert av Big Data og Big Data Analytics, har potensialet til å dekke egenskapene i rammeverket på en tilstrekkelig måte (Appelbaum, 2016, s.24). For å undersøke dette må vi først belyse hvilke utfordringer som knyttes opp til de ulike egenskapene ved bruk av digitale bevis. Utfordringene fremstilles ved å koble egenskapene opp mot høringsutkastet til IAASB (2016) hvor de adresserer utfordringer ved implementeringen av Big Data og Big Data Analytics i revisjonsmetodikken.

2.3.1 Risiko for endring

Først og fremst foreligger det en risiko for at digitale revisjonsbevis kan endres uten et revisjonsspor eller et annet bevis. Det bør være vanskelig å endre digitale bevis ifølge Appelbaum (2016, s.23) og uten spor på endringen vil det gå utover bevisets troverdighet. I rammeverket kalles denne egenskapen *risiko for endring*. Høringsutkastet (2016) presiserer betydningen av at revisor etablerer kvalitetskontrollprosesser både ved bruk av selvutviklede verktøy men også ved bruk av verktøy utviklet av en uavhengig tredjepart. Det er dermed forventet at revisor tar stilling til påliteligheten av teknologien. I kvalitetskontrollprosessen bør revisor, for å sikre påliteligheten av teknologien, se til at revisjonsbeviset ikke kan endres fra sin opprinnelige form uten et revisjonsspor (IAASB, 2016, p.19j).

2.3.2 Troverdighet

Dagens revisjonsstandard hevder at eksterne kilder øker *troverdigheten* når det innhentes uavhengig av kunden og kan bekreftes av revisor (ISA 500.A31). Siden Big Data Analytics kan ta i bruk utradisjonelle kilder utfordrer digitale bevis dette standpunktet (Appelbaum, 2016, s.18). En faktor som kan øke troverdigheten til digitale bevis, generert fra eksterne kilder, er sikker registrering av dataens opprinnelse, registrering av endringer og eventuell annen metadata⁶ som revisor har tilgang til (Appelbaum, 2016, s.21). Punkt 19b i høringsutkastet (2016) presiserer betydningen av at revisor tar stilling til påliteligheten av informasjon brukt som revisjonsbevis hentet internt hos revisjonsklienten. I tillegg må revisor, i henhold til punkt 19c i høringsutkastet

⁶ Metadata er informasjon om data, og kan være en sentral faktor i å sikre opprinnelsen til dataen man skal basere seg på. (Appelbaum 2016, s.21)

(2016), ta i betraktning påliteligheten av data hentet ekstern fra revisjonsklienten. Revisor kan ikke gå ut i fra at data hentet fra en ekstern tredjepart er pålitelig uten å utføre prosedyrer knyttet opp mot validering, fullstendighet og nøyaktighet (IAASB, 2016, p. 19c).

2.3.3 Fullstendighet

Fullstendighet er egenskapen i rammeverket som skal sikre at alle essensielle sider ved et revisjonsbevis er verifiserbart. På lik linje med elektronisk bevis kan det være vanskelig å verifisere fullstendigheten av en transaksjon hvis dataen for transaksjonen ikke er tilgjengelig for revisor eller at dataen foreligger i et annet format enn tekstformat. Det er dermed nødvendig at informasjon produsert både internt og ekstern fra revisjonsklienten er fullstendig for å bli ansett som pålitelig revisjonsbevis (IAASB, 2016, p.19b(i) & 19c).

2.3.4 Bekreftelse på attestasjoner

For at revisor skal kunne utføre test av kontroller er det nødvendig at internkontrollen er transparent og lett å verifisere. Denne egenskapen kalles for *bekreftelse på attestasjoner* i rammeverket. Ved digitale bevis kan attestasjoner for internkontrollen for eksempel registreres som metadata (Appelbaum, 2016, s.31). Spørsmålet er hva minimumsnivået av test av generelle IT kontroller og resultatene av testene må være når revisor anvender Big Data Analytics og planlegger å bygge på de generelle IT kontrollene (IAASB, 2016, p.9a(i)). Hva er virkningen av mangler i generelle IT kontroller og applikasjonskontroller som revisor har til hensikt å bygge på for å konkludere på påliteligheten av data generert fra IT systemet (IAASB, 2016, p.9a(ii)).

2.3.5 Brukervennlighet

Rammeverket (2016) presiserer at digitale revisjonsbevis krever standardiserte rutiner og verktøy for å utvinne og ta i bruk data på en brukervennlig måte. Denne egenskapen kalles *brukervennlighet* og er med på å støtte opp under “compliance” og dermed pålitelighet. Ved å bruke Big Data kan det oppstå utfordringer med hvordan man skal utvinne store datasett på en brukervennlig måte. Noen av disse utfordringene innebærer overføring av data fra kunde til revisor, sikker-

het, personvern og lagring av datasettene (IAASB, 2016, p.18a). Spesialister øker brukervennligheten til Big Data Analytics for sluttbruker ved å bistå revisor i bruken av verktøyet (IAASB, 2016, p.18d).

2.3.6 Klarhet

Den siste egenskapen i rammeverket er *klarhet*. Et bevis anses som klart når det er kompetent nok til at en annen revisor kan gjenskape beviset og trekke samme konklusjon (Appelbaum, 2016, s.21). Den underliggende dokumentasjonen til revisjonsbeviset bør være intuitivt for å kunne gjenskapes (Appelbaum, 2016, s.23). I dagens standard kreves det ikke at revisor skal beholde all informasjon brukt ved utvelgelse av elementer til testing, men krever at revisor dokumenterer identifiserende karakteristikk ved utvalget som testes (ISA 230.9). Dokumentasjonskravet ved bruk av Big Data Analytics trenger ikke å være annerledes, men det oppstår utfordringer rundt anvendelsen. Utfordringene kan blant annet gjelde retning for testing, tilsyn, utførelse og review ved bruk av Big Data Analytics (IAASB, 2016, p.19i). I tillegg må revisjonsteamet ta stilling til om all data og alle detaljer om rutinene, som er blitt testet, skal inkluderes i dokumentasjonen (IAASB, 2016, p.19i).

3. Metode

Som følge av at Big Data og Big Data Analytics i revisjon er et relativt nytt fenomen har vi valgt å anvende casesdesign som kvalitativ metode. Ved å benytte casesdesign gir det oss mulighet til å gå i dybden på studieobjektet og utforske komplekse sider ved fenomenet (Malsch & Salterio, 2016). Hovedsakelig vil vi samle inn og benytte data fra dybdeintervju.

3.1 Intervju

I 2018 utførte vi fire forskjellige intervju med informanter fra forskjellige revisjonsselskap. Tre av intervjuene ble foretatt med kun en informant mens det siste intervjuet ble foretatt med to informanter fra samme revisjonsselskap. Til sammen intervjuet vi fem ulike informanter fra fire forskjellige revisjonsselskap.

Tabell 2

Oversikt over informanter

Informant nr.	Selskap nr.	Stilling	Området	Erfaring	Intervju (ca) tidsbruk
Informant 1	Selskap 1	Partner	IT risikotjenester	19 år	30 min
Informant 2	Selskap 2	Manager	Revisjon	7 år	40 min
Informant 3	Selskap 3	Direktør	Data Analytics i Revisjon	14 år	35 min
Informant 4	Selskap 4	Senior manager	Risk Advisory Services	9 år	60 min
Informant 5	Selskap 4	Manager	Risk Advisory Services	6 år	60 min

Revisjonsselskapene ble valgt på grunnlag av størrelse og er kategorisert som fire av de fem store revisjonsselskapene som råder revisjonsbransjen i Norge. Vi valgte store revisjonsselskap med betydelige markedsandeler for å fange opp generelle tendenser. I tillegg ga ulike revisjonsselskap oss bredde i svarene og mulighet til å fange opp likheter og ulikheter i bransjen. Antall informanter begrunnes ut i fra et ønske om å sammenligne og finne mønstre mellom informantene. Dessuten gir det bedre grunnlag for analyse. I tillegg til at informantene var ansatt i et stort revisjonsselskap stilte vi krav til at de skulle ha erfaring og kompetanse med Big Data og Big Data Analytics i revisjon. Dette kravet ble stilt for å få grundige og informative svar. Alle informantene var norske og hadde bred erfaring med revisjonsbransjen i Norge. Majoriteten av informanter var lokalisert ved selskapets hovedkontor i Oslo slik at intervjuene ble utført via telefon,

Skype eller Google Meet. Et intervju ble derimot avholdt i Stavanger og dermed med personlig oppmøte. De fleste informantene jobbet per dags dato hovedsakelig i en revisjonsavdeling med spesialisering rettet mot IT risiko og dataanalyser. Informant 4 og 5 jobbet ikke i en revisjonsavdeling men i en avdeling som håndterte IT risiko og dataanalyser både for revisjonsformål men også for andre formål. Til tross for at noen av informantene ikke jobbet i en revisjonsavdeling i dag hadde alle informantene tidligere erfaring som revisor. Å kartlegge informantenes erfaring rettet spesifikt mot Big Data og Big Data Analytics var mer utfordrende. Vi kom derimot frem til et gjennomsnitt på ca. to år.

Intervjuene ble gjennomført som dybdeintervju med åpne spørsmål for å gi informanten mulighet til å svare fritt, få frem kompleksitet og nyanser (Brinkmann & Kvale, 2015, s. 47). Vi fulgte et semistrukturert intervju og brukte en intervjuguide med noen overordnede holdepunkt. Det ble også stilt tilleggsspørsmål underveis for å klargjøre diverse svar. Informanten ble garantert anonymitet for å ivareta informantens personvern samtidig som at vi ønsket å hente ut verdifull innsikt. Informanten hadde mulighet til å trekke seg før, under og etter intervjuet. Forespørsel om opptak via diktafon ble også innvilget.

Intervjuguiden startet med en introduksjon av studentene, fagområdet og forskningsspørsmål. Deretter ble informanten spurt om posisjon og erfaring for å kartlegge kompetanse. Videre i intervjuet ble spørsmålene delt inn i to temaer med utgangspunkt i oppgavens forskningsspørsmål. Det første temaet omhandlet definisjonen av Big Data og Big Data Analytics i revisjon, hvor seks spørsmål ble stilt. Deretter tok intervjuguiden for seg det andre temaet, altså pålitelighetskravet til revisjonsbevis. Her ble det stilt syv spørsmål fordelt i underkategorier, ut i fra egenskapene i det teoretiske rammeverket. Til sammen ble 14 hovedspørsmål stilt for å undersøke oppgavens forskningsspørsmål og problemstilling.

3.1.1 Pilottesting

Før de faktiske intervjuene utførte vi pilottesting av intervjuguiden for å evaluere gjennomførbarhet, tidsbruk og andre uønskede hendelser (Lindquist, R., 1991, s. 91-92). Vi utførte pilottes-

tingen på to medstudenter med ca. tre-fire års erfaring i revisjon. Under testen simulerte vi en intervjusituasjon hvor intervjuguiden ble presenterte og hvor tema og spørsmål ble gjennomgått. Under testen fikk vi løpende tilbakemelding på oppsett, formulering av spørsmål og andre uklarheter. Tilbakemeldingene ble brukt til å justere intervjuguiden og klargjøre den til de faktiske intervjuene.

3.1.2 Analyse av intervju og koding

De transkriberte intervjuene var grunnlaget for den etterfølgende analysen, hvor formålet var å identifisere de temaene som var mest fremtredende. Analysen ble gjennomført i to steg. Først ble en grov analyse gjennomført ved å kode og filtrere ut de mest sentrale funnene, etterfulgt av en mer detaljert analyse. I den detaljerte analysen ble utdrag fra intervjuene overført til separate kategorier, organisert i hoved- og underkategorier. Hovedkategoriene tok utgangspunkt i forskningsspørsmål F1 og F2 i intervjuguiden. Underkategoriene i forskningsspørsmål F1, om definisjonen av Big Data og Big Data Analytics i revisjon, ble organisert ut i fra aspekter med definisjonen som vi undersøkte, hvor aspektene tok utgangspunkt i høringsutkastet til IAASB.

I forskningsspørsmål F2, om pålitelighetskravet til revisjonsbevis generert fra Big Data og Big Data Analytics, ble underkategoriene organisert i de ulike egenskapene i rammeverket. Vi ser dermed at andre del av intervjuguiden, og dermed analysen av forskningsspørsmål F2, tar utgangspunkt i det teoretiske rammeverket til Appelbaum (2016). Underkategoriene ble deretter reanalysert flere ganger og etter hver analyse ble utdrag filtrert for å sikre at kun de mest sentrale utdragene ble beholdt. Hvis nødvendig gikk vi igjennom de originale transkriberingene for å kontrollere arbeidet.

Under transkriberingen anvendte vi NVivo. Dette er en programpakke med analyseverktøy for kvalitative forskningsdesign (Alfasoft, 2018). Selve transkriberingen tok i underkant av 30 timer, men programmet ble også brukt under analysen når vi kodet den transkriberte teksten. Programmet er spesielt hensiktsmessig ved prosessering og analyse av mye ustrukturert informasjon. Ved å bruke NVivo ble transkribering og koding gjennomført på en organisert og systematisk måte. Dette ga oss et godt utgangspunkt for den videre analysen.

4. Resultater

Vi vil i dette kapittelet presentere empirisk bevis for hvert av forskningsspørsmålene (F1 & F2). For analyseformål presenteres resultatene i kronologisk rekkefølge, hvor vi først presenterer resultatene vedrørende definisjon av Big Data og Big Data Analytics før vi legger frem resultatene som omhandler pålitelighetskravet til digitale revisjonsbevis.

4.1 Definisjon av Big Data og Big Data Analytics i revisjon, (F1)

Intervjuguiden ble tilsendt og presentert for de respektive informantene, hvor de generelle definisjonene til Big Data og Big Data Analytics, basert på Cao et al. (2015) fremkom. Med utgangspunkt i de generelle definisjonene tilpasset informanten begrepene til revisjon.

4.1.1 Big Data

Gjennom intervjuene fikk vi inntrykk av at informantene var kjent med den presenterte definisjonen av Big Data. På en annen side stilte samtlige informanter spørsmål om definisjonen passer inn i en revisjonssammenheng.

Gjennom intervjuene kom det frem ulike oppfatning av Big Data i revisjon. Blant annet kan Big Data oppfattes som:

... ulike typer transaksjonsdata. Altså data som ligger i ERP-systemer (Informant 5).

I tillegg nevnte Informant 3 ulike typer markedsdata som også kan ansees som Big Data i revisjon. Dette var eksempelvis:

Dun and Bradstreet⁷, Fraud index⁸, markedspremier og BETA data (Informant 3).

I intervjuene var alle informantene enig i at volum er den største driveren bak definisjonen til Big Data. Informant 2 uttalte blant annet at:

Når jeg tenker på Big Data, så tenker jeg på store mengder av informasjon (Informant 2). Hvilken enhet volumet av dataen bør måles i ble derimot problematisert. Det ble blant annet stilt følgende spørsmål:

Er det et kvantum, er det et mål eller er det en million? (Informant 5).

På en annen side ga flertallet av informantene et eksempel i omsetning når de utdypet aspektet volum. I tillegg ble antall transaksjoner og annen finansiell data fra ERP-system nevnt som måleenhet. Det var en felles forståelse blant informantene at mengden data måtte være så stor at tradisjonelle verktøy ikke kunne anvendes. Til tross for dette fremkom det i intervjuene at omfanget ikke er fastsatt i praksis hos noen av revisjonsfirmaene.

Et annet aspekt ved definisjonen som ble problematisert var egenskapen variasjon. Det gjaldt spesielt om variasjon inkluderer både strukturert og ustrukturert data når definisjonen anvendes i revisjon. Informant 3 uttrykte at det er en utbredt misforståelse i bransjen at Big Data:

... kun består av ustrukturert data, men at det også består av strukturert data (Informant 3)

Et godt eksempel som illustrer denne misforståelsen er følgende uttalelse fra Informant 2:

⁷ Dun & Bradstreet har den største globale kommersielle databasen. De tilbyr blant annet informasjon om kommersiell kreditt samt rapporter om virksomheter. De er i hovedsak mest kjent for Data Universal Numbering System (D.U.N.S) som genererer rapporter for mer enn 100 millioner bedrifter over hele verden (Dun & Bradstreet, 2018)

⁸ Fraud-index inneholder en beregning som kvantifiserer frekvensen av svindelforsøk og hvordan det endres over tid. Indeksen undersøker ulike aspekter av bedrageri, for eksempel hvilke land eller regioner som er mest truet av svindel, og hvilke bransjer som bør se nærmere på hva som skjer rundt dem. (PYMNTS, 2017).

Når selskaper har 35 milliarder i omsetning vil det være ekstreme mengder med transaksjoner og det vil jeg også kalle Big Data, men det tror jeg ikke passer inn i definisjonen (Informant 2).

Som dette eksempelet understreker råder det en usikkerhet om strukturert data kan inkluderes i begrepet Big Data.

Intervjuet med informantene i selskap 4 indikerer at troverdighet og variasjon kan gå på kompromiss av hverandre i revisjonssammenheng. Det ble påpekt at troverdigheten ble vanskeligere å verifisere desto høyere grad av variasjon det var i datagrunnlaget. Informant 4 forklarte videre at:

Data med lav troverdighet er ikke i et revisjonsperspektiv så interessant å se på ... men det er det som gjør at revisjonen har vanskelig for å benytte seg av det (Informant 4).

4.1.2 Big Data Analytics

Gjennom intervjuene kom det frem at alle informantene var enig i den presenterte definisjonen av Big Data Analytics. Imidlertid kom det frem mer nyanserte betraktninger angående hva Big Data Analytics er for revisjonsformål gjennom diverse oppfølgingsspørsmål.

Informantene ble blant annet spurt om deres respektive revisjonsselskap genererer revisjonsbevis fra Big Data Analytics. Her var det bare Informant 1 som svarte ja uten noe ekstra å tilføye. De resterende nyanserte svarene hvor de påpekte at det avhenger av definisjonen til Big Data i revisjon. Dette gjenspeiles i en managers uttalelse, hvor svaret var ja til at de genererte digitale revisjonsbevis med forbehold om:

... at store grupper av transaksjoner som hovedbok er Big Data (Informant 2).

Informant 3 presiserte at de bruker Big Data Analytics til å generere revisjonsbevis hvis strukturert data fra ERP-system går inn under definisjonen til Big Data. Videre poengterte Informant 2

at de bruker Big Data Analytics til å generere revisjonsbevis når tradisjonelle verktøy ikke er i stand til å håndtere mengden data som skal analyseres.

Big Data teknologi ble også trukket frem av Informant 3 for å belyse at revisjonsselskapet ikke kun bruker Big Data Analytics til å genere revisjonsbevis, men en helhet av forskjellige typer teknologi knyttet opp mot Big Data. Hadoop⁹ og Spark¹⁰ ble blant annet trukket frem av samme informant som eksempler på teknologi som ble anvendt.

I spørsmålet om hvilke type revisjonshandling informantene anså Big Data Analytics for å være opplyste alle informantene at de anså det primært som substanshandling, da enten som detaljtest eller analytisk handling (Tabell 3). Som vi ser fra tabellen var det kun tre informantene som anså Big Data Analytics som en analytisk handling, og fire informanter anså Big Data Analytics som detaljtest. Dessuten anså informant 1, 3, 4 og 5 Big Data Analytics som test av kontroll.

Tabell 3

Illustrasjon av Big Data Analytics som revisjonshandling

	Substanshandling		Test av kontroll
	Detaljtest	Analytisk handling	
Informant 1		x	x
Informant 2	x	x	
Informant 3	x	x	x(automatisk)
Informant 4	x		x
Informant 5	x		x

⁹ Apache Hadoop er et Big Data rammeverk som tilbyr distribuert behandling av store datasett ved hjelp av enkle programmeringsmetoder. Hadoop tilbyr en kostnadseffektiv løsning for lagring og behandling av strukturerte, semi- og ustrukturerte data uten formatkrav (IBM, 2018).

¹⁰ Apache Spark er en enhetlig analysemotor som behandler store datasett, med innebygde moduler for streaming, SQL, maskinlæring og prosessering (Apache Spark, 2018).

Høringsutkastet (2016) stiller spørsmål ved om Big Data Analytics kan ansees som alle de overnevnte typene av revisjonsbevis, henholdsvis detaljtest, analytisk handling og test av kontroll. Fra tabellen ser vi at det kun er Informant 3 som klassifiserer revisjonsbevis generert fra Big Data Analytics som alle tre. Imidlertid klassifiserer fire av fem informanter Big Data Analytics som både substanshandling, da enten detaljtest eller analytisk handling, og test av kontroll. Noe som indikerer at på tross av at de ikke eksplisitt klassifiserer handlingen som alle tre, klassifiserte majoriteten den i de to overordnede kategoriene.

Høringsutkastet (2016) nevner også at Big Data Analytics kan endre den tradisjonelle forståelsen av revisjonsmetodikken ved at arten av revisjonshandlingene kan endres. Et godt eksempel på at dette også er forståelsen i praksis kan illustreres med følgende sitat fra en manager:

I metodikken vil Big Data Analytics som substanshandling erstatte test av kontroller (Informant 2).

Til tross for uttalelsen anså ikke informanten Big Data Analytics som test av kontroll.

4.2 Pålitelighetskrav til revisjonsbevis, (F2)

Basert på de seks egenskapene i rammeverket til Appelbaum (2016), presentert i teorikapittelet, skal vi nå presentere empirisk bevis for hvordan informantene stiller seg til pålitelighetskravet til revisjonsbevis når beviset er generert av Big Data og Big Data Analytics. For analyseformål presenteres egenskapene i kronologisk rekkefølge sett opp mot rammeverket.

4.2.1 Risiko for endring

Først og fremst skal vi presentere funnene våre sett opp mot egenskapen risiko for endring. For å sikre denne egenskapen trakk informant 1, 4 og 5 frem internkontroll sett opp mot IT. I den sammenheng ble test av tilganger og sporbarhet spesifikt trukket frem, hvor en informant uttalte:

Ved den strukturerte delen fra et ERP-system så ser man på tilganger, endringslogger og hele sikkerheten rundt systemet (Informant 3).

Som høringsutkastet (2016) nevner er revisjonsspor, innebygd i analytiske verktøy, sentralt for å redusere risiko for endring. Potensialet som ligger i fremtidig teknologi for å sikre sporbarhet ble belyst av Informant 4 ved å uttale at:

I en ideell verden kunne du sett for deg at det hadde eksistert noe lignende blockchain eller en digital signatur som sikrer at data ikke har blitt tuklet med (Informant 4).

I følge intervjuene vil også kilden til revisjonsbeviset ha betydning for risiko for endring. Hvis kilden stammer fra en uavhengig tredjepart, som ikke har insentiv til å tukle med informasjonen, vil det gi et mer troverdig bevis. Dette resonnementet bygger også på ISA 500.A31 om generaliseringen at revisjonsbevis er mer pålitelig når det innhentes fra en uavhengig kilde utenfor enheten. På en annen side kan også internt genererte kilder til revisjonsbevis redusere risikoen for at revisjonsbevis er endret. Dette gjelder for eksempel hvis forskjellige sider av et ERP-system kan sees opp mot hverandre og bekrefte samme informasjon. Som forklart av Informant 3:

Så er det fryktelig vanskelig å stole alt på en kilde for da har du ikke kontroll, men får du den samme informasjonen fra flere ulike kilder, så mener jeg at det reduserer risiko for endring (Informant 3).

Denne handlingen kalles avstemming og vil i sin tur være avhengig av god arbeidsdeling hos revisjonskunden.

4.2.2 Troverdighet

På lik linje meg risiko for endring vil egenskapen troverdighet sikres gjennom flere av de samme handlingene. Dette gjelder blant annet avstemming av forskjellige kilder av data og tilgangskontroller. Denne sammenhengen illustreres av en informant ved uttalelsen:

Så lenge risikoen for endring er lav så er jo også troverdigheten til stede (Informant 1).

I forbindelse med internt generert data fra revisjonskunden nevnes kompenserende handlinger sammen med tilgangskontroller som en nødvendighet for å sikre troverdighet. I følge informantene fra selskap 4 er kompenserende handlinger:

Handlinger som utføres når revisor tester et eller annet rundt tilgangskontroller for å sikre at for eksempel økonomidirektøren ikke kan seile under falskt flagg under noen andre anstattes rettigheter (Informant 4 og 5)

Med andre ord skal det sikre integriteten av dataen som genereres fra tilgangene.

Tre informanter understrekte også hvordan troverdigheten til revisjonsbevis trolig kan sikres mer tilstrekkelig i fremtiden. I den sammenheng ble muligheten for å avstemme intern data mot eksternt informasjon til leverandører, “kunders kunder” og offentlige myndigheter trukket frem. Det ble påpekt av ulike informanter at dette først vil være mulig når det åpnes opp for mer utveksling av data som igjen avhenger av regulatoriske myndigheter.

4.2.3 Fullstendighet

For å sikre fullstendighet av data presiserte alle informantene betydningen av å se på helheten av handlingene som ble utført under revisjonen. Dette ble illustrert med følgende eksempel:

Hvis vi går på salg for eksempel. Da lager du en salgsordre, du gjør en levering, du får en faktura, foretar en betaling og får eksternt informasjon fra kunden om salget. I tillegg så ser du at ting har gått ut av lageret. Da har du sett på flere kontroller i helhet og sannsynligheten for at informasjonen er vesentlig feil er ganske liten (Informant 3).

Ved å se kontrollene og kildene til informasjon samlet kan man oppdage hvis noe ikke er reflektert i dataen slik at sannsynligheten for ufullstendig data blir redusert. Dette kom frem ved følgende utsagn:

Vi prøver å gjøre mange handlinger for å finne ut om det er noen transaksjoner som ikke er reflektert i den dataen vi har fått, slik at mangelen ikke reflekteres i analysen (Informant 2).

Intervjuene indikerer at nøkkelen til fullstendighet er å se på hele prosesser i stedet for å se på kontrollen isolert sett. En tre-veis-kontroll ble trukket frem som et eksempel av Informant 4 for å sjekke at tallene fra kildene har konsistens. Avstemming av data ble trukket frem av både Informant 1 og 4 for å sikre fullstendighet. Informantene presiserte utfordringen med å sikre fullstendigheten når noe manglet i datasettet. Som Informant 4 formulerte er:

Fullstendighet er en generell utfordring. Jo mer perifer dataen er ... jo vanskeligere blir det med et vanlig revisjons maindsett å vurdere fullstendigheten, med mindre man har en ekstern kilde (Informant 4).

Informant 1 belyste også muligheter for å sikre fullstendighet bedre i fremtiden ved uttalelsen:

Vi ser større muligheter fremover for å være sikker på at alt har havnet i bøkene hos revisjonsklienten. Jeg tror vi får mer og mer informasjon og sammenstille mot kundens kunder i forhold til å kunne gjøre avstemninger mellom en detalj som er registrert hos vår kunde og den inputen som kommer fra deres kunder igjen. Vi er ikke helt der i dag, men på sikt vil den type informasjon kunne innhentes fra eksterne kunder (Informant 1).

En informant uttalte at de sikret fullstendighet ved å trekke ut ønskede parametere fra et datasett, hvor ulike parametere som kunne brukes for å sikre fullstendigheten av transaksjon var:

Dato for når den ble ført, når den ble fakturert, tekst, hvem som førte og klokkeslett (Informant 2).

Ved å trekke ut forskjellige parametere fra samme datasett og se de ulike datasettene, med ulike parametere, opp mot hverandre er det med på å bygge opp under fullstendigheten. Disse parametere kan også beskrives som metadata.

4.2.4 Bekreftelse på attestasjon

Viktige aspekter ved internkontrollen må være transparent og lett å verifisere for at revisor skal kunne bekrefte attestasjoner. I den sammenheng nevnte alle informantene revisjon av selskapets IT system med ulike innfallsvinkler. Tilgangskontroller, fullstendige endringslogger og autorisasjonsmatriser ble blant annet nevnt som hjelpemidler for å sikre egenskapen. En informant utdyppet svaret ved å forklare at:

Autorisasjonsmatriser og toleransegrenser matches opp mot hvilke personer som faktisk har utført første og andre gangs godkjenninger, og ser dette videre opp mot selskapets policy (Informant 3).

Denne uttalelsen poengterer igjen hvordan ulike kilder med informasjon sees opp mot hverandre for å bekrefte eller avkrefte en egenskap.

4.2.5 Brukervennlighet:

I forbindelse med egenskapen brukervennlighet, kom det klart frem at informantene har ulik grad av innsikt alt etter hvilken stilling de har i selskapene. Intervjuene indikerer at selskapene har ulike innfallsvinkler til sikring av egenskapen brukervennlighet. Informant 1 trakk frem at måten de sikret egenskapen er ved selvutviklede standardiserte maler for å utvinne data likt for alle kunder. På en annen side fremhevet Informant 2 at de tar i bruk egne datasentre, såkalte hubber, som er spesialisert på å utvinne data og gjøre det lesbart. Hubbene utfører i tillegg validerings tester av dataen for å gjøre det mer brukervennlig for revisjonsselskapet.

For å sikre brukervennlighet har revisjonsselskap 3 utviklet egne verktøy som skal være lette å bruke og som skal gi sluttbruker en enkel visualisering av resultatet. Som informanten selv la det fram:

... motoren kan være fryktelig komplisert men det trenger ikke være så vanskelig å kjøre bilen. Vi lagrer alle rutinene i en database som er utrolig kompleks, men i forhold til

sluttbrukeren, så sitter han med en enkel visualisering i for eksempel powia eller click view eller lignende som gjør det lett for brukeren å identifisere avvik (Informant 3).

Utviklingen av de analytiske verktøyene er det spesialister som tar seg av, mens revisor selv skal kunne bruke dataanalyser i sluttrapportene. Videre forklarte Informant 3 at til tross for at mye er standardisert bør revisor ha en viss grunnkompetanse angående verktøyene for å kunne stille gode spørsmål. Med andre ord ligger det mye spesialisert kompetanse og arbeid i å utvikle verktøyene, men på en annen side er det fokus på å gjøre det brukervennlig og lett å navigere i for revisor. Det handler om å sette sammen et team hvor man har IT-spesialister og eksperter i revisjon som jobber sammen.

Informantene fra selskap 4 er spesialister innen IT. De påpekte at de henter ut data på to forskjellige måter gjennom et ERP-system. For det første kan de gå rett på en underliggende database.

Fordelene med dette er:

... at man får en fast struktur og ikke problemer med datamengden (Informant 4).

Ulempen var derimot:

... at det er litt høyere brukerterskel fordi det krevet teknisk kompetanse for å gjøre det (Informant 4).

Den andre måten å hente ut data på er at revisor selv henter et uttrekk ved å “trykke på en standard rapport”. Det som sikrer brukervennlighet, når revisor selv henter et uttrekk, er at selskapets selvutviklede verktøy tar kundens data “kverner, transformerer og presenterer” dataene i et analytisk verktøy som revisoren er kjent med å navigere. Dette er derimot ikke alltid en mulighet, etter som at datamengden kan være for stor. Det er da man må gå direkte til datagrunnlaget ved hjelp av IT eksperter. Altså der hvor IT ekspertene er nødvendige vil ofte brukervennligheten for en ordinær revisor synke.

4.2.6 Klarhet

Den siste egenskapen i rammeverket er klarhet. Som forklart av en informant går dette ut på at:

Alt skal være dokumentert og etterprøvbart for en med lignende kompetanse

(Informant 3).

For å sikre denne egenskapen presiserte alle informantene at de lagret en viss form for dokumentasjon. Hvilken type dokumentasjon som ble lagret varierte derimot mellom informantene. Informant 1 og 2 presiserte at selve datasettet ble lagret slik at det enkelt kun tas opp igjen på et senere tidspunkt for å bli etterprøvd. Informant 3, 4 og 5 lagret ikke selve datasettet, men en kode eller et skript for å kunne gjenskape beviset på et senere tidspunkt. Ved å lagre koder som en del av dokumentasjonen kunne revisor når som helst hente ut det samme revisjonsbeviset fra dataen. Det samme konseptet gjaldt når informantene tok i bruk et skript. Som forklart av Informant 5 så er et skript:

En type logg som teknisk viser hva som har skjedd og som tar vare på alle stegene som er gjennomført for å komme frem til et revisjonsbevis. Loggen blir lagret som et sett med instruksjoner som gir revisor de nødvendige stegene for å gjenskape revisjonsbeviset (Informant 5).

For å bevare integriteten av rådataen ble nødvendigheten av å holde rådataen adskilt fra den prosesserte dataen presisert. I begge tilfeller ble datagrunnlaget til kunden, som ble brukt til å komme frem til revisjonsbeviset, slettet etter signering. Det ble begrunnet ut i fra et ønske om sikkerhet til kunden og kundens plikt til å kunne gjenskape datagrunnlaget på et senere tidspunkt.

5. Diskusjon

For å svare på første forskningsspørsmål: *Hvordan defineres Big Data og Big Data Analytics i revisjon?* har vi utarbeidet definisjoner basert på resultatene i undersøkelsen. Nedenfor vil vi drøfte resultatene og presentere forslag til de revisjonsrettede definisjonene av Big Data og Big Data Analytics.

Som resultatene viser er volum fortsatt hoveddriveren bak definisjonen til Big Data i revisjon og er dermed naturlig å plassere først i definisjonen. Variasjons aspektet av definisjonen må derimot endres for å presisere at Big Data i revisjon per i dag hovedsakelig består av strukturert data. På samme måte er troverdighets aspektet trukket frem for å poengtere at dataen må inneha denne karakteristikken for å kunne anvendes for revisjonsformål. Det ligger i revisjonens natur at revisor ikke kan bygge på kilder som mangler troverdighet hvis det skal anvendes som revisjonsbevis. Ut i fra resultatene kan det også diskuteres om definisjonen bør inneholde en presisering av at data med høy variasjon kan gå på kompromiss av troverdigheten til dataen. Altså at data med høyere grad av variasjon mest sannsynlig innehar lavere troverdighet. På en annen side viser resultatene at man per i dag hovedsakelig anvender strukturert data fra ERP-system med lav grad av variasjon. På grunn av dette har vi valgt å utelate denne presiseringen. Til slutt har vi valgt å ta bort hastighets aspektet ved definisjonen. Som følge av at en revisjon blir foretatt i etterkant av selve datagenereringen ansees ikke hastighetsaspektet å være relevant for definisjonen per i dag. Det kan derimot tenkes en fremtidig situasjon hvor revisjonen foretas i nåtid. Da vil hastigheten av datagenereringen være mer relevant. Fra våre funn foreslår vi dermed følgende definisjon:

Big Data er informasjon kjennetegnet ved høyt volum som krever nye metoder for prosessering og tilrettelegging. Dataen har som oftest liten grad av variasjon hvor det hovedsakelig brukes

strukturert data som transaksjonsdata fra ERP-system. I tillegg må dataen inneha en høy grad av troverdighet for å fungere som grunnlag for bedre beslutninger, økt innsikt og optimalisering av prosesser.

Som resultatene våre også viser var informantene i hovedsak enig i den presenterte definisjonen av Big Data Analytics. Imidlertid manglet definisjonen er presisering av at Big Data Analytics består av nye metoder for å prosessere data, og at det derfor i hovedsak anvendes hvis tradisjonelle analyser ikke er i stand til å håndtere mengden data. Om Big Data Analytics bør defineres som en type revisjonshandling er vanskelig å si. Som følge av ulik klassifisering av revisjonshandlingen i undersøkelsen har vi derimot valgt å utelate dette fra definisjonen. På en annen side anser vi det viktig med videre forskning på dette området. Et annet funn som er interessant å belyse er hvordan Big Data Analytics som revisjonshandling kan erstatte test av kontroller. Funnet viser hvilken betydning Big Data Analytics kan ha for revisjonsmetodikken og endringer som standardsetterne må ta stilling til ved utarbeidelsen av den nye standarden. På en annen side var det kun en informant som trakk denne konklusjonen. Det vil derfor ikke være signifikant og er dermed utelatt fra definisjonen. For revisjonsformål foreslår vi dermed følgende definisjon:

Big Data Analytics er prosessen, som anvendes når tradisjonelle analyser ikke er i stand til å håndtere mengden data, hvor man inspiserer, vasker, transformerer og modellerer Big Data for å oppdage og formidle nyttig informasjon og mønstre, foreslå konklusjoner og støtte opp under beslutningstaking.

For å svare på det andre forskningsspørsmålet: *Hvordan sikres egenskaper til revisjonsbevis ved bruk av Big Data og Big Data Analytics?* har vi utvidet det teoretiske rammeverket til å gjelde

både digitale bevis i dag og digitale bevis for fremtiden. Med utgangspunkt i handlingene som fremkommer i undersøkelsen har vi forsøkt å forklare i hvilken grad egenskapene dekkes.

Tabell 4

Utvidet rammeverk med digitale revisjonsbevis

Egenskaper til revisjonsbevis	Papirbasert bevis	Elektronisk bevis	Digitale bevis i dag	Digitale bevis i fremtiden
Risiko for endring Bevis som kan endres fra sin opprinnelige form mangler troverdighet; revisjonsbevis bør være vanskelig å endre.	Det er vanskelig å endre papirbasert dokumentasjon uten av det blir oppdaget	Endring av elektronisk bevis kan være utfordrende å oppdage uten å gjennomføre nærmere bestemte tester	Generell IT kontroll gjennom tilganger, endringslogger og revisjonsspor gir revisor redusert risiko for endring. Samtidig vil avstemming av ulike kilder som bekrefter samme informasjon bygge opp under denne	Blockchain og digitale signaturer som sikrer at dataen ikke er blitt tuklet med kan bidra til å redusere risikoen for endring av digitale bevis i fremtiden.
Troverdighet Dagens revisjonsstandard hevder at eksterne kilder øker troverdigheten når det innhentes uavhengig av kunden og kan bekreftes av revisor	Eksterne kilder av papirbasert bevis innhentet direkte av revisor øker troverdigheten	Internkontroller øker troverdigheten til elektronisk dokument. Ekstern elektronisk dokumentasjon mangler sikkerheten internkontrollene medbringer; at beviset ikke er falskt eller endret.	Troverdighet har sammenheng med egenskapen ovenfor, hvor lav risiko for endring gir høyere grad av troverdighet for digitale bevis. Det gjelder spesielt gjennom kontroll av tilgangskontroller sammen med kompensierende handlinger som sikrer integriteten av dataen.	I fremtiden kan troverdighet av digitale bevis sikres mer tilstrekkelig ved at lovgivning åpner opp for mer utveksling av data som muliggjør avstemming av intern informasjon mot ny ekstern data.
Fullstendighet Alle essensielle sider av en transaksjon skal være vertifiserbare	Vanligvis er de viktigste vilkårene inkludert i tekst format	Et elektronisk system kan erstatte tekstformat med koder og kryssreferanser til annen data som ikke er tilgjengelig for revisor	Fullstendighet sikres ved å se på helheten av alle kontrollene, samt kontrollere transaksjonens metadata. Derimot er egenskapen vanskeligere å sikre dersom det er eksten data.	På lik linje med egenskapen troverdighet er det større potensialet til å sikre egenskapen fullstendighet mer tilstrekkelig i fremtiden gjennom legalisering av større informasjonsflyt.
Bekreftelse på attestasjon Viktige aspekter ved internkontrollen skal være transparent og lett å vertifisere	Papirdokumentasjon som inkluderer attestasjoner øker dokumentasjonens fullstendighet	Ved elektronisk attestasjon kan det være behov for ytterligere vertifisering	Kontroll av kundens IT system med vekt på autorisasjonsmatriser og toleransegrenser kan bekrefte attestasjoner.	
Brukervennlighet Brukervennlige programmer og tilganger støtter etterlevelse (compliance)	Papirbasert bevis kan vanligvis evalueres uten bruk av tilleggsverktøy og/eller ferdigheter	Elektroniske bevismaterialet kan muligens kreve bistand fra en ekspert for å innhente data	Digitale bevis kan genereres fra standardiserte maler og verktøy. Ved innhenting av data brukes spesialiserte og øker brukervennligheten.	Vi ser indikasjoner på at fremtidig revisjonsteam i større grad vil bestå av IT spesialister som assisterer revisor for å øke brukervennligheten.
Klarhet Et bevis anses som kompetent når en annen revisor kan gjenskape beviset og trekke samme konklusjon	Papirbasert dokumentasjon er klar og dermed intuitiv for en annen revisor	Elektronisk bevis er ikke alltid like klar for en annen revisor, spesielt ved fravær av passende kontroller	Digitale bevis kan gjenskapes ved at koder eller skript av datasettet lagres og gjør beviset etterprøvbart. Dette sikrer klarhet.	

Som tabellen viser dekkes alle egenskapene for digitale bevis per i dag. Hvor godt egenskapene dekkes varierer derimot fra egenskap til egenskap. Troverdighet vil for eksempel dekkes indirekte gjennom flere av de samme handlingene som reduserer risiko for endring. God internkontroll sett opp mot IT er gjennomgående for alle egenskapene men med ulike vinklinger. Autorisasjonsmatriser og toleransegrenser er sentralt for å dekke bekreftelse på attestasjoner mens fullstendighet dekkes hovedsakelig ved å se flere kilder fra internkontrollen samlet. Standardiserte maler og verktøy gjør det lettere for revisor selv å utarbeide digitale bevis. På en annen side anvendes spesialister når revisoren ikke har kompetanse til å handtere større mengder data og har behov for assistanse. Til slutt ser vi at klarhet sikres gjennom koder eller skripter som gjør revisjonsbeviset etterprøvbart.

I sum er det flere funn som reflekterer at tradisjonelle handlinger fortsatt anvendes for å bekrefte egenskapene til digitale revisjonsbevis. Dette kan ha sammenheng med at Big Data som kilde per i dag ikke er vesentlig forskjellig fra tradisjonelle kilder. Forskjellen ligger hovedsakelig i at mengden av data som genereres er større. På grunn av de små forskjellene mellom tradisjonell data og Big Data i revisjon per i dag kan det derfor tenkes at tradisjonelle handlinger fortsatt er tilstrekkelige for å dekke egenskapene i rammeverket.

Det utvidede rammeverket viser også handlinger som kan anvendes for digitale revisjonsbevis i fremtiden. Disse handlingene består blant annet av blockchain og digitale signaturer som reduserer risiko for endring. I tillegg vil økt informasjonsflyt muliggjør avstemming mellom flere forskjellige kilder som i sin tur vil øke troverdigheten og fullstendigheten av digitale revisjonsbevis. Til slutt vil økt bruk av IT spesialister i revisjonsteamene øke brukervennligheten.

Tabell 5

Oppsummering av digitale revisjonsbevis

Egenskaper til revisjonsbevis	Digitale bevis i dag	Digitale bevis i fremtiden
Risiko for endring	✓✗	✗
Troverdighet	✓✗	✗
Fullstendighet	✓✗	✗
Bekreftelse på attestasjon	✓	
Brukervennlighet	✓✗	✗
Klarhet	✓	

- ✓ Egenskapen sikres
- ✓✗ Egenskapen sikres til e viss grad
- ✗ Egenskapen kan sikres mer tilstrekkelig i fremtiden

Som tabell 5 belyser dekkes de ulike egenskapene i rammeverket i varierende grad. Handlingene som vedrører bekreftelse på attestasjoner og klarhet fremstår konkrete nok til at de dekker egenskapene i tilstrekkelig grad. For de resterende egenskapene har undersøkelsen resultert i overordnede revisjonshandlinger. Som følge av at handlingene kun gir indikasjoner på at egenskapene dekkes kan vi si at egenskapene sikres til en viss grad, men ikke absolutt. Til slutt oppsummerer tabellen hvilke egenskaper som har potensialet til å bli dekket mer tilstrekkelig i fremtiden gjennom handlinger som fremkommer i undersøkelsen og som er nevnt i avsnittet ovenfor.

6. Konklusjon

Målet med denne oppgaven er å svare på problemstillingen: *Hvordan sikrer revisor påliteligheten av revisjonsbevis når Big Data og Big Data Analytics anvendes i revisjon?* Gjennom å undersøke hvordan Big Data og Big Data Analytics defineres i revisjon og hvordan egenskaper til revisjonsbevis blir møtt når revisor anvender Big Data og Big Data Analytics skal vi svare på problemstillingen.

Som følge av at Big Data i revisjon per i dag hovedsakelig består av store mengder strukturert data med liten grad av variasjon kan det ansees tilstrekkelig at tradisjonelle revisjonshandlinger tas i bruk for å dekke flere av egenskapene i rammeverket. På en annen side har undersøkelsens natur ført til flere overordnede funn som gjør det vanskelig å konkludere med at egenskapen er sikret tilstrekkelig. For egenskaper hvor funn har resultert i mindre konkrete handlinger vil resultatene kun gi indikasjoner på at de er blitt sikret.

I sum kan vi konkludere med at revisor sikrer påliteligheten til digitale revisjonsbevis, generert fra strukturert data med liten grad av variasjon, hovedsakelig gjennom handlinger som stammer fra tradisjonell revisjonsmetodikk. Dette inkluderer blant annet avstemming, tilgangskontroller og revisjonsspor. Samtidig ser vi starten på anvendelsen av nye typer revisjonshandlinger som kan inkludere endringslogger, skripter og metadata. Til slutt kan det nevnes at påliteligheten til fremtidige digitale revisjonsbevis har potensiale til å bli dekket mer tilstrekkelig gjennom blockchain, høyere informasjonsflyt og spesialkompetanse.

Som følge av oppgavens beskrivende natur kan undersøkelsen bli sett på som en kartlegging til videre forskning. Blant annet vil det være interessant å se videre på hvilken type struktur dataen som anvendes for revisjonsformål består av. Vil revisjonsbransjen bevege seg mot bruk av mer ustrukturert data som innehar høyere variasjon? I så fall, hvilke utfordringer må revisor ta stilling til sett opp mot troverdigheten av dataen. Blir troverdigheten av dataen svekket desto mer ustrukturert dataen er, og hvordan tar revisor stilling til denne utfordringen? Det kan i sin tur knyttes opp mot utfordringene revisor står ovenfor hvis regulatoriske myndigheter åpner opp for en høyere grad av informasjonsflyt. Da vil revisor få tilgang til et høyere antall eksterne kilder som kan brukes til blant annet avstemming. Til tross for at eksterne kilder har sine styrker vil det også

medbringe utfordringer rundt bekreftelse av troverdighet og sannferdighet av data. Til slutt vil det vært interessant å sett på teknologien som ligger bak blockchain og hvor i revisjonsmetodikken revisor eventuelt kan nyttiggjøre seg av denne i fremtiden.

Litteraturliste

Alfasoft. (2018). *Hva er NVivo*. Hentet fra <https://www.alfasoft.com/no/produkter/statistikk-og-analyse/nvivo.html> (lest 04 mai 2018).

Andersen, L. M., Bakkelig, M. (2015). Hva er Big Data, og hva betyr Big Data for deg? Hentet fra <https://www.pwc.no/no/publikasjoner/information-management/big-data.pdf>

Apache Spark. (2018). Hentet fra <https://spark.apache.org/> (lest 13 mai 2018).

Appelbaum, Deniz. (2016) Securing Big Data Provenance for Auditors: The Big Data Provenance Black Box as Reliable Evidence. *Journal of Emerging Technologies in Accounting*, 13(17), 17-36. Doi: 10.2308/jeta-51473.

Appelbaum, D., Kogan, A., & Vasarhelyi, M. (2017). Big Data and Analytics in the Modern Audit Engagement: Research Needs. *Auditing: A Journal of Practice & Theory*, 36(4), 1-27.

Aurstad Thorolf. (2017) Revisjonsbevis i en digital hverdag. *Bilag til Revisjon og Regnskap, 2017* (nr. 7), 26-27. Hentet fra <https://www.revregn.no/asset/blapdf/YmxhcGRm-MjAxNy8wNw/index.html#p=96>.

Brown-Libur, H., Issa, H., & Lombardi, D. (2015). *Behavioral implications of Big Data's impact on audit judgment and decision making and future research directions*. *Accounting Horizons*, 29(2), 451-468.

Brown-Libur, H., and M. A. Vasarhelyi.(2015). Big Data and audit evidence. *Journal of Emerging Technologies in Accounting* 12 (1): 1–16.

Cao, M., Chychyla, R., & Stewart, T. (2015). Big Data analytics in financial statement audits. *Accounting Horizons*, 29 (2), 423-429.

Dun & Bradstreet. (2018). *Our Company*. Hentet fra <https://www.dnb.com/> (lest 1 juni 2018).

Eckblad, B., Husby, M., Sæter K. & Tobiassen M. (2018). Advokat om Tidal-manipulasjon: Kan rammes av bedrageri-paragraf. *Dagens Næringsliv*. Hentet fra <https://www.dn.no/etter-Bors/2018/05/14/1123/Musikk/advokat-om-tidal-manipulasjon-kan-rammes-av-bedrageri-paragraf>

IAASB (2016) Høringsutkast: *Exploring the Growing Use of Technology in the Audit, with a Focus on Data Analytics*. Hentet fra <https://www.ifac.org/publications-resources/exploring-growing-use-technology-audit-focus-data-analytics>.

IBM. (2018) *Apache Hadoop: Built for big data, insights, and innovation*. Hentet fra <https://www.ibm.com/analytics/hadoop> (lest 13 mai 2018).

ISA 200 (2017). *Overordnede mål for den uavhengige revisor og gjennomføringen av en revisjon i samsvar med de internasjonale revisjonsstandardene*. Bergen: Fagbokforlaget

ISA 230 (2017) *Revisjonsdokumentasjon*. Bergen: Fagbokforlaget

ISA 330 (2017) *Revisors håndtering av anslåtte risikoer*. Bergen: Fagbokforlaget

ISA 500 (2017) *Revisjonsbevis*. Bergen: Fagbokforlaget

Kvale, Brinkmann, Anderssen, Rygge, Brinkmann, Svend, Anderssen, Tone Margaret, & Rygge, Johan. (2015). *Det kvalitative forskningsintervju* (3.utg., 2. oppl. ed.). Oslo: Gyldendal akademisk.

Lacoma, Tyler. (2017). What is an RSS feed and how can I use it? Hentet fra <https://www.digitaltrends.com/computing/what-is-an-rss-feed/>

Laney, D. (2001). 3D Data management: Controlling data volume, velocity and variety. *Meta Group*. Hentet fra <https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf>

Liao, C., and A. Squicciarini. (2015). Towards provenance-based anomaly detection in MapReduce. In Proceedings of the 15th IEEE/ACM International Symposium on *Cluster, Cloud and Grid Computing (CCGrid)*, 647–656. IEEE. Hentet fra https://ieeexplore.ieee.org/document/7152530/?tp=&arnumber=7152530&url=http%2F%2Fieeexplore.ieee.org%2Fxppls%2Fabs_all.jsp%3Farnumber%3D7152530

Lillebekk, B., Lyngstad, C. P. (2015). KPMGs Big Data muligheter. s. 2. Hentet fra <https://home.kpmg.com/content/dam/kpmg/pdf/2016/07/BigDataMulighetsstudium-produkt-ark.pdf>

Lindquist, R. (1991). Don't forget the pilot work! *Heart and Lung: Journal of Critical Care*, 20(1), 91-92.

Lukoianova, T., & Rubin, V. L. (2014). Veracity roadmap: Is big data objective, truthful and credible? Hentet fra <https://ir.lib.uwo.ca/fimspub/62/>

Malsch, B., and S. E. Salterio. (2016). “Doing good field research”: Assessing the quality of audit field research. *Auditing: A Journal of Practice & Theory* 35 (1): 1–22. doi:10.2308/ajpt-51170

Russom, P. (2011). Big data analytics. *TDWI best practices report, fourth quarter*, 19(4), 1-34.

Schroeck, M., Shockley, R., Smart, J., Romero-Morales, D., & Tufano, P. (2012). Analytics: the real-world use of big data: How innovative enterprises extract value from uncertain data, Executive Report. *IBM Institute for Business Value and Said Business School at the University of Oxford*.

Taylor, M., Haggerty, J., Gresty, D., & Hegarty, R. (2010). Digital evidence in cloud computing systems. *Computer Law & Security Review*, 26(3), 304-308.

PYMNTS. (2017). Hentet fra <https://www.pymnts.com/global-fraud-index/> (lest 13 april 18).

Quin, L. (2016). *Extensible Market Language*. Hentet fra <https://www.w3.org/XML/>

Vedlegg: Big Data og Big Data Analytics i Revisjon

Intervjuguide for Masteravhandling

Informasjon om gjennomføring av intervju

Vi er to studenter fra HHUiS som studerer master i regnskap og revisjon. Vi skriver masteravhandling om **Big Data og Big Data Analytics i revisjon** som er to viktige temaer på standardsetternes agenda (IAASB Høringsutkast 2016). Formålet med intervju spørsmålene er å dekke 2 temaer:

- 1) Hvordan definerer dere Big Data og Big Data analytics i revisjonssammenheng,
- 2) Hvordan dere håndterer / ville håndtert pålitelighetskravet for digital revisjonsbevis.

- Det er 14 intervju spørsmål (forskningsspørsmål) som er utarbeidet og forankret i revisjonsteori og forskning. Først vil vi stille noen generelle spørsmål om stilling og erfaring med Big Data og Big Data Analytics. Deretter vil vi fortsette med 14 intervju spørsmål om våre to temaer.
- Vi ønsker å benytte lydopptak for å sikre at vi får all nødvendig og relevant informasjon. Lydopptaket skal kun benyttes av undertegnede for gjennomføring av oppgaven. Etter innlevering vil lydopptaket bli slettet.
- Deltakelse er frivillig og informasjon fra intervjuet vil bli anonymisert.
- Vi setter pris på eksempler fra praksis.
- Anslått tidsbruk 30 min.

Studentene, HHUiS

Vårin Bjørnø

Malin Nordhus

Veileder, NHH

Carmen Olsen



Universitetet
i Stavanger

Big Data og Big Data Analytics i Revisjon

Intervjuguide for Masteravhandling

Generelle spørsmål

Posisjon:

Erfaring med Big Data og Big Data Analytics:

Intervjuspørsmål

Tema 1: Definisjon av Big Data og Big Data Analytics i revisjon

Big Data er informasjon som kjennetegnes med høyt volum, høy hastighet (velocity), høy grad av variasjon (variety) og troverdighet (veracity) og som krever nye metoder for prosessering og tilrettelegging for å kunne fungere som grunnlag for forbedrede beslutninger, økt innsikt og optimering av prosesser.

Big Data Analytics er prosessen hvor man inspiserer, vasker, transformerer og modellerer Big Data for å oppdage og formidle nyttig informasjon og mønstre, foreslå konklusjoner og støtte opp under beslutningstaking.

1. Ut i fra definisjonene av Big Data og Big Data Analytics presentert her, er du enig i definisjonen eller ønsker du å tilføye/ta bort noe?
2. Generer ditt revisjonsselskap revisjonsbevis fra Big Data Analytics?
3. Hvilken type revisjonshandling anser dere Big Data Analytics for å være? (substanshandling, analytisk handling, test av kontroller)
4. Hvor i revisjonsprosessen anvender ditt revisjonsselskap Big Data og Big Data Analytics?
5. Når dere innhenter Big Data for revisjonsformål er kilden (altså dataen) lokalisert internt hos revisjonskunden eller eksternt?
6. Hvilke kilder av Big Data anvendes i dataanalysen?

Tema 2: Påliteligheten av revisjonsbevis

Rammeverket til Appelbaum (2016), som er vedlagt, er et teoretisk forslag for seks egenskaper til revisjonsbevis som er ment å dekke pålitelighetskravet og illustrer hvordan papirbasert- og elektronisk bevis kan oppfylle disse egenskapene. Vi vil i vår masteroppgave utvide dette teoretiske rammeverket med revisjonspraksis rundt revisjonsbevis generert fra Big Data, heretter digitale bevis.

Risiko for endring:

Digitale bevis bør ikke kunne endres uten et revisjonsspor eller et annet bevis på denne endringen.

Spørsmål:

7. På hvilken måte sikrer dere egenskapen “risiko for endring”? Gjerne utdyp/konkretiser med et eksempel.

Troverdighet:

Dagens revisjonsstandard hevder at eksterne kilder øker troverdigheten når det innhentes uavhengig av kunden og kan bekreftes av revisor. Digitale bevis utfordrer dette standpunktet. En faktor som kan øke troverdigheten til digitale bevis generert fra eksterne kilder er sikker registrering av dataens opprinnelse, endring av dataen og eventuell metadata som revisor har tilgang til.

Spørsmål:

8. På hvilken måte sikrer dere egenskapen troverdighet av digital dokumentasjon innhentet intern fra revisjonsklienten?
9. På hvilken måte sikrer dere egenskapen troverdighet av digital dokumentasjon innhentet eksternt fra revisjonsklienten?

Fullstendighet:

På lik linje med elektronisk bevis kan det være utfordrende å verifisere fullstendigheten av en alle transaksjon når dataen for transaksjonen ikke er tilgjengelig for revisor eller foreligger i et annet format enn tekstformat.

Spørsmål:

10. På hvilken måte sikrer dere egenskapen fullstendighet?

Bekreftelse på attestasjon:

Ved digitale bevis kan attestasjoner for internkontrollen registreres som metadata.

Spørsmål:

11. På hvilken måte sikrer dere bekreftelse på attestasjoner?

Brukervennlig:

Digitale revisjonsbevis kan kreve standardiserte rutiner og verktøy for å utvinne data på en brukervennlig måte.

Spørsmål:

12. Hvordan utvinner dere Big Data på en brukervennlig måte?

13. På hvilken måte er de analytiske verktøyene som anvender Big Data brukervennlige?

Klarhet:

Digitale bevis anses som kompetente når en annen revisor kan gjenskape det og ut fra beviset trekke samme konklusjon

Spørsmål:

14. Hvordan sikrer dere egenskapen "klarhet"?

Tusen takk for at du har tatt deg tid til å svare

Hvis du i ettertid kommer på noe relevant for vår avhandling, ta gjerne kontakt på e-mail: vaarin.haaland@hotmail.com eller telefon: 98030686.

Rammeverk (Appelbaum, 2016, s.30)

Egenskaper til revisjonsbevis	Papirbasert bevis	Elektronisk bevis
<i>Risiko for endring</i> Bevis som kan endres fra sin opprinnelige form mangler troverdighet; revisjonsbevis bør være vanskelig å endre.	Det er vanskelig å endre papirbasert dokumentasjon uten av det blir oppdaget	Endring av elektronisk bevis kan være utfordrende å oppdage uten å gjennomføre nærmere bestemte tester
<i>Troverdighet</i> Dagens revisjonsstandard hevder at eksterne kilder øker troverdigheten når det innhentes uavhengig av kunden og kan bekreftes av revisor	Eksterne kilder av papirbasert bevis innhentet direkte av revisor øker troverdigheten	Internkontroller øker troverdigheten til elektronisk dokument. Ekstern elektronisk dokumentasjon mangler sikkerheten internkontrollene medbringer; at beviset ikke er falskt eller endret.
<i>Fullstendighet</i> Alle essensielle sider av en transaksjon skal være vertifiserbare	Vanligvis er de viktigste vilkårene inkludert i tekst format	Et elektronisk system kan erstatte tekstformat med koder og kryssreferanser til annen data som ikke er tilgjengelig for revisor
<i>Bekreftelse på attestasjon</i> Viktige aspekter ved internkontrollen skal være transparent og lett å vertifisere	Papirdokumentasjon som inkluderer attestasjoner øker dokumentasjonens fullstendighet	Ved elektronisk attestasjon kan det være behov for ytterligere vertifisering
<i>Brukervennlighet</i> Brukervennlige programmer og tilganger støtter etterlevelse (compliance)	Papirbasert bevis kan vanligvis evalueres uten bruk av tilleggsverktøy og/eller ferdigheter	Elektroniske bevismaterialet kan muligens kreve bistand fra en ekspert for å innhente data
<i>Klarhet</i> Et bevis anses som kompetent når en annen revisor kan gjenskape beviset og trekke samme konklusjon	Papirbasert dokumentasjon er klar og dermed intuitiv for en annen revisor	Elektronisk bevis er ikke alltid like klar for en annen revisor, spesielt ved fravær av passende kontroller