



BI Norwegian Business School - campus Oslo

GRA 19204

Master Thesis in Accounting and Auditing

Thesis Master of Science

Kunstig intelligens i revisjon

En studie av muligheter og utfordringer ved bruk av kunstig intelligens i revisjon

Navn: Alexander Kobberød, Henrik Berge Braaten

Start: 01.01.2018 09.00

Finish: 03.09.2018 12.00

Masteroppgave ved Handelshøyskolen BI

Master i regnskap og revisjon

Kunstig intelligens i revisjon

En studie av muligheter og utfordringer ved bruk av kunstig intelligens i revisjon

Innleveringsdato: 27.08.2018

Veileder: Flemming Ruud

Alexander Kobberød
Henrik Berge Braaten

“Denne oppgaven er gjennomført som en del av Master i regnskap og revisjon ved Handelshøyskolen BI. Dette innebærer ikke at Handelshøyskolen BI går god for de metoder som er anvendt, de resultater som er fremkommet, eller de konklusjoner som er trukket.”

Førord

Denne masteravhandlingen representerer vårt avsluttende arbeid på master i regnskap og revisjon (MRR) ved Handelshøyskolen BI. Dette har vært selvstendig, lærerikt og krevende, men samtidig et interessant arbeid. Prosessen med å avlegge en masteravhandling var en ny og annerledes erfaring for oss, da vi jobbet intensivt med kun ett prosjekt over flere måneder.

Vi ønsket å skrive om et dagsaktuelt tema innenfor revisjon og gjerne digitalisering. Temaet kunstig intelligens (KI) ble forslått av PwC. Vi ønsket å bygge kunnskap utenfor det vi fikk gjennom ordinær studieplan - om hvordan revisjon kan forvente å utvikle seg. Samtidig ser vi på dette som en innledende omstilling vi vil ta med oss inn i arbeidshverdagen, med positiv holdning og innstilling til KI i revisjon. Kunnskap om KI innenfor revisjon vil være nyttig å ta med seg videre, hvor vi i tillegg til å se muligheter, også er klar over flere utfordringer.

Vi ønsker å takke våre intervjukandidater for deres tid, samt bidrag til oppgaven vedrørende deres erfaringer, innsikt og innspill. Vi ønsker også å takke PwC for deres scholarship, hvilket omfattet verdifull veiledning og løpende kontakt underveis. Av den grunn fortjener vår veileder ved PwC, Ronny Lysmen, en takk for oppfølging og gode tips til oppgaven. Andre involverte parter med innspill, råd og motivasjon fortjener også en takk.

Vi må til slutt takke vår veileder Flemming Ruud, som har gitt oss kontinuerlig konstruktiv tilbakemelding på vårt arbeid. Vi takker for all hjelp med blant annet valg av metode, fokus på struktur for oppgaven og å prege oppgaven vår i en riktig retning.

Handelshøyskolen BI, Oslo, 27. august 2018

Henrik Berge Braaten

Alexander Kobberød

Sammendrag

Profesjonen revisjon er i endring. Kunstig intelligens (KI) er et dagsaktuelt tema i flere bransjer hvor man ser potensialet og de mulighetene den type teknologi kan gi. Samtidig er det beheftet utfordringer rundt bruk av KI. Formålet med denne studien er å avdekke muligheter og utfordringer ved bruk av KI i revisjon.

Problemstillingen belyses gjennom 22 forskningsspørsmål forankret i litteratur.

Litteratursøk med hensiktsmessige søkekriterier som tilfredsstillende KI og revisjon har bidratt til å avdekke relevant litteratur og forskning i tråd med studiens formål. En grundig og robust litteraturgjennomgang har vært nødvendig for tilstrekkelig og hensiktsmessig forståelse av fenomenet. Fordelene og mulighetene viser at KI kan gi styrket tillit ved å heve effektiviteten, revisjonskvaliteten og gi økt innsikt gjennom å innhente, analysere og tolke store datasett raskere. Samtidig er motargumenter og utfordringer om vi har tillit til KI, regulatoriske hindringer, frykten for rettstvister, samt mangelfull kompetanse, forståelse og dokumentasjon av den interne strukturen og dermed etterprøvbareheten. Vi avdekket mangelfull litteratur på spesifikke standarder som begrenser KI og KI til å utøve skjønn og skepsis.

Studiens formål belyses gjennom en pragmatisk tilnærming gjennom intervju og spørreskjema, men med hovedfokus på den kvalitative forskningen. Studien er basert på fire store og et mellomstort selskap, samt Den norske Revisorforening. Våre resultater viser at de overnevnte momentene er reelle og relevante. På en annen side utfyller de litteraturen på KI til å utøve skjønn og skepsis, der det påpekes at dette gjelder lengre frem i tid. Vår empiri viser motstridende argumenter for standardenes begrensning på KI. Selv om revisjonsbransjen sakte har tilpasset seg ny teknologi viser våre resultater at det å overse nye innovative teknologiske løsninger kan «utkonkurrere» bransjen. Dette fordrer derimot at man er varsomme og trygge på de verktøyene man tar i bruk. Dette understøttes av vår forskning, hvor studien impliserer at proaktive revisjonsselskap som fremmer sine analytiske og kognitive evner, er de som vil lykkes i markedet.

Vår konklusjon er at det er beheftet en rekke muligheter og utfordringer rundt temaet KI i revisjon. Vi håper derfor at denne studien kan bli et viktig bidrag for akademikerne, utøverne, standardsetterne og tilsynsmyndighetene, i tillegg for oss selv, i forskning på og utvikling av en mer moderne revisjonsprofesjon.

Innholdsfortegnelse

| | |
|--|-----|
| Forside | I |
| Forord | II |
| Sammendrag | III |
| Innholdsfortegnelse | IV |
| Oversikt over figurer, tabeller, formler og forkortelser | VI |
| 1. Innledning | 1 |
| 1.1 Bakgrunn for valg av oppgave..... | 1 |
| 1.2 Problemstilling og formålet med oppgaven..... | 3 |
| 1.3 Oppbygging av oppgaven..... | 3 |
| 1.4 Utvikling og forskning..... | 4 |
| 1.5 Avgrensninger | 5 |
| Litteraturgjennomgang: Kapittel 2-4 | 6 |
| 2. Revisjon | 6 |
| 2.1 Revisors rolle og formål | 8 |
| 2.2 God revisjonsskikk | 9 |
| 2.3 Generelle prinsipper | 10 |
| 2.4 Revisjonsprosessen | 11 |
| 2.5 Revisjonshandlinger | 13 |
| 2.6 Revisjonsbevis | 16 |
| 2.7 Digitalisering av revisjon - En introduksjon til KI | 18 |
| 3. Kunstig intelligens | 20 |
| 3.1 Begrepet intelligens | 21 |
| 3.2 Oppbygging og algoritmer..... | 23 |
| 3.3 Maskinlæring | 24 |
| 3.4 Nevrale nettverk og dyp læring | 27 |
| 4. En revisjon med kunstig intelligens..... | 29 |
| 4.1 Muligheter | 30 |
| 4.1.1 Avansert analyse..... | 32 |
| 4.1.2 Fra datamining til innsikt..... | 34 |
| 4.1.3 Revisjonsprosessen → Sanntidsrevisjon | 37 |
| 4.1.4 Revisjonshandlinger | 40 |
| 4.1.5 Moderne revisjonsbevis | 48 |
| 4.2 utfordringer | 50 |
| 4.2.1 KI-risiko og datahåndtering..... | 50 |
| 4.2.2 Skjønn, skepsis og kompetanseutvikling..... | 54 |
| 4.2.3 Revisjonsstandardene | 57 |

| | | |
|-------|---|-----|
| 4.2.4 | Tilstrekkelig og hensiktsmessig revisjonsbevis | 63 |
| 4.2.5 | Etterprøvbarhet og dokumentasjonskrav | 65 |
| 5. | Forskningsmetode | 71 |
| 5.1 | Forskningsdesign | 71 |
| 5.2 | Valg av metode | 74 |
| 5.3 | Datainnsamling | 75 |
| 5.3.1 | Intervjukandidater | 78 |
| 5.3.2 | Rekrutteringsprosess | 80 |
| 5.3.3 | Gjennomføring av intervju | 80 |
| 5.4 | Forskerens utfordringer | 81 |
| 5.5 | Validitet og Reliabilitet | 83 |
| 6. | Resultat og analyse | 86 |
| 6.1 | Innledende analyse | 86 |
| 6.2 | Forsknings spørsmål 1 | 87 |
| 6.3 | Forsknings spørsmål 2 | 91 |
| 6.4 | Forsknings spørsmål 3 | 96 |
| 6.5 | Forsknings spørsmål 5 | 98 |
| 6.6 | Forsknings spørsmål 6 & 7 | 99 |
| 6.7 | Forsknings spørsmål 8 | 103 |
| 6.8 | Forsknings spørsmål 9 | 104 |
| 6.9 | Forsknings spørsmål 10 | 106 |
| 6.10 | Forsknings spørsmål 11 | 109 |
| 6.11 | Forsknings spørsmål 12 | 111 |
| 6.12 | Forsknings spørsmål 13 | 113 |
| 6.13 | Forsknings spørsmål 14 | 115 |
| 6.14 | Forsknings spørsmål 17 | 116 |
| 6.15 | Forsknings spørsmål 18 | 118 |
| 6.16 | Forsknings spørsmål 19 | 119 |
| 6.17 | Forsknings spørsmål 21 | 125 |
| 6.18 | Forsknings spørsmål 22 | 129 |
| 7. | Avslutning og konklusjon | 131 |
| 7.1 | Oppsummering og konklusjon | 131 |
| 7.2 | Teoretiske og praktiske implikasjoner ved KI i revisjon | 133 |
| 7.3 | Oppgavens begrensninger | 135 |
| 7.4 | Videre forskning | 136 |
| | Referanseliste | 139 |
| | Vedlegg | 151 |
| | Utstedt intervjuguide | 151 |
| | Transkribering av intervju | 155 |

| | |
|----------------------------|-----|
| Transkripsjon: A | 155 |
| Transkripsjon: B | 164 |
| Transkripsjon: C | 174 |
| Transkripsjon: D | 188 |
| Transkripsjon: E | 198 |
| Transkripsjon: F | 211 |
| Utfylte spørreskjema | 217 |
| Stata. | 223 |

Oversikt over figurer, tabeller, formler og forkortelser

Figurliste

| | | |
|---|--|-----|
| Figur 1: Assurance- and nonassurance services (Arens & Beasley, 2013, s. 32).... | 6 | |
| Figur 2: Trepartsforholdet (Arens et al., 2016, s. 32: figur 1-2)..... | 7 | |
| Figur 3: Forenklet modell av revisjonsprosessen (Arens et al., 2016, s. 25) | 11 | |
| Figur 4: Kategoriserte definisjoner på KI..... | 22 | |
| Figur 5: Læringsprosessen under veiledet læring (oversatt) (Liu, 2011, s. 66).... | 24 | |
| Figur 6: AlphaGo Zeros spillprogressjon | 26 | |
| Figur 7: Nevralt nettverk (Datatilsynet, 2018) | 28 | |
| Figur 8: Fra data til innsikt (ACCA, 2015, s. 11) | 36 | |
| Figur 9: Sanntidsrevisjonsprosessen (Chan & Vasarhelyi, 2011, s. 158)..... | 39 | |
| Figur 10: Venndiagram - Relevans, tilgjengelighet & pålitelighet (CPA, 2016, s. 6)..... | 51 | |
| Figur 11: Frekvenstabell, påstand 3..... | 90 | |
| Figur 12: Frekvenstabell, påstand 5..... | 92 | |
| Figur 13: Frekvenstabell, påstand 4 | Figur 14: Frekvenstabell, påstand 6..... | 97 |
| Figur 15: Frekvenstabell, påstand 8..... | 99 | |
| Figur 16: Frekvenstabell, påstand 1 | Figur 17: Frekvenstabell, påstand 2.. | 100 |
| Figur 18: Frekvenstabell, påstand 7..... | 117 | |
| Figur 19: Frekvenstabell, påstand 9..... | 121 | |
| Figur 21: Frekvenstabell, påstand 10..... | 128 | |

Tabelliste

| | |
|---|-----|
| Tabell 1: Analytiske metoder og teknikker | 45 |
| Tabell 2: Fra tradisjonelle til moderne revisjonshandlinger | 47 |
| Tabell 3: Påliteligheten av revisjonsbevis | 65 |
| Tabell 4: Metodisk tilnærming til forskningsspørsmålene | 73 |
| Tabell 5: Intervjukandidater | 79 |
| Tabell 6: Deskriptiv statistikk om påstand 3 | 90 |
| Tabell 7: Deskriptiv statistikk om påstand 5 | 92 |
| Tabell 8: Deskriptiv statistikk om påstand 4 | 97 |
| Tabell 9: Deskriptiv statistikk om påstand 6 | 97 |
| Tabell 10: Deskriptiv statistikk om påstand 8 | 98 |
| Tabell 11: Deskriptiv statistikk om påstand 1 | 100 |
| Tabell 12: Deskriptiv statistikk om påstand 2 | 100 |
| Tabell 13: Deskriptiv statistikk om påstand 7 | 117 |
| Tabell 14: Deskriptiv statistikk om påstand 9 | 121 |
| Tabell 15: Deskriptiv statistikk om påstand 10 | 127 |

Formler

| | |
|---|----|
| Formel 1: Revisjonsrisikomodellen (Gulden, 2016, s. 119)..... | 12 |
|---|----|

Forkortelser

| | |
|-------|---|
| AICPA | American Institute of Certified Public Accountants |
| DnR | Den norske Revisorforening |
| FRC | The Financial Reporting Council |
| IAASB | The international Auditing and Assurance Standards Board |
| ICAEW | The institute of Chartered Accountants in England and Wales |
| IFAC | International Federation of Accountants |
| ISA | International Standards on Auditing |
| KI | Kunstig intelligens |
| NLG | Naturlig språkgenerering |
| NLP | Naturlig språkbehandling |
| PCAOB | Public Company Accounting Oversight Board |
| RADAR | Rutgers AICPA Data Analytics Research Initiative |

1. Innledning

Vi skal i dette kapittelet gi en innledning til det valgte temaet og bakgrunn for valg av oppgave. Vi vil presentere vår problemstilling og hvordan vi ønsker å besvare denne. Pågående forskning og nødvendige avgrensninger blir presentert.

1.1 Bakgrunn for valg av oppgave

Temaet i vår masteravhandling er kunstig intelligens (KI) i revisjon. Den tradisjonelle revisjonen er i endring som følge av teknologi, digitalisering og mer og lettere datatilgang. KI innebærer sofistikert teknologi og er et verktøy med både positive og negative egenskaper. Næringslivet er ofte tidlig ute med å implementere og teste ny teknologi i sine prosesser. Av den grunn ønsker vi å undersøke muligheter og utfordringer ved å benytte KI i revisjon og revisjonsprosessen. Vi ønsker blant annet å utforske hvilke deler av revisjonen som kan utføres av KI og hvor man kan støte på utfordringer. Vår oppgave og presentert litteratur omhandler primært ekstern revisor og den eksterne finansielle revisjonen.

Formålet med revisjon er å beskytte brukerne av regnskapet gjennom å øke tilliten til informasjonen ved å avgi en bekreftelse, etter en kritisk gjennomgang av informasjonen selskapet fremlegger. En kritisk gjennomgang av informasjon betyr at revisor skal utføre revisjonen i henhold til god revisjonsskikk. Herunder skal revisor innhente og evaluere revisjonsbevis, ved å utføre revisjonshandlinger. I tillegg skal revisor bidra til å forebygge og avdekke misligheter og feil, samt påse at den revisjonspliktige har ordnet formuesforvaltningen på en betryggende måte med forsvarlig kontroll.

KI er et overordnet begrep for flere underliggende teknologier. Flere av disse presenteres i denne studien. KI har til hensikt å programmere maskiner til å etterligne menneskelig intelligens. Dette skal gi maskiner, roboter og dataprogrammer intelligent atferd. Implementering av KI har til hensikt å understøtte eller utføre arbeidsoppgaver, som tradisjonelt har blitt utført manuelt.

Teknologiske innovasjoner kan medføre radikale endringer for næringslivet. Trond Morten Lindberg (2017, s. 5) hevder at 40 % av dagens virksomheter i USA vil vært borte om ti år og at 65 % av dagens barn vil arbeide i jobber som enda ikke er oppfunnet. Denne utviklingen omtales som den fjerde industrielle revolusjonen, eller industri 4.0 (Schwab, 2017). En konsekvens ved denne revolusjonen vil være at over fem millioner mennesker globalt vil miste jobbene sine innen 2020 (Forum, 2016, s. 13). Fjørtoft (2018) viser til undersøkelser som sier at revisjonsassistenter med nær 100 % sannsynlighet er utryddet om fem år og KPMG sier de kan automatisere 80 % av et revisjonsoppdrag allerede nå, men at regulatoriske krav setter en begrensning på dette foreløpig (Asklund, 2017). Dette understøttes av Torkil Hindberg ved PwC som mener at regulatoriske krav ikke tillater at KI benyttes direkte i revisjonsprosessen og at en person må kontrollere tallene (Grandal, 2018).

Industri 4.0 omhandler disruptive teknologier og utnyttelsen av synergiene mellom dem. Det handler om å digitalisere industrien, hvor smart og mer effektiv produksjon, samt prosesser står i fokus. Revisjonsbransjen er intet unntak fra denne utviklingen ettersom aktører som ser potensialet i ny teknologi kan ha mulighet til å produsere samme produkt på en smartere og rimeligere måte. Enn så lenge har ekstern revisor gjennom revisorloven, monopol på den endelige bekreftelsen av den finansielle informasjonen (Ruud, 2000, s. 20). Brukerne og andre interessenter stiller høye krav til revisjonsprofesjonen. Derfor er det viktig å følge den teknologiske utviklingen for å opprettholde relevansen og eksistensen i markedet ved å imøtekomme interessentenes forventninger. De største revisjons-selskapene har startet sine investeringer og tar tunge investeringer innen teknologi og heriblant KI.

Ettersom flere av de største revisjonsselskapene har startet sine investeringer og satser på en revisjon med KI, er dette et spesielt interessant tema for oss som nå avslutter MRR-studiet. Ved å avlegge en masteravhandling som omhandler teknologi, digitalisering og fremtidens revisjon med fokus på KI, vil være fordelaktig for oss som skal inntre som revisorer i revisjonsbransjen, ved å kjenne til en rekke muligheter og utfordringer beheftet rundt KI. Denne master-avhandlingen gjenspeiler av den grunn en studie med revisjon som det drivende, men som igjen forsøkes å tilpasses til en revisjon med KI som verktøy.

1.2 Problemstilling og formålet med oppgaven

Vi er på vei inn i en ny teknologisk æra hvor betegnelsene evolusjon og revolusjon er brukt (IAASB, 2016, s. 19; Kinserdal, 2017). KI er her en viktig bakenforliggende faktor. Av den grunn har vi formulert følgende problemstilling:

Kunstig intelligens i revisjon

Hvilke muligheter og utfordringer medfører bruk av kunstig intelligens i revisjon?

For å belyse problemstillingen vil vi utlede forskningsspørsmål fortløpende basert på gjennomgått litteratur. Vårt mål og ønske med denne oppgaven er å belyse muligheter og utfordringer rundt bruk av KI i den eksterne finansielle revisjonen.

1.3 Oppbygging av oppgaven

Dette er blitt en omfattende oppgave da vi har valgt å gå dypt inn i tilgjengelig litteratur på området. Vi ønsker med dette å klargjøre hvor og hvordan KI kan benyttes i revisjon, samtidig som vi forklarer de ulike nøkkelelementene. Vi mener dette er viktig for å sette ting i sammenheng og forstå KI-konteksten bedre, sett i lys av bruken i en revisjonssammenheng.

Denne studien består av syv kapitler. Den første delen gir en innledende oversikt til oppgaven og presenterer studiens problemstilling og formål. Kapittel 2 presenterer relevant revisjonsteori, mens kapittel 3 fokuserer på relevant teori rundt KI. Kapittel 4 ser kapittel 2 og 3 i sammenheng, hvor tidligere forskning og litteratur er presentert. Det er i denne delen snittområdet mellom KI og revisjon gjennomgås. Kapittel 5 forklarer anvendt metode og hvordan våre forskningsspørsmål skal besvares. Her presenteres vår forskningsetikk, datainnsamling og utvelgelse av informanter. Kapittel 6 presenterer og drøfter resultatene i vår forskning gjennom å se funn mot litteratur. I denne delen svarer vi på våre forskningsspørsmål med bakgrunn i funn fra empiri og litteratur. Vi oppsummerer avslutningsvis problemstillingen og vår samlede konklusjon i kapittel 7. Kapittelet impliserer den praktiske betydningen av KI fremover i revisjon, hvor forslag til videre forskning også foreslås.

1.4 Utvikling og forskning

Gjennom grundig research har vi funnet litteratur på KI i revisjon. Vi avdekket at det i noen sammenhenger synes noe i litteraturen mellom KI og andre teknologier som ikke er fullt så intelligente. Dette gjelder eksempelvis Robotic Process Automation (RPA), en teknologi hvor KI ikke direkte anvendes, men kan få stor nytteverdi innen revisjon alene eller sammen med KI.

På enkelte områder eksisterer det mangelfull litteratur. Dette gjelder eksempelvis spesifikke revisjonsstandarder som tillater eller begrenser KI, samt bruk av KI til å utøve profesjonelt skjønn og skepsis. Ettersom relevant forskningslitteratur er presentert i kapittel 4, tildeles denne delen til å presentere igangsatt forskning og prosjekter.

Universitet Rutgers har flere pågående prosjekter innen forskning av den mer moderne revisjonen¹. Rutgers har blant annet et prosjekt som går på utvikling av en kognitiv (intelligent) revisjonsassistent i revisjon (Li & Vasarhelyi, u.å.). I tillegg har de fra 2016, i samarbeid med andre interessenter, igangsatt prosjektet RADAR. Vasarhelyi (2017) presenterer at RADAR består av følgende tre pågående prosjekter:

- 1) MADS – et rammeverk som fokuserer på eksepsjonelle avvik og guider revisor i identifisering av mange avvik og prioriteringen av dem (Rutgers & AICPA, u.å.-a).
- 2) Prosessmining – fokuserer på prosessmining til å evaluere internkontroll (Rutgers & AICPA, u.å.-b).
- 3) Visualisering som revisjonsbevis – et prosjekt hvor visualisering og analytiske handlinger, omtalt som visuell analyse, kan benyttes som revisjonsbevis (Rutgers & AICPA, u.å.-c). Det er besluttet at dette prosjektet skal implementeres i de to overnevnte².

¹ <http://raw.rutgers.edu/carlab/Research-Projects.html>

² <https://www.aicpa.org/interestareas/frc/assuranceadvisoryservices/radarprojects.html>

I tillegg har forskningsrådet bevilget ti millioner kroner til NHH som skal benyttes til å forske på endringene og effektene av digitalisering, herunder KI, innen revisjonsbransjen. Denne forskningen skjer i samarbeid med de fem største revisjonsselskapene (Kinserdal, 2017).

1.5 Avgrensninger

En viktig avgrensning i denne masteravhandlingen er at den omhandler og fokuserer på KI. Andre potensielle viktige teknologier har blitt sett bort fra. Dette gjelder eksempelvis RPA, blockchain og andre automatiseringsteknologier. Vi har ikke sett på anvendelsesområdet for KI i andre bransjer og derfor ikke trukket paralleller mellom disse mot revisjonsbransjen. Dette gjelder tilsvarende for generalisering av vår studie til andre typer revisjon og revisorer. Vår kompetanse strekker ikke til det tekniske og matematiske bak KI. Programmering og koding av algoritmer, statistikk og matematikk er derfor ikke behandlet. Dette er derfor en oppgave i revisjon som reflekteres gjennom hva vi kan som revisorer. Av den grunn er oppgaven strukturert med hensyn til revisjon hvor vi forsøker å se betydningen av KI.

I en studie som fokuserer på muligheter og utfordringer er det ikke anledning til å dekke enhver mulighet og utfordring som kan eksistere. Vi har imøtekommet dette ved å fokusere på de mulighetene og utfordringene som litteraturen og forskningen presenterer opp mot vår empiri. Vi har blant annet valgt å se bort fra utfordringer som omhandler KI opp mot ny GDPR-forordning og andre personvernregler. En annen viktig avgrensning i denne studien er at den omhandler den eksterne finansielle revisjonen. Likevel er mye av presentert litteratur og funn er overførbare til andre typer av revisjon.

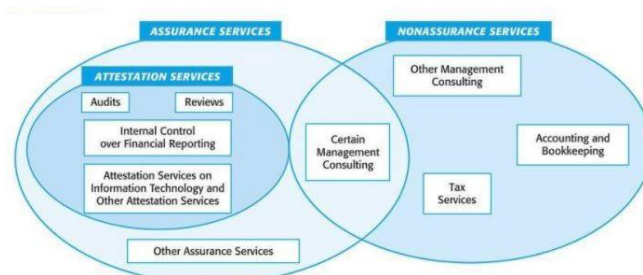
Chatbots drevet av KI benytter ofte maskinlæring, NLP og NLG. Slik lærer maskinen å respondere raskt og effektivt på brukerens behov. Chatbots kan benyttes til intern og ekstern bruk, enten det er administrative, revisjonstekniske eller som revisjonsassistent til å purre klienter, i revisjonshandlinger eller som en kunnskapsbase for revisor. Vi har ikke hatt anledning til å drøfte dette i denne studien. Vi har også valgt å se bort fra trusler vedrørende cyberangrep og andre lignende trusler som eksisterer i en mer teknologisk og digital æra.

Litteraturgjennomgang: Kapittel 2-4

2. Revisjon

Revisjon kommer fra latin og betyr «ettersyn/gjensyn». Begrepet inneholder et element av kontroll, men har allsidig betydning. Utviklingen av revisjon gjenspeiles gjennom samfunnets økende behov for kontroll. Innledningsvis påpekte vi at det er flere ulike former for revisjon og flere som foretar revisjon (Birkeland, 2017, s. 35 & 43; Gulden, 2016, s. 21-23; Johansen, 2015, s. 30-31).

Assurance services er ikke likt med rådgivningstjenester (business advisory/nonassurance services) som er et toveis forhold mellom partene. Assurance services er uavhengige, profesjonelle tjenester som forbedrer informasjonskvaliteten, eller dens sammenheng, for beslutningstakere (Sekundærkilde: Ruud, 2000). Assurance services har blitt et overordnet begrep for ulike revisjonstjenester, eksempelvis operasjonell-, finansiell-, etterlevels-, systemsikkerhets-, miljø- og due diligence-revisjoner med formål å forbedre informasjonskvaliteten eller dens sammenheng (Ruud, 2000, 2001). Attestasjonsoppdrag er en underkategori av assurance services som omfatter den eksterne finansielle revisjonen, hvor revisors hovedoppgave er å gi assurance/bekreftede innholdet i årsregnskapet (Arens, Elder & Beasley, 2016, s. 33; Kristoffersen, 2014, s. 226). Ruud (2000) understreker at begrepet assurance ikke må forveksles med det norske begrepet «assurans» som betyr forsikringsforhold- og ansvar.



Figur 1: Assurance- and nonassurance services (Arens & Beasley, 2013, s. 32)

Formålet med finansiell rapportering er å gi beslutningsnyttig informasjon til ledelsen og brukerne. For at denne skal være relevant må den være tidsriktig og uten vesentlige feil (Chan & Vasarhelyi, 2011, s. 2; FASB, 2006). I denne sammenheng er revisors relevans, rolle og formål av betydning og har en viktig rolle i samfunnet som en leverandør av tillit.

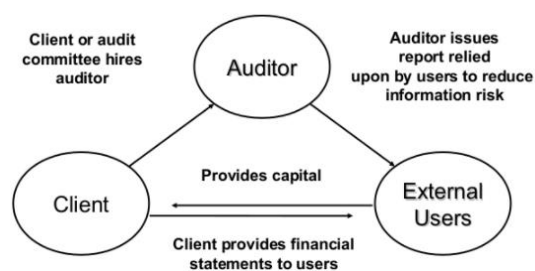
I videre fremstilling av begrepene revisjon og revisor vil vi forholde oss til teori og litteratur rundt ekstern revisor og den eksterne finansielle revisjonen, altså den lovpålagte revisjonen.

Revisjon defineres av Arens et al. (2016, s. 28) som:

Revisjon er å innhente og evaluere akkumulerte revisjonsbevis for å konkludere hvorvidt regnskapet er i overensstemmelse med etablerte kriterier og rapportere i hvilken grad det er korrespondanse mellom bevis og kriterier. Revisjon bør bli utført av en kompetent og uavhengig person (egen oversettelse).

Dette er en av flere definisjoner av revisjon da revisorloven ikke inneholder noen legaldefinisjon av begrepet. Gulden (2016, s. 28 flg.) understreker at det er et viktig moment at den som utfører ekstern revisjon er uten tilknytning til foretaket. Ekstern revisjon er underlagt strenge lovkrav, både til revisors egenskaper og ved utførelse av revisjon. Revisorloven kapittel 5 regulerer revisors oppgaver ved revisjon av årsregnskapet.

Etterspørselen til ekstern revisjon kan forklares av to faktorer. Den første faktoren er vilkårene for lovpliktig revisjon i revisorloven § 2-1. Den andre faktoren definerer Arens et al. (2016, s. 30) som informasjonsrisiko. Revisjon bidrar til verdifull innsikt for eierne og er med på å redusere revisjonsklientens agentkostnader, herunder asymmetrisk informasjon (Agent-prinsipal-teorien³), og informasjonsrisikoen i kapitalmarkedet. En uavhengig ekstern revisjon av et foretak bidrar til en effektiv overvåkning av hele organisasjonen og skaper tillit i kapitalmarkedet. Revisjon er et attestasjonsoppdrag og kjennetegnes som et trepartsforhold bestående av praktiserende revisor, ansvarlig part og tiltenkte brukere (figur 2). Dette er med på å belyse formålet med revisjon og rollen til revisor.



Figur 2: Trepartsforholdet (Arens et al., 2016, s. 32: figur 1-2)

³ (Bøhren, 2011, s. 28-31)

2.1 Revisors rolle og formål

Revisors rolle er definert i revisorloven § 1-2 som allmennhetens tillitsperson, og stammer fra EU direktivets artikkel 21 (Ot.prp.nr.78, 2008-2009, s. 24). Bedrifter avgir finansiell informasjon av stor betydning for brukerne, som er bakgrunnen for revisors rolle. For å ivareta rollen som allmennhetens tillitsperson kreves det at revisor har høy troverdighet, reflektert gjennom revisorloven, ISA-standardene og de yrkesetiske prinsippene.

Revisors primære oppgave er i henhold til revisorloven § 5-1 å vurdere hvorvidt årsregnskapet er utarbeidet i samsvar med lov og forskrifter. Revisor skal gjennom revisjonen bidra til å forebygge og avdekke misligheter og feil, i tillegg til å påse at den revisjonspliktige har ordnet formuesforvaltningen på en betryggende måte med forsvarlig kontroll. Revisor skal ytterligere utføre revisjonen etter beste skjønn, jf. § 5-2, og i samsvar med god revisjonsskikk. Lovgivningen er lite konkret vedrørende krav til utførelsen av en revisjon og hva den innebærer (Gulden, 2016, s. 30; Revisorutvalget, 1997, s. 27 flg).

Formålet med revisjon er definert i ISA 200 punkt 3:

Formålet med revisjon er å øke de tiltenkte brukernes tillit til regnskapet. Dette oppnås ved at revisor gir uttrykk for en mening om hvorvidt regnskapet i det alt vesentlige er utarbeidet i samsvar med det gjeldende rammeverket for finansiell rapportering. I de fleste rammeverk med generelt formål gir denne meningen uttrykk for hvorvidt regnskapet i det alt vesentlige gir en dekkende fremstilling, eller gir et rettvisende bilde i samsvar med rammeverket. En revisjon gjennomført i samsvar med ISA-ene og relevante etiske krav, gjør det mulig for revisor å danne seg en slik mening.

Revisors overordnede mål ved revisjon supplerer formålet med revisjon. Revisors formål og overordnede mål reflekteres i revisjonsberetningen, i form av en gjengivelse av revisor ansvar i forbindelse med revisjonen. ISA 200 punkt 11 utleder revisors overordnede mål på følgende vis:

- (a) å oppnå betryggende sikkerhet for at regnskapet totalt sett ikke inneholder vesentlig feilinformasjon, verken som følge av misligheter eller feil, og dermed gjøre det mulig for revisor å gi uttrykk for en mening om hvorvidt regnskapet i det alt vesentlige er utarbeidet i samsvar med et gjeldende rammeverk for finansiell rapportering; og
- (b) å rapportere om regnskapet, og å kommunisere slik det kreves i ISA-ene, i samsvar med revisors funn.

2.2 God revisjonsskikk

Revisorloven § 5-2 andre ledd krever at revisjon skal utføres i samsvar med god revisjonsskikk. NSRFs (Norges Statsautoriserte Revisorers Forening) norm om grunnleggende prinsipper for revisjon § 2, definerer god revisjonsskikk slik:

God revisjonsskikk er å utføre revisjonsoppdrag i overensstemmelse med den oppfatning av etiske og revisjonstekniske prinsipper som til enhver tid er alminnelig anerkjent og praktisert av dyktige og ansvarsbevisste utøvere av yrket.

Denne normen ble avløst av de nye revisjonsstandardene, International Standards on Auditing (ISA), men definisjonen har fortsatt relevans. ISA-ene kan sies å være representative for hva dyktige og ansvarsbevisste utøvere av yrket ville gjort (Knudsen, Siebke, Cordt-Hansen & Revisorforeningen, 2010, s. 189). Kravet til god revisjonsskikk omfatter all virksomhet som revisor utfører for sin klient, jf. revisorloven § 1-1 tredje ledd. God revisjonsskikk har en supplerende funksjon til lovens øvrige bestemmelser (Revisorutvalget, 1997, s. 116), og ikke til hensikt at loven kan tolkes innskrenkende når det gjelder revisjonens innhold og utførelse etter god revisjonsskikk (Ot.prp.nr.75, 1997-1998, s. 35). Den er ment å være som en rettslig standard som innebærer at innholdet er under kontinuerlig utvikling og forbedring (Ot.prp.nr.78, 2008-2009, s. 37). Den rettslige standarden er primært forankret til utførelsen og innholdet av et revisjonsoppdrag. Revisjonsstandardene er utarbeidet av IAASB og videre oversatt av DnR. Standardene har stor betydning for bransjen da de har høy anerkjennelse og mye anvendt i praksis. Standardsetterne utformer innholdet til profesjonen og skal sikre et dynamisk og oppdatert innhold i tråd med samfunnsutviklingen.

Standardenes betydning i rettspraksis er derimot noe omstridt. Domstolene er ikke bundet av slike normer selv om de gir uttrykk for hva ansvarlige og dyktige utøvere av yrket mener er forsvarlig (Nygaard, 2007, s. 206). Dette er gjeldende på tross av at god revisjonsskikk er forankret i loven. Innholdet avgjøres derfor i siste instans i domstolene. Revisjonsstandardene er tilsynelatende det som blir mest og hyppigst vektlagt av domstolene i forståelsen av begrepet god revisjonsskikk. Dette betyr derimot ikke at standardene er uttømmende for innholdet av begrepet god revisjonsskikk. I domstolene har også andre kilder blitt lagt til grunn for forståelsen av innholdet i den rettslige standarden (Birkeland, 2017, s. 154). Innholdet i god revisjonsskikk omfatter også uskrevne normer og praksis som dyktige og ansvarsbevisste utøvere av yrket fastsetter eller etterlever (Birkeland, 2017, s. 153). Uttalelser fra Finanstilsynets tilsynsmetoder er også med på å avgjøre forståelsen av begrepet god revisjonsskikk.

2.3 Generelle prinsipper

Revisor er underlagt strenge rammevilkår for å bevare tilliten i markedet. Revisorloven har flere detaljbestemmelser knyttet til god revisorskikk, i utførelsen av sitt yrke. Revisjonsstandardene definerer og utdyper innholdet av de generelle prinsippene og er nærmere beskrevet i DnRs regler om etikk. Regler om etikk og etterlevelse av de grunnleggende etiske prinsippene er essensielle for å kunne opprettholde tilliten til revisorer og deres bekreftelser. Ifølge DnRs regler om etikk skal revisor etterleve følgende grunnleggende prinsipper; integritet, objektivitet, faglig kompetanse og tilbørlig aktsomhet, konfidensialitet og profesjonell atferd (Revisorforeningen, 2009c, s. 4). I tillegg er uavhengighet sentralt for å kunne utføre oppdraget med integritet, objektivitet og opprettholde en profesjonelt skeptisk holdning. Revisor plikter etter revisorloven § 5-2 å gjennomføre revisjonen etter beste skjønn, men loven nevner ikke begrepet skepsis. ISA 200 har definert disse begrepene på følgende måte:

Punkt 13 (k): Profesjonelt skjønn

Anvendelsen av relevant opplæring, kunnskap og erfaring innenfor rammen av standarder for revisjon, regnskap og etikk når det fattes informerte beslutninger om hvilke handlingsplaner som er hensiktsmessige etter omstendighetene ved revisjonsoppdraget.

Punkt 13 (l): Profesjonell skepsis

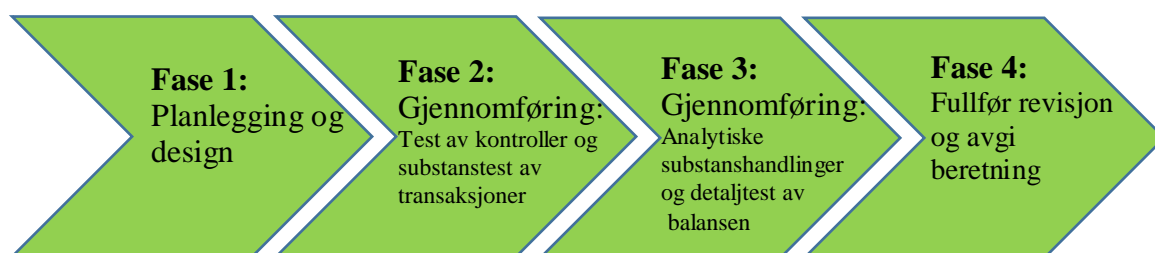
En holdning som innebærer at revisor stiller spørsmål og er oppmerksom på forhold som kan indikere mulig feilinformasjon som følge av feil eller misligheter, og foretar en kritisk vurdering av revisjonsbevis.

Disse egenskapene er avgjørende for å opprettholde en kritisk holdning ved bekreftelse av forelagt informasjon. Begge er viktige å anvende i hele utførelsen av et revisjonsoppdrag da presentert informasjon kan inneholde både bevisste og ubevisste feil eller mangler. Det stilles derfor krav om at revisor anvender mer enn bare ren kunnskap, hvor revisor ikke kan ukritisk følge standardene, i sin utførelse (Birkeland, 2017, s. 150). Begrepet profesjonelt skjønn må ses i sammenheng med revisors overordnede mål, i vurdering av revisjonsrisikoen og ved utforming og gjennomføring av revisjonshandlinger for å redusere revisjonsrisikoen til et akseptabelt lavt nivå (ISA 200 A25). Revisors profesjonelle skjønn baseres på vurderingen av risiko og den interne kontrollen på området og andre forhold revisor er kjent med, for å oppfylle kravene i revisorloven § 5-1 og § 5-2 (Birkeland, 2017, s. 56). Profesjonell skepsis kan derfor sies å ligge underbetont i profesjonelt skjønn, som er å anse for å være overordnet for god revisjonsskikk (Birkeland, 2017)

2.4 Revisjonsprosessen

Revisjonsprosessen er kjernen av profesjonen, hvor metodikken og utførelsen av revisjon, blir utviklet av bransjen selv (Birkeland, 2017, s. 52). Denne selvreguleringen beskriver Birkeland som en metodefrihet, og gjelder utførelsen av et revisjonsoppdrag. Oppdraget utøves med revisors profesjonelle skjønn og skepsis i tråd med god revisjonsskikk. Birkeland deler revisjonsprosessen inn i tre faser: planlegging, gjennomføring og konklusjon/rapportering (2017, s. 52).

Revisjonsprosessen i figur 3 er utledet fra Arens et al., (2016).



Figur 3: Forenklet modell av revisjonsprosessen (Arens et al., 2016, s. 25)

Etter oppdragsvurdering og aksept av klientforholdet starter planleggingsfasen. I denne fasen fastsettes en overordnet revisjonsstrategi, herunder en vurdering og planlegging av vesentlighet, risiko og angrepsvinkel. En revisjonsplan utarbeides for de ulike revisjonsområdene på bakgrunn av den overordnede revisjonsstrategien. Revisjonsplanen er en kontinuerlig prosess som oppdateres fortløpende gjennom hele revisjonen. I utførelsesfasen gjennomfører revisor revisjonshandlinger for å innhente tilstrekkelig og hensiktsmessig revisjonsbevis som legger grunnlaget for revisors konklusjon. I følge Arens et al. (2016) er denne fasen todelt, med fokus på revisjon av internkontroll (IK) og resultatposter gjennom året og balansepostene rundt balansedagen. I den avsluttende fasen avgir revisor sin uavhengige konklusjon, revisjonsberetningen, etter å ha opparbeidet seg betryggende sikkerhet på grunnlag av en kritisk vurdering av innhentet revisjonsbevis (Gulden, 2016, s. 55-56). I kapittel 4.1.3 kommer vi tilbake til hvordan KI kan endre og effektivisere revisjonsprosessen.

ISA 300-serien inneholder krav til planleggings- og gjennomføringsfasen. For å fastslå risiko for vesentlig feilinformasjon foretar revisor risikovurderingshandlinger i henhold til ISA 315 punkt 6. Revisjonsrisikomodellen kan utledes indirekte fra ISA-standardene og benyttes i revisjonen som et virkemiddel for å planlegge og gjennomføre revisjonen på en måleffektiv måte i tråd med ISA 300 punkt 4. ISA 200 punkt 13 c, sier at revisjonsrisiko er «risikoen for at revisor gir uttrykk for en uriktig mening i revisjonsberetningen når regnskapet inneholder vesentlig feilinformasjon. Revisjonsrisiko er en funksjon av risikoene for vesentlig feilinformasjon og oppdagelsesrisiko». Revisor må vurdere risiko for vesentlig feilinformasjon (iboende- og kontrollrisiko) og derav styre oppdagelsesrisikoen for å redusere revisjonsrisikoen til et akseptabelt lavt nivå. Iboende risiko er risikoen for at en regnskapspåstand kan inneholde vesentlig feilinformasjon når man ser bort fra selskapets internkontroll. Revisjonsrisikomodellen består av følgende fire risikokomponenter:

| |
|---|
| $\text{Revisjonsrisiko} = \text{Iboende risiko} * \text{Kontrollrisiko} * \text{Oppdagelsesrisiko}$ |
|---|

Formel 1: Revisjonsrisikomodellen (Gulden, 2016, s. 119)

2.5 Revisjonshandlinger

For å konkludere hvorvidt regnskapet er fri for vesentlig feilinformasjon må revisor utføre revisjonshandlinger for å innhente revisjonsbevis. Revisjonshandlingene benyttes til å innhente bevis for å gjøre revisor i stand til å bekrefte at revisjonsmålsettingene (ledespåstandene), iht. ISA 315 punkt A131, er uten vesentlige feil. Revisjonshandlingene styrer oppdagelsesrisikoen og omtales som risikohåndteringshandlinger og angrepsvinkel, med hensikt å redusere revisjonsrisikoen til et akseptabelt lavt nivå. Dette gjøres ved test av kontroller og substanshandlinger, herunder detaljtester og analytiske substanshandlinger. ISA 330 regulerer og fastsetter krav til gjennomføringen av revisjonshandlinger for å håndtere anslåtte risikoer. Den ideelle revisjonen er å kontrollere alle transaksjonene gjennom detaljtester, men ville medført et høyt ressursbruk og høyt honorar. Detaljtester omfatter revisors fysiske inspeksjon, granskning av dokumentasjon mot underliggende forhold, forespørsler og gjentakelse. Gjennom substanshandlinger reduserer revisor oppdagelsesrisikoen (Gulden, 2016, s. 158).

ISA 500 punkt A14-A22 nevner følgende revisjonshandlinger for hva slags bevis som kan innhentes. I delkapittel 4.1.4 blir disse tradisjonelle revisjonshandlingene mer diskutert i en KI-kontekst hvor de til slutt blir sammenstilt mot en mer moderne utførelse av dem.

- Inspeksjon – består i å undersøke regnskapsmateriale eller dokumenter, interne eller eksterne, i papirform, elektronisk form eller annet medium. Omfatter også fysisk undersøkelse av en eiendel.
- Observasjon – består i å betrakte en prosess eller en rutine som utføres av andre, for eksempel revisors observasjon av varetelling eller annen kontrollaktivitet.
- Ekstern bekreftelse – er et direkte skriftlig svar til revisor fra en tredjepart. For eksempel bekreftelse av kontosaldo, avtalevilkår eller transaksjoner inngått med tredjepart i papirform, elektronisk eller annet medie.
- Etterregning – består i å kontrollere den matematiske nøyaktigheten av dokumenter eller regnskapsmateriale og kan utføres manuelt eller elektronisk.

- Gjentakelse – revisors uavhengige utførelse av rutiner eller kontroller som opprinnelig ble utført som en del av enhetens interne kontroll.
- Analytiske handlinger – består i å evaluere finansiell informasjon ved å studere mulige sammenhenger mellom både finansielle og ikke-finansielle data. Omfatter også undersøkelse av identifiserte fluktuasjoner og avvik.
- Forespørsler – består i å innhente informasjon fra kyndige personer, interne eller eksterne. Kan være formelle og skriftlige eller uformelle og muntlige.

Kontrollaktiviteter er en av komponentene i COSO-modellen for internkontroll⁴. Kontrollaktiviteter består av retningslinjer og rutiner som skal sikre at ledelsens former for risikohåndtering blir gjennomført og integreres med risikovurderingen for at selskapet skal nå sine målsettinger. En slik aktivitet kan utføres av et menneske eller ved hjelp av datateknologi (Kristoffersen, 2014, s. 51-52). Arens et al. (2016, s. 384) grupperer kontrollaktiviteter i følgende fem typer:

- Adekvat arbeidsdeling
- Autorisasjon av transaksjoner og aktiviteter på riktig nivå
- Tilfredsstillende dokumentering og registrering
- Fysisk kontroll over eiendeler og registreringer
- Uavhengig kontroll på de ulike aktivitetene

Test av kontroller har til hensikt å vurdere måleffektiviteten av virksomhetens internkontroll og kontrollaktiviteter for å forebygge, avdekke eller korrigere vesentlig feilinformasjon (ISA 330 punkt 4b). Denne risikoen omtales som kontrollrisiko. Det er et vilkår at kontrollene har fungert etter sin hensikt og er fulgt, for å understøtte revisors vurdering av kontrollrisikoen som lav/middels. Derav kan revisor redusere omfanget av substanshandlingene, som vil bidra til en måleffektiv revisjon (Gulden, 2016, s. 180). Kontrollhandlingene ved test av kontroller omfatter vanligvis inspeksjon, observasjon, gjentakelse og forespørsel (Arens et al., 2016, s. 450).

⁴ Kan lastes ned fra: <https://iia.no/produkt/internkontroll-et-integrert-rammeverk-sammendraget/>

Analytiske handlinger er analyse av plausible sammenhenger og definert i ISA 520 punkt 4. God revisjonsskikk krever at revisor benytter analytiske handlinger i planleggings- og avslutningsfasen i revisjonen. Analytiske handlinger i gjennomføringsfasen avhenger av revisors skjønn og vurderinger. De er mye anvendt i praksis, da de er kostnadseffektive og kan gi oversikt over hvor det kan foreligge vesentlige feil i informasjonen. Samtidig dekker analytiske substanshandlinger de fleste påstandene (Arens et al., 2016, s. 228; Gulden, 2016, s. 170). Svakheten er at de er relativt overfladiske og lite presise, da underliggende tall kan motregne hverandre. De gir lite bevis på påliteligheten av tallene mot underliggende bevis.

Formålet med analytiske handlinger varierer med revisjonsfasen de benyttes i, og kan omfatte både finansiell og ikke-finansiell informasjon. Yoon (2016, s. 53) presenterer fire steg for analytiske handlinger, understøttet av ISA 520:

1. Sett en forventningsverdi
2. Sett et tolererbart forventningsavvik (akseptabelt avvik)
3. Sammenlign forventede og faktiske verdier - undersøk betydelige avvik
4. Følg opp og vurder ledelsens kommentar på avvik - evaluér og konkluder

Analytiske handlinger kan brukes som revisjonsbevis dersom avviket mellom revisors forventning og klientens faktiske tall ligger innenfor vesentlighetsgrensen (Arens et al., 2016; Yoon, 2016, s. 50). Appelbaum, Kogan og Vasarhelyi (2018, s. 84-85) og Yoon (2016, s. 52) påpeker at analytiske substanshandlinger vil være mer effektivt og hensiktsmessig enn detaljtester i enkelte tilfeller. De fremhever at analytiske substanshandlinger vil være mer tilstrekkelig og hensiktsmessig revisjonsbevis og mindre kostbart enn detaljtester hvor det er store datasett med mange transaksjoner eller hvor risiko for vesentlig feilinformasjon er lav.

ISA 500 punkt A52 oppgir tre metoder for utvelgelse av elementer til å innhente revisjonsbevis. Utvelgelse av alle elementer (100% testing), utvelgelse av spesifikke elementer og stikkprøver. Valg av metode er basert på revisors skjønn, herunder identifisert risiko, internkontroll og vesentlighet knyttet til regnskapspostene, samt den praktiske gjennomførbarheten og effektiviteten av de ulike metodene. Den tradisjonelle revisjonen har vært stikkprøvebasert grunnet iboende begrensninger, kapasitet, tilgjengelighet, og krav om en måleffektiv revisjon. Det

er derimot nevneverdig at valg av metode for utvelgelse av elementer er en valgmulighet og at valget derfor avgjøres på bakgrunn av revisors vurdering.

2.6 Revisjonsbevis

Revisjonsbevis er definert i ISA 500 punkt 5 bokstav c:

«Informasjon brukt av revisor for å komme frem til konklusjonene som revisors mening bygger på. Revisjonsbevis omfatter både informasjon som finnes i regnskapsmaterialet som underbygger regnskapet, og informasjon innhentet fra andre kilder».

Revisjonsbevis innhentes primært gjennom revisjonshandlinger og består av to elementer. Regnskapsmateriale (ISA 500 punkt 5a) og informasjon innhentet fra andre kilder. ISA 500 definerer ikke «informasjon innhentet fra andre kilder», men gir noen antydninger på hva det innebærer i «veiledning og utfyllende forklaringer». Informasjon innhentet fra andre kilder er generelt ment til å underbygge regnskapsmaterialet og ledelsens påstander som innhentes fra tidligere revisjoner, kvalitetskontrollrutiner, analyser, avstemminger, møtereferat, uttalelser fra ledelsen, bekreftelser, rapporter og benchmarking data. Eksterne skriftlige bevis nær balansedagen gir sterke revisjonsbevis, mens interne muntlige vil gi vesentlig svakere bevis (Gulden, 2016, s. 173; ISA 500 A31).

Innsamlet revisjonsbevis danner grunnlaget for revisors konklusjon i revisjonsberetningen, hvorvidt regnskapet i det alt vesentligste er avlagt etter gjeldende rammeverk. ISA 500 punkt 6 krever at revisor utformer og utfører revisjonshandlinger som er hensiktsmessige ut fra omstendighetene, for å innhente tilstrekkelig og hensiktsmessig revisjonsbevis. Veiledningspunkt A1 - A25 gir lite veiledning i formuleringen “revisjonshandlinger som er hensiktsmessige ut fra omstendighetene”. Arens et al. vektlegger tre sentrale momenter i denne overnevnte vurderingen. Eliminer overlappende handlinger, utfør alle nødvendige handlinger ved selve innhenting og utfør handlingene i en hensiktsmessig rekkefølge for best mulig perspektiv (2016, s. 503). ISA 500 og 200 bygger dette rundt revisors skjønn og vurderinger av vesentlighet, risiko, internkontroll og utvelgelse, samt kost/nytte-betraktning rundt gjennomførbarheten og effektiviteten rundt innhenting, sett mot bevisets relevans og

pålitelighet. Nye avanserte dataverktøy kan være en ny omstendighet etter ISA 500 punkt 6. Slik teknologi kan bidra til mer effektive og hensiktsmessige revisjonshandlinger fremfor tradisjonelle handlinger.

ISA 500 punkt 6 krever at revisor innhenter tilstrekkelig og hensiktsmessig revisjonsbevis gjennom revisjonshandlinger. Tilstrekkelig er utdypet i veiledningspunkt A4 som et mål på kvantiteten av revisjonsbevis, altså omfanget av innhentet bevis. Veilednings-punkt A5 beskriver hensiktsmessighet som et mål på revisjonsbevisets kvalitet og består av relevans og pålitelighet. Relevans omhandler påstanden revisor ønsker å få bekreftet gjennom revisjonshandlingen og kontrollretningen (A27). Revisjonsbevisets relevans må ses mot den enkelte påstand ved vurdering av feilinformasjon da det kan være relevant for én påstand, men ikke for en annen. De fleste revisjonsbevis dekker flere, men ikke alle revisjonsmålsettingene. Pålitelighet omhandler hvorvidt revisor kan stole på beviset, bevisets troverdighet, dette avhenger av kilden, type og omstendighetene rundt innhenting (A31). En bekreftelse på hvorvidt regnskapet er uten vesentlige feil avhenger av at innhentet revisjonsbevis er overbevisende nok (Arens et al., 2016, s. 222). For å avgjøre om det er innhentet tilstrekkelig og hensiktsmessig revisjonsbevis som underbygger revisors konklusjon, må revisor anvende sitt profesjonelle skjønn (Birkeland, 2017, s. 56).

Revisorloven § 5-3 forplikter revisor til å dokumentere hvordan revisjonen er gjennomført og resultatet på en tilstrekkelig måte for å kunne underbygge og etterprøve revisors konklusjon, herunder bevisene. Dette innebærer at revisor må dokumentere på en måte som muliggjør etterprøving helt fra aksepteringsfasen til avslutningsfasen med avlagt revisjonsberetning (Ot.prp.nr.75, 1997-1998, s. 44). Gulden (2016, s. 403) beskriver at hovedkravet til form, innhold og omfang av arbeidspapirene skal være tilstrekkelig omfattende og detaljerte for å gi en totalforståelse av revisjonen. Dokumentasjonen skal være så omfattende at en annen dyktig revisor, uten kjennskap til oppdraget, skal få en forståelse av det utførte arbeidet og revisors beslutninger.

Kravet til etterprøvbarehet (re-performance) er viktig i revisjon. Dette innebærer at en annen profesjonell revisor skal kunne gjøre de samme handlingene og komme frem til samme konklusjon. ISA 230 angir prinsipielle krav til dokumentasjonen.

Forhold av vanskelig karakter eller hvor skjønn er utøvd stiller høyere krav til god dokumentasjon, inkludert de kjente omstendighetene. Udokumentert forståelse og kunnskap revisor har opparbeidet seg gjennom revisjonen om enheten er ikke tilstrekkelig revisjonsbevis ved etterprøving (Knudsen et al., 2010, s. 219-226). Dokumentasjonen skal derfor være tilstrekkelig for at en annen revisor, uten kjennskap til oppdraget, skal forstå type, tid, omfang og resultat av utførte revisjonshandlinger, samt vesentlige funn, nådde konklusjoner og profesjonelt skjønn utøvd i å trekke disse beslutningene fram til revisors konklusjon.

Dokumentasjon er det viktigste når det kommer til en erstatningssak og hvor retten legger til grunn at det som ikke er dokumentert er ansett for å ikke være utført (Gulden, 2016, s. 418). En gjennomgang av Finanstilsynets merknader og tilsyns- og kvalitetssaker viser at det stadig er mangel på eller utilfredsstillende dokumentasjon og bevis som ligger til grunn for revisors konklusjon. I delkapittel 4.2.5 belyser vi relevant litteratur rundt og problematiserer dokumentasjonskrav ved bruk av teknologiske verktøy med bruk av KI. utfordringer rundt hvordan en annen dyktig revisor kan komme fram til samme konklusjon, ved å anvende data- og KI-verktøy vil også diskuteres.

2.7 Digitalisering av revisjon - En introduksjon til KI

Presentert revisjonsteori reflekterer den tradisjonelle måten å utføre en revisjon på. I dag har vi den teknologiske utviklingen som påvirker mange ulike bransjer, og dermed i hvilken grad revisjonsbransjen har tilpasset seg dette. Vi vil i den anledning undersøke hvordan digitalisering og teknologi, herunder spesielt KI kan påvirke revisjon.

“Å digitalisere er ikke noe nytt” (Heggernes, 2017, s. 36). Likevel mener Heggernes at begrepet å digitalisere har endret seg de senere årene. I dag snakkes det oftere om digitalisering av forretningsdrift. Et eksempel er fra Financial Times som skriver at Nordea-konsernet kutter 6000 arbeidsplasser som følge av det teknologiske skiftet (Milne, 2017). Analysebyrået Gartner (2009) har utledet følgende definisjon av digitalisering:

«Digitalisering er å bruke digitale teknologier for å endre forretningsmodell og å finne nye måter å skape verdi og skaffe seg inntekter på, det er prosessen med å flytte til en digital forretning» (Oversatt av Heggernes, 2017).

Digitalisering handler om at informasjonsteknologi (IT) skal gå fra å være et verktøy, til å bli en del av bedriftens DNA. En undersøkelse utført av Forbes (2014) viser at bedriftsledere ikke tror at digitalisering er viktig for deres bedrift, fordi det har gått så bra hittil. Disse lederne kan bli offer for Amara's lov som antyder at mennesker har en tendens til å overvurdere teknologi på kort sikt, og undervurdere teknologien på lang sikt (Partington, 1996).

I dag er store deler av tilgjengelig informasjon digital. Denne informasjonen kan være tall, symboler eller bokstaver sammensatt i kombinasjoner. Transaksjoner i en virksomhet blir dokumentert digitalt og lagret i databaser (Heggernes, 2017, s. 16). Slike data inngår i samlebetegnelsen stordata (big data) og er sammensatt av strukturerte og ustrukturerte data. Stordata er store mengder data med tilnærmet uendelig tilgang på og kommer fra mange ulike kilder (Andersen & Bakkeli, 2015). Det kreves kunnskap for å utnytte slike data hensiktsmessig. I dag tilbys regnskapssystemer som er direkte integrert med bank, for eksempel Visma eAccounting og 24sevenoffice. I disse systemene blir transaksjoner direkte bokført i regnskapet, hvor bildegjenkjenning av faktura kan benyttes. Slik kan det manuelle arbeidet minimeres (Rammen, 2017). Trond M. Lindberg forteller at det eksisterer gode skyløsninger i regnskapsbransjen, mens revisjonsbransjen har mer effektiv bruk av stordata og dataanalyseverktøy (Asklund, 2016). Store endringer er forventet innen revisjonsbransjen som følge av digitalisering (Kinserdal, 2017).

På bakgrunn av økt digitalisering i flere bransjer, men også innen revisjon, vil vi i kapittel 3 se nærmere på hva KI er og består av, samt hvordan det fungerer. Videre i kapittel 4 presenterer vi relevant litteratur på hvor og hvordan KI kan benyttes i revisjon, og hvilke muligheter og utfordringer dette gir.

3. Kunstig intelligens

I dette kapittelet skal vi definere begrepet KI med utgangspunkt i begrepet intelligens. Vi vil se på oppbyggingen og hva KI består av. Vi skal deretter undersøke noen av de bakenforliggende teknologiene som muliggjør KI, herunder maskinlæring og nevralt nettverk (dyp læring).

Andrew Burgess (2018, s. 3) utleder åtte kjerneområder for bruk av KI: bildegjenkjenning, språkgjenkjenning, søk, clustering, naturlig språkbehandling, optimalisering, prediksjon og forståelse. De fire førstnevnte handler om å fange opp informasjon, for eksempel ved å lage strukturerte data ut fra ustrukturerte data. Naturlig språkbehandling, optimalisering og prediksjon forsøker, basert på informasjonen, å forklare hva som skjer. Forståelse handler om å forstå hvorfor noe skjer. I videre litteraturgjennomgang vil vi komme tilbake til flere av disse.

Ideen med å lage intelligente maskiner er gammel og forskningen startet i kjølvannet av den andre verdenskrigen. Alan Turing utformet Turing-testen⁵ i 1950 som fikk stor oppmerksomhet. Gjennom historien har det vært flere forsøk på å utvikle KI. Et eksempel er Japans forsøk på å utvikle KI på 80- og 90-tallet, hvor det ble investert to mrd. over en periode på ti år, uten store gjennombrudd (Tørresen, 2013, s. 19). Det arrangeres konkurranser innen KI, for eksempel den årlige konkurransen “Loebner Prize” ved Cambridge, med det formål å lage den beste maskinen som består Turing-testen (Nilsson, 2009, s. 62).

Vi har de siste årene sett omtrentlig en årlig dobling av data og stadig bedre program- og maskinvarer. Billigere lagring og maskinkraft, samt økt tilgjengelighet av data, sammen med konstant tilkobling til nettverk, har bidratt til i større grad å modne denne type teknologi (Burgess, 2018, s. 13, 17 og 19; Sandberg, 2017).

⁵ Rolstadås et al., 2017, s. 44; Russell, Norvig & Davis, 2016, s. 2

3.1 Begrepet intelligens

Forståelse av KI krever fortrolighet med begrepet naturlig intelligens. Det finnes flere definisjoner på intelligens. En av disse er utarbeidet av Gottfredson (1994):

”Intelligens er en svært generell mental evne som blant annet innebærer evnen til å resonnerer, planlegge, løse problemer, tenke abstrakt, forstå komplekse ideer, lære raskt og lære av erfaring...” (egen oversettelse).

“Definisjonene av intelligens har til felles at de ikke gir noen klare anvisninger om hvordan intelligens kan observeres og måles” (Sundet, 2015, s. 21). Måling av intelligens har blitt utført gjennom intelligenstester. Charles Spearman mente at prestasjoner på intelligenstester skyldes underliggende faktorer som krever intelligens, benevnt som g-faktor. De viktigste parameterne i g-faktoren er evnen til deduksjon, induksjon, problemløsning, forstå sammenhenger og regelmessigheter, i tillegg til å kunne se likheter og ulikheter (Referert i Sundet, 2015, s. 43). Howard Gardner har utvidet intelligensbegrepet til syv varianter av intelligens hos mennesker; språklig, logisk/matematisk, visuell, musikalsk, kroppslig og sosial intelligens og selvinnsikt. Med KI forsøkes det å inkorporere Gardners varianter av intelligens i maskiner, hvor fokuset primært har vært språklig, logisk/matematisk og visuell intelligens. KI kan derfor sies å være en samlebetegnelse på en rekke teknologier som skal få maskiner til å fremstå som intelligente (Gardner, 1999; Rolstadås, Krokan, Dyrhaug & Norges tekniske, 2017, s. 43).

Fagterminologien skiller mellom svak og sterk KI. Maskiner som simulerer intelligens benevnes som svak, mens maskiner som faktisk tenker (ikke simulerer tenkning) karakteriseres som sterk KI (Russell, Norvig & Davis, 2016, s. 1020). Kokina og Davenport (2017, s. 119) påpeker at det ikke er kognitiv teknologi med mindre maskinen er selvstyrende og evner å lære over tid. Rolstadås et al. (2017, s. 50) sier at vi er langt unna å lage maskiner med sterk KI, altså maskiner som har intelligens til å lære seg det meste. Flere av dagens tjenester har implementert KI, herunder maskinlæring og dyp læring. Eksempler på dette er IBM Watson, Google Deepmind, bilde-, lyd- og tekstgjenkjenning, selvkjørende kjøretøy og robotene Alice og Amelia. Disse kategoriseres som svak KI, da de kun utfører spesifikke oppgaver like bra eller bedre enn mennesker (Copeland, 2016).

Russell et al. (2016, s. 2) har samlet og kategorisert åtte definisjoner av KI (figur 4). De to øverste boksene reflekterer tankeprosesser og resonnement, mens de to nederste skildrer oppførsel. Venstre side gjelder menneskelig atferd, mens høyre side illustrerer den ideelle atferden, rasjonalitet.

| | |
|---|---|
| <p>Thinking Humanly</p> <p>“The exciting new effort to make computers think . . . <i>machines with minds</i>, in the full and literal sense.” (Haugeland, 1985)</p> <p>“[The automation of] activities that we associate with human thinking, activities such as decision-making, problem solving, learning . . .” (Bellman, 1978)</p> | <p>Thinking Rationally</p> <p>“The study of mental faculties through the use of computational models.” (Charniak and McDermott, 1985)</p> <p>“The study of the computations that make it possible to perceive, reason, and act.” (Winston, 1992)</p> |
| <p>Acting Humanly</p> <p>“The art of creating machines that perform functions that require intelligence when performed by people.” (Kurzweil, 1990)</p> <p>“The study of how to make computers do things at which, at the moment, people are better.” (Rich and Knight, 1991)</p> | <p>Acting Rationally</p> <p>“Computational Intelligence is the study of the design of intelligent agents.” (Poole <i>et al.</i>, 1998)</p> <p>“AI . . . is concerned with intelligent behavior in artifacts.” (Nilsson, 1998)</p> |

Figur 4: Kategoriserte definisjoner på KI

En intelligent maskin må beherske følgende fire funksjoner (Russell et al., 2016, s. 2):

- Naturlig språkbehandling for å kunne kommunisere på en god måte
- Kunnskapsrepresentasjon til å lagre informasjon
- Automatisert resonnement for å utnytte lagret informasjon for å løse oppgaver og problemer, samt trekke nye konklusjoner
- Maskinlæring for å tilpasse seg nye omstendigheter for å oppdage og utlede nye mønster

Innledningsvis definerte vi at KI skal gi maskiner, roboter og dataprogrammer intelligent atferd med hensikt å programmere de slik at de etterligner menneskelig intelligens. Det er denne definisjonen vi sikter til når vi snakker om KI i revisjon.

3.2 Oppbygging og algoritmer

Datatilsynet (2018) beskriver KI som en kompleks maskin uten særlig gjennomskiktighet av hva som faktisk foregår inne i den. For å forstå det som foregår i systemet er det viktig med riktig kompetanse. Kompetanse om oppbyggingen og bestanddelene til maskiner med KI vil fremover bli mer verdsatt (ICAEW, 2017a, s. 10). Oppbygging av KI består av en eller flere algoritmer med hensikt å reflektere naturlig intelligens. En algoritme kan defineres som “en prosess eller sett av regler som benyttes i kalkulasjoner eller andre problemløsningsoppgaver, særlig av en datamaskin” (Oxford, 2018 (egen oversettelse)).

Oppbygging og algoritme satt i sammenheng er definert som modellering. Modellering krever teknologisk kompetanse og statistikk ("NTNU Kveld: Kunstig intelligens: fantastisk eller farlig? [videoklipp]," 2018). Modellering kan eksempelvis utføres med matematisk metode (Fenstad, 2017, s. 183):

- 1: Analyse av et problem eller en oppgave med hensikt å danne en struktur
- 2: Språk og regler i form av ligninger, til å analysere strukturen og forstå hvordan den utvikler seg over tid
- 3: Algoritmer og implementering, med formål å gi presis prediksjon og beskrive handlinger som skal til for å løse problemet eller oppgaven

En struktur kan være et eller flere objekter, konkret eller abstrakt med visse relasjoner og egenskaper mellom objektene. Det kan være ord, tall, figurer og/eller mønstre. Språk er et verktøy for å beskrive og nyttiggjøre strukturene og sette disse i perspektiv (Fenstad, 2017, s. 183-184). «Algoritmer er ganske enkelt en samling, skritt-for-skritt instruksjoner som skal utføres rent mekanisk for å oppnå et ønsket resultat...» (Chabert et al., 2012, s. 1 (egen oversettelse)).

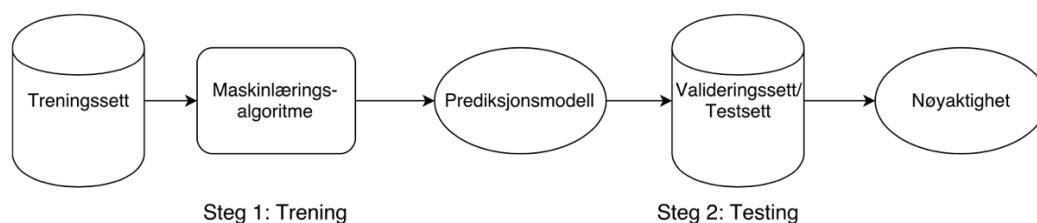
Algoritmen må være korrekt, som vil si at den har potensiale til å løse oppgaven. Algoritmer som produserer gale svar eller ikke fungerer hensiktsmessig har liten eller ingen verdi. Algoritmer er verktøy for å transformere en oppgave og struktur til et resultat, altså til handling eller innsikt (Fenstad, 2017, s. 185). Det er viktig å implementere og etablere en eller flere hensiktsmessige algoritmer som presterer bra på både testdata og nye input-data. Ytterligere må systemet overvåkes for å sikre kvalitet og forbedring, ettersom output sjeldent inneholder null feilmargin (Goodfellow, Bengio & Courville, 2016, s. 409-410).

3.3 Maskinl ring

Maskinl ring er en tiln rming for   oppn  KI gjennom systemer som kan l re av erfaring for   finne m nstre i ett sett data (Mayes, 2017, s. 10 (egen oversettelse)). Copeland (2016) skriver at maskinl ring handler om   benytte algoritmer til   analysere data, l re og deretter ta en beslutning eller predikere et utfall, basert p  informasjonen. Evnen til   l re og utf re en bestemt oppgave skjer ved at maskinen benytter l ringsalgoritmer og trenes p  store datasett, kalt treningsdata. Maskinl ring kan benytte store mengder data til   lage modeller for uthenting av beslutningsnyttig informasjon, l se en bestemt oppgave eller predikere utfall.

I delkapittel 2.7 poengterte vi  kningen i tilgjengelig data vi har hatt de siste  rene. Dette er informasjon fra ulike kilder som for eksempel kunder, nyheter, v rmeldinger, informasjon tilgjengelig p  internett og en rekke andre kilder. Desto mer data tilgjengelig vil kunne gi maskinen et bedre prediksjonsgrunnlag (Marsland, 2015, s. 5; Nilsson, 2009, s. 496).

Vi mennesker har en begrenset evne til   motta informasjon, mens maskiner har tiln rmet en ubegrenset evne til dette (Marsland, 2015, s. 2). Ved maskinl ring kan maskiner endre og tilpasse sine handlinger, og dermed bli enda mer nøyaktige underveis som de utsettes for ny informasjon. En handling kan som nevnt v re prediksjon. Nøyaktighet i prediksjon m les etter hvorvidt prediksjonen er i tr d med virkeligheten (Marsland, 2015, s. 4). Input-verdier og attributter (features) er egenskaper hos det vi fors ker   l re og benyttes til   trene et maskinl rings-system (Mayes, 2017, s. 16). Bing Liu (2011) forklarer prosessen for   evaluere nøyaktigheten til en prediksjon. Figur 5 viser hvordan man kan kontrollere nøyaktigheten i en prediksjon ved veiledet l ring.



Figur 5: L ringsprosessen under veiledet l ring (oversatt) (Liu, 2011, s. 66).

Steg 1 handler om å klargjøre treningsdata og utforme læringsalgoritmer som et grunnlag for trening og læring. Steg 2 består i å evaluere maskinens resultater ved å benytte valideringsdata som testgrunnlag. Dette skal hindre memorering av data og sikre gode resultater på data maskinen ikke har sett før. Hensikten med dette er å måle nøyaktigheten i maskinens prediksjoner.

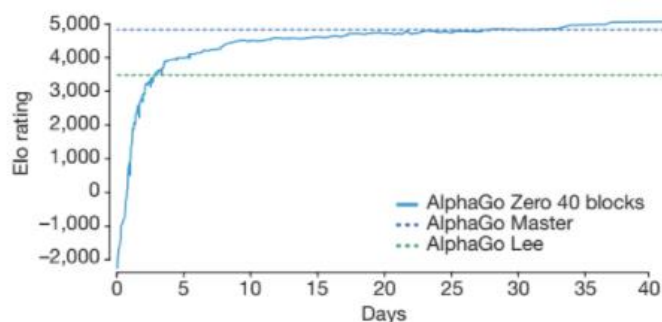
Den enkleste formen for maskinlæring er lineær regresjon. Regresjonsanalyse benyttes til å analysere innsamlede data. Formålet med analysen er å konstruere en modell til å kunne predikere verdier (y) basert på informasjon om en eller flere variabler (x). Dette skaper den lineære likningen $y = ax + b$ (Kassambara, 2018, s. 6; Montgomery, Peck & Vining, 2015, s. xiii). Lineær regresjon forsøker å plassere en rett linje i testdataen, for å videre bruke den til prediksjon (Mayes, 2017, s. 41). Innenfor revisjon kan dette benyttes til å estimere eller predikere verdier og anslag på regnskapsposter eller spesifikke eiendeler. Videre kan regresjon benyttes til å teste betydningen av forklaringsvariabelen, estimere y med et konfidensintervall, predikere nye verdier av y , evaluere om modellen passer dataene og teste for avvik innad i datasettet. I tabell 1 på side 45 har vi listet flere mulige teknikker som kan benyttes innen maskinlæring og dermed KI.

Marsland (2015) stiller spørsmålet; hvordan vet maskinen om den blir flinkere, og hvordan vet den hvordan den forbedrer seg? Marsland forsøker å besvare dette ved at vi kan fortelle algoritmen det korrekte svaret på forhånd, slik at maskinen får rett ved neste forsøk. Ved å gi noen korrekte svar, ønsker vi at maskinen klarer å finne ut hvordan den skal komme frem til korrekt svar for andre problemer, altså evnen til å generalisere. En alternativ fremgangsmåte er å gi tilbakemeldinger på om svaret maskinen avgir er galt eller rett. Dersom vi ikke har et fasitsvar, kan algoritmen brukes til å finne muligheter og sammenhenger. Dette er to metoder innen maskinlæring kategorisert som veiledet og ikke-veiledet læring, altså oppgaver med- og uten tilbakemeldinger (Burgess, 2018, s. 7; Marsland, 2015, s. 5-6). Veiledet læring benytter kategoriserte treningsdata, mens ikke-veiledet læring benytter ikke-kategoriserte data (Mayes, 2017, s. 25-26). Marsland (2015) hevder at veiledet læring er mest utstrakt i dag. Yann LeCun, Yoshua Bengio og Geoffrey Hinton (2015) spår at ikke-veiledet læring vil bli den dominerende i fremtiden ettersom vi mennesker lærer på denne måten. Vi får ikke konstante tilbakemeldinger på alt vi gjør, vi lærer underveis og ved å observere.

Eksemplene nedenfor viser hvorfor ikke-veiledet læring vil kunne bli den mest fremtredende. I praksis kan maskinlæring illustreres ved Googles AlphaGo, et dataprogram som spiller det kinesiske brettspillet Go. Utviklerne bak AlphaGo viste systemet titusenvise av tidligere Go-spill utført av profesjonelle spillere, og med dette grunnlaget lærte programmet å spille som en erfaren spiller. I 2016 vant AlphaGo fire av fem spill over verdensmester Lee Sedol i Go (BBC, 2016).

AlphaGo blir sammenlignet med sjakkprogrammet Deep Blue som slo verdensmester Garry Kasparov i sjakk helt tilbake i 1997 (IBM, 2018). Dette er veiledet læring, hvor maskinen får tilgang til masse data og tilbakemeldinger underveis på sine rette eller gale trekk. Denne læringsmetoden kalles en top-down fremgangsmåte.

Den andre læringsmetoden er en bottom-up fremgangsmåte. Metoden innebærer å kun gi maskinen nødvendig informasjon, som maskinen selv prosesserer og lærer fra, uten veiledning (Allen, Smit & Wallach, 2005). Etter AlphaGo har Google utviklet AlphaGo Zero. Dette programmet fikk kun tilgang til spillereglene til Go. Underveis fikk programmet tilbakemelding av seg selv ved tap, men ikke hvilke trekk som var dårlige. Denne metoden er en hybrid av veiledet og ikke-veiledet læring, som i fagterminologien defineres som læring med forsterkning (reinforced learning). I treningsfasen spilte AlphaGo Zero millioner av spill mot seg selv. Figur 6 illustrer hvilken utvikling maskinen hadde (Silver et al., 2017). Læringskurven måler ytelsen til læringsalgoritmen (Russell et al., 2016, s. 757). Etter tre dager slo AlphaGo Zero den versjonen av AlphaGo som slo verdens beste Go-spiller, som første maskin uten menneskelig hjelp.



Figur 6: AlphaGo Zeros spillprogressjon

I følge Datatilsynet (2018) er det et fundamentalt problem knyttet til KI, *den svarte boksen*. Hva har skjedd inne i maskinen og algoritmene før vi får utlevert et resultat? Modellen gir ofte ikke grunnlag for sine resultater. Datatilsynet stiller spørsmål om det er mulighet for å studere innholdet i den svarte boksen. I deres rapport er det gjort intervju med skatteetaten knyttet til en prediktiv modell som skal hjelpe etaten å velge ut hvilke selvangivelser de skal undersøke nærmere. De har følgende svar om den svarte boksen:

«Når vi bygger modellen på denne måten, vet vi ikke nødvendigvis hva det er som gjør at en bestemt skatteyter blir rangert til å ha stor risiko for feil. Rangeringen er et resultat av komplekse sammensetninger av dataene i modellen» (2018, s.12).

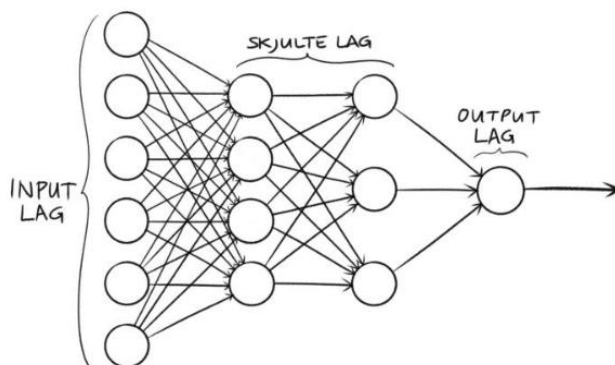
Denne uttalelsen underbygger kompleksiteten rundt KI ved bruk av maskinlæring. Problematikken rundt den svarte boksen blir nærmere diskutert i kapittel 4.2.5.

3.4 Nevrale nettverk og dyp læring

Kunstig nevralt nettverk modellerer en forenklet del av hjernen og er en viktig del for læring og resonnering innen KI (Tørresen, 2013, s. 58). Nevrale nettverk, en underkategori av maskinlæring, er inspirert av hvordan vi tror hjernens funksjoner og funksjonalitet er.

Et kunstig nevralt nettverk består av kunstige nevroner som kommuniserer sammen. Et nevron er en celle som formidler signaler til og fra andre nevroner. Et perceptron modellerer en forenklet utgave av et nevron, en algoritme, ved å bruke binære input og output. Kunstig nevralt nettverk består av en eller flere perceptroner, som kan lede til svært store og komplekse nettverk. Perceptronet har en vektor, et tall som vektlegger hvor viktig hver input er, og påvirker derfor output (Telle, 2017). Vektoren endres underveis som modellen trenes og forbedrer seg i læringsprosessen. Under læring kan det derfor dannes nye og sterkere forbindelser mellom nevronene, hvor vektoren blir forbedret underveis. For eksempel slik som AlphaGo Zero trente og lærte av sine trekk.

Et nevralt nettverk består av tre hoveddeler; input, et eller flere skjulte lag, og output. Ved mer enn ett skjult lag i det nevrale nettverket, er det å anse som dyp læring (Datatilsynet, 2018, s. 13). Flere skjulte lag gir mer dybde, mens flere nevroner gir mer bredde. Figur 7 nedenfor består av fire lag med henholdsvis seks, fire, tre og en nevroner. Nevronene samhandler på en bestemt måte, hvor nevron i lag en blir input for nevroner i lag to og tilsvarende for etterfølgende lag.



Figur 7: Nevralt nettverk (Datatilsynet, 2018)

Dyp læring er et begrep innenfor nyere metoder av maskinlæring. Dyp læring bruker nevrale nettverk på nye måter og i flere nivåer enn tidligere (Heggernes, 2017, s. 43). Dette muliggjør at datasystemer kan forbedres med erfaringer og data (Goodfellow et al., 2016, s. 8). Dype læringsmodeller er konstruerte systemer inspirert av den biologiske hjernen, men selv i dag er det en kompleks oppgave å forstå hvordan hjernen fungerer. Dyp læring kan oppdage kompliserte strukturer i store datasett gjennom oppbygging av ulike lag og algoritmer, og være en bidragsyter i krevende beslutningsprosesser. Dette gir mulighet til å finne mønstre utenfor datasettene de trenes på (LeCun et al., 2015, s. 440), hvor et eksempel er å identifisere svarte svaner, når maskinen er trent på å se hvite svaner. Dyp læring har vært et viktig teknologisk gjennombrudd og brukes for eksempel i medisin, stemme-, bilde- og mønstergjenkjenning, naturlig språkbehandling og i chatbots som Apples Siri (Miralles-Pechuán, Rosso, Jiménez & García, 2017, s. 655).

Det eksisterer flere ulike modeller innenfor nevrale nettverk, men disse forblir utenfor oppgavens omfang. De kan grovt sett deles inn i to overordnede modeller, vanlige nevrale nettverk (fremoverrettet flerlags perceptron) og tilbakevendende nevrale nettverk (nettverk med tilbakekoblinger) (Tørresen, 2013, s. 59).

4. En revisjon med kunstig intelligens

Dette kapittelet består av to deler. Den første delen fokuserer på muligheter med KI i revisjon og i revisjonsprosessen, mens den andre delen ser på utfordringer.

Fra et revisjonsperspektiv kan KI betraktes som en teknologi med hensikt å komplementere og endre den tradisjonelle revisjonen (Issa, Sun & Vasarhelyi, 2016). Revisjonsprosessen er under kontinuerlig påvirkning av tilgjengelig teknologi. Burgess (2018, s. 6) hevder at ved hjelp av KI kan revisor bruke mer tid på å analysere informasjon, fremfor research og innhenting av data. ICAEW (2017a) fremhever at KI de neste tiårene gradvis vil ta over beslutningstakende oppgaver.

De største selskapene har startet arbeidet med å implementere kognitiv teknologi i revisjon, det benyttes enda ikke sterk KI, men svak KI veiledet av mennesker (Kokina & Davenport, 2017, s. 119). En undersøkelse utført av Forbes Insights avdekket at 58 % av revisorer tror teknologi vil ha stor påvirkning på profesjonen, og at KI innen 2020 vil være et prioritert investeringsområde (Macaulay, 2016b). Revisjonsselskapene som forstår, overvåker og forbedrer sine analytiske- og kognitive systemer og prosesser er de som vil lykkes (Davenport, 2016).

Prosjekter innen KI, maskinlæring og dyp læring blir ofte outsourcet til ledende selskaper innen denne type teknologi, slik som IBM Watson og Microsoft (D. Appelbaum, A. Kogan & M. Vasarhelyi, 2017a, s. 7). KPMG samarbeider med IBM Watson og har lansert sin revisjonsplattform Clara, som de hevder vil bidra til en helt ny generasjon revisjon (KPMG, 2016a, 2017a). Revisjonsselskapet EY jobber med å implementere KI i en rekke tjenester de leverer (Mazzei & Duffy, 2018). Deloitte har startet et samarbeid med Kira Systems Inc. om et verktøy for å analysere dokumenter (Deloitte, 2016). PwC har vunnet prisen “Audit Innovation of the Year” i både 2016 og 2017. I 2016 for deres verktøy Halo som muliggjør datafangst, visualisering av mønstre og trender og sette store datamengder i sammenheng. I 2017 fikk de prisen for verktøyet GL.ai, et samarbeid mellom PwC og H2O.ai, som benytter maskinlæring (Barghini, 2017; PwC, 2016b, 2017c).

4.1 Muligheter

Revisjon kan deles inn i to komponenter; en rekke repetitive- og automatiserbare arbeidsoppgaver og beslutninger ved bruk av profesjonelt skjønn. Tradisjonelt sett har revisjonsbransjen vært strukturert med mange revisormedarbeidere som utfører de manuelle repetitive oppgavene. For eksempel å innhente bankbrev, mens deres overordnede trekker konklusjoner basert på dette arbeidet, ofte ved bruk av skjønn (Issa et al., 2016, s. 11-12). Strukturerte og repetitive oppgaver kan automatiseres ved hjelp av blant annet RPA, programmering, systemintegrasjoner og API-kall. KI kan understøtte eller løse oppgaver som krever en tankeprosess eller intensivt manuelt arbeid (Rapoport, 2016). KI kan bistå revisor med å analysere store mengder data, komplekse datasett og hele populasjoner, både finansielle og ikke-finansielle, raskere (Macaulay, 2016a). Det kreves full tiltro til prediksjonskraften for at KI kan brukes til å utføre manuelle oppgaver.

Effektiv utnyttelse av disruptive teknologier vil kunne bidra til økt effektivitet, heve revisjonskvaliteten, redusere risiko og styrket tillit, gitt at de blir grundig og riktig implementert (EY, 2017c, s. 17). Dette medfører at revisor får en dypere innsikt i klientens virksomhet. Dette underbygger ACCA (2017) som forventer at dataanalyser med bruk av prediktive algoritmer og mønstergjenkjenning for å vurdere og gjennomgå ledelsens estimer og verdsettelse, vil heve revisjonskvaliteten. Macaulay (2016a) eksemplifiserer dette med at KI kan benytte informasjon fra utradisjonelle kilder som sosiale medier, TV, radio og internett til å vurdere ekstern informasjon som kan påvirke revisjonen, direkte eller indirekte. Macaulay oppsummerer følgende muligheter ved å anvende KI i revisjon:

- Å gjennomføre en detaljert vurdering av intern og ekstern finansiell og ikke-finansiell informasjon.
- Oppdage ekstremverdier og unormaliteter i datasettene.
- Gi innsikt i og identifisere risikofaktorer knyttet til rapporterings-, prosess- eller operasjonelle aktiviteter.

KPMG (2017a) benytter KI og maskinlæring i sin revisjonsplattform. Dette gir revisor mulighet til å analysere ustrukturerte data fra flere kilder og oppnå en dypere forståelse av klientens virksomhet og risiko. Deloitte bruker teknologien

NLP til å analysere tekstdokumenter, eksempelvis leiekontrakter. Denne teknologien leser og forstår nøkkelinformasjon i dokumenter (Deloitte, 2016). Maskinlæring bidrar til at systemet lærer seg å identifisere og trekke ut nøkkelinfo fra slike dokumenter (ICAEW, 2017b). Automatisering av oppgaver gir revisor anledning til å bruke mer av sin kompetanse og tid på områder som krever vurderinger, profesjonell skepsis og skjønn (Raphael, 2017). Dette medfører at revisjonen utvikler seg i retning mot å bli mer innsiktsfull og hensiktsmessig (Raphael, 2015).

Kognitive systemer som KI, maskinlæring og datamining innebærer sofistikert teknologi hvor utviklingen skjer i et høyt tempo. Systemene bidrar med nøyaktige outputs og i mange tilfeller bedre enn menneskelige forsøk (ICAEW, 2017a). I følge PwC (2017a) kan kognitive systemer bidra til at revisjonen utføres raskere, mer effektivt og redusere risiko for feil. En maskin er billig i drift, blir ikke utmattet og så lenge input er korrekt gjør maskinen sjeldent feil, gitt at den er riktig programmert (ICAEW, 2017a, s. 5).

ICAEW (2017a, s. 6) lister noen eksempler på muligheter hvor KI kan bidra:

- Maskiner kan prosessere store mengder strukturerte og ustrukturerte data.
- Svake eller komplekse sammenhenger kan plukkes opp. For eksempel i transaksjonsrekker hvor vi mennesker finner lite forutsigbarhet.
- Maskiner kan bidra til konsistente beslutninger. De blir ikke påvirket av kjedsomhet eller utmattelse, og i utgangspunktet ikke bias, gitt riktig programmert.
- Maskinlæring er endringsdyktig, den kan lære av feil og ny informasjon.

Ken Tysiac (2017) påpeker noen kritiske suksessfaktorer knyttet til KI; en god idé, kilder til teknologi, et enkelt brukergrensesnitt, rekruttering av dyktige ansatte, innovasjonsevne og tilstrebe konkurransefortrinn. Revisjonsbransjen er i kontinuerlig utvikling, ledet av de største selskapene. Små selskaper uten mulighet til å gjøre store teknologiske investeringer vil sannsynligvis få muligheten til å benytte seg av teknologien gjennom nyutviklede skyløsninger og betalingsmuligheter (Burgess, 2018, s. 18-19; Vasarhelyi & Sun, 2017).

På basis av det vi har skrevet vil det være interessant å vite mer om dette. Vi utleder av den grunn følgende to forskningsspørsmål:

F1. Hvilke muligheter kan KI medføre i revisjonsprosessen?

F2. Ved å benytte KI i revisjonsprosessen, hvilke fordeler og ulemper medfører dette?

4.1.1 Avansert analyse

Dataanalyse kan bidra til en effektiv revisjon, tilføre verdifull innsikt og økt lønnsomhet (KPMG, 2016b). AICPA (2017) har følgende definisjon på dataanalyse i revisjon (egen oversettelse):

«En teknikk som kan oppdage og analysere mønster, uregelmessigheter og utvinning av annen nyttig informasjon i data som støtter oppunder eller kan relateres til gjenstanden for en revisjon gjennom analyse, modellering og visualisering med det formål å planlegge eller gjennomføre revisjonen».

Dataanalyse kan analysere og teste hele populasjoner fremfor et begrenset utvalg. Revisor kan dermed fokusere på avvik, ekstremverdier og andre uregelmessigheter i datastrømmen som naturlig intelligens kan overse. Dette gir verdifull innsikt i klientens transaksjoner, prosesser og kontroller. Dataanalyse kan benyttes innen risikovurdering, substanstesting, test av kontroller og analytiske handlinger, og i større grad avdekke misligheter. Kombinasjon av dataanalyse og KI vil vi betegne som avansert analyse. For at revisor skal kunne benytte slike verktøy i revisjonsprosessen kreves det at denne type teknologi og teknikker er i samsvar med revisorlov og god revisjonsskikk (Earley, 2015; IAASB, 2016). EY har utviklet en prediktiv modell for uredelig regnskapsrapportering som ved hjelp av avansert analyse skal bidra til å forbedre revisjonskvaliteten ved å kalkulere sannsynligheten for feilinformasjon (Persico & Sidhu, 2017).

Hensikten med avansert analyse er å trekke ut informasjon for å understøtte en beslutning. Fordelen er at revisor har mulighet til å gjennomgå både strukturerte og ustrukturerte data. Avansert analyse kan bidra til analyser som predikerer hendelser, risiko, årsak og tidspunkt for inntreffelse, og dermed gi revisor økt treffsikkerhet og relevans (ICAEW, 2017b). På denne måten kan misligheter og

feil identifiseres før de inntreffer. Kognitive analyseverktøy kan gi revisor bedre grunnlag til å vurdere eksempelvis justeringer i forhold til avsetninger for miljø- og opprydningskostnader. Prediktiv analyse kan beregne dette ved å innhente all nødvendig informasjon, slik som etterspørsel av produksjon, tidligere nedleggelse av produksjonssted, selskapets og markedets utvikling og eventuelle andre relevante forhold (Aurstad, 2017).

Torkil Hindberg (2015) påpeker noen utfordringer. Hvilke data som skal innhentes og hvordan de skal innhentes og analyseres. Revisjonsselskapene bruker 80 % av tiden sin på å innhente data og 20 % på å tolke data (Pedersen, 2016). De ansatte må ha kompetanse til å tolke dataene korrekt. Kombinering av stordata, dataanalyse, visualiseringsverktøy og KI i analyse av interne og eksterne data, vil frigjøre mye tid knyttet til innhenting og analyse av data.

Eija Koskivaara (2004) bruker begrepet “analytical review (AR)”, som en overordnet betegnelse på en rekke analytiske handlinger. Hensikten med AR er å innhente revisjonsbevis. Ved å anvende maskinlæring i analyseverktøy lærer maskinen å gjenkjenne mønstre og egenskaper for å kunne oppdage avvik. Dette kan underbygge revisors konklusjon om hvorvidt det eksisterer vesentlig feilinformasjon i regnskapet og revisjonsteamet kan justere angrepsvinkelen basert på informasjonen og potensiell risiko (Persico & Sidhu, 2017).

Analyseverktøy kan eksempelvis i en populasjon med transaksjoner kjøre en Benfords analyse kombinert med KI og ekstern informasjon, hvor dette kan identifisere potensielle feil i regnskapet (Raphael, 2017, s. 31). Analyseverktøy og visualiseringsverktøy kan hjelpe revisor med å velge fokusområder for videre revisjonshandlinger. Ved å identifisere mange risikoområder og avvik kan det medføre utfordringer i form av en omfattende og kostbar revisjon (Appelbaum et al., 2017a; D. Appelbaum, A. Kogan & M. A. Vasarhelyi, 2017b, s. 11). Begrepet analytiske handlinger har i vår studie en bredere betydning enn det som refereres til i ISA 520 om analytiske substanshandlinger.

4.1.2 Fra datamining til innsikt

En definisjon av datamining er «å benytte ulike teknikker for å behandle, analysere og søke i store datamengder for å identifisere ukjente mønstre, relasjoner, trender, utlede beslutningsregler eller konstruere prediktive modeller» (Sekundærkilde: Brown-Libur, Issa & Lombardi, 2015, s. 452 (egen oversettelse)). Datamining er en undergruppe av dataanalyse (Gee, 2014, s. 11). Datamining kan styrke og forbedre effektiviteten og hensiktsmessigheten av revisjonshandlinger, hvor effektivitet er et mål på forholdet mellom input og output, mens hensiktsmessighet sammenstiller mål mot resultat.

Datamining i revisjon kan benyttes til å gjennomgå ekstern informasjon i vurdering av klientens foretaksrisiko, risiko for misligheter og andre risikovurderinger. Datateknikken kan også identifisere trender og mønstre, samt se dette mot fortsatt drift-forutsetningen. På denne måten kan revisor lettere identifisere risiko for feil (Brown-Libur et al., 2015, s. 452). Brown-Libur et al. (2015, s.453) trekker frem datamining på en detaljert analyse av alle kontantstrømmene. Dette kan avdekke hvitvasking gjennom å identifisere ikke-godkjente, hyppige eller transaksjoner nær terskelverdien til godkjennelse. Videre kan disse transaksjonene være av interesse for ytterligere granskning.

Gray og Debreceny (2014, s. 360) belyser og understøtter bruken av resultatene fra datamining, hvor revisor kan rette revisjonsinnsatsen mot identifisert forventningsgap og «red flags» mellom faktiske-, forventede- og predikerte verdier. Studien viser til konkrete eksempler ved å bruke to ulike tilnærminger innenfor datamining; veiledet (top-down) og ikke-veiledet (bottom-up). Veiledet datamining inkluderer klassifisering, estimering og prediksjon. Ikke-veiledet datamining inkluderer affinitetsgruppering, clustering, beskrivelse og visualisering.

Affinitetsgruppering finner sammenhenger mellom variabler i et datasett, ved å se mønster i bilagene mellom eksempelvis anleggsmidler og vedlikehold. Clustering grupperer transaksjoner som følger kjente mønstre. For eksempel salgsinntekter, varekostnader og kundefordringer basert på relasjoner eller likhetstrekk i datasettet. På denne måten får revisor en oversikt over de transaksjonene som ikke følger et etablert mønster. Avvik tilhører sjeldent en gruppe og er ofte

sideliggende fra tyngdepunktet. Revisor må gjennomgå og vurdere avvik fra etablerte mønstre. Vurdering av avvik innebærer å forstå årsak og virkning, samt å tilrettelegge for at de ikke skal inntreffe igjen. Klassifisering handler om å klassifisere innenfor ulike kategorier gjennom å utvikle ett sett med regler fra et treningssett (Gray & Debreceeny, 2014, s. 360-362). Fordelen ved clustering er at den segmenterer data i grupper med likhetstrekk. Samtidig krever det at revisor tenker kreativt i hensiktsmessig tolkning av analysene, da de ikke er selvforklarende (McKee, 2018b). Summerte avvik kan deretter vurderes mot fastsatt vesentlighetsgrense, for å vurdere om noen kan være årsak til vesentlig feilinformasjon i regnskapet. En alternativ fremgangsmåte er å få internrevisor til å revidere avvikene eller clusterne, for å identifisere kontrollsvakheter (McKee, 2018c).

Paul E. Byrnes (2015) har studert clustering og datamining innen revisjon, hvor han konkluderer med at revisor kan bruke datamining til å lettere avdekke misligheter og dermed tilføre revisjonsbransjen merverdi. Analysene kan finne sammenhenger, ekstremverdier, avvik og uregelmessigheter i transaksjoner som ikke er i overensstemmelse med forventede transaksjonsflyt eller suspekterte transaksjoner. Analysene kan ikke kategorisk identifisere misligheter. Videre må revisor evaluere maskinens resultater og avgjøre behovet for videre undersøkelser (Brown-Liburd et al., 2015, s. 455; Gray & Debreceeny, 2014, s. 363).

Datamining kan brukes til å teste eksisterende hypoteser eller for å utvikle nye hypoteser. Det er foretrukket å utføre en veiledet tilnærming for å bekrefte revisjonsmålsetningene etter ISA 315 punkt A131. Denne tilnærmingen kan klassifisere, estimere og predikere risiko. Ikke-veiledet tilnærming kan bli en fremtidig praksis, på bakgrunn av det vi skrev i kapittel 3.3. Ikke-veiledet tilnærming muliggjør analyse av data uten bias eller veiledning (Gray & Debreceeny, 2014). Brown-Liburd et al. (2015, s. 455) understreker revisors behov for å forstå dataen, dens kvalitet og relevans, for å kunne trekke hensiktsmessige konklusjoner. I en analytisk revisjon er det viktig at output fra datamining og avansert analyse tolkes og vurderes i tråd med tiltenkt formål. Dette har betydning for og påvirkes av hvilken del i revisjonsprosessen og for hvilke påstander, output knyttes mot.

Revisorloven § 5-1 siste ledd og § 5-2 første ledd krever at revisor skal vurdere risikoen for at det kan foreligge feilinformasjon som følge av misligheter og at revisor gjennom revisjonen skal bidra til å forebygge og avdekke misligheter. ISA 240 punkt 6 påpeker at misligheter kan innebære avanserte og nøye planlagte opplegg for å holde det skjult og fremhever konsekvensen av et for tett samarbeid mellom klient og valgt revisor. Gray og Debreceny (2014, s. 376) presenterer i deres figur 7 hvordan datamining kan brukes til å avdekke mislighetsmetoder tilknyttet spesifikke regnskapsposter, relatert til ISA 240 punkt A3-A5.

Komplekse og avanserte IT-systemer som klienten benytter understøtter revisors behov i å forstå, kartlegge og evaluere etablerte IT-kontroller og systemets loggføring. Mislighetsoppdagelse vil være en viktig del i den moderne revisjonen (2014, s. 377). Appelbaum et al. (2017b, s. 5) påpeker at mer data ikke betyr mer informasjon og innsikt (se figur 8). Ledelsen kan fremdeles overstyre interne kontroller, og dataens volum og kompleksitet kan utgjøre en barriere i vurderinger og analyser revisor foretar. Undersøkelser i Norge og i utlandet antyder at det eksisterer et forventningsgap vedrørende mislighetsoppdagelse mellom brukerne og revisor. Mange brukere er av oppfattelsen at revisor skal forhindre og/eller avdekke vesentlige misligheter (Gulden, 2016, s. 292). Revisor kan i større grad møte disse forventningene ved å ta i bruk mer avanserte dataverktøy og KI. Dette fordrer at teknikkene er akseptert i tråd med ISA 240 og testet før bruk.

Informasjon fremkommer når strukturerte og ustrukturerte data er prosessert og analysert.

Tolking og strukturering av informasjon leder til kunnskap som kan benyttes til å oppnå innsikt. Innsikt bidrar til at man kan ta gode forretningsbeslutninger (ACCA, 2015, s. 11). Økt innsikt og verdifull informasjon

opparbeidet gjennom revisjonen kan presenteres for ledelsen, i den grad det ikke strider mot uavhengigheten eller annen innsynsrett.



Figur 8: Fra data til innsikt (ACCA, 2015, s. 11)

På bakgrunn av det vi har skrevet ønsker vi å utdype dette gjennom følgende tre forskningsspørsmål:

- F3.** Hvordan kan KI bidra til mer effektiv analyse, herunder å predikere presise anslag og verdier?

F4. Hvilke muligheter eller utfordringer med avansert analyse eksisterer i stegene data tilgjengelighet, -innhenting, -vask, -transformasjon, -analyse og tolkning?

F5. Kan avansert analyse mer effektivt avdekke misligheter, og vil mislighetsoppdagelse bli et større og viktigere område for revisor fremover?

4.1.3 Revisjonsprosessen → Sanntidsrevisjon

I kapittel 2.4 presenterte vi en forenklet revisjonsprosess bestående av fire steg. Birkeland (2017) påpeker revisors metodefrihet innenfor begrepet god revisjonsskikk. Gulden (2016, s. 54) beskriver denne prosessen med at enhver handling utført i forbindelse med et revisjonsoppdrag er et ønske om å oppfylle revisors formål på en så lite ressurskrevende måte som mulig.

Vi har funnet støtte for at manuelle, strukturerte og repetitive oppgaver er de første som vil bli automatisert, ved hjelp av automatiseringsteknologi med og uten KI. Dette vil medføre færre tradisjonelle tester av underliggende transaksjoner og med større fokus på test av oppbyggingen av prosesser, internkontroll og IT-kontroller i systemene (Brørs & Sellæg, 2015, s. 310). Abdolmohammadi (1999) har kartlagt 332 forskjellige typer revisjonsoppgaver. Majoriteten av disse er strukturerte og semi-strukturerte oppgaver, henholdsvis 131 og 135 oppgaver (266 av 332). Brennan, Baccala og Flynn (2017) anerkjenner at KI gjør det bra i oppgaver som omhandler datainnsamling og datamining, mens på andre områder kan KI være underlegen i forhold til mennesker.

Issa et al. (2016) har referert til en syvtrinns revisjonsprosess utledet av Louwers, Ramsay, Sinason, og Strawser (2015), og videre beskrevet hvordan KI kan implementeres i hver av disse fasene (egen oversettelse):

1. Innledende planlegging
2. Kontraktsfase
3. Opparbeide forståelse av internkontroll og identifisere risikofaktorer
4. Kontrollere risikohåndtering
5. Substanshandlinger
6. Evaluere revisjonsbevis
7. Revisjonsberetning

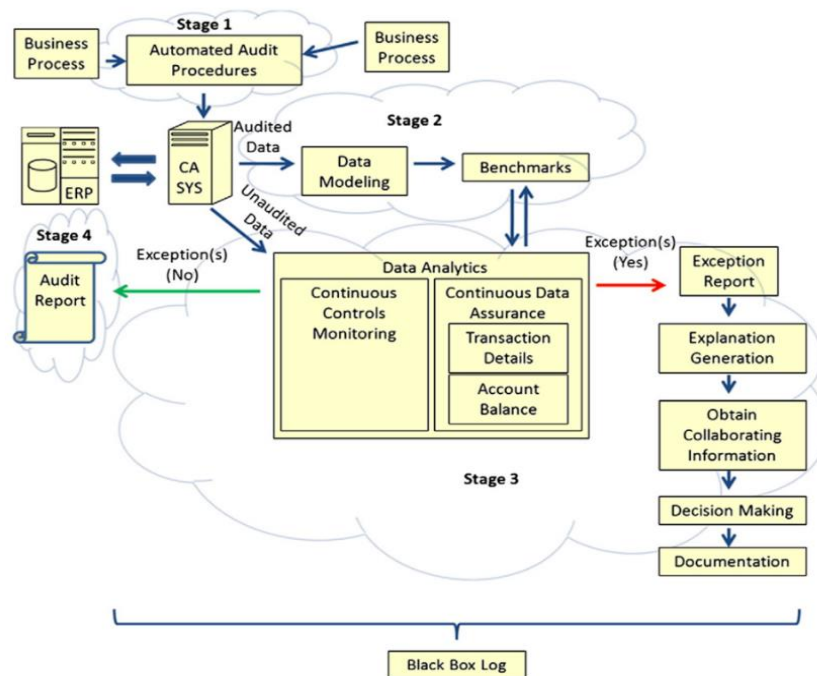
I fase en kan KI brukes til å innhente og analysere data fra en rekke kilder; både fra klient, bransje og andre eksterne kilder. KI vil i neste omgang inkorporere klientens strukturer, metoder og systemer og dermed foreta en risikovurdering, før inngåelse av kontrakt på bakgrunn av analysert informasjon (Issa et al., 2016). Innhentet data kan supplere revisors risikovurdering og påvirke revisjonsstrategi og revisjonsplan for utførelse av revisjonen (Appelbaum, 2016, s. 18). I kontraktsfasen kan KI benyttes til å estimere antall revisjonstimer, honorar og utforme klientspesifikk engasjementbrev klar til signering. KI kan også gi en anbefaling eller score på klientaksept basert på analysert informasjon i fase en. I tredje fase anvendes teknikker innen datamining og avansert analyse for å identifisere uvanlige forhold, risikofaktorer og svakheter i internkontrollen gjennom mønstergjenkjenning, maskinlæring og visualisering. Både fase tre og fire knytter seg til internkontroll.

Issa et al., (2016) foreslår å implementere KI i internkontrollsystemet for kontinuerlig overvåking. KI vil i disse fasene benytte prosessmining til å avdekke svakheter og brudd på internkontrollen. Systemet genererer automatisk en logg-fil for å sikre dataintegriteten slik at den kan benyttes som revisjonsbevis (Issa et al., 2016). Rezaee, Sharbatoghlie, Elam og McMickle (2002, s. 150) skriver at revisor må forstå transaksjonsflyten og relevante kontroller for å sikre validiteten og reliabiliteten til informasjonen i et papirløst informasjonssystem.

Issa et al. (2016) hevder at den største forskjellen mellom tradisjonell revisjon og revisjon ved bruk av KI vil forekomme i den femte fasen. Samtidig som KI innhenter data, kontrollerer den også dataens opphav (proveniens) og kvalitet. De sier at dette kan foregå i sanntid hvor revisor får testet 100 % av populasjonen. Dette gjelder både transaksjoner og detaljtester av balansen. Dette muliggjøres ved hjelp av mønstergjenkjenning, visualiseringsverktøy, benchmarking, samt avanserte analyser for å oppdage ekstremverdier. Fase seks utføres derfor kontinuerlig med fase fem.

I siste fase kan KI med en prediktiv modell estimere identifiserte risiko. KI kan også revolusjonere revisjonsberetningen i form av en kontinuerlig gradert rapport (score 0 - 100) som erstatter den tradisjonelle kategoriserte beretningen (Issa et al., 2016).

Den teknologiske omformingen av den tradisjonelle revisjonsprosessen underbygges av Chan og Vasarhelyi (2011), hvor de setter sanntidsrevisjon opp mot den tradisjonelle revisjonen (2011, s.153). De utdyper fremtidsscenarioet sanntidsrevisjon i fire steg, illustrert i figur 9 nedenfor.



Figur 9: Sanntidsrevisjonsprosessen (Chan & Vasarhelyi, 2011, s. 158)

Sanntidsrevisjon, også benevnt som kontinuerlig revisjon (continuous auditing), er definert av AICPA som: “En metode benyttet av ekstern revisor med mulighet til å gi en bekreftelse i nærmest sanntid på informasjon fremlagt av de med overordnet ansvar, selskapets ledelse...” (Sekundærkilde: AICPA, 2015, s. 4-5 (egen oversettelse)).

Sanntidsrevisjon underbygges av Kogan, Alles og Vasarhelyi (2010, s. 4). Rezaee et al. (2002, s. 151) sier at det i en sanntidsrevisjon vil være et viktigere fokus på å teste IT-kontroller fremfor substanshandlinger for å påse adekvat og hensiktsmessige kontrollaktiviteter rundt og i IT-systemet. Av den grunn kan det bli mer hensiktsmessig å revidere i større grad gjennom klientens IT-system fremfor rundt IT-systemet. En sanntidsrevisjon vil tilsa en kontinuerlig og automatisk overvåking av selskapets transaksjoner, hvor transaksjonene sammenstilles mot observerte og forventede benchmarks. Dermed identifiseres avvik fortløpende.

EYs revisjonsplattform Canvas gir klienten mulighet til å overføre informasjon direkte inn i Canvas og klienten kan bli revidert nærmest i sanntid. Revisor kan med automatiseringsteknologier kontinuerlig og automatisk overvåke alle transaksjoner, som i tillegg varsler når unormale transaksjoner finner sted (Kogan et al., 2010, s. 4). Chan og Vasarhelyi (2011, s. 156) påpeker at fremtidens revisjon kan gå fra å være assurance til insurance av vesentlige feil i regnskapet, på grunn av sanntidsrevisjon.

I lys av litteraturen vil vi stille følgende forskningsspørsmål:

F6. Hvordan kan revisjonsprosessen endres som følge av KI?

F7. Hvordan kan KI bidra til en sanntidsrevisjon?

4.1.4 Revisjonshandlinger

Brown-Liburd og Vasarhelyi (2015, s. 6) antyder at revisjonsmålsetningene mest sannsynlig ikke vil endre seg, men at ny teknologi vil føre til endringer i revisjonshandlingene. Dette omhandler handlingens formål, art, type, omfang og tidspunkt for utførelse. Eksempel på endringer i revisjonshandlinger ved bruk av KI er en fullautomatisering for test av gyldighet og nøyaktighet av betalingstransaksjoner. Revisjonsklientene kan også redusere tiden til å svare på dokumentforespørsler fra revisor. Eksterne bekreftelser rettet mot leverandører og kunder er revisjonshandlinger som også kan automatiseres og utføres med KI (Brennan et al., 2017, s. 2).

KI kan bistå revisor i å revidere raskere og med redusert revisjonsrisiko. Ved bruk av NLP og avanserte maskinlæringsteknikker kan revisor få raskere tilgang til, identifisere og trekke ut nøkkelinformasjon fra elektroniske dokumenter (ICAEW, 2017b). KI kan identifisere kunder, produsere utkast til engasjementbrev og andre dokumenter. Slike andre dokumenter kan være revisjonsberetning, e-post, sende påminnelser og signerte svar til revisorer (Meek, 2017). EY (2017c) indikerer at dagens pilotprogrammer med KI muliggjør en gjennomgang av 70-80 % av innholdet i elektroniske leieavtaler, mens resterende gjennomgås av revisor. For avanserte leieavtaler indikerer EY at programmet kan analysere 40 %, men at andelen vil øke etter hvert som maskiner blir mer avansert og lærer å analysere og tolke data bedre.

EY jobber med å implementere KI i droner. Revisor kan benytte droneteknologi integrert med KI, herunder maskinlæring, i forbindelse med varetelling og kontroll av varelager. Formålet er å øke revisjonskvaliteten gjennom å forbedre nøyaktigheten og hyppigheten i varetellingen. Teknologien kan også benyttes i produksjon, lagerstyring og i leverandørkjeden (EY, 2017a). Revisjon med bruk av droneteknologi understøttes av PwC Polen. Polen har et utviklet juridisk rammeverk for bruk av droner, hvor droner brukes i forbindelse med overvåkning (Foy, 2016; PwC, 2016a). Appelbaum og Nehmer (2017b) har gjennomgått anvendelsesområdet for droner i revisjon. Droner i utførelse av revisjonshandlinger kan særlig understøtte påstandene eksistens og verdsettelse av eksempelvis varelager. Droner kan relateres til revisjonshandlingene observasjon og inspeksjon. Videre fremhever Appelbaum og Nehmer at droner vil være godt egnet til varelagertelling siden revisor vil kunne inspisere alle varelager. Tilsvarende også for de omstendighetene hvor varetelling er utrygt. Det utelukkes ikke droner så små som fluer eller som kan åpne kasser. Ved å benytte droner til inspeksjon kan påstanden verdsettelse bekreftes ved at dronen flyr rundt og kartlegger tilstand og forhold ved varer, bygninger og andre fysiske eiendeler (Appelbaum & Nehmer, 2017b, s. 101-103). Bruk av droner gjør det teoretisk mulig å tilfredsstillere kravene i ISA 501 punkt 4, sammen med påstander som eksistens og verdsettelse. Et spørsmål blir om kravet for å «være til stede ved den fysiske varetellingen ...» kan tilfredsstilles gjennom droner. Dette må ses i tråd med delkapitlene 3.6, 4.1.5, 4.2.4 og 6.7.

Informasjon innhentet via droner samles og brukes for videre analyse og tolkning (Appelbaum & Nehmer, 2017a, s. 41). Frem til KI er integrert i droner slik at de kan fly av seg selv eller gjøre egne vurderinger gjennom mønstergjenkjenning på sanseintrykk den mottar, for å klassifisere objekter, tilstand og antall fra bilder, video og sensorinfo (IoT), finnes noen alternativer. Revisor kan samtidig som dronen filmer, være tilkoblet å se «real-time» ved å bruke smartbriller (Appelbaum & Nehmer, 2017a, s. 43), eller analysere i etterkant av innspillingen.

På grunnlag av dette utleder vi følgende forskningsspørsmål:

- F8.** Kan revisor innhente revisjonsbevis ved å benytte droner, eller må standardene oppdateres for å være i samsvar med god revisjonsskikk?

Nevrale nettverk er egnet til bruk i revisjon når det finnes data fra tidligere revisjoner og kjente input- og output verdier (Brown, Wong & Baldwin, 2007, s. 17; Fisher, Garnsey & Hughes, 2016, s. 172). Egenskapene til nevralt nettverk er egnet for bruk i sanntidsrevisjon da det raskt oppdager unormale transaksjoner. Banker, forsikrings- og andre kredittselskaper har lenge brukt nevralt nettverk til å «screeene» transaksjoner for mislighold og svindel (Brown et al., 2007, s. 18). Ved å anvende dyp læring kan bransjen automatisere flere tradisjonelle manuelle oppgaver, gitt at maskinen får tilstrekkelig treningsdata på hvordan revisor tar sine beslutninger. Dette kan være til å kontrollere og klassifisere varelager, behandle arbeidspapirer og klientens elektroniske dokumenter, gjennomgå kontrakter, utarbeide revisjonsberetning (Vasarhelyi & Sun, 2017, s. 26), i tillegg til samtlige revisjonshandlinger opplistet i ISA 500 A2. Videre er nevralt nettverk egnet til å predikere misligheter, modifiserte beretninger og fortsatt drift-forutsetningen.

De største revisjonsselskapene utvikler KI som utnytter tidligere revisjonserfaring og bransjekunnskap. Dyp læring suppleres for å lære maskinen å utøve skjønnvurderinger basert på en analyse og gjennomgang av revisors arbeidspapirer. Eksempelvis estimering av tapsavsetninger, klassifisere leiekontrakter og identifisere unormale kontrakter (Appelbaum et al., 2017a). Appelbaum et al. (2017a) argumenterer med at dyp læring og nevralt nettverk er godt egnet da disse bruker kunnskap representert i mange lag til å representere hendelser og forhold. På denne måten får man svært komplekse strukturer og prosesser som maskinen kan lære fra, behandle, prosessere og utføre. Issa et al. (2016) supplerer anvendelsesområdet til dyp læring med å gjennomgå bilag, underliggende dokumenter og utradisjonelle data. Bruk av maskinlæring og dyp læring gjennom flere årsregnskap muliggjør en mer presis vurdering av hvor de største risikoene ligger. KI kan dermed predikere hvor fremtidig risiko ligger og forbedre revisjonsteamets angrepsvinkel basert på prediktive risikoanalyser (EY, 2017c, s. 16). Maskinlæring vil derfor være en viktig bidragsyter til en måleffektiv revisjon, påkrevd etter ISA 300 punkt 4. Det er essensielt å tilføre riktig treningsdata for at maskinlæring skal kunne bidra med avanserte handlinger og nøyaktige outputs (Brennan et al., 2017, s. 2). En utfordring ved maskinlæring er faren for overtrening (overfitting) i dyp læring. Ved å tilføre for mye unyansert treningsdata vil nettverket memorere og ikke se sammenhenger og støy i nye data (Srivastava, Hinton, Krizhevsky, Sutskever & Salakhutdinov, 2014).

Et annet bruksområde for KI er i tekstgjenkjenning og tekstanalyse (tekstmining) som inspeksjon, bekreftelse og analytisk handling. Dette gjelder tilsvarende for inkorporering av stemmegjenkjenning i forbindelse med intervju og handlingene inspeksjon og forespørslar. Videre kan KI transkribere og analysere dette. Observasjon, inspeksjon og forespørsel kan maskinen analysere gjennom dype nevralt nettverk og tolke bilder og video. Videre kan det benyttes til 3-ways matching av digitale eller elektroniske ordre, fraktsedler og fakturaer - sett opp mot innbetaling, som vi ser bruk av teknologi kan utføre (PCAOB, 2017, s. 3). Nevrale nettverk kan videre produsere tekst på bakgrunn av et bilde, ved å trene på en rekke fakturaer. Maskinen kan deretter klassifisere dato, type kjøp og dermed riktig bokføring. Maskinlæring muliggjør mønstergjenkjenning og kan benyttes til det nevnte ovenfor, samt gjenkjenne fakturanummer, leverandørnavn, mva-beløp, totalbeløp og tilsvarende. PCAOB (2017, s. 3) fremhever kognitiv teknologi til å sammenstille inntektsføring mot betingelser og innhente eksterne data for å evaluere risiko og sammenligne klientens data med eksterne data.

Bedrifter kan implementere chip- eller radiofrekvensidentifikasjonsbrikker (RFID) i varer og andre eiendeler. Ren automatisering, ved bruk av RPA-teknologi, kan gjennomføre varelagertelling slik at det oppnås en bekreftelse på flere av påstandene etter ISA 315 punkt A131 (b) automatisk. Tradisjonelt sett utføres dette fysisk i tråd med ISA 501 (Brown-Liburd & Vasarhelyi, 2015, s. 10; Yoon, Hoogduin & Zhang, 2015, s. 434). Brown-Liburd et al. (2015, s.13) påpeker endringen mot digitale revisjonsbevis gjennom digitale og elektroniske spor ved hjelp av RFID, GPS og IoT. Yoon et al. (2015, s.433) mener dette kan underbygge påliteligheten av tradisjonelle bevis. Fordelen ved å implementere brikker som inneholder digitale spor er at disse kan ses i sammenheng med relaterte dokumenter som ordre, fraktseddel, faktura og innbetaling (Brown-Liburd et al., 2015, s. 11). Med digitale spor kan revisor også bekrefte varer som er under frakt. Appelbaum og Nehmer (Sekundærkilde: 2017b, s. 102) påpeker at dette ikke vil tilfredsstille som direkte bevis, da det kreves direkte observasjon, med mindre det eksisterer etablerte kontroller og disse er testet og bekreftet. KI kan analysere og vurdere de digitale sporene mot det som skal understøttes.

I delkapittel 4.1.3 presenterte vi at KI kan benytte prosessmining til å kartlegge og vurdere IK. Et informasjonssystem registrerer aktiviteter og hendelser et selskap

foretar. Slike system loggfører spor i hendelsesloggen (event log), når en maskin eller person har utført en handling (van Der Aalst, 2012, s. 76). Formålet med prosessmining er å utnytte denne dataregistreringen for å opparbeide en forståelse, overvåke og/eller forbedre selskapets prosesser. Ved å ta utgangspunkt i hendelsesloggen kan man forstå hvordan arbeidsoppgaver og interaksjonen mellom personer utføres i selskapet, og i spesifikke prosesser (Ferreira, 2017, s. 1; van der Aalst, 2016, s. 31). Prosessmining står mellom maskinlæring og datamining på den ene siden og prosessmodellering og analyse på den andre (van der Aalst, 2016, s. 31). Tre prosessutvinningsteknikker kan utføres basert på hendelsesloggen (van Der Aalst, 2012, s. 77 flg.):

- Oppdagelse – til å oppdage og forstå utførelsen av de faktiske prosesser
- Overensstemmelseskontroll – til å avstemme utførelsen av eksisterende prosessmodell mot hendelsesloggen
- Forbedring - til å evaluere og forbedre eksisterende prosesser

Jans, Alles og Vasarhelyi (2014, s. 1751-1755) har gjennomført en feltstudie på bruken av prosessmining som en analytisk handling i revisjon. De forklarer prosessmining som en systematisk analyse av automatisk lagret informasjon i et moderne informasjonssystem. I denne prosessen kan ulike typer algoritmer benyttes. Ved å undersøke digitale sporene i informasjonssystemet fant de revisjonsrelevant informasjon ved å ta utgangspunkt i selve transaksjonsprosessen hvor de fant brudd på interne kontroller og avvikende transaksjoner. Fordelen med prosessmining er at den kan komplementere eksisterende revisjonsmetodikk og være et verktøy til test av kontroller. En kritisk gjennomgang og forståelse av klientens informasjonssystem er nødvendig, også over tid. Alternativt kan revisor implementere kontinuerlige overvåkelsesaktiviteter, med avviksalarmar. Prosessmining fokuserer primært på transaksjonsflyten fremfor verdiene og benyttes på hele populasjonen (2014, s. 1769). Boer, Eimers og Elsas (2014, s. 16) argumenterer for at teknikken overensstemmelseskontroll også kan være en substanshandling, i tillegg til å være test av kontroll ved gjentakelse, og derfor være «test med to formål» etter ISA 330 punkt A23. Chiu, Vasarhelyi, Alrefai og Yan (2018) har presentert et rammeverk, hvor revisor kan benytte prosessmining i sin risikovurdering som et nytt revisjonsbevis. Chiu og Jans (2017) har fokusert på hvordan prosessmining kan bistå revisor i å evaluere selskapets internkontroll. Vi utleder på bakgrunn av dette følgende forskningsspørsmål:

F9. Hvordan benytte KI i prosessmining som en revisjonshandling?

Tradisjonell revisjon har vært stikkprøvebasert. ISA 500 punkt A53 antyder at utvelgelse av alle elementer, altså 100 % undersøkelse, bare gjelder under visse omstendigheter. Siste kulepunkt i A53 lyder følgende: “En gjentakende beregning eller andre prosesser som utføres automatisk av et informasjonssystem gjør det kostnadseffektivt å velge en 100 % undersøkelse”. Appelbaum et al. (2017b, s. 8) påpeker at dagens teknologi kan understøtte kravet om automatisering av generelle revisjonshandlinger, i tillegg til å håndtere store datamengder, kan bidra til å innhente og generere bevis på en helt ny måte og utfordre tradisjonelle revisjonshandlinger. Deres presentasjon av analytiske metoder og teknikker er oppsummert i tabell 1 (2017b, s. 9). Nye teknikker gjør det mulig for revisor å levere et produkt av høyere kvalitet gjennom økt innsikt og forståelse av virksomhetens prosesser, kontroller og regnskapsgrunnlag.

| Metode | Teknikker |
|-----------------------|---|
| Revisjonsundersøkelse | Transaksjonstester, nøkkeltallanalyse, utvalg, bekreftelser, gjentakelse og automatiserte IT-baserte revisjonsteknikker |
| Ikke-veiledet | Gruppering (clustering), tekstmining, visualisering, prosessmining (oppdagende) |
| Veiledet | Prosessmining (optimalisering), støttevektormaskin, nevralt nettverk, genetiske-, boosting- og bagging-algoritmer, ekspertsystemer, beslutningsstøtteverktøy (eks. beslutningstrær), klassifikatorer, Bayesian-teori og nettverk, Dempster-shafer teori og andre sannsynlighetsteori-modeller |
| Regresjon | Lineær, logistisk, tidsserie, ARIMA, univariat og multivariat statistikk |
| Annen statistikk | Benfords lov, fler-kriterier beslutningsstøtte, beskrivende statistikk, strukturelle modeller, og andre statistiske modeller og metoder |

Tabell 1: Analytiske metoder og teknikker

Russell et al. (2016) diskuterer det tekniske bak flere av disse teknikkene, men forblir utenfor oppgavens omfang. Appelbaum et al. (2018) har sett på hvor det er behov for mer forskning rundt disse metodene og teknikkene i tabell 1, implementert i de ulike fasene av revisjonsprosessen. Flere av disse teknikkene kan benyttes til å predikere sannsynligheten for misligheter (Sekundærkilde Rutgers, 2017, s. 13).

Standardene gir anvisninger på oppgaver knyttet til analytiske handlinger rundt hver fase, men definerer ikke eksakte typer og teknikker (Appelbaum et al., 2017b, s. 8; Appelbaum et al., 2018). Teknikkene kan baseres på revisors skjønn, erfaring og kunnskap, og mindre på standardene, også hvorvidt tradisjonelle analytiske handlinger skal vike for mer komplekse analyser (Appelbaum et al., 2018, s. 83). Appelbaum et al. (2017b, s. 12) påpeker at AICPA har konkludert med at standardene ikke forbyr bruken av moderne og nye analyser. De fremhever at standardene og kostnaden ved slike analyser er vanskelige og upraktiske, om ikke umulig, sett i sammenheng mot krav om tradisjonell metodeutførelse. Dette legger ikke til rette for å ta i bruk ny revisjonsmetodikk.

Analytiske handlinger benyttes vanligvis når transaksjonene er aggregert og samlet i kontosaldoer. Dette hindrer revisor i å oppdage potensielle avvik og problemer fortløpende, samtidig som man mister oversikt over underliggende transaksjoner. Aggregerte saldoer muliggjør identifisering av generelle mønstre. Aggregerte data i et større tidsperspektiv vil påvirke evnen til å oppdage avvik fordi disse vil være innbakt og glattet ut, og alternativt være etablert som bias i maskinen. Problemene kan løses ved å revidere underliggende transaksjoner i sanntid (Kogan et al., 2010, s. 7 & 18). Analytiske handlinger i ISA-ene krever at revisor har etablert en forventning. KI med intern og ekstern informasjon kan bidra til å sette presise forventningsverdier og kan gi ytterligere presisjon i analysene og derav en mer robust konklusjon.

Som følge av mulighetene som ligger i KI, vil det være av interesse å undersøke dette nærmere ved å stille følgende forskningsspørsmål:

F10. I hvilke eksisterende revisjonshandlinger kan KI benyttes, og kan nye revisjonshandlinger utvikles?

Innhenting av eksterne bekreftelser er sterke revisjonsbevis, men har tradisjonelt sett vært en tidkrevende prosess. Gjennom å digitalisere revisjonsprosessen, vil dette være en revisjonshandling hvor både innhenting, avstemming og kontroll kan automatiseres (Aurstad, 2017). Denne muligheten er tilsynelatende i tråd med ISA 500 punkt A18, som nevner at bevis kan innhentes gjennom «annet medie». Dette kan gjøres uten KI, men KI kan bidra til å understøtte vurderinger i denne prosessen. Tabell 2 er hentet fra Kinsersdal (2018), hvor vi har supplert den med gjennomgått litteratur på hvordan tradisjonelle revisjonshandlinger kan utføres som moderne revisjonshandlinger. Tabell 2 er ikke uttømmende og hvor KI i moderne revisjonshandlinger må ses mot relevant teori i øvrige kapitler. Dette gjelder hvilken teknologi bak KI og teknikk som er mest egnet. Dette er ikke særskilt behandlet i denne oppgaven.

| Tradisjonell | | Moderne |
|--------------------------|---|--|
| Inspeksjon/Observasjon | ⇒ | KI, droner, kamera, video, datamining (tekst & prosess) nevrale nettverk, dyp læring |
| Ekstern bekreftelse | | KI, blockchain, RPA, Chatbot |
| Etterregning/Gjentakelse | | Integrerte systemer, KI, dyp læring, datamining avansert analyse og roboter |
| Forespørsel | | KI, automatisk innhenting, chatbot |
| Analytiske handlinger | | Avansert analyse og KI til å sette presis forventningsverdi, datamining & tabell 1 |
| Skjønn og skepsis | | KI – herunder nevrale nettverk og dyp læring |

Tabell 2: Fra tradisjonelle til moderne revisjonshandlinger

Det er av interesse å undersøke det vi har gått gjennom i de overnevnte delkapitlene gjennom å stille følgende forskningsspørsmål:

F11. Kan KI benyttes i risikovurderings- og risikohåndteringshandling, og bør disse skilles eller kombineres?

4.1.5 Moderne revisjonsbevis

Ny teknologi kan føre til at tradisjonelle revisjonsbevis ikke lenger vil være tilstrekkelig eller hensiktsmessig for utførelsen. Helen Brown-Liburd og Miklos A. Vasarhelyi (2015) påpeker at stordata og ny teknologi vil være en driver for at flere bedrifter digitaliseres. Macaulay (2016a) hevder fremvoksende teknologi kan generere mer innholdsrik og detaljerte revisjonsbevis. I følge ISA 500 kan revisjonsbevis utarbeides fra informasjon som ikke er av finansiell karakter. Revisjonsbevis omfatter all informasjon om enheten som enten bekrefter, motsier eller er nøytrale angående ledelsens påstander rundt regnskapet og internkontrollen (Appelbaum et al., 2017b, s. 3). Fravær av informasjon kan også utgjøre revisjonsbevis, jf. ISA 500 punkt A1. En utfordring er å utlede verdifulle bevis fra stordata som understøtter revisors skjønn og beslutninger, og som er relevant og pålitelig. Informasjon med innhold av støy eller som er irrelevant kan medføre en dårligere skjønnsvurdering av beslutningstaker (Brown-Liburd et al., 2015, s. 455-456). Dette må revisor forstå og ta til betraktning når påstander skal kontrolleres. Maskiner påvirkes derimot mindre av følelser, fordommer og støy enn oss mennesker og kan av den grunn foreta bedre objektive skjønnsmessige vurderinger.

Revisor kan bruke en ekspert etter ISA 620 for å innhente tilstrekkelig og hensiktsmessig revisjonsbevis. Tilsvarende kan revisor benytte KI-verktøy som en agent for å løse en bestemt oppgave. Revisor er ansvarlig for å sikre relevans, pålitelighet og effektivitet ved å benytte slike verktøy (Omoteso, 2012, s. 8491). Revisjonsselskapet må derfor etablere et kvalitetskontrollsystem hvor påliteligheten og integriteten til disse systemene kontrolleres og overvåkes kontinuerlig. PCAOB (2017, s. 2) påpeker at nyutviklede verktøy kan påvirke revisjonsrisikoen hvor revisors evne til å avdekke vesentlig feilinformasjon enten kan øke eller reduseres. De sier dette avhenger av hensiktsmessigheten til nye revisjonsmetoder ved hjelp av dataanalytiske verktøy og andre teknologier, samt revisjonsfirmaets kvalitetskontroller rundt implementering. Kontrollsystemet må ytterligere kunne vurdere utøvd skjønn, utført av eller sammen med KI.

Fremtidsscenarioet hvor bedrifter implementerer brikker i varer og eiendeler med sporingssignaler, samt smarte enheter i bygg og miljø, kan videre inkorporeres i et informasjonssystem for analyse. Dette er utviklingen innen IoT, hvor informasjon

innhentes via smarte enheter, og hvor disse kontinuerlig kommuniserer sammen gjennom et nettverk (Burgess, 2018, s. vii; Dai & Vasarhelyi, 2016, s. 6). Dai og Vasarhelyi (2016, s. 5 flg.) går nærmere inn på at teknologiene muliggjør intelligens, fleksibilitet, sammenkobling og samhandling. De utdyper hvordan dette fremtidsscenarioet kan understøtte neste generasjon revisjon. Gjennom chip, RFID, smarte enheter og andre typer sensorer blir e-sporing en mulighet (Brown-Liburd & Vasarhelyi, 2015; Dai & Vasarhelyi, 2016). Dette kan benyttes som bevis til å verifisere transaksjoner (Dai & Vasarhelyi, 2016, s. 6; Moffitt & Vasarhelyi, 2013, s. 9). Dette kan benyttes som selvstendig eller underbyggende bevis og kan endre hvordan revisor innhenter tilstrekkelig og hensiktsmessig revisjonsbevis. Appelbaum og Nehmer (2017b, s. 102) peker på to potensielle problemer ved bruk av chip og RFID, hvor endelig utfall avhenger av revisors vurdering av etablerte kontroller, at hver chip representerer det faktiske produktet og at hvert produkt er representert og merket. Dette blir to kontrollpunkt for revisor å teste (2017b, s.112). Revisor kan aggregere bevis fra automatiserte revisjonshandlinger for å danne seg et helhetlig bilde på om det er innhentet tilstrekkelig og hensiktsmessig revisjonsbevis (2017b, s. 107).

KI i revisjonshandlinger er ingen ansvarsfraskrivelse, men kan gi en effektivitets- og produktivitetsgevinst ved å oppnå høyere sikkerhet for at regnskapet er uten vesentlige feil. Informasjon fra smarte enheter og sporingssignaler kan, alene eller med andre bevis, kan underbygge påliteligheten av underliggende dokumenter og innhentede bevis. Underliggende dokumenter kan være ordre, pakksedler, fraktdokumenter og fakturaer mot betalinger. Moffitt og Vasarhelyi (2013, s. 9) benytter multi-modale revisjonsbevis som betegnelse på bevis innhentet fra utradisjonelle kilder som stordata, bilder, video, digitalisert medie, RFID, GPS, IoT og tilsvarende. Nye former for revisjonsbevis vil komplementere eller erstatte eksisterende revisjonsbevis (Moffitt & Vasarhelyi, 2013, s. 13-14). Revisor må derfor videre tolke og evaluere resultatene fra informasjonen mot påstandene. For å trekke hensiktsmessige beslutninger må revisor forstå dataens integritet, tilgjengelighet, kvalitet, relevans, sikkerhet og konfidensialitet i IT-miljøet. På basis av delkapittel 4.1.5 om revisjonsbevis har vi valgt å utarbeide følgende to forskningsspørsmål:

F12. Hvordan kan multi-modale bevis integreres og vil de være selvstendige revisjonsbevis eller kun underbyggende?

F13. Kan KI benyttes til å evaluere informasjon mot påstandene, og kan revisor bygge på disse vurderingene?

4.2 utfordringer

Presentert litteratur ovenfor viser mange muligheter med KI i revisjon. Vi har fortløpende påpekt utfordringer hvor dette har vært hensiktsmessig.

Revisjonsbransjen har sakte tilpasset seg ny tilgjengelig teknologi. Viktige bakenforliggende faktorer har vært følgende; generelt høye kostnader knyttet til direkte investeringer og implementering, utdanning og trening av revisorer. Slike kostnader kan gå på bekostning av utbytte i revisjonsselskap, som er basert på en partnermodell. Revisjonsprofesjonen er høyt regulert og standardorientert og henger av den grunn etter fremskrittene i samfunnet. Juridiske problemstillinger og klient-bekymringer vedrørende ubegrenset datatilgang, har også vært viktige faktorer. Revisjonsrisikoen - problemer når feil i dataene ikke avdekkes, eller mangelfull tolkning og/eller håndtering av dataen som revisjonsbevis, har også vært drivere for den langsomme utviklingen i bransjen (Alles & Gray, 2016; Asklund, 2016; Issa et al., 2016). I det kommende vil vi gå nærmere inn på spesifikke utfordringer rundt KI i revisjon. Revisjon sin historiske utvikling gjør det interessant å vite mer om de fremtidige konsekvensene, ergo spør vi:

F14. Hvilke konsekvenser ser revisjonsbransjen ved å ikke investere i KI?

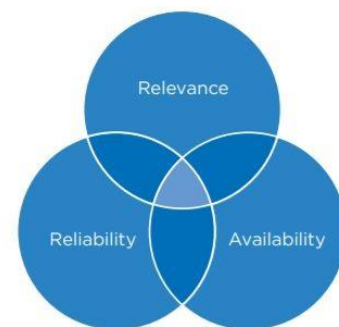
4.2.1 KI-riisiko og datahåndtering

Deloitte (2017) presenterer ulike typer risiko som følge av disruptiv teknologi. Identifiserte risikoer er gruppert inn i operasjonelle, finansielle, regulatoriske, organisatoriske og teknologiske. En utfordring er å sikre at kognitiv teknologi blir godt designet, utviklet og implementert. Konsekvensene kan bli store ved å gjøre feil, og utgjør en finansiell- og omdømmerisiko. Frykten for rettsvister kan hindre innovasjon i revisjonsprosessen (Zhang, Pawlicki, McQuilken & Titera, 2012, s. 200). Særlig gjelder dette usikkerheten rundt en erstatningssak hvor bruk av KI i revisjon kan være å anse som brudd på god revisjonsskikk. I vurderingen av om revisor har opptrådt erstatningsbetingende må normer som beste skjønn, god revisjonsskikk og skepsis legges til grunn (Birkeland, 2017). Det eksisterer alltid

en viss revisjonsrisiko ved et revisjonsoppdrag. Er et eventuelt brudd på god revisjonsskikk, gjennom nye teknologiske verktøy og metoder, grunnlag for erstatningsansvar? En sentral vurdering er hva dyktige ansvarsbevisste revisorer ville gjort. Det vil si om feilen ville blitt avdekket gjennom tradisjonell revisjonsmetode og -metodikk. Problemstillingen går på om brudd på god revisjonsskikk er en nødvendig og tilstrekkelig betingelse for ansvar (Birkeland, 2017, s. 155). Brudd på god revisjonsskikk er en nødvendig, men ikke nødvendigvis tilstrekkelig betingelse for ansvar og hvor Birkeland (2017, s. 156) sier «Etter mitt skjønn er det likevel ikke hensiktsmessig å si at brudd på god revisjonsskikk er en nødvendig betingelse for ansvar». Tilstrekkelig forståelse og testing av systemet er derfor kritiske suksessfaktorer for å kunne ta i bruk KI i første omgang. Kinserdal (2017) påpeker at det er vanskelig for tilsynsmyndighetene og standardsetterne å godkjenne ny revisjonsmetodikk før den er oppfunnet og testet.

Internt i selskapet kan det oppstå uro og umoral som følge av KI. Det er derfor kritisk at bransjen er proaktiv med å videreutvikle revisjonsstaben (Deloitte, 2017). I tillegg til risiko med utvikling og implementering, må de ansvarlige også påse at maskinen etterlever relevante krav for sensitive data, slik som konfidensielle data og personopplysninger. IAIS (2017, s. 2) fremhever to fremvoksende risiko knyttet til data. Påliteligheten i datamining og forståelsen av algoritmene som ligger til grunn. ACCA (2017, s. 3) mener mer avanserte verktøy kan forårsake mindre kontakt mellom revisjonsteamet og klienten. Dette kan medføre at revisjonsteamet får en redusert forståelse av klienten og påvirke revisjonskvaliteten negativt. Selv om KI kan bidra til å redusere risiko på noen områder, vil den samtidig bringe nye som revisor må håndtere.

En annen utfordring er tilgjengelighet og innhenting av data. Dette gjelder tilsvarende dataens integritet, relevans og kvalitet (figur 10). Dette er kjerneegenskapene til data og påvirker revisors beslutninger (CPA, 2016, s. 4). Disse egenskapene blir viktige risikofaktorer i en datarik revisjon. Dette påpeker flere av høringsinnspillene til IAASB sin gruppe DAWG. High, Harriott, Sporrer, Davies og



Figur 10: Venn diagram - Relevans, tilgjengelighet & pålitelighet (CPA, 2016, s. 6)

Poh (2016, s. 4) skriver at fullstendigheten og nøyaktigheten av data er viktige påstander å teste. For å få riktig output må man først verifisere input av dataen. ISCA (2017, s. 2) er enige i det overnevnte, men peker på en utfordring vedrørende datainnhenting hvor klientene ikke har integrerte systemer eller mange selvstendige systemer.

KPMG (2017b, s. 6) sier at endring i koding, kontroller eller prosesser av ledelsen vil direkte påvirke revisors evne til å benytte teknologi til å trekke ut og analysere data effektivt, til å fremskaffe revisjonsbevis. Integriteten til data og populasjoner kan i stor grad tilfredsstilles med automatiske bekreftelser og bruk av elektronisk data-prosessering (Moffitt & Vasarhelyi, 2013, s. 14). Det viktigste for revisor er å forstå hva man skal med analysen og hva analysen skal bekrefte. Revisor må derfor vurdere om han har trukket slutninger som er holdbare.

ISA 330 punkt A16 og A27 anerkjenner IT-baserte revisjonsteknikker (CAAT) til omfattende elektronisk testing av hele populasjoner. IAASB (2016, s. 8) påpeker at CAAT-referansen i ISA 330, var designet og utformet i en periode da revisors programmer ikke var så avanserte. På bakgrunn av presentert litteratur, gir imidlertid ISA-ene rom for avansert analyse, datamining, stordata og 100 % testing. Stordata har derimot motsatt virkning av, eksterne revisjonsbevis innhentet fra uavhengige kilder utenfor enheten (ISA 500 A31) og eksterne bekreftelser (ISA 505) - som er mer pålitelige. Stordata kan med andre ord mangle, og som påpekt i delkapittel 4.2.4, proveniens og verifiserbarhet og derav være mindre pålitelig (Appelbaum, 2016, s. 18). CPA-Canada (2017b, s. 10) er usikre på om det er noen, og i så fall hvilke, ulikheter det er mellom CAAT og dataanalyse.

ISA 315 punkt A21 krever at revisor skal opparbeide seg en forståelse for hvordan enheten har håndtert risiko som følge av IT. Basert på dette må revisors ta stilling til om han vil revidere;

- Rundt IT-systemet, eller
- Gjennom IT-systemet
 - Egne data med kjente utfall gjennom foretakets system
 - Foretakets data gjennom eget utprøvd system

Når strukturerte data fra klientens ERP-system og ustrukturerte data fra andre kilder overføres til revisors system, må revisor vurdere hvorvidt man kan stole på dataene. Revisor må vurdere om dataen er pålitelig, fullstendig og nøyaktig, eller om det foreligger risiko for manipulasjon. Har revisor fått alt eller kun revisorversjonen? Klientene vil ha ulikt syn på hvorvidt og hvilken del av informasjonen de vil gi revisor tilgang til (Earley, 2015, s. 498). En klient vil sjeldent gi revisor tilgang til alle sine databaser, men forberede kopier til revisor. Dette vil være en «setback» for revisor i å kunne benytte eksempelvis datamining hensiktsmessig i klientens databaser for å avdekke misligheter (Gray & Debreceny, 2014, s. 378). Earley (2015, s. 498) reflekterer over problemstillingen rundt pålagt klientrotasjon eller det å miste en klient. I slike tilfeller må revisor sannsynligvis kvitte seg med klientens datainformasjon og hindrer revisor i å bygge industridatabaser.

IAASB (2016, s. 12) fremhever økt fokus når det gjelder generelle IT-kontroller (ITGC) og applikasjonskontroller ved dataanalyse. Dette må utelukkende også tilsvarende gjelde ved bruk av maskiner med KI. Revisor må i større grad enn før evaluere og konkludere på etablerte IT-kontroller i IT-miljøet når revisor anvender IT-baserte revisjonsteknikker med KI. Dette står i motsetning til det ICAEW mener. De mener at noen IT-kontroller er irrelevante ettersom revisor tester hele populasjonen fremfor et begrenset utvalg og at validiteten og integriteten til dataen er forholdvis kontrollert når dataen er innhentet og transformert. De viser også til at en analyse av en hel populasjon med få eller ingen avvik, kan indirekte si noe om effektiviteten til prosesser og kontroller på området. Samtidig sier de at en helhetlig forståelse av internkontrollen er nødvendig for å forstå selskapet og for å gjennomføre en god risikovurdering (ICAEW, 2016, s. 19). Det er også ulikt syn på dette under høringsinnspillene IAASB og DAWG mottok fra de ulike interessentene.

Tradisjonelt sett har revisor forholdt seg til utvalgsrisiko. Nå kan vi se at denne går mer mot dataanalyse-risiko. Dette handler om at mange avvik avdekkes og i hvilken grad revisor må kontrollere alle eller hvordan prioritere dem. Det er upraktisk for revisor å kontrollere alle disse avvikene (KPMG, 2017b, s. 9), og særlig dersom de egentlig ikke representerer såkalte eksepsjonelle avvik. Et tilfeldig utvalg av avvikene vil ikke være å anse som tilstrekkelig (Rutgers, 2017,

s. 10). En metode som effektivt bistår revisor i å håndtere disse avvikene er derfor nødvendig. Tilsvarende er det behov for en slags prioritering av avvik, til såkalte eksepsjonelle avvik. KI kan eksempelvis gjennomgå og klassifisere akseptable og bemerkelsesverdige avvik som revisor må kontrollere nærmere.

IDW (2017, s. 4) mener avvik bør behandles som før, også i en datarik revisjon. Det er ikke akseptabel fremgangsmåte å ignorere identifiserte avvik, men å forsøke å forstå hva som forårsaker avvikene. De foreslår å analysere et utvalg av avvikene for å omklassifisere analyseparameterne for videre analyse. MADS-rammeverket presentert i delkapittel 1.4 handler om å kategorisere og lære mer om avvikene i ulike risikokategorier basert på valgte kriterier. Deretter kan de prioriteres ved å anvende revisjonshandlinger til hver kategori. Risiko varierer mellom klientene og gjør det nødvendig for revisor å benytte profesjonelt skjønn i forståelsen av enheten ved identifiserte avvik⁶. Flere høringsinnspill, heriblant ISCA (2017, s. 3), peker på et forventningsgap som kan oppstå. De viser til en forskjell mellom å teste og analysere 100 %, hvor klienter kan misforstå og forvente høyere mislighetsoppdagelse eller sikkerhet, på den informasjonen som bekreftes.

Det ville på bakgrunn av dette være nyttig å undersøke følgende forskningsspørsmål nærmere:

F15. Hvordan sikre datakvaliteten i data som revisor innhenter og tolker?

F16. Kan standardenes henvisning til CAAT benyttes med revisjons-selskapets KI-verktøy, og må revisor forstå klientens IT-miljø og tilhørende kontroller i større eller mindre grad enn før?

4.2.2 Skjønn, skepsis og kompetanseutvikling

Kan KI erstatte revisjonsstaben eller er den kun komplementær? Andrew Griggs (2017) mener det er stor sannsynlighet for at revisor vil bli erstattet av KI. Mye av gjennomgått litteratur fremhever muligheten for at revisor kan bruke mer tid på å utøve sitt profesjonelle skjønn, men kan intelligente maskiner også gjøre denne jobben for oss? ISA 200 punkt A29 krever at profesjonelt skjønn utøves gjennom

⁶ Se video – MADS: <https://www.aicpa.org/interestareas/frc/assuranceadvisoryservices/radar.html>

hele revisjonen og dokumenteres på en tilfredsstillende måte. Persico og Boillet (2017) hevder at vi er et godt stykke unna at maskinene kan erstatte menneskelig skjønn, skepsis og personlig erfaring. Reuters (2017) sier det utvikles maskiner og modeller med KI som kan bistå revisor i beslutninger. KI som et beslutningsstøtteverktøy komplementerer revisors skjønn og gjør det mer effektivt og verdifullt (Reuters, 2017, s. 10). Ettersom skjønn tas på grunnlag av flere faktorer og ikke er enkelt programmert i en algoritme, bør KI derfor i første omgang brukes som et beslutningsstøtteverktøy fremfor å erstatte revisors skjønn. Selv om KI kan erstatte menneskelige beslutninger og vurderinger, ligger dette lengre frem i tid. Fremtidens revisjon kan med stor sannsynlighet på bakgrunn av dette leveres med høyere kvalitet ved å utnytte fordelene ved KI. I motsetning til litteraturen ovenfor hevder Agrawal, Gans og Goldfarb (2016, s. 17) at KI kan utføre skjønn, gi vurderinger og intelligente anbefalinger ved å lære og etterligne menneskelig atferd og intelligens. Dette krever derimot at maskinen trenes på eksempler av utøvd skjønn. Agrawal et al. (2016, s. 17) mener videre at i noen tilfeller kan maskinens prediksjoner erstatte skjønn, men peker samtidig på at KI ikke kan erstatte all skjønn (2016, s. 1). Maskinen kan predikere menneskelig skjønn og på denne måten gjøre skjønn overflødig. Skjønn er som påpekt vanskelig koding og krever objektivitet for at maskinen ikke skal lære menneskelige bias.

Når vi beveger oss mot sterkere eller bedre versjoner av KI vil den i større grad være egnet til å brukes og stoles på i revisjonen i mer krevende prosesser. Å stole på de valgene maskinen tar krever innsikt og evaluering av systemene og resultatene deres. Dette diskuteres nærmere i delkapittel 4.2.5. Noen studier argumenterer for at økt bruk av IT vil redusere profesjonell skepsis (JICPA, 2017, s. 3). Dette er begrunnet i at revisjonsmedarbeidere ikke lærer å lese mennesker, tenke kritisk og sondere for svar (Sekundærkilde: Brazel & Schaefer, 2015, s. 3).

Et interessant spørsmål blir derfor om KI kan gjøre dette for oss som er bakgrunnen for forskningsspørsmål sytten:

F17. Kan kunstig intelligente maskiner utøve profesjonelt skjønn og profesjonell skepsis, og kan revisor konkludere på bakgrunn av disse vurderingene?

KI i revisjonsverktøy innebærer sofistikert teknologi, men er lite fleksibel. Den er ofte fokusert mot spesifikke arbeidsoppgaver som utføres på bakgrunn av bestemte treningsdata. Utfordringen er tilfeller hvor en bestemt arbeidsoppgave ikke har et tilgjengelig datasett som maskinen kan lære fra. I slike tilfeller kan revisor ikke erstattes (ICAEW, 2017a). PwC har benyttet en metode som simulerer data fra andre kilder i tilfeller hvor man lite eller ikke tilstrekkelige gode data. Dette har vist seg å være nødvendig rundt mislighetsoppdagelse. Det er en kompleks prosess å innhente transaksjoner og data for konkrete misligheter, som har attributter som forklarer mislighetene (Milner & Berg, 2017).

Studier har vist at disruptiv teknologi kan automatisere og erstatte arbeidsplasser, men samtidig gi grobunn for nye arbeidsplasser (Economist, 2016; Griggs, 2017; McKinsey, 2017). Nagar og Malone (2011) konkluderer med at mennesker komplementert med maskiner fører til de mest presise prediksjonene. ICAEW (2017a) antyder at når systemer og maskiner blir mer trent og enda kraftigere, vil de kunne benyttes på mer komplekse arbeidsområder og erstatte flere arbeidsoppgaver.

Vasarhelyi og Sun (2017) antyder at revisor kan benytte KI i revisjon, herunder teknologien dyp læring, uten å ha ekspertkompetanse, ved å benytte skybaserte løsninger. Forbedring av dype læringsalgoritmer fordrer at revisorer og IT-personell samhandler i modellutvikling og datasettene slik at de designes spesifikt til revisjonsoppgavene de er ment å løse. Kokina og Davenport (2017, s. 120) trekker frem en utfordrende faktor, at intelligente maskiner reflekterer bias fra involverte parter. Dette kan oppstå ved at maskinen trenes på data med bias eller at maskinene lærer biasene fra de som trener dem. Selv om en maskin ofte trenes med tanke på evnen til generalisering, vil KI aldri bli bedre enn den treningsdataen den i utgangspunktet ble trent på.

Revisjonsselskapene må som konsekvens utvikle og øke sin kompetanse og kunnskap innen statistikk, IT og analytiske ferdigheter. Dette er nødvendig for å kunne ta i bruk KI og andre teknologiske innovasjoner på en hensiktsmessig måte i revisjon (CPA-Canada, 2017a, 2018; PwC, 2017b). Dette støttes av Rutgers (2017, s. 3) som fremhever at det er essensielt at selskapene fortsetter å trene revisorer i å forbedre ferdigheter innen kritisk tenkning og profesjonell skepsis.

Griggs (2017) legger til nødvendigheten med kompetanse og egenskaper på områder som krever kreativitet, etisk tenkning, lederskap og ferdigheter ved mellommenneskelig interaksjon. KPMG er proaktiv og samarbeider med flere universiteter med det formål om å utdanne flere innen data og analyse (Barghini, 2017). Ytterligere har ICAS (2016) trukket frem tre ferdigheter et revisjonsteam må besitte for å møte morgendagens forventninger; avanserte forretningsmessige ferdigheter, teknologi- og data forståelsesferdigheter og myke ferdigheter.

Nye kriterier til revisjonsmedarbeidere gjør det av interesse å undersøke følgende forskningsspørsmål:

F18. Hvilken kompetanse bør nyutdannede besitte for å møte bransjens forventninger?

4.2.3 Revisjonsstandardene

Standardorientert revisjonsprofesjon gjør det komplisert å utvikle og benytte ny teknologi og metode uten aksept fra standardsetterne (Issa et al., 2016, s. 9). Samtidig, som påpekt tidligere, er det vanskelig for standardsetterne å godkjenne ny revisjonsmetodikk og –verktøy før det er oppfunnet. Ved implementering av og for hensiktsmessig bruk av KI i revisjon, kan det oppstå barrierer i standarder og et fremprovoserende behov for å oppdatere eksisterende standarder, påpekt av Torkil Hindberg (Grandal, 2018, s. 18). Vi har de siste årene sett eksempler på fremprovoserte lovendringer som følge av disruptiv teknologi; Uber og Airbnb.

Issa et al. (2016) påpeker at sanntidsrevisjon blir møtt med motvilje fordi standardene baserer seg på tradisjonell revisjonsmetodikk. De hevder at dagens standarder var hensiktsmessige da databasene var små, men er lite hensiktsmessige i dag hvor databasene er store og økonomien eksisterer digitalt i sanntid. Krahel og Titera (2015, s. 416) understøtter at revisjonsstandardene er basert på et papirsamfunn. Zhang et al. (2012, s. 199) fremhever at standardene henger etter fremskrittene i teknologien og at mange grunnleggende revisjonshandlinger ikke er oppdatert i henhold til utviklingen. Standardene må tillate og oppfordre revisjonsbransjen til å utnytte mulighetene og fordelene ved bruk av KI i revisjon (2012, s.200). Issa et al. (2016) og Bizarro og Dorian (2017) støtter en oppdatering av eksisterende revisjonsstandarder for å møte disse

forventningene. Bizarro og Dorian (2017, s. 24) antyder derimot at KI kan bidra til synergier i revisjonen på sikt, men møter utfordringer i policyer, prosedyrer og standarder underveis. Revisjonsstandardene må oppdateres for å oppfordre revisjonsselskapene til å benytte KI i revisjonshandlingene (Issa et al., 2016, s. 10).

KPMG har teknologien til å automatisere omtrent 80 % av et revisjonsoppdrag, men sier at begrensningen ligger i standardene. De påpeker eksempelvis usikkerhet ved å benytte prosessmining i kombinasjon med dataanalyse i revisjon (Asklund, 2017). IAASB (2016, s. 10-11) sier at på tross av dette er tilfelle, så må revisjonsselskapene være modige i å utforske nye måter å revidere på, uten støtte i revisjonslitteraturen. Brukerne bemerker at revisor og standardene må gjenspeile gjeldende praksis og utvikling. For å opprettholde relevansen og møte forventningene, kreves effektiv bruk av ny teknologi. Dette kan på den andre siden medføre en maktkonflikt og uenighet mellom utøverne av yrket og tilsynsmyndighetene. IAASB sier videre at revisjonsstandarder bør fremme økt revisjonskvalitet og imøtekomme fremtidige utviklinger, uten å kontinuerlig oppdateres. ICAS (2016, s. 15) mener revisjonsstandardene burde være mer prinsippbaserte eller objektive som ville hindret at standardene henger etter utviklingen i teknologi og praksis.

Det kan sies at noen standarder henger mer etter når det kommer til å legge en begrensning på KI. Noen av de kommende momentene stammer fra IAASB sin gruppe DAWG og deres høringsinnspill vedrørende bruk av dataanalyse, men er likeledes relevant for avansert analyse, datamining, maskinlæring og KI⁷.

ISA 230

CPA-Canada (2017b, s. 12) sier det er et behov for å klargjøre dokumentasjon vedrørende krav til innhold i revisjonsfila (audit file) når mer avanserte analyser utføres av revisor. De mener en oppdatering av ISA 230 vil oppmuntre til økt bruk og føre til mer konsistens rundt dokumentasjon av analysene. AUASB (2017, s. 5)

⁷ <https://www.ifac.org/publications-resources/exploring-growing-use-technology-audit-focus-data-analytics>

mener standarden bør regulere oppbevaringskrav for interne og eksterne data og re-performance av tester basert på data på et spesifikt tidspunkt. Rutgers (2017, s. 12) mener det er behov for standarden å anbefale revisor i å gi en begrunnelse for verifisering og validering av dataegenskapene. Eksempelvis input og systeminnstillinger som benyttes til å utvikle og gjennomføre analyser og valideringen av disse teknikkene. De mener også at standarden bør ta høyde for å anbefale kvalitetskontrollprosesser på de teknikkene revisor benytter.

ISA 315 og 330

FRC (2017, s. 4) mener standardene bør reflektere bruk av ulike teknikker hvor dataverktøy benyttes. Rutgers (2017, s. 8) mener at risikovurdering og risikorespons bør tillate bruk av prediktive analyser til å undersøke hele populasjoner. De mener prediktive analyser gir revisor god forståelse av enheten samtidig som det kan benyttes som relevant og pålitelig revisjonsbevis. Prediktive analyser kan teste eksistens, fullstendighet og nøyaktigheten av populasjonen (2017, s. 9). De trekker frem at prediktive analyser kan benyttes som test med to formål og i noen tilfeller test med tre formål. Dette kan føre til at skillet mellom standardene blir overflødig, selv om de representerer to ulike deler av revisjonen. AUASB (2017, s. 3-4) er enige i det overnevnte. De mener det er behov for å klargjøre om analytiske handlinger gir revisjonsbevis. Ytterligere ønsker de en klarlegging på når test av kontroller og substanshandlinger er påkrevd å utføres og om dette er hensiktsmessig å utføre når nye verktøy gir mulighet for å analysere og teste hele populasjoner. EY (2017b, s. 3 og 7) mener det er behov for å inkludere dataanalyse som en spesifikk revisjonshandling, da det til tider er vanskelig å plassere teknikken inn under eksisterende handlinger. ISA 330 punkt 10 bør reflektere muligheten for prosessmining til å evaluere IK og test av kontroller (Rutgers, 2017, s. 6). Selv om standardene var basert på et mindre teknologisk samfunn, som ikke kunne forutsett den teknologiske utviklingen, er IAASB positive i sine uttalelser. Deres reviderte forslag til ISA 315 ser ut til å støtte oppunder denne utviklingen. Revidert ISA 315 benytter begrep som automatiserte verktøy og teknikker, visualiseringsteknikker og ved bruk av teknologi (IAASB, 2018b).

ISA 500:

Standarden slik som den er i dag gir revisor en rekke valg, også til avansert analyse som bevis. Punkt 6 i standarden sier at revisor skal utforme og utføre revisjonshandlinger som er hensiktsmessige ut fra omstendighetene... Punkt 10 sier revisor skal fastsette hvilke metoder ... som effektivt oppfyller formålet med revisjonshandlingen. Veiledningspunkt A2 utleder at revisjonshandlinger for å innhente revisjonsbevis kan omfatte ... og nevner eksempelvis analytiske handlinger. “Kan omfatte” antyder at listen ikke er uttømmende og hvor nye avanserte revisjonshandlinger kan utvikles og benyttes sammen med de syv tradisjonelle handlingene. Veiledningspunkt A52 og A53 viser også at utvelgelse av alle elementer (100 % undersøkelse) er mulig når revisor har besluttet at dette vil være mest hensiktsmessig (Kinserdal, 2018). Dagens veiledning i ISA 500 nevner at inspeksjon og ekstern bekreftelse kan undersøkes og innhentes gjennom «annet medium/medie» og at etterregning kan gjøres elektronisk.

CPA-Canada (2017b, s. 13) mener veiledning vedrørende påliteligheten av data produsert av enheten selv bør klargjøres. En presisering av hvilke handlinger revisor er forventet å utføre vedrørende presisjon, fullstendighet og nøyaktigheten til informasjonen, er også ønskelig. Dette kan være en viktig klargjøring ettersom avansert analyse, datamining og bruk av KI ofte legger interne data produsert av enheten selv som grunnlag for sine analyser og konklusjoner.

Rutgers (2017, s. 8) mener punkt A8 bør revideres med tanke på at informasjon fra uavhengige kilder er mer pålitelig. De sier dette kun gjelder under visse omstendigheter hvor handlinger knyttet til dataintegriteten er utført. Ved fravær av disse integritetskontrollene vil opprinnelsen (proveniensen) og sikkerheten medføre tvilsomme data. De mener veiledning på handlinger til å fastslå dataintegriteten til interne og eksterne data bør foreligge, ettersom kvaliteten og påliteligheten av stordata vil variere. CPA-Canada (2018, s. 5) sier det er flere nøkkelementer som må oppdateres i standarden. Eksempelvis påliteligheten av data fra ulike kilder, samt type og omfang av revisjonsbevis og revisors håndtering ved identifisering av mange avvik. AUASB (2017, s. 5) støtter det overnevnte, men mener også en henvisning til relevante påstander som dekkes vil være hensiktsmessig. Rutgers (2017, s. 6-7) mener standarden bør gi veiledning til hvordan revisor kan vurdere fullstendigheten, nøyaktigheten, påliteligheten og

proveniensen av stordata og multi-modale bevis. De sier dette vil gi innsikt og kanskje også revisjonsbevis. Dette gjelder tilsvarende for punkt 9, informasjon som er generert av enheten selv. KPMG (2017b, s. 9) mener standardsetterne bør ta hensyn til hvordan handlinger kan automatiseres. De eksemplifiserer dette med om en programvare kan inspisere et styremøte, om en drone kan observere varelager og om oppdragsansvarlig revisor kan dirigere og overvåke kognitiv teknologi, som lærer seg å revidere.

ISA 520:

Punkt A6 påpeker at hensiktsmessigheten av en bestemt analytisk kontrollhandling avhenger av revisors vurdering, og hvor det i punkt A7 står at i enkelte tilfeller kan en enkel prognosemodell være en effektiv analytisk kontrollhandling. Punkt A15 har implisitt en forventning/hypotese – test – samsvar. Ved 100 % testing har man sjeldent hypoteser eller forventninger (Kinserdal, 2018).

IDW (2017, s. 3) trekker frem at forventningsverdier bør dokumenteres på en eksplisitt måte i anvendelse av dataanalyse som analytisk substanshandling. Tilsvarende blir dette et sentralt moment i avansert analyse og hvor KI setter forventningsverdier. Kan disse eksplisitt dokumenteres og forstår revisor de bakenforliggende faktorene avgjørende for forventningsverdien. Dokumentasjon, forståelse og skjønn blir derfor avgjørende faktorer i denne sammenheng. Rutgers (2017, s. 8) sier standarden ytterligere bør reflektere prediktive analyser. Prediktive analyser kan bruke data generert av enheten selv sammen med ikke-tradisjonelle data til å sette mer nøyaktige forventningsverdier.

ISA 530

Som påpekt ved ISA 500 har revisor et valg fremfor et krav, selv om gjeldende praksis er bygd på utvalgtesting og stikkprøver. Punkt 1 sier at standarden gjelder når revisor har besluttet å ... Punkt 12 og 13 indikerer derimot at revisor skal kontrollere og følge opp alle avvik. Dette understøttes av ISA 450 punkt 3. AUASB (2017, s. 5) presiserer et behov for en oppdatering vedrørende krav og veiledning angående avvik. Særlig når dataanalyser benyttes og mange avvik identifiseres. Det er også påpekt usikkerhet om standarden gjelder når dataanalyse gjennomføres på eksempelvis 90 % av populasjonen.

CPA-Canada (2018, s. 5) sier kvalitetskontrollstandardene må mer eksplisitt håndtere kvalitetskontroll av mer moderne verktøy og metoder benyttet av revisor. Det er også påpekt at flere av metodene og teknikkene i ISA 240 er utdaterte. ACCA (2017, s. 2) peker på at standardene bør reflektere følgende to scenarier:

- 1) Dataanalyse til 100 % testing av transaksjoner
- 2) Blockchain-teknologien, hvor transaksjoner ikke behøver å testes

Moffitt og Vasarhelyi (2013, s. 14) sier at standardene må gå fra tradisjonell stikkprøvebasert revisjon til å forstå implikasjonene ved stordata og analyse av hele populasjoner. Rutgers (2017, s. 5) mener en nøkkelfordring i et mer teknologisk revisjonsmiljø er om revisorer kan trenes opp til en ny og annerledes tankemåte når bevis skal evalueres fra mer moderne analyser. Det er også bevist at mer skjønnsbaserte standarder frembringer mer kritisk tenkning og resonering, fremfor normative standarder. Av den grunn spiller standardene en viktig rolle med hensyn til revisors tankegang (Sekundærkilde: Rutgers, 2017, s. 5).

En potensiell svakhet i standardene er definisjonen i ISA 200 punkt 13 d. Her defineres revisor som den eller de personene som gjennomfører revisjonen, vanligvis oppdragsansvarlig revisor eller andre medlemmer av revisjonsteamet. ISA-ene stiller en rekke krav og retningslinjer til hva revisor skal utføre i revisjonen hvor spørsmålet blir om en kognitiv intelligent maskin som lærer å revidere kan tas i bruk og erstatte eller kun komplementere revisjonsmedarbeidere.

Så langt virker det som ISA-ene ikke hindrer eller forbyr stort sett, men anbefaler og stimulerer heller ikke, til KI i revisjon. Endelig sluttresultat vil derfor være opp til oppdragsansvarlig revisor å avgjøre. Dette tyder på at ikke vesentlige endringer må til, men at standardene i større grad bør oppdateres for å reflektere et mer datarikt og papirløst samfunn med fremvoksende teknologier.

På bakgrunn av litteraturen og diskusjonen rundt revisjonsstandardene vil vi ytterligere i denne studien innhente empiri for å besvare følgende forskningsspørsmål:

F19. Hvilke revisjonsstandarder utfordrer bruk av KI i revisjon?

4.2.4 Tilstrekkelig og hensiktsmessig revisjonsbevis

Ny teknologi påvirker revisors skjønn i vurderingen av tilstrekkelig og hensiktsmessig bevis. Tilstrekkelig revisjonsbevis vil ikke være den primære utfordringen ved bruk av ny teknologi, ettersom effektive verktøy tillater en full gjennomgang av transaksjonene. Målet for tilstrekkelighet vil skifte mot tilgjengelighet og revisors evne til å innhente, analysere, tolke og evaluere informasjonen effektivt og meningsfullt. Hensiktsmessig revisjonsbevis kan være mer utfordrende fordi relevans sannsynligvis vil være gjenstand for en skjønnsmessig vurdering, slik som i dag (Brown-Liburd & Vasarhelyi, 2015, s. 6). Revisor må evaluere beviset, uavhengig av om det er innhentet automatisk ved hjelp av KI eller manuelt. Automatiske prosesser vil gi høyere grad av pålitelighet fremfor manuelle prosesser (Brown-Liburd & Vasarhelyi, 2015, s. 6). Thorolf Aurstad (2017) er noe kritisk til påliteligheten av automatisk innhentet revisjonsbevis. Han påpeker at vi kan miste kontrollen på bevisets kilde og andre underliggende klausuler. Vurdering av påliteligheten i samsvar med standardene blir en viktig faktor, og særlig i de tilfellene hvor revisor bygger sin konklusjon på maskinens vurderinger.

En annen utfordring er at informasjon kan mangle proveniens og verifiserbarhet, altså mangelfull opphavsoversikt og ufullstendige loggfiler. Dette kan være informasjon vesentlig for revisjonen, men svakere revisjonsbevis på grunn av bevisets pålitelighet (Appelbaum, 2016, s. 17). Bevis basert på digitalt materiale, automatiske prosesser eller utledet fra stordata må vurderes med høyere profesjonell skepsis. Revisor må forholde seg til et kontrollspor, hvor sporbarhet mellom filer og dokumentasjon fungerer tilfredsstillende, for å sikre pålitelighet. Nearon (2005, s. 34) reiser tretten spørsmål som en skeptisk revisor bør ha en formening om, og kan bistå revisor i videre evaluering av påliteligheten av digitale bevis.

Ekstern informasjon ledelsen har benyttet i sine vurderinger av en regnskapspost, vil ved handlingen gjentakelse være hensiktsmessig å inkludere også for revisor (Appelbaum, 2016, s. 6). Videre vil det ved revisjon av estimer etter ISA 540 være fordelaktig for revisor å ta stordata til betraktning i sine prediksjoner til eksempelvis ukuransvurderinger – sammen med informasjon fra klienten (Kinserdal, 2018). Slike data kan være med på å gi revisor en bedre mening om

rimeligheten av regnskapsposten, ved å utarbeide en prediktiv forventningsverdi før eller etter ledelsens fremleggelse av regnskapet. I forhold til ISA 540 punkt 8 har det blitt uttrykt ønske om at det bør anbefale bruk av analytiske datateknikker (Rutgers, 2017, s. 13). Disse teknikkene kan effektivt bistå revisor i bekreftelse av fullstendighet, nøyaktighet og periodisering.

Appelbaum og Nehmer (2017b, s. 101) sier man må søke i akademisk litteratur vedrørende bilder, video og sensorinformasjon som revisjonsbevis. Ved å benytte KI i denne sammenheng kan bilder, video og sensorinformasjon prosesseres, analyseres og tolkes av maskinen. Videre kan KI vurdere og ta en beslutning på bakgrunn av informasjonen. Pålitelighet ved maskinens utfall og prediksjoner må i den forbindelse ses i sammenheng med relevant revisjonsteori og annen teori beskrevet. Basert på ISA 500 punkt A31 er det ikke utelukkende at slik informasjon kan benyttes som revisjonsbevis, men at man må påse effektiviteten av etablerte kontroller, ettersom punkt A31 sier at bevis i form av video, digitalisert eller annen elektronisk form er mindre pålitelig enn originaldokumenter. Påliteligheten er svakere når selskapet har produsert og presentert informasjon enn at revisor selv har innhentet informasjonen direkte. Boer et al. (2014, s. 17) har sett på påliteligheten av revisjonsbevis innhentet i en datarik revisjon opp mot ISA 500 punkt A31 og oppsummert følgende (egen oversettelse og utheving):

| ISA 500 punkt A31: Indikasjoner på bevisets pålitelighet | Effekt på påliteligheten i en datarik revisjon |
|--|--|
| Revisjonsbevis er mer pålitelig når det innhentes fra uavhengige kilder utenfor enheten | Påliteligheten øker når revisor benytter eksterne bekreftelser på beregningsnøkler |
| Revisjonsbevis som er generert internt , er mer pålitelig når de tilknyttede kontrollene pålagt av enheten, herunder kontroller knyttet til utarbeidelse og vedlikehold, er effektive | Påliteligheten øker når nøkkelkontroller tilknyttet vare- og tjenesteflyten er integrert i ERP-systemet |
| Revisjonsbevis som innhentes direkte av revisor er mer pålitelig enn revisjonsbevis som innhentes indirekte eller gjennom følgeslutninger | Påliteligheten øker når revisor utformer og utfører beregningene |
| Revisjonsbevis er mer pålitelig når det foreligger skriftlig , det være på papir, elektronisk eller annet medium, enn revisjonsbevis som er innhentet muntlig | Påliteligheten øker når revisor observerer flyten av varer og tjenester fremfor å lese dokumenter og utføre walkthrough-tester i |

| | |
|--|--|
| | forståelse av relevante kontroller |
| Revisjonsbevis i form av originaldokumenter er mer pålitelig enn revisjonsbevis i form av kopier eller faksimiler, eller dokumenter som er blitt filmet, digitalisert eller på annet vis omgjort til elektronisk form, hvis pålitelighet kan avhenge av kontrollene knyttet til deres utarbeidelse og vedlikehold | Uklart i et digitalisert miljø hvor originale (papir) dokumenter blir mer og mer fraværende |

Tabell 3: Påliteligheten av revisjonsbevis

Appelbaum (2016, s. 19) påpeker at man må innhente, gjennomgå og evaluere mer bevis dersom det er usikkerhet knyttet til påliteligheten eller verifiserbarheten av underliggende informasjonen. ISA 330 snakker om å innhente overbevisende (persuasiveness) revisjonsbevis. Det er ikke snakk om avgjørende (conclusive) revisjonsbevis. Informasjon varierer i hvilken grad det overbeviser revisor. Tilstrekkelighet og hensiktsmessighet er de to faktorene for om revisjonsbeviset overbeviser. Arens et al. (2016, s. 223) fremhever seks faktorer som påvirker om beviset er overbevisende; kildens uavhengighet, internkontrollens effektivitet, revisors direkte kunnskap, kvalifikasjoner til de som fremskaffer informasjonen, graden av objektivitet og tidsriktigheten (nærme balansedagen).

Et aktuelt forskningsspørsmål vil på bakgrunn av dette være:

F20. Vil tilstrekkelig og hensiktsmessig revisjonsbevis endres som følge av KI?

4.2.5 Etterprøvbarehet og dokumentasjonskrav

En maskin kan finne svar på oppgaver den er trent til å løse, mens en forklaring på hvordan og hvorfor maskinen ga et spesifikt utfall er mindre vanlig. Datatilsynet (2018, s. 12) har påpekt at man kan ha vanskeligheter med å forstå hva som har skjedd i maskinen, omtalt som den svarte boksen (black box). Ved bruk av nevralt nettverk og dyp læring med mange lag av variabler er det mindre transparens i hva som skjer i maskinen, relativt til enkle algoritmer (2018, s. 23). Dette underbygger Nugent og Cunningham (2005, s. 163) som påpeker at algoritmer og nevralt nettverk har nådd et nivå av kompleksitet hvor de ikke lenger har transparens og åpenhet til begrunnelsen av deres prediksjoner. Dette

kan være et problem når maskinen genererer sine egne algoritmer og lærer fra treningssett og output. Det er ingen helhetlig forståelse av dype nevrale nettverk til tross for innsatsen gjort i å forsøke og forstå hva som skjer i den interne strukturen. Derav problematikken vedrørende den svarte boksen. Mangelfull evne til å tolke den interne strukturen ser ut til å ha blitt kompensert for evnen de presterer å yte på oppgaver. En årsak for denne tolkningsproblematikken skyldes at hvert lag i det nevrale nettverket ikke nødvendigvis trenger å ha fornuftig mening isolert sett (Alain & Bengio, 2016, s. 1). Kokina og Davenport (2017, s. 120) sier at manglende transparens kan gjøre det svært vanskelig å overlate beslutninger og skjønn til KI.

Andrew Clark (2017) underbygger vanskeligheten med å forstå den interne strukturen i maskinlæringsalgoritmer, ved å benytte et eksempel fra Carlos Guestrin. Eksempelet var en modell utviklet for å avgjøre om det var avbildet en husky eller ulv. Problemet med denne algoritmen var at den var trent på bilder hvor ulver var plassert i snø-bakgrunn og derfor klassifiserte den huskyer som ulv, når et bilde inneholdt snø. Dette understøtter problematikken med å stole fullt på maskiner, selv om feilen skyldtes menneskelig programmering. Slike små feil kan få alvorlige konsekvenser for et revisjonsselskap som lever av å levere tillit.

For å løse problematikken rundt den svarte boksen bør man søke etter en input-output forståelse hvor etterprøvbarhet er mulig for å kontrollere effektivitet og pålitelighet, i tillegg til å kunne verifisere algoritmene. Det blir mer utfordrende å forstå, kontrollere og etterprøve den interne strukturen i maskinene etter hvert som de avanseres. Særlig når de lærer av seg selv på grunnlag av ny informasjon og erfaring, altså gjennom maskinlæring. Implementering av algoritmer som forklarer hvordan og hvorfor maskiner har resonnert som de gjør er en vanskelig prosess. Dype nevrale nettverk er bygd opp på en kompleks og lite forklarbar måte. På samme måte som det er flere aspekter ved menneskelig atferd som er vanskelig å forklare i detalj, heriblant profesjonelt skjønn. Kanskje vil det ikke bli mulig for kunstig intelligente maskiner å forklare alle sine handlinger. Mennesker besitter også en slik svart boks, ved at vi ikke klarer å gi en detaljert forklaring på alle våre handlinger (Datatilsynet, 2018), inkludert hva dyktige og ansvarsbevisste revisorer ville gjort. Spesielt hvordan skjønn og skepsis forbedres hos dyktige revisorer kan betraktes som en svart boks. Det kreves ikke at revisor må kartlegge

hvordan hans læring, kunnskap og erfaring har ført til sin skjønnsmessige vurdering. Hvorfor skal vi da stille disse kravene til en maskin? Et annet moment er om algoritmene og nevrale nettverk sin svarte boks og sluttresultatet kan bero på revisors overordnede skepsis og skjønnsmessige vurdering. Brazel og Schaefer (2015, s. 8) sier at profesjonell skepsis i praksis er en svart boks.

Ved å inkorporere avansert teknologi i revisjonsprosessen oppstår et spørsmål om mindre transparent revisjonsspor rundt revisjonsbevis. Teknikker innen prosess-mining kan brukes til å utlede revisjonslogger og etablere datasystemer som produserer logger som til enhver tid kan gjennomgå og kontrolleres (Brown-Liburd & Vasarhelyi, 2015, s. 6). Et tilbakemeldingssystem som evaluerer utfall og resultater over tid vil være hensiktsmessig og kan benyttes til å evaluere ytelsen på revisjonssystemet som brukes (2015, s. 7). Ved å benytte en plattform som inneholder forklaringsbaserte sikkerhetssystemer kan man kontinuerlig vurdere hvordan systemet fungerer og kontrollere at systemet fungerer etter sin hensikt og etterlever relevante krav.

En annen mulig løsning kan innebære statistiske analyser for å spore modellens prediksjoner gjennom læringsalgoritmene og tilbake til treningssettet. Deretter kan man ta utgangspunkt i de punktene som er ansvarlige for en prediksjon gjennom å søke innsikt i hvordan dette punktet har brukt input og derav utledet output. Poenget er å forstå modellens oppførsel og prediksjoner ved å se på hvordan det ble utledet fra treningsdata (Wei Koh & Liang, 2017). En annen fremgangsmåte er å benytte lineære klassifikatorer til å forstå dynamikken innenfor dype nevrale nettverk (Alain & Bengio, 2016). Å reversere algoritmene og gå motsatt vei for å forstå den interne strukturen, kan også være en mulig metode. Forsøk på å forstå den interne strukturen i et system gjennom disse metodene, har sine utfordringer. Flere fremgangsmåter har forsøkt å implementere forklaringsystemer som forsøker å tolke hva maskinen gjør underveis, gjennom beslutningsregler eller -trær. Nugent og Cunningham (2005, s. 164) påpeker at disse metodene har sine praktiske begrensninger og mener en bedre forståelse kan skapes gjennom et case-basert resonnement. Her ligger fokuset på maskinens treningsdata og tidligere erfaring (tidligere caser). Ulempen er at selv denne metoden er ikke fullstendig transparent (2005, s. 166).

En god modell kan fort bli verdiløs dersom man ikke forstår hva den gjør eller hvorfor den gjør det den gjør, da det ofte er nødvendig å forstå modellens oppbygging eller algoritmene bak. Den svarte boksen må derfor være mulig å åpnes. Generelt ligger det lite logikk i den svarte boksen, ofte kun tall og regler. Utgangspunktet blir ofte å studere algoritmene sammen med datagrunnlaget (treningsdataen). Dette ble presentert under en konferanse i regi av Datatilsynet, 12. januar 2018. Det foreligger ingen standard og generelt lite veiledning på dette feltet. Det interessante er hva som kreves dokumentert og hvor langt ned i systemet man må forstå og kunne dokumentere. Man kan også stille spørsmål om en intelligent maskin kan få en sertifisering av tilsynsmyndighetene eller andre til å utføre en revisjon.

Dokumentasjon er som påpekt tidligere en viktig del av god revisjonsskikk. Kravet til etterprøvbarehet understreker viktigheten av at revisor i sin dokumentasjon inkluderer verktøy og teknikker som er benyttet. Tabeller, grafer og scripts benyttet og hvordan de ble generert med screenshots kan være viktig dokumentasjon. Særlig i de tilfeller hvor underliggende data ikke er inkludert i dokumentasjonen. Ytterligere bør stegene fra tilgang på data, og hvis mulig hvordan data ble hentet ut og transformert for revisjonsbruk, til analyse, evaluering, tolking og videre oppfølging dokumenteres. Til slutt vil det være viktig å dokumentere evalueringen av resultatene og hendelser identifisert gjennom avansert analyse, datamining og KI. Dette gjelder også handlinger tatt ved disse resultatene og hendelsene. Avvik og all feilinformasjon identifisert, utenom klart uvesentlige, bør dokumenteres, samt hvorfor de er avvik (McKee, 2018a).

FRC (2017, s. 3) trekker frem at revisors analyser og klientens data på et senere tidspunkt kan være i et annet format som vanskeliggjør etterprøvbarehet. De legger derfor til grunn at benyttet metode inkludert scripts for å uthente data bør være tilgjengelig i revisjonsloggen/-fila. De mener tilsvarende bør gjelde for utøvd skjønn under en slik omstendighet. Kvalitetskontrollsystemene bør også være åpne for tilsynsmyndighetene ved kontroller for å kontrollere integriteten til systemene. Vi kan derfor reise to vesentlige spørsmål vedrørende etterprøvbareheten ved KI. Hvordan etterprøve skjønn foretatt av en intelligent

maskin og i hvilken grad innsynsrett og dokumentasjon av koder og algoritmer i systemene kan hemmeligholdes eller må offentliggjøres?

ISA 200 punkt A29 krever at en annen revisor skal kunne forstå det betydelige skjønnet som er utøvd for å komme fram til konklusjoner om vesentlige forhold. Punkt 26 sier at det som kjennetegner skjønnet som forventes av revisor er at opplæring, kunnskap og erfaring har bidratt til å utvikle ferdighetene som er nødvendige for å utøve rimelig skjønn. Som kartlagt tidligere, vil opplæring, kunnskap og erfaring til en maskin med KI som benytter maskinlæring, kunne imøtekomme dette kravet.

Det er viktig å ha tillit til de systemene som tas i bruk. Tillit krever at det føres trygg og etisk styring av kunstig intelligente systemer. Herunder ansvar for algoritmene og strukturene som ligger til grunn, etterlevelse av lover og interne retningslinjer, bekreftelse på integriteten av data, beskyttelse av personvern og personlig informasjon (Banavar, 2016). Tilsyn med systemene er derfor viktig for å overvåke at det ikke forplanter seg bias i systemet. Dette underbygger Kolbjørnsrud (2017) som har foretatt en undersøkelse som viser at kun 6 % internasjonalt og 14 % i Norden fullstendig avviser muligheten til å stole på rådene fra kunstig intelligente maskiner. Tillit til disse systemene fordrer følgende tre vilkår:

1. Forståelse av hvordan systemet virker og gir sine prediksjoner
2. Systemet leverer relevante og pålitelige resultater over tid
3. Systemet har evne til å gi en forklaring på hvordan den har resonnet seg frem til resultatet.

Disse tre vilkårene er essensielle også i et revisjonsperspektiv, sett i tråd med ISA 230. Det må derfor etableres en logg av et revisjonsspor som følger datainnhenting, datavask, transformering, analyse og avvik til visualisering, evaluering, tolking av informasjonen og mot handling. Denne loggen må tilfredsstillende krav i ISA 230 og kan forenkle dokumentasjonsprosessen. Chan og Vasarhelyi (2011, s. 156) mener en logg-fil kan fungere som et revisjonsspor og være et bevis på at revisjonshandlinger ble utført i tråd med revisjonsstandardene.

Det vil av den grunn være relevant å reise følgende to forskningsspørsmål:

F21. Hvilke krav bør stilles til KI sin etterlevelse av standarder, dokumentasjon av den interne strukturen og etterprøvbarehet?

F22. Hva må en revisjonslogg inneholde for å tilfredsstillere ISA 230?

5. Forskningsmetode

Forskning kan defineres som: «en prosess som mennesker foretar seg på en systematisk måte for å finne ut av ting, og derav øke deres kunnskap» (Saunders, Lewis & Thornhill, 2015, s. 5). I dette kapittelet vil vi presentere vår metodiske tilnærming til vår forskning. Her vil vi gjennomgå hvordan vår forskning er gjennomført, vurderinger rundt valg av metode, datainnsamling og behandling av dataene. Studiens reliabilitet og validitet vil også bli diskutert. Våre funn, drøftinger og resultat ved forskningen vil bli presentert i kapittel 6 «Resultat og analyse», og videre forankret opp mot presentert litteratur.

5.1 Forskningsdesign

Et forskningsdesign kan anses som en logisk og systematisk plan for å styre forskningsprosessen. Det er nødvendig å beskrive studiets målsetninger, hvilken metode og hvilke teknikker som benyttes for å besvare problemstillingen og forskningsspørsmålene. Planen må inkludere kildene til informasjon. I praksis vil forskningsdesignet være et hjelpemiddel for at forskeren skal holde et målrettet fokus (Krishnaswami & Satyaprasad, 2010, s. 40; Zikmund, Babin, Carr & Griffin, 2010, s. 66 flg.). En beskrivelse av vår forskningsprosess er nødvendig for at den skal fremstå som pålitelig og gyldig. Dermed blir det er mulig for andre å se hvilke valg vi har tatt og eventuelt etterprøve det (Jacobsen, 2005, s. 371).

Vår forskning er en studie av muligheter og utfordringer ved å benytte KI i revisjon. Vi har brukt mye tid på litteraturgjennomgangen for å fremskaffe relevant litteratur, samt å oppnå en tilstrekkelig og hensiktsmessig forståelse av temaet. Studien omhandler en teknologi som lenge har vært umoden, men som vi ser modnes mer og mer. Derfor har en robust litteraturgjennomgang vært nødvendig. Temaet vårt har begrenset litteratur på noen områder hvor vi ønsker å tilføre ny empiri gjennom vår forskning, samtidig som vi kartlegger de praktiske implikasjonene. På andre områder ønsker vi å supplere, bekrefte eller avkrefte litteraturen. Vi har valgt å utforme en eksplorerende problemstilling fremfor en enkel problemstilling. Dette skal undersøkes gjennom spesifikke forskningsspørsmål. En eksplorerende tilnærming med undersøkende design har til hensikt å utdype det vi vet lite om eller ønsker å vite mer om (Jacobsen, 2005, s. 62).

KI er på et tidlig stadium, derfor er det hensiktsmessig med et undersøkende design fremfor et deskriptivt eller kausalt design. De alternative designene innebærer å beskrive eller forklare et fenomen. Begrunnelsen er at det eksisterer noe forskning, men lite konkret teori om problemstillingen vår. Av den grunn supplerer vi oppgaven avslutningsvis med forslag til videre forskning.

Vi har utarbeidet 22 forskningsspørsmål som skal belyse problemstillingen vår: «Kunstig intelligens i revisjon - *Hvilke muligheter og utfordringer medfører bruk av kunstig intelligens i revisjon?*». Tabell 4 nedenfor presenterer hvordan disse skal besvares og utforskes.

| Forskningsspørsmål | Metode |
|---|--|
| F1. Hvilke muligheter kan KI medføre i revisjonsprosessen? | Intervju Litteraturstudie Spørreskjema |
| F2. Ved å benytte KI i revisjonsprosessen, hvilke fordeler og ulemper medfører dette? | Intervju Litteraturstudie Spørreskjema |
| F3. Hvordan kan KI bidra til mer effektiv analyse, herunder å predikere presise anslag og verdier? | Intervju Litteraturstudie Spørreskjema |
| F4. Hvilke muligheter eller utfordringer med avansert analyse eksisterer i stegene data tilgjengelighet, -innhentning, -vask, -transformasjon, -analyse og tolkning? | Videre forskning |
| F5. Kan avansert analyse mer effektivt avdekke misligheter, og vil mislighetsoppdagelse bli et større og viktigere område for revisor fremover? | Spørreskjema Litteraturstudie |
| F6. Hvordan kan revisjonsprosessen endres som følge av KI? | Intervju Spørreskjema Litteraturstudie |
| F7. Hvordan kan KI bidra til en sanntidsrevisjon? | Intervju Litteraturstudie |
| F8. Kan revisor innhente revisjonsbevis ved å benytte droner, eller må standardene oppdateres for å være i samsvar med god revisjonsskikk? | Intervju Litteraturstudie |
| F9. Hvordan benytte KI i prosessmining som en revisjonshandling? | Intervju Litteraturstudie |
| F10. I hvilke eksisterende revisjonshandlinger kan KI benyttes, og kan nye revisjonshandlinger utvikles? | Intervju Litteraturstudie |

| | |
|--|--|
| F11. Kan KI benyttes i risikovurderings- og risikohåndteringshandlinger, og bør disse skilles eller kombineres? | Intervju Litteraturstudie |
| F12. Hvordan kan multi-modale bevis integreres og vil de være selvstendige revisjonsbevis eller kun underbyggende? | Intervju Litteraturstudie |
| F13. Kan KI benyttes til å evaluere informasjon mot påstandene, og kan revisor bygge på disse vurderingene? | Intervju Litteraturstudie |
| F14. Hvilke konsekvenser ser revisjonsbransjen ved å ikke investere i KI? | Intervju Litteraturstudie |
| F15. Hvordan sikre datakvaliteten i data som revisor innhenter og tolker? | Videre forskning |
| F16. Kan standardenes henvisning til CAAT benyttes med revisjonsselskapets KI-verktøy, og må revisor forstå klientens IT-miljø og tilhørende kontroller i større eller mindre grad enn før? | Videre forskning |
| F17. Kan kunstig intelligente maskiner utøve profesjonelt skjønn og profesjonell skepsis, og kan revisor konkludere på bakgrunn av disse vurderingene? | Intervju Spørreskjema Litteraturstudie |
| F18. Hvilken kompetanse bør nyutdannede besitte for å møte bransjens forventninger? | Intervju Litteraturstudie |
| F19. Hvilke revisjonsstandarder utfordrer bruk av KI i revisjon? | Intervju Spørreskjema Litteraturstudie |
| F20. Vil tilstrekkelig og hensiktsmessig revisjonsbevis endres som følge av KI? | Videre forskning |
| F21. Hvilke krav bør stilles til KI sin etterlevelse av standarder, dokumentasjon av den interne strukturen og etterprøvnbarhet? | Intervju Spørreskjema Litteraturstudie |
| F22. Hva må en revisjonslogg inneholde for å tilfredsstillere ISA 230? | Intervju Litteratur |

Tabell 4: Metodisk tilnærming til forskningsspørsmålene

Videre forskning innebærer at vi har kartlagt noe litteratur, men hvor det er behov for mer forskning. Det må poengteres at det kan eksistere annen litteratur eller forskning som vi ikke har avdekket eller som er igangsatt. Dette kan ikke utelukkes. Derfor benyttes videre forskning som en samlebetegnelse for de spørsmålene som vi enten ikke har hatt kapasitet eller data til å belyse.

5.2 Valg av metode

En metode er nødvendig for å gjennomføre en undersøkelse på riktig måte og for å sikre gode resultater som er reliable (Jacobsen, 2005, s. 20). Ang (2014, s. 7) forteller at forskning vanligvis klassifiseres som en induktiv eller deduktiv studie. Dette omtaler Jacobsen (2005, s. 28) som forskningsstrategi, en strategi for datainnsamling.

En induktiv studie søker å utlede teori fra empiri, altså «fra empiri til teori» (Ang, 2014, s. 7; Jacobsen, 2005, s. 29). Empiri er data om virkeligheten (2005, s. 14). Den kvalitative metoden er i utgangspunktet induktiv. En induktiv studie har en åpen metode, der forskeren forsøker å legge så få føringer som mulig på den informasjonen som skal samles inn. Dette er fordelaktig ved forskning på ukjente fenomen og når man har interesse av å få frem ulike forståelser. Kvalitative åpne tilnærminger vil gi bedre data i den forstand at de har høyere relevans for dem som blir undersøkt (2005, s. 29). Kvalitative data består av empiri i form av ord og meninger (Jacobsen, 2005, s. 124). Et slikt intensivt design går i dybden ved å få fram nyanser og detaljer rundt et fenomen (Jacobsen, 2005, s. 89).

En alternativ fremgangsmåte vil være å utføre en kvantitativ metode. Dette er en deduktiv fremgangsmåte, altså «fra teori til empiri» (Jacobsen, 2005, s. 28). Fordelen med ekstensivt design, altså en deduktiv og kvantitativ metode, er at vi kan innhente tilbakemeldinger på spørsmål fra enda flere enheter ved å forske bredt fremfor dypt (2005, s. 93-94). Kvantitative data består av empiri i form av tall, størrelser og symboler (2005, s. 124). Denne metoden ville gitt oss for lite dybde og forståelse om temaet til å kunne skape ny teori eller forstå de praktiske implikasjonene godt nok.

Flere utforskende forskningsspørsmål vil ved en kvantitativ studie derfor ikke alene være hensiktsmessig for formålet (Jacobsen, 2005, s. 62). Den kan derimot være supplerende, til å besvare eller støtte oppunder, det kvalitative ved enkelte av våre forskningsspørsmål. En kvantitativ deduktiv metode er en alternativ metode for oss, men våre funn ville ikke vært pålitelige ettersom en vanlig revisor har lite eller ingen kunnskap om KI. I så fall måtte hele undersøkelsesopplegget blitt vesentlig lagt om med hensyn til en slik fremgangsmåte.

Vi ønsker å besvare majoriteten av våre forskningsspørsmål gjennom en kvalitativ studie. Formålet er å belyse mulighetene og utfordringene rundt KI ved å tilføre ny teori på området, bekrefte eller avkrefte litteraturen og kartlegge de praktiske implikasjonene. Studien vil altså være en empiribasert induktiv og intensiv studie. En kvalitativ metode vil hjelpe oss med nyanserte data, evne til å gå i dybden og er følsom for uventede forhold, og dermed åpen for kontekstuelle forhold (Jacobsen, 2005, s. 62). Saunders et al. (2015) underbygger at et undersøkende design derfor vil være relevant og benyttes ofte ved semi-strukturerte intervju. Vi benytter denne metodiske fremgangsmåten ettersom vi ønsker å klargjøre et fenomen hvor det er viktig å få frem nyanser og meninger.

Selv om vi i utgangspunktet har en kvalitativ studie, har vi benyttet muligheten for metodetriangulering. Vi utarbeidet kvantitative spørsmål som skulle supplere problemstillingen. Hver respondent fikk tilsendt én side med elleve kvantitative påstander de skulle ta høyde for. Dette var et avkryssningsskjema med standardiserte påstander som skulle tolkes likt av respondentene. Dette benyttes som et supplement i vår forskning rundt konkrete påstander av KI i revisjon hvor vi ikke har behov for nyanser eller meninger. Det var likevel plass til kommentarer ved hver påstand. Kombinasjon av en kvalitativ og kvantitativ studie kalles for metodetriangulering (Jacobsen, 2005, s. 124). Dette innebærer en kombinasjon hvor kvantitativ og kvalitativ metode kan utfylle hverandre. Vi vil derfor ha en pragmatisk tilnærming i vår studie (2005, s. 41). Kvalitative og kvantitative metoder representerer alternative, men ikke gjensidig utelukkende forskningsstrategier hvor de svært ofte kan supplere hverandre (Sekundærkilde, 2005, s. 41). Dette gir ofte økt validitet og reliabilitet, ettersom det som er en ulempe i den ene tilnærmingen ofte er en fordel i den andre (2005, s. 229).

5.3 Datainnsamling

Hvordan intervjuene skal gjennomføres er et valg for studien. Vi har valgt et åpent individuelt intervju med hver kandidat. Dette kan utføres personlig (ansikt-til-ansikt), via telefon eller internett (Jacobsen, 2005, s. 142). Alle disse teknikkene vektlegger detaljer, nyanserikdom og det unike ved hver enkelt respondent (2005, s. 129). Jacobsen (2005, s.144) påpeker at personlig intervju virker som det beste med hensyn til gyldighet og pålitelighet.

Et personlig intervju har til fordel med å skape et miljø av fortrolighet og hvor intervjueren har mulighet til å lese kroppsspråk. Lesing av kroppsspråk kan gi en bedre forståelse av hva respondenten mener. Det vil også være vanskeligere for kandidaten å snakke usant, fordi det skapes mer personlig kontakt, enn ved telefonintervju eller via internett. En ulempe er at forskeren kan negativt påvirke intervjukandidatene til å opptre unormalt på grunn av sitt nærvær, omtalt som intervju-effekten. Dette handler om at intervjukandidat og intervjuer kan påvirke hverandre uten å være klar over det selv, gjennom ordene som brukes, kroppsspråk og ansiktsuttrykk (Jacobsen, 2005, s. 143-144). Oppfattelsen intervjukandidaten har av intervjueren kan også påvirke svarene.

Teknikken vi skal benytte for å oppnå dybdeforståelse og samle inn primærdata vil være ved å gjennomføre semi-strukturerte intervju. Dette gir oss fleksibilitet til å stille åpne spørsmål som kan føre til mer detaljert informasjon og ta høyde for uforutsette forhold, samtidig som vi har en intervjuguide å forholde oss til. Dette omtales ofte for dybdeintervju (Saunders et al., 2015). Slike intervju innebærer at tema og spørsmål som stilles er bestemt på forhånd og hvor respondentene har stor frihet til å svare slik de ønsker. Fordelen ved semi-strukturerte intervju er at vi som intervjuere lett kan stille oppfølgingsspørsmål der vi føler behov for det. Fleksibilitet og struktur bidrar i stor grad til å sikre at spørsmålene vi stiller dekkes.

Jacobsen (2005, s. 149) hevder en tid mellom 1 - 1,5 time anses som optimalt for at respondenten skal få utdypet sine meninger, samtidig kort nok til å ikke bli utmattet. Gjennom pilotintervju har vi kommet frem til en intervjulengde på omtrent en time. Ved gjennomføring av intervju har vi benyttet lydopptak, for å sikre at vi får med all nødvendig og relevant informasjon. Dette er en viktig faktor for at vi skal kunne transkribere og analysere dataene hensiktsmessig. Fordelen med dette er at vi kan opprettholde naturlig samtalekontakt og sikre at intervjuet får best mulig flyt (Jacobsen, 2005, s. 148). På en annen side kan en ulempe ved lydopptak være at informantene blir tilbakeholdne med informasjon eller reagere negativt.

Vi har avholdt seks intervjuer. Opprinnelig ble ni kandidater forespurt, men tre takket nei. Likevel mener vi at vi har klart å nå vårt metningspunkt, hvor flere

intervjuer ikke ville vært hensiktsmessig med tanke på ny informasjon. Dette knytter seg til loven om gradvis avtakende informasjon (Jacobsen, 2005, s. 142). Dybdeintervju er tidkrevende datainnsamlingsmetode hvilket medfører at utvalget må begrenses, med hensyn til tid. Vi har gjennomført fire personlige intervjuer og to via telefon. Alle intervjuene ble gjennomført i eller fra Oslo.

Det finnes flere alternative kvalitative datainnsamlingsmetoder. En alternativ metode ved vår forskning er gruppeintervju. Gruppeintervju er å intervjuere en gruppe samtidig (Jacobsen, 2005, s. 154). Her kunne vi samlet flere med kunnskap om temaet til en samtale og diskusjon rundt muligheter og utfordringer. Denne datainnsamlingsmetoden er derimot utfordrende å gjennomføre ettersom revisorer er travle mennesker med dyrebar tid. Andre potensielle metoder ville vært å foreta en observasjon av dette i praksis eller gjennomført et case. De sistnevnte metodene er tidkrevende hvor mange av våre funn måtte blitt hemmeligholdt, ettersom revisjonsbransjen består av konkurrerende virksomheter. Vi ønsket derimot ikke en konfidensiell oppgave.

Responsraten på spørreskjemaet vårt var høy da dette ble avtalt samtidig som intervjuet. Vårt kvantitative spørreskjema har kun seks respondenter som gjør det svært vanskelig å generalisere funnene. Vi hadde fokus på ordlegging og spørsmålsrekkefølgen, det visuelle designet og utformingen av spørsmålene er viktig (Statistisk sentralbyrå, 2006). Formålet er å få noen flere konkrete synspunkter på KI i revisjon, som enklere lar seg gjennomføre gjennom et avkrysningsskjema fremfor et intervju. Samtidig er det satt av rom for kommentar til hver påstand, dersom de ønsker å utdype. Ettersom vår studie primært forsker kvalitativt med et lite antall respondenter, har vi valgt å ikke gjennomføre noen avanserte analyser av den kvantitative forskningen. Den vil kun benyttes som et supplement under våre forskningsspørsmål. Det er likevel interessant å avdekke om det er samsvar mellom respondentene eller ikke på de ulike påstandene. Til analyse av spørreskjemaet benyttes analyseprogrammet Stata. Målenivået ble på forhånd bestemt til å være en kontinuerlig linje med lik avstand. Etter utført undersøkelse benyttet vi en likert-skala. Svaralternativene ble inndelt etter en likert-skala på fem nivåer, hvor 1 er helt uenig, 2 delvis uenig, 3 verken eller, 4 delvis enig og 5 helt enig. Likert-skala måler holdningen og respondentens styrke i henhold til påstanden. I ytterpunktene måler vi de ekstreme, mens verdi 3

tilsvarende en nøytral posisjon (Schmee & Oppenlander, 2010, s. 14). Målenivået vil være diskrete variabler og ordinalnivå. Disse er gjensidig utelukkende og kan rangeres logisk (Tjønndal, 2018, s. 28). Schmee og Oppenlander (2010, s. 15) påpeker at selv om disse måleskalaene tilsynelatende er ordinale, behandles de ofte som kontinuerlige.

5.3.1 Intervjukandidater

Det er nødvendig å redusere problemer knyttet til gyldighet. Vi må sikre at vi har fått tak i den informasjonen vi ønsket å få tak i. For å imøtekomme dette bør våre respondenter ha den kompetansen som behøves. Et vanlig problem er å ikke få tak i de riktige kildene, i tillegg til å være oppmerksomme på forhold som kan virke inn på respondentens motivasjon til å gi riktig informasjon. Dette forsøker vi å løse ved å intervjuere personer fra samme bransje, samt distribuere oppgaven og dens funn til de intervjukandidatene som var villige til å stille opp.

Populasjonen kan sies å være både stor og begrenset. Aktuelle kandidater har vært praktiserende revisorer, IT-revisorer og andre datascientister som jobber i revisjon. Mange praktiserende revisorer har derimot lite eller ingen kunnskap om KI. Det ville derfor vært uheldig for vår studie å intervju disse. En optimal intervjukandidat for denne forskningen er derfor en person som har tilknytning til begge temaene. Enten gjennom opplevelse og erfaring av KI eller med kunnskap og interesse. Vår utvalgsstrategi har vært å henvende oss til personer som kan tilfredsstille disse kriteriene. Av praktiske hensyn ble utvalget vårt tatt i et begrenset geografisk område, med ett unntak. Flere av de vi henvendte oss til har hovedkontor og flest ansatte her i Oslo. Det ble på forhånd satt et mål på cirka seks intervjukandidater, mens vi henvendte oss til ni potensielle.

Vi tok kontakt med de fem største revisjonsselskapene, et mellomstort revisjonsselskap, Revisorforeningen og Finanstilsynet for å skape størst mulig grunnlag for sammenligning og analyse. Vi etterspurte kandidater med kompetanse eller erfaring innen både KI og revisjon. Vår forskning baserer seg på en homogen gruppe, altså respondenter som har en rekke fellestrekk. De har derimot noe ulik bakgrunn og erfaring. I tabell 5 presenteres kort hver av de seks kandidatene som stilte opp til intervju.

Vi har intervjuet følgende seks kandidater:

| Informant | Selskap | Dato | Sted | Varighet | Type |
|--------------------|-------------------|-------------|-------------|-----------------|-------------|
| Siv Irene Aasen | BDO | 26.04.2018 | BDO | 49 min | Personlig |
| Jo Sigurd Pedersen | KPMG | 26.04.2018 | KPMG | 65 min | Personlig |
| Reidar Ludvigsen | Deloitte | 02.05.2018 | Deloitte | 66 min | Personlig |
| Torkil Hindberg | PwC | 02.05.2018 | BI | 52 min | Telefon |
| Ruben Bjerketveit | DnR | 03.05.2018 | DnR | 67 min | Personlig |
| Rolf Kåre Sæther | Nordmøre Revisjon | 07.05.2018 | BI | 35 min | Telefon |

Tabell 5: Intervjukandidater

Siv Irene Aasen er leder for IT-risiko tjenester i BDO. Hun er utdannet revisor og har nærmere tyve års erfaring med revisjon. Hun begynte hos BDO i 2013 og før det EY. Nå er hun partner i sin avdeling hvor hun til daglig jobber med IT-revisjon.

Jo Sigurd Pedersen er leder av «Data and Analytics» i revisjon hos KPMG Oslo. Han er direktør og vært i KPMG i tolv år, tidligere innom EY. Han har bakgrunn fra IT og finans. Til vår oppgave bringer han med seg teknologisk kompetanse.

Reidar Ludvigsen er partner i «Audit and Assurance» hos Deloitte og vært der siden 2004. Han har opprinnelig en siviløkonom-utdanning fra NHH og senere MRR fra BI. Reidar har mest erfaring fra store internasjonale børsnoterte selskaper, men har de siste årene jobbet med innovasjon og analytics. Han er «Audit Innovation Leader» i Deloitte Norge og «Analytics Leader» i Norden.

Torkil Hindberg er partner i «Risk Advisory Services» og «Data Analytics» hos PwC, en del av PwC Assurance. Han er ansvarlig for «Digital Assurance» som er revisjon av og med IT. Torkil har vært i PwC siden 2012, med tolv år i KPMG.

Ruben Bjerketveit er rådgiver i Den norske Revisorforening, hvor han har vært i 7 år. Han jobber med bransjens rammebetingelser, er rådgiver for medlemmer og kursholder. Han er statsautorisert revisor fra NHH. Ruben har erfaring fra Revisorkollegiet, i tillegg til nesten syv år i KPMG.

Rolf Kåre Sæther er statsautorisert og oppdragsansvarlig revisor hos Nordmøre Revisjon. Han har tretti års erfaring fra yrket, styremedlem i Revisorforeningen og har en rolle ved MRR-studiet hos NTNU, med oppstart 2018.

5.3.2 Rekrutteringsprosess

Vi henvendte oss til alle de fem store revisjonsselskapene. Potensielle kandidater og de vi hadde kjennskap til ble forespurt over e-post. Noen henviste oss videre til en mer egnet kandidat. Alle de aktuelle kandidatene fikk deretter e-post med intervjuguide, god tid i forveien, slik at det var mulighet for forberedelse eller spørsmål ved eventuelle uklarheter. Videre ble tidspunkt for intervju avtalt. Vi fikk opplevelsen av at de fleste var positive til vår henvendelse. Vi er veldig takknemlige for de seks som valgte å stille til intervju. Samtidig har vi forståelse for de som ønsket å stå over.

5.3.3 Gjennomføring av intervju

Samtlige personlige intervjuer ble avholdt i respondentenes egne lokaler. Dette er et sted intervjukandidaten er godt kjent og fortrolig med. Konteksten rundt intervjuet har en tendens til å påvirke innholdet i intervjuet og omtales for konteksteffekten (Jacobsen, 2005, s. 147). De ble på forhånd, i intervjuguiden, opplyst om en intervjutid på omtrent en time. Til lydinnspilling ble telefonopptak anvendt, slik at vi som intervjuere kunne følge intervjuet og stille oppfølgings-spørsmål. Dette ble både opplyst i intervjuguiden og rett før intervjuet startet. Årsaken til dette var å sikre all nødvendig informasjon, muligheten for å benytte sitater og for å hindre at informasjon ble oppstykket eller tatt ut av kontekst.

Gjennom kvalitetssikring av intervjuguiden fikk vi tilbakemelding på at noen av spørsmålene kunne være krevende for kandidatene å svare på. Dette tok vi til etterretning hvor vi forsøkte å hjelpe kandidaten når vedkommende stod fast. «Tre-med-grener-modellen» er benyttet som grunnlag. Intervjuguiden presenterer hovedtema med grener til de enkelte temaene innenfor. Dette er en hensiktsmessig modell dersom forskeren vet hvilke tema som skal utforskes (Thagaard, 2013, s. 102-103). Intervjuguiden startet med noen generelle spørsmål for deretter å gå innpå de mer krevende. Vi avsluttet intervjuet på en mild måte gjennom noen enklere avslutningsspørsmål.

5.4 Forskerens utfordringer

Jacobsen (2005, s. 17) påpeker at det finnes ingen perfekt forskningsprosess og at prosessene er beheftet med feil, svakheter eller manglende presisjon. Noe av formålet med forskningsmetode er derimot å kunne gjøre rede for potensielle svakheter rundt disse resultatene. Jacobsen (2005, s. 18 & 131) trekker fram undersøkelseeffekt. Det er en viss mulighet for at våre resultater er skapt av undersøkelsen vår fremfor å måle hvordan respondenten opplever fenomenet.

Postholm (2010, s. 40-41) presenterer en rekke problemer som en kvalitativ forsker bør være oppmerksom på og forsøke å imøtekomme. For eksempel bør forskeren være oppmerksom på at en kvalitativ studie krever mye tid og man bør sette seg grundig inn i temaet på forhånd. Gjennom studien må man være åpen for forhold som man på forhånd ikke har vurdert, i tillegg til at man ikke skal være forutinntatt. Forskeren skal kunne, uten problemer, være i stand til å innhente nødvendige data for å besvare problemstillingen på en beskrivende og detaljert måte (Postholm, 2010, s. 40-41). Jacobsen (2005, s. 130-131) beskriver kvalitativ forskning som ressurskrevende. Videre supplerer han med at prioritering og generalisering er utfordrende når man prioriterer mange variabler med få enheter. En annen ulempe er at det er vanskelig å tolke innhentet informasjonen på grunn av sin nyanserikdom. Dataen er kompleks og vi kan ubevisst sile vekk viktig informasjon. Fleksibiliteten kan skape nærhet mellom forskeren og informantene. Dette kan gi en negativ effekt med hensyn til forskerens personlige verdier og evne til å påvirke undersøkelsen. To av våre intervjuer ble avholdt via telefon. Jacobsen (2005, s.144) trekker frem en svakhet ved at man mister muligheten til å observere hvordan intervjukandidaten opptrer når det benyttes telefonintervju.

Kvalitativ forskning har også blitt kritisert for en rekke forhold presentert av Jonker (2007, s. 88-89). Eksempler på kritikken er manglende teori i forkant av forskningen, vanskeligheter med å replisere studiet og mangel på objektivitet. Vi har forsøkt å imøtekomme disse utfordringene på en hensiktsmessig måte. Vi har sørget for robust litteratur som grunnlag for videre forskning. Vi har forsøkt å opprettholde objektivitet ved å belyse begge sider rundt problemstillingen. Forskningen har også vært et selvstendig arbeid uten særlig påvirkning fra andre parter. Vi er ytterligere åpne for ethvert avdekket forhold i vår studie.

I all forskning er det nødvendig med god forskningsetikk. Forskningsetikk innebærer en drøftelse av potensielle «trusler» en forsker kan stå overfor. I vurderingen mellom forskerens ønske om fullstendig og best mulig informasjon på den ene siden, har vi følgende tre på den andre siden (Jacobsen, 2005, s. 43):

- respondentens krav på privatliv, integritet og personvern
- oppdragsgivers ønsker og preferanser
- samfunnets normer og verdier

Det er tre grunnleggende krav for god forskningsetikk i Norge knyttet til forskeren og dem det forskes på. Informert samtykke, krav på privatliv og krav på å bli korrekt gjengitt (Jacobsen, 2005, s. 45). Vi ønsker å ivareta informantene på en så god måte som mulig. I vår forespørsel til potensielle intervju kandidater er det tilbud om anonymitet, respondenten kan unnlate å svare på spesifikke spørsmål, samt avbryte intervjuet underveis. Samtidig blir all informasjon fra intervjuene strengt behandlet, kun benyttet i forbindelse med denne forskningen og ikke spredd under noen omstendighet.

Vi har oppfattet at revisjonsselskapene er forsiktige med å røpe selskaps-hemmeligheter og annen intern informasjon. Dette hensynet er forsøkt ivarett ved å generalisere spørsmålene for best mulig dialog. Ingen av intervju-kandidatene uttrykte et ønske om anonymitet. For å øke oppgavens kredibilitet ble intervju kandidatene presentert i delkapittel 5.3.1. I kapittel 6, resultat og analyse, har vi valgt å anonymisere intervju kandidatene gjennom å representere hver kandidat med egen bokstav.

Denne studien skrives med hjelp av et scholarship fra PwC. Dette kan skape problematikk for studiens reliabilitet. PwC var med på å velge tema og ønsket en studie om KI i revisjon. Et eksempel på den type problematikk kan være for sterke avhengighetsbånd mellom «oppdragsgiver» og forsker. For eksempel kan forskeren unnlate å publisere funn som er negative for «oppdragsgiver». Vi vil understreke at vi er bevisste på denne problematikken og ønsker ikke å bidra til upubliserte funn. Med et utvalg på seks personer må man være forsiktig med å generalisere funn som er gjort. Vår utvalgsmetode var subjektiv og påvirket av hvem selskapene valgte ut som intervju kandidater. Vi mener likevel at de vi har intervjuet er med på å gi et godt bilde på problemstillingen vår.

5.5 Validitet og Reliabilitet

For å kunne utføre en kvalitativ studie har vi i litteraturgjennomgangen forsøkt å forstå dagens situasjon gjennom perspektivet til revisjonsbransjen (Postholm, 2010, s. 17). Jacobsen (2005, s. 19) forteller at empirien vi ønsker å samle inn må tilfredsstillende to krav:

1. Empirien må ha gyldighet og relevans (validitet)
2. Empirien må ha pålitelighet og troverdighet (reliabilitet)

Validitet handler om at vi måler det vi ønsker å måle, at det oppfattes som relevant og at vi kan generalisere det vi måler (Jacobsen, 2005, s. 19). Jacobsen (2005, s.19-20) skiller mellom intern og ekstern gyldighet. Intern gyldighet er hvorvidt vi har dekning i vår empiri (data vi har samlet inn) for de konklusjonene vi trekker. Ekstern gyldighet dreier seg om resultater kan generaliseres, altså om resultater fra vårt forskningsprosjekt er gyldige også i andre sammenhenger. Begrepsgyldighet og relevans går på om vi faktisk måler det vi tror vi måler.

Reliabilitet betyr at forskningen må være til å stole på og gjennomført på en troverdig måte (Jacobsen, 2005, s. 20). Reliabilitet kan styrkes ved at flere forskere samarbeider, slik som i vår studie (Jacobsen, 2005, s. 19-20; Thagaard, 2013, s. 203). Jacobsen (2005, s. 225) påpeker derimot at undersøkelsesmåten kan påvirke resultatene vi kommer fram til. Intern reliabilitet omhandler konsistens, hvorvidt en annen forsker ville kommet frem til tilsvarende resultat ved samme datamateriale. Ekstern reliabilitet er i hvilken grad denne forskningen kan gjentas på et annet datamateriale og oppnå samme resultat som oss (Zohrabi, 2013, s. 260).

Åpenhet og eksplisitt avklaring av metodiske valg er en viktig forutsetning for å sikre studiens gyldighet og pålitelighet (Jacobsen, 2005, s. 371). Vårt formål med denne forskningen er som tidligere nevnt å kartlegge muligheter og utfordringer rundt KI i revisjon. Vi har tatt utgangspunkt i forskningsspørsmålene i utarbeidelsen av intervjuguiden. Dette skal sikre begrepsgyldighet og relevans slik at vi får svar på det vi lurer på. Våre funn og vurderinger fra intervjuene reflekterer problemstillingene bransjen står ovenfor og derfor relevante for virkeligheten.

Vi har forsøkt å intervju nøkkelpersonell som sitter på kunnskap om både KI og revisjon. Dette skal sikre oppnåelse av troverdig og pålitelig informasjon. Ettersom dette har vært generell forskning for bransjen og ikke selskapsspesifikk, har vi valgt å ikke anonymisere selve informantene innledningsvis, men gjøre dette i kapittel 6. Dette bidrar til å sikre at våre resultater og funn er pålitelige og troverdige. Intervjukandidatene tilfører studien innhold ved at de har ulik bakgrunn og erfaring, hvor de på mange måter komplementerer hverandre. Jacobsen (2005, s.142) understreker at individuelle intervjuer fremskaffer den enkeltes respondens holdninger, oppfatninger og individuelle synspunkter og kan derfor ha problemer med gyldigheten når vi forsøker å få tak på hva en gruppe mener - eksempelvis hele bransjen.

Vi har intervjuet gode fremstående representanter for bransjen med relasjon til teknologi, analytics, IT, KI og/eller revisjon med innsikt i temaet. Disse intervjukandidatene kan gi oss en indikasjon på muligheter og utfordringer som er beheftet rundt bruk av KI. Dette kan utdypes ved at vi mener intervjukandidatene har kompetanse og erfaring til å kunne uttale seg om vår problemstilling. Selv om de kan besitte god kunnskap om vår problemstilling, utgjør et utvalg på seks personer en utfordring rundt muligheten for generalisering til hele bransjen, omtalt som statistisk generalisering (Jacobsen, 2005, s. 222). Statistisk generalisering har ikke vært hovedformålet i denne studien. På den andre siden supplerer informantene nyanser som kan benyttes til teoretisk generalisering og kartlegging av de praktiske implikasjonene. Videre påpeker Jacobsen at teoretisk generalisering er en styrke ved kvalitative metoder. Den kan derimot ha noe, men begrenset overførbarhet til andre typer revisjon og revisorer ettersom vårt fokusområde for empirien har vært den eksterne finansielle revisjonen.

Vår empiri er innhentet fra fem ulike revisjonsselskap, samt Revisorforeningen. Dette er et praktisk valg da det ofte er de største selskapene som er først ute med å teste og utvikle nye avanserte teknologiske verktøy. Disse revisjonshusene sitter også på mye kompetanse ved å være globale og ha mange ansatte. Hvor mye erfaring eller kjennskap informantene hadde til KI var noe sprikende, og dermed en grunn til at vi har innhentet data fra flere kilder.

Validiteten og reliabiliteten kan påvirkes negativt gjennom at intervjukandidatene ikke har satt seg godt nok inn i temaet på forhånd. Det kan også være at vi som intervjuere ikke har formidlet problemstillingen eller spørsmålene tydelig nok. Intervjueffekten gjennom hvordan spørsmålene stilles og oppfattes vil påvirke svarene. Alle disse momentene kan svekke studiens validitet og reliabilitet ved å ikke ta høyde for slik problematikk tidlig nok. Vi har forsøkt å imøtekomme dette gjennom en grundig, ryddig og tidsriktig rekkefølge, samt kvalitetssikring av intervjuguiden. Jacobsen (2005, s. 228-229) sier intervjueren kan påvirke troverdigheten i negativ retning. Dette kan forekomme gjennom slurv hos intervjueren i nedtegning og analyse av dataen. En annen faktor er tolkningsfeil. Tolkingsfeil eller feilslutning er når vi har tolket mer inn i dataen enn det vi hadde datagrunnlag til (Jacobsen, 2005, s. 374-375). Vi mener derimot at vi har håndtert dette på en god måte ved å benytte lydopptak. Dette gjør at vi kan spole og sikre nedtegningen. Samtidig har vi også vært to personer som har etterprøvd hverandre og kryssjekket informasjonen i etterkant. Vi mener derfor at nevnte muligheter og utfordringer ellers er både valide og reliable, det vil si at undersøkelsens totale gyldighet er god, selv om det finnes ingen perfekt undersøkelse (Jacobsen, 2005, s. 388).

Vår presentasjon av en robust litteraturgjennomgang skal være med å støtte opp under eller avdekke ulikheter i vårt datamateriale, i kapittel 6. Vi har vektlagt at oppgaven skal reflektere nøyaktighet, troverdighet og pålitelighet. Ved å holde fokus på ordlegging og relevant oppbygging av påstandene i spørreskjemaet, har vi forsøkt å sikre typiske tilfeldige og systematiske målefeil og dermed opprettholde høy reliabilitet og validitet (Tjønndal, 2018, s. 30). Dette gjelder til tross for et lite antall respondenter og at generalisering vil være problematisk. Reliabiliteten til våre funn fra spørreskjemaet kan påvirkes av at vi har benyttet en likert-skala i etterkant av utførelse.

Vi mener også at løpende kontakt med vår eksterne veileder ved PwC har gitt mulighet for raske og fleksible svar underveis. Dette har vært med på å sikre innholdet, samt at problemstillingen og våre forskningsspørsmål har blitt tilstrekkelig og hensiktsmessig besvart.

6. Resultat og analyse

Primærdata ble innhentet gjennom intervju og spørreskjema. Dette danner grunnlaget for våre konklusjoner på hvert forskningsspørsmål, sett i lys av litteraturen. Hvert enkelt forskningsspørsmål blir drøftet i rekkefølge med en påfølgende konklusjon. Vi har systematisert, kategorisert og sammenbundet data, hvor all informasjon om hvert enkelt forskningsspørsmål kan anses for å være en egen kategori. Intervjuguiden er på mange måter koblet opp mot flere av forskningsspørsmålene, hvilket gjorde at mye ble kategorisert automatisk og reduserte vårt arbeid i etterkant. Derfor er kapittel seks på mange måter bygd opp på tilsvarende måte som forskningsspørsmålene og intervjuguiden. Vi har benyttet analyseprogrammet Stata til analyse av de kvantitative påstandene, ved hjelp av en likert-skala. Først presenteres intervjukandidatenes erfaring og deres selskaps holdning til KI. Deretter analyserer og drøfter vi hvert enkelt forskningsspørsmål fortløpende etter nummerering. Det er også i denne rekkefølgen litteraturgjennomgangen i kapittel fire i stor grad er bygd opp på.

6.1 Innledende analyse

Innledningsvis ble kandidatene forespurt deres erfaring rundt KI. Majoriteten av kandidatene har en form for erfaring med KI. Både informant B, C og D har testet revisjonsverktøy deres selskaper selv har utviklet, eller i samarbeid med eksterne aktører. E forteller det er lite utstrakt bruk av KI i revisjon, men trekker frem at regnskapsbransjen har kommet ganske langt på området. Dette understøtter F. Også A bekrefter dette med at de ikke er i drift på KI i dag, men benytter datamaskiner som et støtteverktøy i revisjon. Her benytter de stordata i utstrakt grad. A er enig med E og F, om at også hos de er det regnskapsavdelingen som har kommet lengst på bruk av KI. B og D forteller de har gjennomført noen piloter og prosjekter med bruk av KI i revisjon og hevder de har kommet et godt stykke på vei. Informant C viser til at man ikke har gode nok verktøy som aktivt kan brukes i revisjon i dag, men at de benytter maskinlæring til å analysere dokumenter hos store klienter.

Dette er i overensstemmelse med litteraturen fra kapittel 4. Bransjen investerer både internt og med eksterne aktører hvor KI testes, utvikles og implementeres.

På spørsmål om hvilken holdning selskapene har til KI fikk vi et entydig svar fra alle, positiv. Det er bred enighet om at alle er åpne for ny teknologi, herunder KI. De største revisjonsselskapene har startet sine investeringer innen KI. Flere av informantene påpekte at dette var et satsingsområde. Person B og C påpeker at de største investeringene skjer internasjonalt og utviklede verktøy senere distribueres til Norge. En utfordring C påpeker, med støtte fra B, er at verktøyene ofte er utviklet med hensyn til store multinasjonale selskaper og er derfor ofte ikke like anvendbare på norske klienter. B legger til at kundene ønsker dette i større grad fra revisor. At investering og utvikling skjer internasjonalt bryter med prinsippet om at bransjen må være mer dynamisk, mener D.

6.2 Forskningsspørsmål 1

«Hvilke muligheter kan KI medføre i revisjonsprosessen?»

Forskningsspørsmål 1 favner bredt og sammenfaller til en viss grad med noen av de øvrige forskningsspørsmålene. Denne delen må derfor ses i sammenheng med disse. Det er bred enighet i svarene fra informantene om mulighetene ved bruk av KI i revisjonsprosessen. Dette gjelder spesielt i innsamling av data, analyser og risikovurdering. Vi finner derimot en uenighet mellom informantene om hvorvidt KI vil benyttes i standardiserte rutineoppgaver.

Majoriteten av informantene mener KI vil være et viktig verktøy for å analysere data og til å utføre overordnede analyser i planleggingsfasen for å effektivt identifisere risiko. Samtidig trekker D fram at KI kan gjøre en analyse av regnskapet med hensikt å se hvordan regnskapet henger sammen. Dette er nyttig for å identifisere unormale transaksjoner og avvik. F legger til at analytiske handlinger og avviksanalyser kan gjøres mer effektivt ved hjelp av KI. C hevder man vil spare masse tid i planleggingsfasen med hensyn til honoraret man forholder seg til. Videre forteller C at KI kan prosessere store mengder data og vil kunne bidra til å identifisere risiko og røde flagg, med støtte fra E som supplerer med at maskinen kan gjenkjenne type bransje og område og dermed foreta en risikoanalyse ved å sammenligne og binde stordata inn i dette. Selskapet til C benytter internasjonalt et verktøy med mulighet for å samle inn store mengder informasjon om et selskap, som deretter kan bygge en risikoprofil av selskapet.

“Det tror jeg kan være nyttig, for det er så mye informasjon der som vi ikke klarer å prosessere” sier C. KI kan strukturere og sette ting i sammenheng, innhente informasjon, benytte tidligere erfaring om kunde, bransje og inntrufne hendelser, og gi et bedre grunnlag enn det vi har i dag, eksemplifiserer B. Videre trekker A frem KI som et velegnet verktøy til å håndtere stordata og bidra i analyser. KI vil kunne bearbeide mye mer informasjon om populasjonene og inkludere alle forhold, forteller B, med støtte fra de øvrige informantene.

F mener KI kan innhente tallmaterialet på en annen måte enn i dag og lage handlinger som er mer effektive for å få fram tallmaterialet. KI kan benyttes til å sammenligne virksomheten opp mot bransjen og trender i markedet og gi bedre innfallsvinkel til revisjonsstrategien, ved å foreslå handlinger for å dekke identifisert risiko, ifølge A. Person B er enig i dette synspunktet, hvor man kan få mye bedre input ved å bruke KI til å innhente informasjon og gi forslag til videre handlinger. D mener KI kan understøtte det meste så lenge man har tilgang til data. Videre hevder D at maskinen kan lese intern og ekstern informasjon for å forstå virksomheten og på bakgrunn av dette kan KI utarbeide en revisjonsplan. Videre poengteres det at KI vil være i stand til å identifisere hendelser basert på erfaring med regnskapet til det enkelte selskap. I tillegg nevner D muligheten til å trene algoritmen(e) på mange tilsvarende virksomheter i en bransje og dermed få et bedre beslutningsgrunnlag.

Informant B understreker at man ikke har behov for KI til enkle type ting. B mener KI kan automatisere mye av de strukturerte repetitive oppgavene revisor gjør, men presiserer at KI vil ha størst nytte for de vanskelige problemstillingene. Dette gjelder eksempelvis estimater. KI vil kunne ha et større grunnlag for sine vurderinger ettersom den tar for seg hele populasjoner og lærer av erfaring. Dette vil gi revisor et bedre grunnlag. Av den grunn ser B ingen begrensning på hva KI kan gjøre, men heller hvor KI gir størst nytte. C er også inne på muligheten til å analysere store mengder data for å finne avvik, som å gjennomgå hele populasjoner fremfor stikkprøver. Disse analysene vil kunne gi støtte ved skjønnsmessige vurderinger, eksempelvis estimater, ettersom det blir mulig å samle inn mye mer informasjon og deretter lage en forventning.

På spørsmål om hvilke arbeidsoppgaver KI kan utføre i revisjonen svarer E følgende: “Ja, hvilke er det han ikke kan utføre eller støtte?”. E viser derimot til at dagens lov ikke er direkte lagt opp til at KI kan utarbeide og signere revisjonsberetningen akkurat nå. I denne sammenheng poengterer E at ansvarlig revisor fremdeles må signere beretningen, som også D er enig i. Person B ser heller ingen begrensning på hva KI kan gjøre og mener det kan benyttes overalt. C supplerer med at KI kan gjøre mange typer standardiserte oppgaver vi gjør i dag, men hevder at noen oppgaver, for eksempel bankavstemminger kan automatiseres uten KI. Informant F mener derimot at KI vil være best egnet til å håndtere store datamengder og rutinepregede oppgaver.

Samtlige informanter er positive til å benytte KI i planleggingsfasen og risikovurderingen. Informant F mener KI også vil være godt egnet i avslutningsfasen. Dette gjelder i forhold til oppdatering av justeringer og endringer av tall. Dette kan være endringer som er krevende å fange opp, sett mot dokumentasjon, resonnement og enkeltposter, men vil bli enklere med et smartere saksbehandlingssystem. Ytterligere kan KI uttrykke levedyktigheten til selskapet og forutse flere faktorer enn tidligere, sier A. Informant D hevder derimot at revisjonsfasene kan bli kunstige fremover, fordi man på sikt kommer til å gjøre alt på en gang.

I kapittel 3 presenterte vi Burgess (2018) som utleder åtte kjerneområder for bruk av KI. Disse omhandler innhenting av informasjon, forklaring og forståelse med bakgrunn i informasjonen. Særlig de to førstnevnte støtter våre funn oppunder. Videre må informantenes svar ses i sammenheng med kapittel 4.1. Litteraturen fremhever at rutinepregede oppgaver kan automatiseres uten hjelp av KI og støtter oppunder C sitt synspunkt. Manuelle oppgaver, samt oppgaver som krever tankeprosesser, kan understøttes av KI. For eksempel innhenting og analyse av store mengder data slik som våre kandidater har vært inne på. Dette er også en av mulighetene ICAEW (2017a) og Macaulay (2016a) har identifisert. Våre kandidater har også vært enige om at KI er egnet til å identifisere avvik og unormale transaksjoner som flere kilder i litteraturen skriver, og er en av tre muligheter Macaulay (2016a) har oppsummert. I denne sammenheng eksemplifiserte ICAEW (2017a) med at svake eller komplekse sammenhenger lettere kan identifiseres.

Både informant B og D snakker om muligheten til å lese intern og ekstern informasjon om en virksomhet og dermed gi en vurdering, slik som Macaulay fremhever som en annen mulighet med KI. ICAEW (2017a, s. 5) fremhever fravær av følelser som en mulighet med KI (for eksempel kjedsomhet eller utmattelse). Dette er person B og E enig i, at KI ikke besitter disse myke verdiene. B sier at KI lærer etterhvert, som kan tolkes dithen at maskinen lærer av feil og ny informasjon og dermed er endringsdyktig. At en kunstig intelligent maskin er endringsdyktig er også en mulighet i følge ICAEW (2017).

Vi ba videre informantene ta stilling til følgende påstand:

Påstand 3: «KI er egnet til å identifisere risiko»

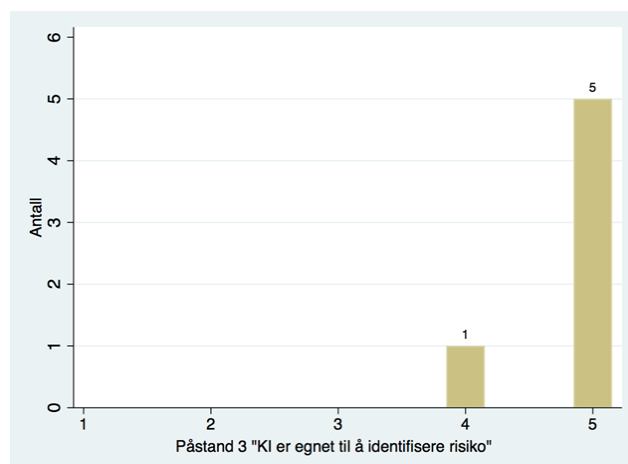
| n | Gj.snitt | Std.avvik | Median | Min | Max | Var.bredde |
|---|----------|-----------|--------|-----|-----|------------|
| 6 | 4,83 | 0,408 | 5 | 4 | 5 | 1 |

Tabell 6: Deskriptiv statistikk om påstand 3

Resultatene fra tabell 6 og figur 11 viser at samtlige respondenter har svar helt eller delvis enig med påstand 3. Det er derfor stor enighet om at bruk av KI er egnet til å identifisere risiko. Dette synspunktet reflekteres både i litteratur, intervju og spørreskjema. Respondent A

tilføyer med at KI har større grunnlag for å identifisere risiko på tvers av bransjer og trener med mer, hvor respondent D legger til styrken med KI er å behandle sammenhengen i store datamengder.

I lys av litteraturstudien støtter resultatene fra våre funn oppunder de mulighetene bruk av KI kan gi når det gjelder å innhente, analysere og vurdere data, fra interne og eksterne kilder. Videre kan det benyttes til å identifisere røde flagg og risiko. Dette gjelder både strukturerte og ustrukturerte data som KI mer hensiktsmessig kan analysere. Deretter kan KI utarbeide utkast til revisjonsplan med forslag til



Figur 11: Frekvenstabell, påstand 3

videre revisjonshandlinger. Andre identifiserte muligheter er at KI er endringsdyktig og ikke lar seg påvirke av følelser ved at den ikke besitter myke verdier. Det er også verdt å merke seg at de øvrige identifiserte mulighetene fra litteraturgjennomgangen ikke blir møtt med motstridende argumenter fra våre intervjukandidater.

6.3 Forskningsspørsmål 2

«Ved å benytte KI i revisjonsprosessen, hvilke fordeler og ulemper medfører dette?»

Dette forskningsspørsmålet overlapper enkelte andre forskningsspørsmål hvor disse må ses i sammenheng. På spørsmål om fordeler ved å ta i bruk KI i revisjon fremmer majoriteten av informantene økt revisjonskvalitet og effektivitet. To informanter trekker også frem økt innsikt som fordel. Innledningsvis kartla vi at samtlige informanter var positive til KI i revisjon. Dette kan påvirke argumentene de har mot å benytte KI i revisjon. Selv om ingen hadde rene argumenter mot KI, presenterte flere av informantene utfordringer eller bekymringer rundt KI, som kan betraktes som ulempe eller argument mot KI i revisjon.

Fordeler

Vedrørende fordeler med KI i revisjon mener A at det blir en mer riktig og kvalitativ revisjon. A begrunner høyere revisjonskvalitet ved at man har fokus på de riktige tingene og ikke bruker tid på unødvendige ting. Altså at man gjør en mer hensiktsmessig revisjon. Informant B og D er enige i at revisjonskvaliteten generelt vil øke. B trekker fram undersøkelse av hele populasjoner og andre typer kilder som faktorer for dette. Informant A antyder man vil få mer innsikt i kundens drift og bli i bedre stand til å gi mer nyttig informasjon tilbake til kunden og input til hva de bør gjøre bedre. Økt innsikt kan oppnås gjennom at KI utarbeider en rapport på selskapet, konkurransesituasjon, marked, utvikling, benchmarking i bransjer, nøkkeltall, med mer, tilføyer C. A poengterer at det vil være positivt at kvaliteten og effektiviteten øker.

C tenker den viktigste fordel med KI er at den kan være en god beslutningsstøtte for revisjonsteamet. Bedre grunnlag i beslutninger trekker også B frem. Bruk av denne beslutningsstøtten i krevende oppgaver er en viktig fordel ved KI, i tillegg til effektivitetsperspektivet. C sier KI kan «hjelp oss til å ta riktig konklusjon, hjelpe oss å utfordre ledelsen og selskapet på en god måte, basert på fakta». C viser til at dette er utfordringen i hverdagen, hvor ledelsen har god dokumentasjon på sine vurderinger og estimater. Informant D er enig i dette, og mener KI i revisjon vil gi et bedre grunnlag for å utøve profesjonelt skjønn.

Informant F mener KI kan medføre høyere presisjon og redusere risikoen for menneskelige feil. Informant B mener KI også vil kunne redusere risiko. Vedrørende risiko mener D at det er beheftet ny risiko som revisor må ta høyde for ved bruk av KI, for eksempel at algoritmene som benyttes er riktige og trent på riktig data.

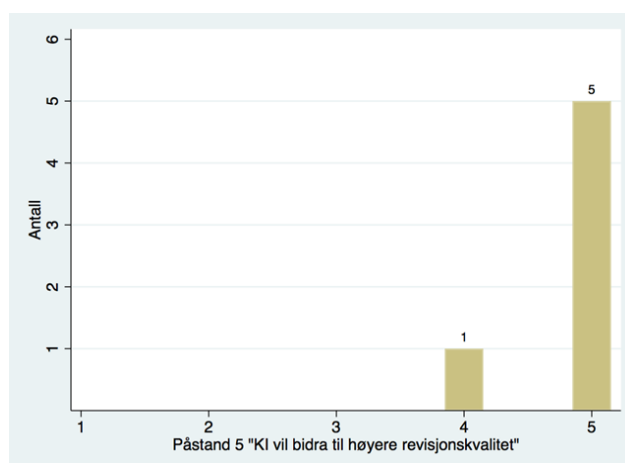
På påstand 5 fra det kvantitative skjemaet var svarfordelingen som følger:

Påstand 5: «KI vil bidra til høyere revisjonskvalitet»

| n | Gj.snitt | Std.avvik | Median | Min | Max | Var.bredde |
|---|----------|-----------|--------|-----|-----|------------|
| 6 | 4,83 | 0,408 | 5 | 4 | 5 | 1 |

Tabell 7: Deskriptiv statistikk om påstand 5

Resultatene fra tabell 7 og figur 12 viser at samtlige respondenter har svart helt eller delvis enig med påstand 5. De er derfor enige om at KI i revisjon vil heve revisjonskvaliteten. Tilsvarende funn har vi også gjort i litteratur og intervju.



Figur 12: Frekvenstabell, påstand 5

Respondent D tilføyer at KI gir mulighet for å analysere mer komplekse årsakssammenhenger enn ved manuelt, og derav høyere kvalitet.

I litteraturgjennomgangen finner vi støtte for at KI kan bidra til økt effektivitet og heve revisjonskvaliteten. Dette understøttes av at KI har stor prosesseringskraft og kan gjøre en raskere gjennomgang hvor også svake sammenhenger kan fanges opp. På bakgrunn av dette har vi kartlagt at KI kan gi styrket tillit, redusere risiko og bidra til at man får dypere innsikt i klienten. ICAEW (2017a) forklarer med at maskinen er endringsdyktig, ikke vil gjøre feil gitt riktig programmering og bidra til mer nøyaktige outputs. Raphael (2017) presenterer at dette vil frigjøre tid på enkelte områder hvor revisor kan bruke sin kompetanse og tid på mer krevende oppgaver og vurderinger. Dette understøttes av våre funn. Våre informanter tilføyer derimot at KI kan gi god beslutningsstøtte og et bedre grunnlag for beslutninger. De utdyper også viktige faktorer som ligger til grunn for økt effektivitet og revisjonskvalitet som ikke nødvendigvis litteraturen fremhever.

Ulemper

Fire av informantene sier de ikke har noen argumenter mot KI i revisjon. Flere av disse trekker derimot frem hva de anser som utfordrende, hvilket de to resterende informantene også gjør.

Undersøkelsen viser noe skepsis til å bruke KI i revisjon. Informant D påpeker at bruk av KI til å gjøre vurderinger i revisjon vil være brudd på revisjonsstandardene slik disse er i dag. Dette vil være en direkte hindring, men sier samtidig at standardene er generiske og åpner for det meste. Informant B trekker frem noen synspunkter angående usikkerhet rundt skyldspørsmål med å benytte KI i revisjon. De øvrige peker på utfordringer som at teknologien er umoden og dermed vanskelig å se ulemper. Informant E presiserer at utfordringen kommer dersom revisjonsbransjen ikke blir med på den teknologiske utviklingen, men i stedet satser på klassisk revisjon. B er enig, men påpeker at en del ting må være på plass før bruk, og trekker frem det regulatoriske. Samtidig understreker informant E at man ikke kan slippe løs noe man ikke er helt trygg på. C mener man vil slite i en liten periode, fra man begynner å bruke KI til man faktisk tør å stole på det.

Informantene har utdypet nærmere hva de mener kan være utfordringer med å ta i bruk KI i revisjon. Noen av momentene som kom fram er: Sikring av fortløpende revisjonsspor når KI gir grunnlag for og støtter risikovurderingen, gir forslag til

videre handlinger og tar egne beslutninger. Argumentasjon for at tankegangen og de vurderingene er riktige. Det etiske perspektivet med kontrollrutiner som sikrer at den gjør det den skal gjøre. Godt governance og det regulatoriske rundt. Forklarbarhet, etterprøvbarhet, pålitelighet av output og revisors ansvar ved feil eller ikke avdekket feil. Fullstendigheten og nøyaktigheten til data. Algoritmer, deres struktur og dermed sikkerhet, regulering og verifisering av KI-verktøyet. Kompetanse og personvern er også trukket frem.

Både informant A og B påpeker revisjonsdokumentasjon som en sentral usikkerhetsfaktor. Informant A forteller det kan bli utfordrende å dokumentere den utførte risikovurderingen når systemene selv tar beslutninger. B spør "Hvordan skal du forklare AI til en i Finanstilsynet?" hvor også informant C sier at det er ikke entydig at Finanstilsynet er positive til veien det går - og det hindrer innovasjonsevnen. Dette forklarer informant D med at det nesten er umulig å redegjøre for hvordan algoritmen gjør ting. Litteraturen vår problematiserer også denne problemstillingen om manglende forståelse av den interne strukturen som vi kommer tilbake til i forskningsspørsmål 21.

«Vi må sjekke at det brukes riktig input, at kalkulasjon er riktig og outputen benyttes på en riktig måte» sier D, i tillegg til å presisere at kvalitetssikring er nødvendig. Informant C er også opptatt av denne kvalitetssikringen og mener noen må inn å verifisere dersom KI benyttes som støtte i vurderinger. C er videre noe kritisk til nøyaktighet, fullstendighet og påliteligheten av KI-verktøyet, samt hvilket ansvar man har ved feil og spør om man da revidert etter revisjonsstandardene. C supplerer med:

«Jeg kommer nok i starten til å slite med å stole helt på det og det er det som er litt av utfordringen. Hvis du ikke stoler helt på det så kommer du til å gjøre det du gjorde tidligere uansett. Så da kjører du alle disse kontraktene gjennom, men i tillegg så tar du disse stikkprøvene. For det er jo det revisjonsstandardene sier, at du skal ta et utvalg.»

C mener det er de regulatoriske myndighetene som hindrer utviklingen og at utviklingen kan skje raskt. Revisjonsstandardene sin begrensning på KI diskuteres under forskningsspørsmål 19.

Flere nevnte problemstillinger i avsnitt to og tre under ulemper vil drøftes i senere forskningsspørsmål. Ulemper vi har avdekket i litteraturen går også på om man kan stole på beslutninger og vurderinger gjort av KI, i tillegg til å være kostnadskrevende investeringer. Det hevdes også at KI kan medføre uro og umoral internt i et revisjonsselskap, dette begrunnes med at KI er fryktet for å overta oppgaver og arbeidsplasser. Det er også en ulempe og utfordring hvor ansatte ikke har tilstrekkelig kompetanse vedrørende koding og programmering, eller manglende forståelse av det KI gjør. Videre kan KI bidra til at det blir mindre kontakt mellom revisor og klient, som kan anses som en ulempe gjennom enten redusert forståelse av klienten eller dårligere revisjonskvalitet. KI kan også avdekke mange avvik eller risikoområder som kan medføre en mer omfattende revisjon. Andre ulemper avdekket i litteraturen er dersom KI medfører brudd på det etiske, regulatoriske, personvern eller hvor feil ikke avdekkes og man får et erstatningsansvar. Dette hindrer bransjens innovasjonsevne. Videre presenterte Hindberg (2015) og Pedersen (2016) utfordringer rundt innhenting og analyse av data. De fleste identifiserte ulempene i litteraturen understøttes av våre funn, hvor disse ulempene også utdypes. I vår empiri finner vi også argumenter mot KI i revisjon som vi har avdekket og som utfyller litteraturen.

Våre informanter trekker frem revisjonsdokumentasjon og sikring av revisjonsspor, sikkerhet, regulering og verifisering av KI-verktøyet, samt påliteligheten av output. Dette er viktige faktorer for at man skal benytte KI hensiktsmessig og ha tillit til beslutninger og vurderinger maskinen tar for revisor. Vi betrakter våre funn som i samsvar med litteraturen, men at informantene har fått forklart og utdypet sine meninger noe mer enn teorien på visse områder. Selv om våre informanter er positive til KI, betyr ikke dette at alle deres medarbeidere entydig også er det. Litteraturen hevder KI potensielt kan overta en rekke arbeidsplasser, som også person E var inne på, og dette kan potensielt føre til umoral internt blant ansatte.

Vi har funnet støtte både i litteratur og gjennom intervju for at KI i revisjon vil medføre en rekke fordeler, særlig økt revisjonskvalitet og effektivitet. I tillegg vil KI være et godt verktøy i beslutningsstøtte, samt gi en dypere innsikt i klientens virksomhet. Revisjonen vil også kunne bli mer hensiktsmessig gjennom å fokusere på de riktige tingene. Studien viser at det ligger mange fordeler for

revisjon å implementere KI i revisjonsprosessen. Det er derimot kritiske forhold revisjonsselskapet må ta høyde for å sikre underveis. Dette gjelder avdekkede argumenter mot eller ulemper som nevnt ovenfor. Det er i første omgang forbundet høye kostnader til investeringer og videreutvikling av ansatte. Videre er det knyttet utfordringer til regulering, dokumentasjon og verifisering av KI. I denne sammenheng påpekte informant E at den største utfordringen vil oppleves blant selskaper som ikke er innovative og tar høyde for ny tilgjengelig teknologi.

6.4 Forskningsspørsmål 3

«Hvordan kan KI bidra til mer effektiv analyse, herunder å predikere presise anslag og verdier?»

Dette forskningsspørsmålet består av to deler. Hvordan KI bidrar til mer effektiv analyse og KI til å predikere anslag og verdier. Førstnevnte drøftet vi dels gjennom de øvrige forskningsspørsmålene. Ved hjelp av KI har revisor mulighet til å innhente informasjon fra mange flere kilder, enn det man tradisjonelt har hatt mulighet til. I tillegg kan KI gjennomgå hele populasjoner mer effektivt og hensiktsmessig. Altså kan KI utfra et effektivitetsperspektiv bidra til å utføre en raskere revisjon med mindre ressurser og med mulighet til å oppnå en høyere grad av sikkerhet gjennom sin populasjonsgjennomgang. Det er gitt at det er en hensiktsmessig gjennomgang av data ettersom effektiv, men ikke hensiktsmessig analyse gir lite eller ingen verdi.

Flere av intervju kandidatene var enige i å benytte KI til å innhente og analysere informasjon og støttes i vår litteraturgjennomgang. B og D presenterte en metodikk hvor informasjon kontinuerlig innhentes fra en rekke eksterne kilder, som hver for seg ikke trenger å være 100 % pålitelig, men som sammen kan være overbevisende nok. Dette kan benyttes til å få en bekreftelse på regnskapet eller et forhold. E mener dersom man innhenter informasjon fra en rekke kilder, så er det mulighet for å argumentere for at anslaget man har fått er tilnærmet riktig svar.

På del to av forskningsspørsmål tre, ba vi respondentene ta stilling til følgende påstander:

Påstand 4: «KI kan sette en presis forventningsverdi i tråd med ISA 520»

| n | Gj.snitt | Std.avvik | Median | Min | Max | Var.bredde |
|---|----------|-----------|--------|-----|-----|------------|
| 6 | 4,5 | 0,836 | 5 | 3 | 5 | 2 |

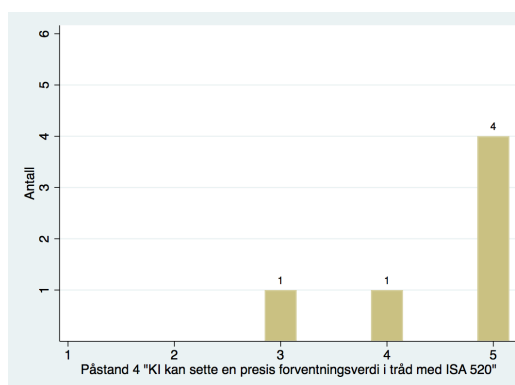
Tabell 8: Deskriptiv statistikk om påstand 4

Påstand 6: «Jeg stoler på prediksjoner gjort ved hjelp av KI»

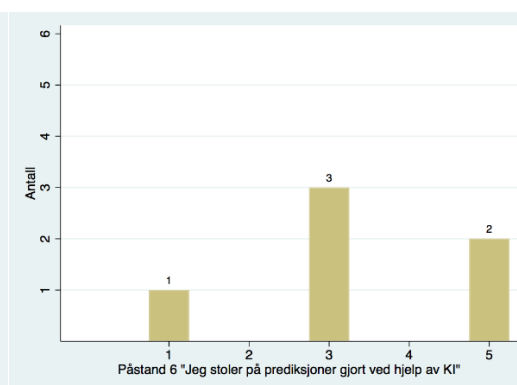
| n | Gj.snitt | Std.avvik | Median | Min | Max | Var.bredde |
|---|----------|-----------|--------|-----|-----|------------|
| 6 | 3,33 | 1,505 | 3 | 1 | 5 | 4 |

Tabell 9: Deskriptiv statistikk om påstand 6

Resultatene fra tabell 8 og figur 13 viser en gjennomsnittlig respondent svarer seg enig. Det var en respondent som forble nøytral, en som var delvis enig og fire som var helt enige. Dette indikerer at de fleste er enige i at KI kan sette en presis forventningsverdi i tråd med ISA 520. Det vesentligste er at ingen sa seg uenige. Resultatene fra tabell 9 og figur 14 viser delte meninger. Halvparten av respondentene har et nøytralt synspunkt til påstanden, samtidig som vi ser at den andre halvparten avviker i hver sin retning. Uenigheten på påstand 6 gjør det vanskelig å trekke en endelig konklusjon utover at det er beheftet delte meninger rundt det å stole på prediksjoner utført av KI. Respondent D forklarer at KI er egnet til å sette presise forventningsverdier basert på sammenhengen i datasettet, men at man ikke kan stole på prediksjonene uten at algoritmene er verifisert.



Figur 13: Frekvenstabell, påstand 4



Figur 14: Frekvenstabell, påstand 6

ICAEW (2017b) trekker frem at KI kan benyttes i analyser og prediksjoner for å gi økt treffsikkerhet. Videre kan dette gjøres ved å benytte KI til å innhente all nødvendig informasjon, relevante forhold og predikere fremtidige hendelser. I delkapittel 4.2.3, vedrørende ISA 520, kom vi fram til dokumentasjon, forståelse og skjønn er viktige faktorer når KI setter forventningsverdier. Videre har vi

kartlagt at Rutgers (2017) mener at prediktive analyser kan benyttes som relevant og pålitelig revisjonsbevis hvor de mener slike analyser kan benyttes som test med opptil tre formål (trippel-test). Informantene sier derimot lite om KI sin egnethet i prediktive analyser for å predikere fremtidige hendelser, årsak og som bevisgrunnlag.

Våre data viser KI kan bidra til mer effektiv analyse, gjennom å benytte KI i stegene innhenting, analyse og tolkning av informasjon. Flere kilder med informasjon kan bidra til høyere presisjon i prediksjon av anslag og verdier. Vi oppfatter at pålitelighetsbegrepet kan endre seg ved at hver kilde til informasjon ikke trenger å være 100 % pålitelig, men hvor flere samstemte kilder kan være overbevisende nok.

6.5 Forskningsspørsmål 5

«Kan avansert analyse mer effektivt avdekke misligheter, og vil mislighetsoppdagelse bli et større og viktigere område for revisor fremover?»

En lovpålagt oppgave som påhviler revisor er å forebygge og avdekke misligheter. Teori og litteratur viser det er stor forventning fra brukerne om at revisor skal avdekke misligheter. Litteraturen fremhever også at mislighetsoppdagelse vil bli enda viktigere i en mer moderne revisjon. Ved å benytte avansert analyse, dataanalyse og KI, vil revisor opparbeide seg en dypere innsikt som kan forhindre og avdekke misligheter på en mer effektiv måte. Ved hjelp av avansert analyse kan sammenhenger, avvik, uforventede transaksjoner og andre uregelmessigheter avdekkes i større grad. Selv om enkelte fremmer analysene ikke kategorisk kan benyttes til å identifisere misligheter, kan revisor benytte maskinens resultater til å foreta en selvstendig vurdering.

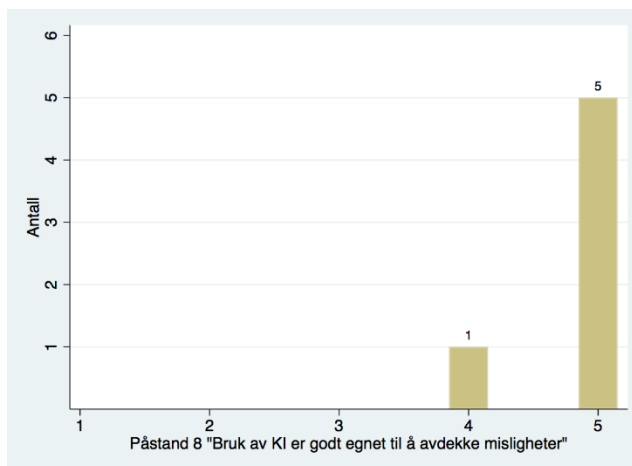
Vi ba respondentene vurdere nærmere følgende påstand:

Påstand 8: «Bruk av KI er godt egnet til å avdekke misligheter»

| n | Gj.snitt | Std.avvik | Median | Min | Max | Var.bredde |
|---|----------|-----------|--------|-----|-----|------------|
| 6 | 4,83 | 0,408 | 5 | 4 | 5 | 1 |

Tabell 10: Deskriptiv statistikk om påstand 8

Resultatene fra tabell 10 og figur 15 viser at fem av seks respondenter svarte helt enig. Det er stor enighet om bruk av KI er godt egnet til å avdekke misligheter. Dette samsvarer med våre funn i litteraturen og intervjuene. Respondent D forklarer dette med at KI kan identifisere anomaliteter i



Figur 15: Frekvenstabell, påstand 8

datasettet og dermed avdekke mislighetstransaksjoner. Samtidig vil KI-analysene være mer omfattende enn ved manuell analyse.

Vi har funnet konkrete synspunkter i litteraturen for hvordan avansert analyse kan benyttes for å bedre avdekke misligheter. Byrnes (2015) finner at datamining lettere kan avdekke misligheter og tilføre merverdi. Avanserte analytiske teknikker kan i større grad identifisere unormale transaksjoner, enn det vi mennesker har evnen til. Misligheter innebærer ofte avanserte og nøye planlagte metoder for å skjule transaksjonene.

Vi har gjennom spørreskjemaet på påstand 8 avdekket et entydig svar. Våre data indikerer bruk av KI er godt egnet til å avdekke misligheter. Dette er basert på litteratur og spørreskjema. Avansert analyse definerte vi innledningsvis som KI og dataanalyse i kombinasjon. På bakgrunn av presentert litteratur og empiri kan vi si at nye avanserte teknikker og metoder kan i større grad avdekke misligheter. Litteraturen fremhever også at mislighetsoppdagelse vil bli viktigere for revisor fremover for å møte forventningsgapet som eksisterer.

6.6 Forskningsspørsmål 6 & 7

«Hvordan kan revisjonsprosessen endres som følge av KI?»

«Hvordan kan KI bidra til en sanntidsrevisjon?»

I kapittel 2.4 presenterte vi revisjonsprosessen med bakgrunn i Birkeland (2017) og Arens et al. (2016). Førstnevnte beskrev revisors metodefrihet innen denne

selvreguleringen av revisjonsprosessen. Dette medfører at det eksisterer ulik oppfattelse og inndeling av revisjonsprosessen. I litteraturen trakk vi fram Issa et al. (2016) sin syvtrinns revisjonsprosess med bruk av KI. Utover det som er skrevet i de øvrige forskningsspørsmålene går ingen av informantene, med unntak av D, inn på hvordan revisjonsprosessen kan endres. Informant D sier følgende:

«Jeg tror det at fasene i revisjon kommer til å bli litt kunstige framover, fordi at du kommer til å gjøre alt i en smekk. Men hvis du tenker litt på kort sikt så vil du i planleggingen kunne bruke kunstig intelligens til å gjøre en analyse av regnskap, også se hvordan det regnskapet henger sammen. Om det er noen ting man må se nærmere, ved å identifisere outliers for å følge opp og da definere hva man har tilstrekkelig revisjonsbevis på, per i dag, og hva du må gjøre mer på.» Issa et al. (2016) støtter dette synet på revisjonsfasene delvis, for eksempel ved at enkelte faser kan overlape hverandre i større grad.

Vi ba respondentene ta stilling til påstandene nedenfor og fikk følgende svar:

Påstand 1: «KI bør vurdere aksept av klientforhold»

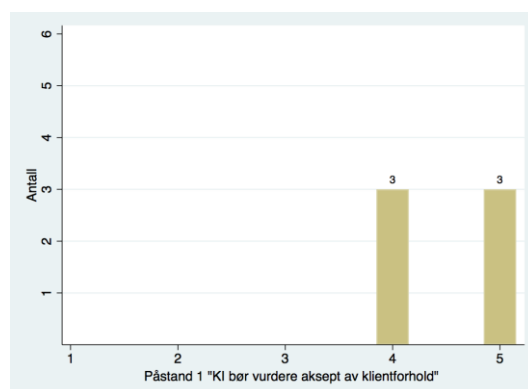
| n | Gj.snitt | Std.avvik | Median | Min | Max | Var.bredde |
|---|----------|-----------|--------|-----|-----|------------|
| 6 | 4,5 | 0,547 | 4,5 | 4 | 5 | 1 |

Tabell 11: Deskriptiv statistikk om påstand 1

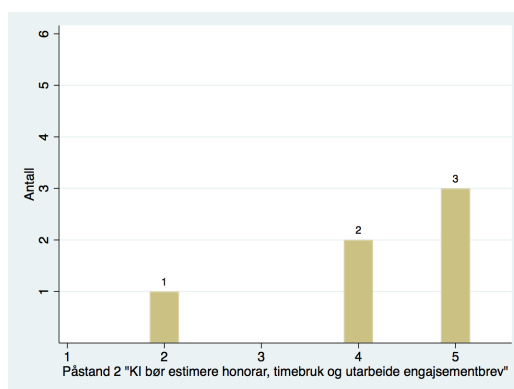
Påstand 2: «KI bør estimere honorar, timebruk og utarbeide engasjementbrev»

| n | Gj.snitt | Std.avvik | Median | Min | Max | Var.bredde |
|---|----------|-----------|--------|-----|-----|------------|
| 6 | 4,16 | 1,169 | 4,5 | 2 | 5 | 3 |

Tabell 12: Deskriptiv statistikk om påstand 2



Figur 16: Frekvenstabell, påstand 1



Figur 17: Frekvenstabell, påstand 2

Tabell 11 og figur 16 viser samtlige respondenter har svart helt eller delvis enig med påstand 1. Det er derfor stor enighet om at bruk av KI er egnet til å vurdere aksept av klientforhold. Dette samsvarer med litteraturen. Respondent A mener fordelene er en mer uavhengig og objektiv vurdering, hvor D legger til at informasjonsinnhentingen kan automatiseres, samt avanserte analyser av eksempelvis uavhengigheten.

Tabell 12 og figur 17 viser at majoriteten av respondentene har svart helt eller delvis enig med påstand 2. Det er derfor enighet om at bruk av KI er egnet for det formål beskrevet i påstand 2, med unntak av en respondent. Majoritetens svar samsvarer med litteraturen. Respondent D forklarer at KI kan produsere et beste estimat på dette som deretter blir menneskelig vurdert. A mener andre faktorer også bør inngå i vurderingen

I planleggingsfasen ser vi et samsvar mellom intervju og litteratur. Dette gjelder bruk av KI til å innhente og analysere data fra en rekke kilder, sette dette i sammenheng med klientens systemer og utføre risikovurdering. Issa et al. (2016) hevder at den største forskjellen vil skje i den femte fasen, substanshandlinger. Flere av intervjukandidatene har nevnt testing av hele populasjoner. “Det blir mer substanshandlinger enn det vi har i dag, tror jeg. Fordi du kan ta fullstendighet av transaksjoner, i stedet for systemtester” sier informant F og understøttes av ICAEW (2016). Dette synet strider derimot med informant A og Brørs og Sellæg (2015) som mener kontrollmiljøet vil være et større fokusområde enn tradisjonelle tester av underliggende transaksjoner. Dette gikk vi innpå i delkapittel 4.1.3 og 4.2.1 og er nevneverdig ettersom det er beheftet med usikkerhet og delte meninger, som er bakgrunnen for at dette er forslag til videre forskning.

Tradisjonelt har revisor forholdt seg til utvalgsrisiko ved å benytte seg av stikkprøver. I delkapittel 4.2.1 påpekte vi at revisor fremover kan forventes å måtte forholde seg til dataanalyserisiko fremfor utvalgsrisiko. I tillegg til å sikre, avstemme og verifisere input, kan dataanalyse med eller uten KI avdekke utallige avvik. Dette vil også innebære avvik som opprinnelig ikke er avvik eller såkalte eksepsjonelle avvik.

Litteraturen presenterer og utdyper konseptet sanntidsrevisjon som et framtidsscenario. Både intervjukandidat A og B sier at de allerede gjør dette i dag, enten på enkelte områder eller at de kan gjøre det. C og E tror dette kommer etterhvert, mens D og F mener sanntidsrevisjon er nødt til å komme.

A trekker frem klienter med skyløsninger med oppdaterte databaser slik at revisjonen kan utføre i sanntid og hevder dette er noe vi vil se mer av fremover. A mener dagens årsoppgjør frem til juni utgjør en utfordring og påpeker relevansen av dette for kunden. E presiserer at regnskapet til kundene oftere blir tilgjengelig via internett, men føler ikke bransjen i stor grad har utnyttet dette fullt ut enda, men sier vi over tid vil komme dit. A forklarer man må være «on time» for å være en verdifull bidragsyter. F påpeker også at tallene blir mindre og mindre relevant etterhvert som dagene går.

B sier man kan programmere inn avviksalarmen med en robot som svarer på dette og samtidig ha en smart-samtale med kunden over e-post. B trekker også fram et scenario hvor også kunden har en robot som svarer. C supplerer med at KI kan overvåke og bidra til at dette skjer i sanntid og gjennomføre oppdaterte analyser fortløpende. Dette synspunktet deler også informant D som fremhever at sanntidsrevisjon vil være godt støttet av KI, fordi informasjon må vurderes kontinuerlig. Dette vil kreve avanserte analyser som vi som personer ikke klarer. “Hvorfor skal man sitte og holde på med historiske tall, det er situasjonen her og nå som betyr noe”, sier D. F mener også at sanntidsrevisjon vil understøttes av KI og sier at man må utvikle seg i denne retningen hvis revisor skal være relevant. D tror en slik revisjon ville vært mer utbredt dersom man lot markedet bestemme fremfor standarder og lovgivningen. På en annen side påpeker D at det likevel er fornuftig å ha en balansedato 31.12 for å få sammenlignbare tall. E deler ikke synspunktet til D som ikke ser at standardene hindrer sanntidsrevisjon og understreker at kunden kan få en raskere oppdatering og mer verdifull partner av en slik revisjon.

Litteraturgjennomgangen som omtaler sanntidsrevisjon har poengtert at man kan implementere overvåkningsaktiviteter og avviksalarmen i klientenes systemer for å lettere håndtere identifiserte avvik underveis. Ved å ta utgangspunkt i figur 9 på side 39, mener vi KI vil være en av flere viktige faktorer som kan bidra til en

effektiv og hensiktsmessig sanntidsrevisjon. Det er også dette synet majoriteten av intervjukandidatene reflekterer. Chan og Vasarhelyi (2011) antyder at sanntidsrevisjon kan føre til at revisjon går fra å være assurance til å bli insurance.

Vi har for lite empiri til å konkludere på hvordan revisjonsprosessen og fasene vil endres som følge av KI. Det er usikkerhet knyttet til hvordan teknologi generelt sett vil påvirke prosessen og fasene. Vi har derimot god støtte for at fremtidens revisjon vil bevege seg mot en sanntidsrevisjon og hvor KI er en av flere viktige faktorer for dette.

6.7 Forskningsspørsmål 8

«Kan revisor innhente revisjonsbevis ved å benytte droner, eller må standardene oppdateres for å være i samsvar med god revisjonsskikk?»

Vi har fått klare indikasjoner på at det er anledning til å benytte droneteknologi i revisjon. Gjennom intervju har person C, D, E og F vært inne på denne muligheten. Det er imidlertid krav om at revisor selv har kontroll over dronen. Dersom dette gir overbevisende revisjonsbevis, som F hevder, handler det om sannsynliggjøring. Både C og D legger til at droneteknologi og KI kan for eksempel beregne volum, mengde, antall og dermed også verdi.

Litteraturen presenterer ulike droneprosjekter som er i utvikling. Flere revisjonsselskap internasjonalt utvikler og benytter dette. Enten til overvåking eller i forbindelse med varetelling, hvor påstandene eksistens og verdsettelse bekreftes. Dette kan understøttes gjennom droner med KI. Hvorvidt en drone kan benyttes i varelagertelling, er beheftet med delte meninger ettersom standarden krever at revisor er fysisk tilstede. Det kan være omstendigheter hvor droner er bedre egnet til å kartlegge varelager, enn det som lar seg gjennomføre fysisk av revisor.

Begrunnet i en positiv sammenheng mellom intervju og litteratur, vil det være anledning til å benytte droner ved innhenting av revisjonsbevis. Droner til å bekrefte balanseposten varelager og i varelagertelling er beheftet med delte meninger. Likevel er droner særlig egnet i tilfeller hvor revisor er mindre egnet til å foreta en fysisk varetelling på grunn av omstendighetene. Det vil i følge de

informantene som drøftet problemstillingen ikke være nødvendig å oppdatere standardene. Dette har heller ikke vært omdiskutert i litteraturen. Formål, påstander og regulatoriske krav legger derimot begrensninger på hva dronen kan benyttes til og må vurderes i hvert enkelt tilfelle av revisor. Det er likevel nevneverdig at revisor selv må styre dronen for å sikre påliteligheten av revisjonsbeviset.

6.8 Forskningsspørsmål 9

«Hvordan benytte KI i prosessmining som en revisjonshandling?»

Prosessmining er et kjent begrep hos de fleste informantene. Flesteparten mener at prosessmining er teoretisk fornuftig å implementere i revisjon. Samtidig trekker flere frem kompleksiteten rundt dette i praksis. Selv om hovedstrømmen av transaksjoner går der de skal, eksisterer det mange avvik hvor transaksjoner ikke følger den tiltenkte utformingen. Dette begrenser den praktiske bruken av prosessmining ettersom revisor må ta hensyn til disse avvikene.

A påpeker at prosessmining kan samle informasjon og teste internkontrollen. En alternativ fremgangsmåte, i følge A, er å se på innstillingene i systemet for å preventivt hindre at feil skal inntreffe. A sier prosessmining kan være å anse som bevis på hva som skjedde. A er av oppfattelsen av at revisjonsbransjen har vært for opptatte av å gjøre den type analyse for virksomheten når denne kontrollen burde ligge ute hos virksomheten selv, fremfor hos revisor. A ser for seg at disse type kontroller og analyser vil være en integrert del i regnskapssystemene eller tilbys eksternt. A poengterer derfor at revisors investering og fokus på prosessmining kan bli overflødig på sikt.

C poengterer at selskapets flytkart ikke alltid stemmer overens med det flytkartet de får frem basert på deres verktøy. Videre forteller C at det er vanskelig å benytte prosessmining på en god måte ettersom prosessene ikke er standardiserte nok og at papirvarianten avviker fra virkeligheten. Det er mye støy, selv om hovedstrømmen av transaksjonene går slik de er tiltenkt på papiret, som gjør det vanskelig å benytte prosessmining hensiktsmessig, fastslår C. Mange avvik

medfører mye oppfølging og er bakgrunnen for at C mener prosessmining ikke er særlig enkelt å benytte i dag.

D poengterer at prosessmining er i vinden for å forstå hvordan prosesser gjennomføres basert på en logg. Dette vil automatiseres av en lærende algoritme, mer nyansert og finner hva som fremstår som en vanlig transaksjonsflyt. D forklarer den største utfordringen er at man i dag har for dårlige forventninger til og ikke nok forståelse av virksomheten. Dette medfører dårlige forventninger og resultater fra prosessmining som man ikke forutså. Selv om disse ikke er feil, blir de regnet som avvik. Dette vil være resultatet når man ikke har nok forståelse og ikke har hensyntatt kompleksiteten i virksomheten man reviderer.

E trekker frem at små kunder har lite formaliserte regler og er derfor bedre egnet på de større virksomhetene. Prosessmining kan være effektivt å benytte dersom man kan stole på internkontrollen og at den har gode systemer som de ansatte etterlever. Da er sannsynligheten for vesentlige feil lavere, selv om man ser på de samme avvikene. F tenker det er fornuftig at KI har en arbeidskapasitet som kan kontrollere samtlige transaksjoner og ikke bare enkeltvis. Samtidig forklarer F at dette kan gjøres gjennom automatiseringsteknologi hvor avviksrapporter implementeres. B forklarer man kan basere seg på resultatene fra prosessmining for deretter å la KI vurdere resultatene og tolke de.

I delkapittel 4.1.4 presenterte vi prosessmining som en handling for å opparbeide seg en forståelse av virksomhetens prosesser ved å ta utgangspunkt i hendelsesloggen. To av teknikkene som van Der Aalst (2012) presenterte, oppdagelse og overensstemmelseskontroll, kan være egnede revisjonshandlinger for å få en mening om internkontrollen. Dette understøttes av våre informanter. Hvorvidt prosessmining utgjør en test av kontroll eller substanshandling har vi ikke dekning i våre data til å svare på. Litteraturen fremhever derimot at prosessmining kan være en test med to formål. Videre utdyper to av våre informanter problemer knyttet til bruk av prosessmining i praksis hvor man vil få mange avvik.

Vi kan ikke konkludere på hvordan KI kan benyttes i prosessmining. KI kan derimot benyttes til å undersøke, følge opp og forklare avvikene som følger fra prosessmining og resultatet bak denne.

6.9 Forskningsspørsmål 10

«I hvilke eksisterende revisjonshandlinger kan KI benyttes, og kan nye revisjonshandlinger utvikles?»

Eksisterende revisjonshandlinger

Samtlige intervjukandidater er enige om at KI er godt egnet til å utføre analytiske handlinger under eksisterende revisjonshandlinger. Både B og C fremhever inspeksjon av dokumenter, for eksempel kontrakter, som allerede brukes i dag. KI benyttes til å finne avvik, men kan også innhente ekstern informasjon for å bekrefte forhold. A gir et åpent svar og sier KI kan benyttes innen de fleste revisjonshandlinger, herunder analytiske handlinger, substanshandlinger og test av kontroller, men med noe ulik modenhet innen de forskjellige. Samtidig sier A, «jeg tror kanskje substanshandlinger som vi tenker på det i dag, det tenker jeg kanskje ikke har så lang levetid fremover», og da heller i større grad bygge på internkontrollen og teste den. Dette synspunktet er ikke F enig i, som mener det vil bli mer substanshandlinger enn det vi har i dag, fordi man kan kontrollere fullstendigheten av alle transaksjoner fremfor systemtesting. Dette må ses i tråd med forskningsspørsmål 6 og 7, samt 16 (videre forskning), da det er beheftet delte meninger rundt denne diskusjonen.

D mener også KI kan benyttes i de fleste revisjonshandlingene, men i de som innebærer bruk av data, og mindre grad i de handlingene som er manuelle. Både B og C fremhever KI til kontraktsgjennomgang. KI kan finne avvik, spesielle betingelser og kombineres med dataanalyse på ulike typer prosesser. Videre kan KI vurdere dette opp mot bransjen forklarer de. B tenker KI kan vurdere informasjonen den har innhentet og tolket på en mer objektiv måte enn et menneske. KI handler på grunnlag av erfaring og kan få bedre informasjon tolket raskere utdyper B.

E trekker frem typiske avstemmingsoppgaver, SAFT-fila som kommer, kontrollsummering, inngående balanse og utgående balanse. E antyder derimot at KI vil slite med å gå mot grunnlagsmaterialet. Selv om den kan få til noe, er det fortsatt mye som gjøres manuelt i dag. Etterhvert vil mer av dette avstemmes desto mer som blir elektronisk. Informant C mener andre mindre smarte systemer vil gjøre avstemminger, men at KI kan gi mer støtte i estimater og vurderinger.

Brown-Liburd og Vasarhelyi (2015, s. 6) antyder at revisjonsmålsettingene ikke vil endre seg ettersom de drives av den finansielle rapporteringen. Revisjonshandlinger for å håndtere disse vil istedenfor påvirkes av hvordan teknologien virker inn på type (formål og art), omfang og tidspunkt for utførelse. Dette støttes av C som mener at teknologien vil erstatte eksisterende revisjonshandlinger. Som påpekt under forskningsspørsmål 6 og 7 kan tidspunkt for utførelse bli mer fleksibel, hvor man går mot en sanntidsrevisjon. Omfang som påpekt i forskningsspørsmål 1 vil kunne gjennomføres på hele populasjoner raskere og mer effektivt. KI kan også benyttes i de ulike type revisjonshandlingene, altså både i test av kontroller og substanshandlinger. I tabell 2, på side 47, oppsummerte vi hvor og hvilke av revisjonshandlingens art som KI kan benyttes i. Dagens tradisjonelle beskrivelse av deres art ble beskrevet i delkapittel 2.5. De moderne revisjonshandlingene kan styrke og forbedre effektiviteten og hensiktsmessigheten av dagens revisjonshandlinger.

I litteraturen fant vi støtte for bruk av KI i inspeksjon og i forbindelse med dokumentgjennomgang. Videre åpner litteraturen for å benytte KI i de fleste revisjonshandlingene. Flere kilder hevder KI i første omgang vil kunne effektivisere handlinger som er manuelle, strukturerte og repetitive. Ved bruk av datamining har IAIS (2017) påpekt to risikofaktorer. Påliteligheten i datamining og forståelsen av algoritmene lagt til grunn. Ved å benytte KI og herunder maskinlæring, nevralt nettverk og dyp læring i revisjonshandlinger kan disse handlingene avanseres og løse mer komplekse arbeidsoppgaver.

Vi har funnet støtte for at KI egner seg i analytiske handlinger og risikovurderingshandlinger gjennom de øvrige forskningsspørsmålene. KI er egnet til å inspisere dokumenter, men kan potensielt benyttes i de fleste eksisterende revisjonshandlinger. Fra forskningsspørsmål 11 vet vi at eksterne bekreftelser, gjennom innhenting av informasjon om et forhold fra en rekke kilder kan utføres ved hjelp av KI. Vi er komfortable med å fortelle at våre funn tilsier at KI kan benyttes i flere ulike revisjonshandlinger. Hvorvidt bruk av KI i revisjonshandlinger er begrenset for praktisk utøvelse er ikke behandlet i denne delen.

Nye revisjonshandlinger

Intervjukandidatene fant det vanskelig å redegjøre for nye typer revisjonshandlinger enn de vi allerede har i dag. På en annen side mener A at enkelte av dagens handlinger kan bli overflødige på sikt. A påpeker også at transaksjoner mellom parter vil kunne være direkte avstembart mot eksterne kilder.

C og D har ingen nye revisjonshandlinger å tilføre, men C hevder at eksisterende handlinger kan erstattes av teknologi. D tilfører at eksterne bekreftelser kan få en utvidende betydning. For eksempel ved verdsettelse kan revisor basere seg på mange kilder, hvor man da ikke er avhengig av at hver kilde er 100% pålitelig, men et sammenstilt gir god nok sikkerhet. Dette beskriver D som en «wikipedia-tankegang». Dette støttes av B som sier at KI kan behandle mye mer ekstern informasjon ved å sammenstille børsinformasjon, råvarepriser, indekser, industrier, globale kunder og dermed avdekke avvik som bør undersøkes. B ser for seg at man vil benytte utenforstående eksterne data for å bekrefte den interne dataen, for dermed å si at det som ligger i ERP-systemet er riktig.

F tenker det vil vokse frem noen nye obligatoriske revisjonshandlinger, uten konkrete eksempler. E ser for seg kontrollsummering av hovedbok eller andre ting gjort veldig fort, men ikke særlig revolusjonerende ting. E legger til at KI kan gjøre mye mer siden den er tidseffektiv, men også potensielt finne mer som krever oppfølging.

Litteraturen fremhever teknikker innen KI, oppsummert i tabell 1 side 45, som potensielt kan benyttes som selvstendige revisjonshandlinger. Dette er eksempelvis datamining, prosessmining og visualisering. I tillegg hevder EY (2017b) at det er behov for å inkludere dataanalyse som en spesifikk revisjonshandling, mens Rutgers (2017) fremhever prediktiv analyse som en egen revisjonshandling.

Våre funn støtter oppunder litteraturen. Det kan potensielt vokse frem nye revisjonshandlinger. Enten i form av en samlebetegnelse for ulike analyser eller bruk av datateknologi/dataverktøy, eller som separate teknikker og handlinger. Empirien gir grunnlag for at eksterne bekreftelser vil bli viktig i revisjon fremover på en ny og annerledes måte.

6.10 Forsknings spørsmål 11

«Kan KI benyttes i risikovurderings- og risikohåndterings handlinger, og bør disse skilles eller kombineres?»

Intervju spørsmålet bak dette forsknings spørsmålet ble ulikt oppfattet av informantene og vil derfor påvirke våre funn. Noen forstod ikke spørsmålet, mens andre hadde klare formening. I revisjonsteorien presenterte vi at risikovurdering og risikovurderings handlingene i utgangspunktet blir gjort før risikohåndterings handlingene. Når man implementerer KI i revisjon kan det velges hvorvidt KI skal benyttes i hver enkelt del separat, eller om KI skal kombineres som en test med opptil tre formål. I kombinasjon kan KI medføre at revisjonsfasene blir mer kunstige og overlappende som vi diskuterte i delkapittel 4.1.3 og 6.6. Som avdekket i forsknings spørsmål 1 er samtlige informanter enige i at KI er godt egnet i risikovurdering. I forsknings spørsmål 10 gjennomgikk vi bruk av KI i revisjonshandlinger.

Informant A er ikke kjent med noen vurderinger i forhold til de ulike fasene og svarer «jeg tror ikke vi har tenkt at mange av disse handlingene er integrert på en gang». B ser potensialet med å kombinere disse handlingene og forklarer at man som revisor ønsker å få oversikt over alt. C støtter B sitt synspunkt og sier en kombinasjon av dette er mulig. Videre utdyper C at de i størst mulig grad prøver å se dette i sammenheng gjennom å kombinere test av kontroller og test av detaljer samtidig. En test med både to og tre formål kan denne type teknologi også gjøre, sier C. Derimot understreker C at revisor må ha nok tid for å kunne verifisere seg enig i risikovurderingen, dersom den er utført av den kunstig intelligente maskinen, før resterende prosesseres. På tross av dette er mulig, mener C at dersom vi skal si oss enige i risikovurderingen, så må denne skje på et tidligere tidspunkt enn risikohåndterings handlingene, når KI står bak den.

Informant D tror ikke vi vil snakke om revisjonshandlinger i fremtiden på samme måte som i dag. D sier algoritmene vil kunne sammenstille data og regnskapet i sin helhet og gjøre alt på en gang. Uten å være inndelt i faser. Videre fokuserer D på datamining i form av å innhente eksterne bekræftelser fra en rekke kilder, som sammenstilt gir et rettviseende bilde. Også informant E sier datamaskinen kan kontrollere alt. Den kan kontrollere eksempelvis autorisasjoner på ordre og si

internkontrollen er god, ved å teste 100 %, samtidig som den kontrollerer faktura mot ordre, og dermed også kostnaden. På denne måten kontrolleres både substansen og internkontrollen. E forventer derimot et samlet, men et litt annerledes arbeidspapir. E understreker at test av kontroll ikke er substanshandling, selv om maskinen gjør alt samtidig. Videre stiller E spørsmål om hvorfor man skal teste internkontrollen når den kan gjøre substanshandlinger samtidig. I denne sammenheng påpeker E at substansen ofte ikke er godt nok linket til hovedboken og regnskapet, da det underliggende ofte kan ligge utenfor. Selv om man gjennomfører kontrollrutinetesting, så er fremdeles ikke substansen under gjort, poengterer E, og legger til at KI etterhvert kan foreslå arbeidspapirer.

F fremmer bruk av avanserte analyser og spesielt hvor datamining kan benyttes til å innhente tallmaterialet. Vedkommende mener det vil bli mer effektivt samtidig som det vil være en handling som får frem det vesentlige i tallmaterialet og avvik. B trekker frem det å vurdere en prosess i sin helhet og hva slags kontroller de har, og samlet sett i de kontrollene - vil gi ganske stor trygghet på at det ikke har vært noen vesentlige feil. B sier «jeg har en følelse i dag, at vi gjør mer og mer ulike typer kontrollhandlinger og ser på de isolert sett, istedenfor en større sammenheng».

Rutgers (2017) mener prediktive analyser som revisjonshandling kan være en test med både to og tre formål, og som kan kombineres eller utføres separat. Vi påpekte i delkapittel 4.1.3 at revisor ønsker å oppfylle sitt formål på en så lite ressurskrevende måte som mulig. I kapittel 2 påpekte vi Arens et al. (2016) sine tre anbefalinger i utførelse av revisjonshandlinger og at de ulike fasene i revisjonsprosessen kontinuerlig oppdateres på bakgrunn av ny informasjon. Dette støtter oppunder at KI i revisjonshandlinger (risikovurderings- og risikohåndteringshandlinger) antakeligvis bør kombineres i større grad, for å se ting i større sammenheng. Andre handlinger presentert i litteraturen kan også benyttes adskilt eller i kombinasjon. IAASB (2018b) skriver i punkt A2 at informasjon fremskaffet gjennom risikovurderingshandlinger kan i tillegg gi revisor revisjonsbevis på internkontroll, transaksjoner eller saldoposter med relaterte påstander, selv om de ikke tilsiktet var test av kontroll eller detaljtest.

Litteratur og empiri gir støtte for at KI er egnet i risikovurderingen, men at det også kan benyttes innenfor revisjonshandlinger. Dette betyr at revisjonshandlingene kan skilles. Det er også funnet støtte for at KI kan se, vurdere og revidere ting i sammenheng og gjøre alt samtidig. Dette støtter oppunder at revisjonshandlingene kan kombineres. C poengterte i denne sammenheng at det vil være fordel for revisor å verifisere denne risikovurderingen før KI går videre. Det ble også påpekt at fasene kan forvente seg å endres, støttet av intervjukandidat D som mener fasene vil bli kunstige fremover.

I forskningsspørsmål 1 fant vi positivitet ved KI til å identifisere risiko, og i risikovurderingshandlinger. I forskningsspørsmål 10 fant vi at KI kan benyttes i eksisterende eller nye fremvoksende risikohåndteringshandlinger. Vi kan si med støtte i litteratur og empiri at det er anledning til å både skille eller kombinere disse når KI tas i bruk, slik som dagens praksis. Våre funn tyder på at disse i større grad bør kombineres for å se ting i sammenheng og med tanke på effektivitetsperspektivet. Samtidig er det essensielt at revisor er enig i den risikovurderingen som er utført før man reviderer videre. Bruk av KI muliggjør at mer av revisors arbeid kan utføres samtidig. Dette gjelder eksempelvis at KI følger opp med risikohåndteringshandlinger etter utført risikovurdering. Vi vil understreke at revisor bør kontrollere og avstemme denne risikovurderingen, samt overprøve den inntil man kan stole nok på at systemet har tatt hensyn til alle faktorene.

6.11 Forskningsspørsmål 12

«Hvordan integrere multi-modale revisjonsbevis i revisjon og vil de være selvstendige eller kun underbyggende revisjonsbevis?»

Ingen av intervjukandidatene var familiære med begrepet multi-modale revisjonsbevis. Etter en presisering og klargjøring av begrepet belyste de nyttige synspunkter rundt hvordan dette egner seg som revisjonsbevis.

Informant E forteller at det ikke er utstrakt bruk av dette enda, men at flere revisjonsselskap fremsnakker det. E sier videre det ikke står at det er ulovlig i henhold til standardene. Samtidig tror vedkommende at Finanstilsynet ville slått

ned på dersom det lå vedlagt en lang film eller lydklipp i dokumentasjonen. Videre hevder E at RFID handler om risikovurdering og hvorvidt man stoler på systemet. E presiserer at dersom materialet er innhentet av revisor selv fra slike kilder, så kan det ikke sies å ha lavere pålitelighet. Dette er også F enig i.

A mener det er visse forutsetninger for at multi-modale revisjonsbevis skal kunne benyttes som selvstendige bevis, og peker på sikkerheten rundt bevisene og cybersecurity. Informant A understreker at innhentingene har stor betydning, da det er knyttet stor usikkerhet til ting som er tilkoblet internett. B påpeker at teknologien må testes og etterprøves i starten, og eksemplifiserer med en drone til å telle trær sammen med en algoritme som klassifiserer antall trær, som vil være et eget revisjonsbevis. B understreker at det er revisor selv som må ha kontroll gjennom å eksempelvis styre dronen selv.

C mener multi-modale revisjonsbevis kun er underbyggende bevis, med mindre det er blitt verifisert. C benytter eksempel med et GPS-signal fra et skip og hvordan man vet at det ikke er en tilfeldig GPS-sender som imiterer et skip. Informant D supplerer med at det er summen av revisjonsbevis som betyr noe, ikke det enkelte bevis. Derfor kan mange forskjellige data sammenstilt være tilstrekkelig og hensiktsmessig. Slike sammenstilte data tror D vi kommer til å se mer av fremover.

I delkapittel 2.6 belyste vi relevant teori vedrørende revisjonsbevis. Her påpekte vi at ISA 500 sin definisjon av revisjonsbevis består av to elementer. Multi-modale revisjonsbevis kan kategoriseres under «informasjon innhentet fra andre kilder». I delkapittel 4.1 og 4.1.5 presenterte Macaulay (2016a) at KI kan benytte informasjon fra utradisjonelle kilder og at fremvoksende teknologi kan generere mer innholdsrik og detaljerte revisjonsbevis. På en annen side trekker Brown-Liburd et al. (2015) frem en utfordring rundt det å utlede verdifulle bevis fra stordata og påpeker konsekvensen ved å bygge på irrelevant eller informasjon med innhold av støy. Som påpekt i delkapitlene 4.1.4 og 4.1.5 kan informasjon innhentet fra utradisjonelle kilder benyttes som selvstendige eller underbyggende bevis. Dette gjelder under forutsetning av at man har kontroll på den informasjonen man mottar faktisk er reell. Videre er det en vurdering om hvorvidt det er innhentet tilstrekkelig og hensiktsmessig revisjonsbevis. Litteraturen er

entydig positiv til denne utviklingen av revisjonsbevis, hvor Moffitt og Vasarhelyi (2013) påpeker at nye former for revisjonsbevis vil komplementere eller erstatte tradisjonelle revisjonsbevis. Litteraturen sier ingenting om hvorvidt disse bevisene kan benyttes selvstendig eller kun som underbyggende bevis.

Bruken av multi-modale revisjonsbevis er forventet å øke fremover. Det er imidlertid usikkert hvordan dette vil integreres og sikres i revisjonen. En alternativ fremgangsmåte er å benytte flere kilder til informasjon for å bekrefte et forhold. På denne måten får revisor et revisjonsbevis som er sammensatt fra en rekke kilder til informasjon. Empirien, samt litteratur, støtter at bevis innhentet fra utradisjonelle kilder utelukkende kan benyttes som underbyggende bevis.

Hvorvidt det kan benyttes som selvstendig eller kun underbyggende revisjonsbevis vil være gjenstand for en helhetlig vurdering av revisor gjennom anvendelse av profesjonelt skjønn og skepsis. Revisor skal oppnå betryggende (høy grad av) sikkerhet for at regnskapet er uten vesentlige feil gjennom å innhente tilstrekkelig og hensiktsmessig revisjonsbevis. Dette støtter D sin formidling av at det er summen av revisjonsbevis som betyr noe, og at handlingsrommet for hvordan dette gjennomføres for innhenting og bevisene som ligger til grunn kan variere.

6.12 Forskningsspørsmål 13

«Kan KI benyttes til å evaluere informasjon mot påstandene, og kan revisor bygge på disse vurderingene?»

Det er bred enighet blant intervjukandidatene på å evaluere informasjon mot påstander. Samtlige informanter er positive til å benytte KI til å innhente og evaluere bevis mot påstander, og dermed benytte KI til å avgjøre revisjonsbevisets relevans. B sier «jeg tror det er veldig lite, hvis noe i det hele tatt, blir produsert revisjonsbevis i dag med KI». B tenker at det ikke vil være problem for KI å innhente og evaluere bevis mot påstandene. Dette støtter også de øvrige intervjukandidatene. Informant F legger til at påstandene vil gå på mye av det samme som før.

A viser til at vi må kjenne koden, forstå logikken bak vurderingene og kunne dokumentere det. D sier at alle revisjonshandlinger er koblet direkte mot en

påstand, slik at dette prinsippet ikke vil være et problem. Informant C mener dette er mulig gjennom gjenkjenning og avstemming, men presiserer behovet for en form for verifisering og sikring av kilden, slik at informasjonen er autentisk. C utelukker ikke et system hvor programmet automatisk logger seg inn i systemet og bekrefter informasjonen, med hensyn til påstandene. E trekker frem at maskinen kan etterkontrollere mye bedre, men påpeker at det er tilfeller der det er vanskelig, eksempelvis når man ikke har gode nok data.

B mener eksterne data fra en ekstern kilde ikke vil være overbevisende nok som bevis, men dersom du får samme informasjon fra fem kilder, så har du større trygghet på at det er riktig. Det er i denne sammenheng B mener KI har stort potensiale, gjennom å innhente denne type informasjon og vurdere samme bevis mot flere ulike kilder for å verifisere at det faktisk er riktig, fremfor å benytte interne data som bevisgrunnlag. Dette ble beskrevet tidligere som en wikipedia-tankegang. D mener vi antakeligvis vil se annerledes på revisjonsbevis fremover, og trekker frem denne sannsynliggjøringen av ekstern informasjon. Informant F ser for seg at avsender- og mottakersystemet kan kommunisere sammen og dermed unngå fiktive hendelser. Når de kommuniseres ut fra avsender vil de samtidig bli akseptert av mottaker.

I kapittel 2 påpekte vi at revisjonsbevis innhentes for at revisor skal kunne konkludere med at regnskapet ikke inneholder vesentlige feil. Dette gjennomfører revisor ved å oppnå en bekreftelse på at samtlige revisjonsmålsettinger, i tråd med ISA 315 punkt A131, er uten vesentlige feil. I følge ISA 330 punkt A27 omhandler bevisets relevans formålet med revisjonshandlingen og påstanden som vurderes. Hvorvidt et revisjonsbevis dekker en eller flere påstander er skjønnsvurdert. Brown-Liburd et al. (2015) påpeker at informasjon med støy eller som er irrelevant medfører en dårligere skjønnsvurdering. Derfor kan maskiner foreta bedre objektive vurderinger. Brown-Liburd og Vasarhelyi (2015) antyder at revisjonsbevisets relevans sannsynligvis vil være gjenstand for skjønnsmessig vurdering slik som i dag. Litteraturen utelukker ikke KI til dette, men understreker at revisor bør kontrollere eller avstemme bevisets relevans og kvalitet.

Litteraturen har også påpekt en voldsom økning i tilgjengelig data og hvor revisors evne til å håndtere disse kan være problematisk. KI kan vurdere og

håndtere stordata og annen intern og ekstern informasjon på en bedre og mer effektiv måte. KI kan videre knytte dette opp mot revisjonsmålsettingene bedre enn revisor. Dette gjelder særlig i de tilfeller hvor smarte enheter kommuniserer sammen.

Empirien taler for, sammen med litteratur, at KI kan benyttes til å tolke informasjon og bevis mot påstandene som skal bekreftes. Enten gjennom en selvstendig vurdering eller i form av en automatisk oppkobling mellom revisjonshandling og revisjonsmålsettinger. Vi vil poengtere at det bør utøves med en viss varsomhet i oppstartingsfasen, fram til man har et sikkert og trygt system. Vi mener likevel at KI og revisor i samhandling vil kunne behandle og vurdere mye mer informasjon og bevis mot påstandene, enn det som gjøres i dag.

6.13 Forskningsspørsmål 14

«Hvilke konsekvenser ser revisjonsbransjen ved å ikke investere i KI?»

Våre informanter er samstemte i det faktum at revisjonsbransjen er nødt til å endre seg. Informant D svarer «da tror jeg revisjonsbransjen dør, rett og slett», på spørsmål om hvilke konsekvenser det vil ha å ikke investere i KI. De øvrige informantene er enige i at de som ikke er med på utviklingen vil få det tøft og dermed kunne bli utkonkurrert. B sier «jeg tenker jo egentlig at det blir de som setter seg mot dette her, de vil tape ganske fort. De som er med på dette her, de vil vinne». C hevder klientene er opptatt av innovasjon, effektivisering og økt innsikt. Derfor blir nye teknologiske investeringer essensielt for bransjens konkurransevne. Både C og F diskuterer muligheten for nye aktører som kan gi den samme typen uavhengige bekreftelse, men med en bedre løsning. Litteraturen er enig med våre informanter om at det er de selskapene som forstår og forbedrer sin bruk av kognitiv teknologi, er de som vil lykkes (Davenport, 2016).

Empirien samsvarer med litteraturen. Dataene viser at revisjonsbransjen og de enkelte selskapene må være med på utviklingen for ikke å bli utkonkurrert. Enten av satsende revisjonsfirma eller nye aktører. Derfor blir vedlikehold og videreutvikling av ansatte, dataverktøy og profesjonen essensielt.

6.14 Forskningsspørsmål 17

«Kan kunstig intelligente maskiner utøve profesjonelt skjønn og profesjonell skepsis, og kan revisor konkludere på bakgrunn av disse vurderingene?»

Informantene hadde ulike synspunkter hvorvidt KI kan utøve skjønn og skepsis. Informant B er overbevist om at KI kan utøve skjønn og benytte et intervall til skepsis slik at man kan oppnå en ønsket treffsikkerhet. For å lære maskinen skjønn mener B det er mulig å trene maskinen på tidligere utøvd skjønn. På den andre siden svarer informant C nei på dette spørsmålet. Dette begrunner C med det vil være vanskelig for maskinen å fange opp ledelsens agenda og insentiver. A og D er noe forsiktig og sier maskinen ikke klarer dette per i dag, men er en potensiell evne på sikt. E mener KI kan utøve mer skepsis enn oss fordi det ikke er følelser involvert og den vil heller ikke la seg lure, dersom den er riktig programmert. Dette blir også diskutert i litteraturen. F har et eksempel fra regnskapsbransjen, hvor maskinen kan justeres etter hvilken sannsynlighet for riktig svar man ønsker. Gjennom å velge 90 %, har man på en måte valgt hvilket skjønn den skal utvise.

For å konkludere på bakgrunn av disse skjønnsmessige vurderingene ønsker D å vite hva som ligger til grunn i algoritmene. Herunder hvilke treningsdata som foreligger, verifisering av datasett og algoritmene, kalibrering og understreker behov for innsikt i strukturen. Tilsvarende sikkerhet ønsker informant B, uavhengig av hva KI skal utføre. F har tiltro til maskinene og hevder det må være mulighet for å konkludere på bakgrunn av skjønn utført av KI.

Definisjonen og anvendelsen av profesjonelt skjønn og skepsis ble presentert i delkapittel 2.3. Disse egenskapene må anvendes gjennom hele revisjonen og dokumenteres. Raphael (2017) presenterer at revisor kan benytte mer av sin tid og ferdigheter på områder som krever skjønn og skepsis når slike teknologier implementeres. I denne sammenheng ville vi avdekke om KI også kan gjøre denne type oppgaver. I delkapittel 3.4 og 4.1.4 poengterte vi at nevrale nettverk og dyp læring kan understøtte krevende arbeidsoppgaver gjennom å lagre informasjon og kunnskap i mange ulike lag og nivåer. Samtidig avdekket vi i delkapittel 4.2.2 litteratur som hevder at KI er for umodent til å kunne erstatte menneskelig skjønn og skepsis. I stedet ser man på KI som en mulighet i

oppgaver hvor revisor kan benytte mer av sin tid på å utøve profesjonelt skjønn og skepsis. KI bør altså benyttes som et støtteverktøy i første omgang. Agrawal et al. (2016, s. 17) hevder derimot at KI kan utføre skjønn, skepsis og etterligne menneskelig adferd, i likhet med informant D, gjennom at maskinen trenes på tidligere utøvd skjønn. Samtidig understreker de at KI ikke kan erstatte all skjønn. Skjønn er sammensatt av flere faktorer og kan ikke enkelt programmeres i en algoritme. Samtidig utøves skjønn ulikt. En av mulighetene ICAEW (2017a) lister er at KI kan bidra til konsistente beslutninger gjennom å ikke påvirkes av følelser. Samtidig som KI ikke besitter myke følelser, mener Agrawal et al. (2016) at KI kan erstatte skjønn med prediksjoner i noen tilfeller. På en annen side er profesjonelt skjønn en viktig oppgave revisor utøver. Derfor bør dette ses i sammenheng med ISA 200-definisjon under forskningsspørsmål 19.

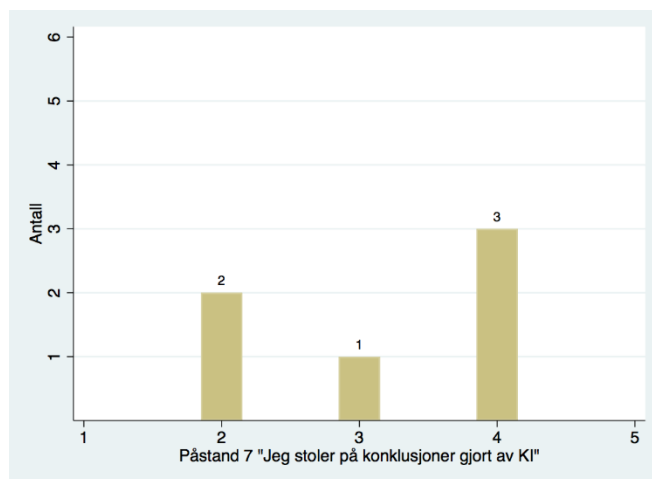
Videre ba vi respondentene ta stilling til følgende påstand:

Påstand 7: «Jeg stoler på konklusjoner gjort av KI»

| n | Gj.snitt | Std.avvik | Median | Min | Max | Var.bredde |
|---|----------|-----------|--------|-----|-----|------------|
| 6 | 3,16 | 0,966 | 3,5 | 2 | 4 | 2 |

Tabell 13: Deskriptiv statistikk om påstand 7

Tabell 13 og figur 18 viser at respondentene har delte synspunkter rundt å stole på konklusjoner gjort av KI. Vår gjennomsnittlige respondent stiller seg nøytral til påstand 7. Ingen av informantene avviser det helt eller stoler helt på konklusjoner gjort av KI. Respondent D mener algoritmene også her må verifiseres hvor A tilføyer med forhåpentligvis etter hvert.



Figur 18: Frekvenstabell, påstand 7

Både informantene og litteraturen har delte meninger rundt dette temaet. Det er derfor vanskelig å konkludere eksplisitt på forskningsspørsmålet. Selv om funnene ikke entydig antyder at KI kan utøve profesjonelt skjønn og skepsis per i

dag, kan dette antakeligvis endre seg på sikt. Vi har understøttende empiri til å anta at dette kan bli en realitet i fremtiden. På bakgrunn av våre data kan vi ikke si hvorvidt revisor kan stole på konklusjoner av KI da det er mange avgjørende underliggende faktorer.

6.15 Forskningsspørsmål 18

«Hvilken kompetanse bør nyutdannede besitte for å møte bransjens forventninger?»

Informantene har lang erfaring fra bransjen, samtidig som de har ledende roller i sitt respektive selskap. Deres uttalelser om fremtidig kompetanse blir derfor viktige funn sammen med funn fra litteraturen.

Det er bred enighet om interesse og engasjement rundt ny teknologi som en viktig faktor. Det blir viktigere med kompetanse innen KI og maskinlæring blant annet, hvor bransjen har et ønske om personer med sterkere kompetanse innen matematikk og teknologi. D sier også at de vil ha god nytte av domene-eksperter med funksjonell kompetanse, altså personer som evner å forklare et forretningsmessig problem til den som utvikler algoritmene. Slik sikres det at denne personen forstår hvordan algoritmen skal fungere. C forteller at det også blir nødvendig med teknisk kompetanse i kombinasjon med de gamle standardene for å benytte analyseprogrammer til å utføre revisjonshandlinger. A sier seg enig, og etterspør mer «hands-on» kompetanse på teknologi. Det etterspørres et bredere kompetansespekter enn tidligere, hvor bransjen ansetter flere personer med IT-kompetanse enn tidligere.

Teknologisk kompetanse er også etterspurt gjennom litteraturen, herunder IT og analytiske ferdigheter. Analytiske ferdigheter er viktig når det kommer til å kunne tolke resultater og vurderinger gjort av KI. Evnen til å forstå hvordan systemet fungerer, vurdere analysene hensiktsmessig og tolke output korrekt vil være verdsett kunnskap hos nyansatte. Spisskompetanse innen KI vil også være nødvendig, både vedrørende oppbygning og bruk. Rutgers (2017, s. 3) fremmer viktigheten av trening og forbedring i kritisk tenkning og skepsis hos revisor.

Vår empiri og litteratur etterspør mer teknisk og teknologisk kompetanse og analytiske ferdigheter, i tillegg til kompetanse innenfor regnskap og revisjon. Gjennom øvrige forskningsspørsmål har vi også funnet at myke verdier, skjønn og skepsis er egenskaper revisor må videreutvikle og forbedre.

6.16 Forskningsspørsmål 19

«Hvilke revisjonsstandarder utfordrer bruk av KI i revisjon?»

Det er delte synspunkter for hvorvidt revisjonsstandardene begrenser KI. Informant E mener det ikke er behov for å oppdatere standardene, med delvis støtte fra F. Resterende fire informanter mener standardene åpenbart må oppdateres, men utelukker ikke at KI kan benyttes enkelte steder uten at standardene oppdateres. Samtlige informanter uttrykker derimot et ønske om mer veiledning på området. De ønsker ytterligere at Finanstilsynet skal bli mer aktive og ikke forbli passive rundt dette.

Informant C tror ikke vi er innenfor standardene og hevder disse bør oppdateres for å hensynta KI i revisjon. På motsatt side står informant E, som ikke klarer å se noen regulatoriske hindringer for KI i revisjon. Samtidig påpeker informant E at det i henhold til standardene må være en ansvarlig revisor på toppen. Mellom disse synspunktene sier D at standardene er generiske og åpner for det meste. D uttrykker derimot at standardene på noen områder kan hindre bruk av KI. A og B støtter C sitt synspunktet om at standardene behøver en oppdatering før KI kan benyttes fullt ut, med støtte fra D. Kandidat F har delte meninger, men tror standardene kanskje ikke må oppdateres.

A peker innledningsvis på at lover og standarder ikke nødvendigvis er oppdaterte på den teknologiske utviklingen. «... ordbruk i revisjonsstandarder trenger jo å oppgraderes ganske betraktelig, eller utformes på en annen måte. Jeg opplever de som veldig tradisjonelle, og ikke tar høyde for ny teknologi». Informant C mener også standardene må utvikles og gjenspeile samfunnsutviklingen i dag, med all tilgjengelig data og teknologi. C tenker KI ikke vil ta av før standardene endres og påpeker at det vil ta tid. Dette poengterer informant D også som anfører at endring av standarder er en tidkrevende prosess, hvor man isteden kan få interesse-

organisasjoner som blir enige om gjeldende praksis. Enten i form av en bransje- eller emnestandard. D mener dette kan være en løsning på «problemet» ved å få en form for veiledning på området, ettersom det vil ta for lang tid å oppdatere standardene.

«Jeg er ikke helt med på at den hindringen er så stor som de hevder», sier E når det kommer til at regulatoriske krav hindrer implementering av KI. Videre hevder E at «det er jo ikke det at standardene sier at dette får du ikke lov å bruke, men det hadde nok vært bra med litt mer veiledning på området, det er jeg enig». E uttrykker at det kunne vært noen flere punkter for veiledning eller eget veiledningsdokument som sier noe om skikk og bruk, dersom det er slik at bransjen er nervøse for å ta i bruk KI. E påpeker, som vi presenterte i litteraturgjennomgangen, at IAASB og DAWG har utgitt en white paper. Her fant de ikke bevis for at standarder hindrer bruk av digital revisjon, herunder KI, men presiserer et behov for klargjøring og veiledning. Dette betyr at standardene åpner for dette i dag. E presiserer ytterligere at det vil være en vanskelig prosess å tilføre veiledningspunkt utover å si at standardene åpner for dette fullt ut, men legger til at standardene ikke fremmer dette. Eksempelvis gjelder dette ved å benytte begrepet «revisor skal...».

Informant D sier «ISA-ene er veldig generisk og åpner egentlig for det meste, også er det noen ting som er direkte motstridende, men det tror jeg mer er noe som er skrevet i en tidsalder der du ikke forutså hva som kom». C trekker frem det utfordrende er der standardene er klare på hva revisor spesifikk skal gjøre. D eksemplifiserer med at det står i standarden at ansvarlig revisor er ansvarlig og skal vite om hva som foregår i en revisjon. Når KI gjør vurderinger vil det være et brudd på standarden slik den tolkes nå. Selv om D mener ISA-ene åpner for det meste, mener vedkommende at standardene åpenbart må oppdateres. I hvert fall må de fortolkes på en annen måte, gjennom eksempelvis en guide. Informant F mener det er et stort rom for skjønnsutøvelse fra revisor og dermed bruk av KI som et hjelpemiddel, på samme måte som Excel, for å komme raskere til målet. Dette støttes av E som benytter et eksempel på Excel sin autosummeringsknapp, opp mot KI. Informant F legger til at standardene er laget for praktisk utøvelse av yrket. Informant A mener det er utfordrende når KI benyttes i skjønnsmessige

vurderinger. For eksempel i risikovurdering, konklusjonsfasen og vurderinger som er vanskelig å dokumentere.

I denne sammenheng ba vi respondentene vurdere følgende påstand:

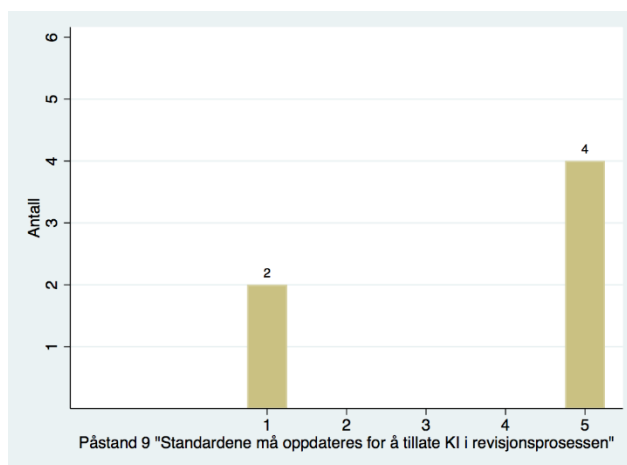
Påstand 9: «Standardene må oppdateres for å tillate KI i revisjonsprosessen»

| n | Gj.snitt | Std.avvik | Median | Min | Max | Var.bredde |
|---|----------|-----------|--------|-----|-----|------------|
| 6 | 3,66 | 2,065 | 5 | 1 | 5 | 4 |

Tabell 14: Deskriptiv statistikk om påstand 9

Tabell 14 og figur 19

reflekterer tilsvarende funn fra intervjuene. Fire er helt enige med at standardene må oppdateres, mens to er helt uenige. Respondent A mener standardene må bedre reflektere virkeligheten hvor respondent D tilføyer at der vi er i dag er ikke standardene et problem. Likevel sier D at jo mer autonome revisjonsverktøy blir basert på KI, desto mindre vil standardene passe.



Figur 19: Frekvenstabell, påstand 9

Flere av kandidatene mener standardene ikke tar høyde for ny teknologi. Dette støttes av IAASB (2018a, s. 9) som sier at standardene ikke fremmer bruk av ny teknologi, men heller ikke står i veien for det. Vi avdekket mangelfull litteratur på hvilke spesifikke standarder som begrenser bruk av KI i revisjon. Dette reflekteres også delvis gjennom mangelfulle tilbakemeldinger fra informantene på spesifikke standarder som begrenser bruk av KI i revisjon. Flere av informantene hadde lite kjennskap til spesifikke standarder, mens noen hadde noe, men ikke full kunnskap. Om mangelfull informasjon om spesifikke standarder støtter oppunder E sitt synspunkt eller skyldes andre faktorer er vanskelig for oss å avgjøre. Informant E sier bransjen er veldig defensiv og at det sikkert er mange gode grunner for å ikke utfordre standardene eller tillate seg å være litt mer frempå. Videre sier E at mangelfulle uttalelser fra Finanstilsynet er uheldig for bransjen.

På konkret spørsmål om ISA 500, stordata og andre eksterne data som bevisgrunnlag svarer C «...ingen fra fagmiljøet som tør å si at det er godt nok». Informant C legger til at det er mye opp til selve partneren på oppdraget hvorvidt ønske om en effektiv og innovativ revisjon står seg mot den risikoen partneren sitter igjen med på om dette står seg. C legger til at bruk av KI i planlegging og risikovurdering er innenfor, men at KI i revisjonshandlinger er verre, med unntak av visse rekalkuleringer. E trekker frem ISA 700 som problematisk dersom KI skal gjøre hele revisjonen og utstede beretningen, hvor ordet revisor problematiserer med hensyn til ansvarlig revisor. Alt under dette, forespørsler etc., da har revisor sendt ut dette ved å benytte KI.

Som påpekt i delkapittel 2.2 var hensikten med god revisjonsskikk og ISA-ene at det skulle sikre et dynamisk og oppdatert innhold i tråd med samfunnsutviklingen. God revisjonsskikk er ment for å være en rettslig standard, hvor innholdet er under kontinuerlig utvikling og forbedring. Dette har derimot ikke vært tilfelle, understøttet av informantene som er usikre rundt bruk av ny teknologi i revisjon. Av den grunn vil et syn være å anse det som at begrepet god revisjonsskikk går bredere enn det standardene sier, ettersom en profesjon må holde tritt med samfunnsutviklingen for å sikre relevans og eksistens. IAASB (2018a) sier også at de ikke finner støtte for at teknologien er begrenset i standardene. Kinserdal (2017) skriver som påpekt i delkapittel 4.2.1 at det er vanskelig å godkjenne ny metodikk før den er oppfunnet og testet. Momentene trukket fram i delkapittel 4.2.3 er derimot høyst relevante.

Vi avdekket at ISA 200 punkt 13d kan oppfattes som problematisk. På bakgrunn av at ISA definerer revisor som en person har vi spurt samtlige informanter om hvordan dette stiller seg når revisor «trykker på KI-knappen». Er det da revisor eller KI som har utført oppgaven og produsert resultatet? Fire av informantene svarer med at dette er revisor, mens to av informantene er mer usikre på definisjonsspørsmålet.

«Jeg tenker at en revisor kan være både maskinell og manuell tenker jeg, at det er en sånn samlebetegnelse på at det peker mer tilbake på selskapet, som har utviklet maskinen og som eier den maskinen», sier informant A. Dette er bakgrunnen for at man må dokumentere handlinger KI utfører og logikken i maskinen. Dette

synet støtter B, som sier for at revisor skal har gjort dette, så må algoritmene være godkjent, låst, testet og dokumentert med hva de gjør. B supplerer med at når algoritmene har blitt godkjent fra riktige instanser, så har revisor gjort det han ble fortalt, slik som dagens metodikk med å teste et utvalg. B mener dette vil være tunge investeringer og vil være hvert enkelt selskaps intellecutal property. C tror derimot vi ikke er innenfor slik standarden er i dag og mener standardene må utvikles for å hensynta dette. «Jeg synes det er kjempevanskelig å argumentere for at bruk av dette er innenfor standardene, og det tror jeg ikke Finanstilsynet heller mener», sier informant C. På grunn av standardene tenker C at man kanskje vil ta i bruk dette nye, men at man også vil gjøre det gamle, for man stoler ikke på at det nye er godt nok og vil holde i en rettsak. Informant E sier:

«Men, jeg ser ikke noen regulatoriske hindringer for kunstig intelligens i revisjon. Og jeg ser ikke at standarden definerer til at du ikke kan gjøre, men det betyr at standarden sier at du skal ha en ansvarlig revisor på teamet og når man benytter ordet revisor i standarden betyr det ansvarlig revisor og de menneskene som jobber på oppdraget. At du bytter ut et par av de menneskene med en datamaskin så har du fortsatt han på toppen som står ansvarlig for det som den datamaskinen gjør, og det funker».

E sier videre at ansvarlig revisor ser på alt medarbeiderne har gjort og går god for dette. Dermed kan medarbeiderne lett skiftes ut med en intelligent datamaskin. Dette fordrer at du forstår programmeringen eller at den får en slags sertifisering. Informant E opplever at mange er skeptiske til å ta dette i bruk, men påpeker at vi ikke må overdramatisere.

Vår samlede oppfattelse og forståelse av hvordan standardene stiller seg til bruk av KI er presentert i de neste avsnittene. Innholdet er basert på gjennomgått litteratur og utførte intervjuer. Dette kan derimot som utgangspunkt ikke brukes ukritisk, hvor man må vurdere eksplisitte krav i standardene mot den delen KI har til hensikt å utføre. Likevel kan innholdet gi en samlet forståelse for hvordan vi oppfatter at KI kan benyttes innunder de nevnte standardene. Vi ser at de informantene som er av mer prinsipiell interessert kanskje fortolker dette på en annen måte enn praktikerne som vil ha klarere retningslinjer for hva de skal gjøre.

Våre data indikerer at det er størst handlingsrom og muligheter for KI innen ISA 240, 315 og selve planleggingsfasen til å begynne med. Dette er begrunnet i at KI kan gjennomføre en innledende analyse, avdekke røde flagg, avvik og misligheter bedre og utføre en risikovurdering men hensyn til mye mer informasjon. Samtidig er det viktig at revisor etterprøver, kontrollerer og supplerer risikovurderingen med sitt skjønn og egne preferanser ettersom vedkommende er selv ansvarlig for den. KI kan også sette forslag til vesentlighetsgrenser basert på den informasjonen, i tråd med ISA 320.

Videre presenterer våre data et viss handlingsrom, men samtidig flere forhold innen nevnte standarder som må vurderes særskilt. Følgende standarder er ISA 230, 330, 520, 530 og definisjonen i ISA 200. Flere av nevnte standarder kan integreres med hensyn til KI, men ikke blindt. KI i gjennomføringsfasen blir neste steg, etter forrige avsnitt. KI kan foreslå revisjonshandlinger på bakgrunn av underliggende informasjon og risikovurdering, samt produsere arbeidspapirer opp mot kravene i ISA 230. I forhold til ISA 500 kan KI benyttes innen flere ulike revisjonshandlinger for å fremskaffe bevis hvor vi i forskningsspørsmål 10 og 13 så at samtlige informanter var enige om KI til å innhente bevis. Selve bevisinnsamlingen er ofte tidkrevende for revisor, men samtidig der det er mest å hente ved hjelp av KI. Det som gjør ISA 500 utfordrende er dataene lagt til grunn (datakvaliteten) og bevisets pålitelighet, med andre nevnte momenter i delkapittel 4.2.3. Revisor skal oppnå høy grad av sikkerhet som kan oppnås ved å teste alt, altså hele populasjoner. Så lenge revisor innhenter tilstrekkelig og hensiktsmessig revisjonsbevis på denne måten, kan ISA 530 og stikkprøver anvendes der revisor mener dette vil være mer hensiktsmessig enn hele populasjonstester av KI.

Avslutningsfasen kan også KI benyttes i, men hvor ISA 700 som nevner ansvarlig revisor til signering av revisjonsberetningen vil i så fall ligge mye lengre fram i tid. Definisjonen i ISA 200 gjør at vi mener KI kan benyttes ettersom man kan skifte uten noen av medarbeidere med KI, samtidig som andre medarbeidere og ansvarlig revisor arbeider sammen på oppdraget. Dette fordrer andre sentrale momenter trukket frem i denne studien. Dataene impliserer ikke forbud, men flere usikre forhold og vurderinger som kreves før KI utnyttes fullt ut innen flere standarder.

Våre funn viser at selv om standardene ikke reflekterer dagens samfunnsutvikling, med hensyn til data og teknologi, mener vi at vi har tilstrekkelig grunnlag for å si at flere standarder åpner for bruk av KI. Dette på bakgrunn av at standardene inneholder flere prinsipper og valg, samtidig som de er basert på praktisk utøvelse av yrket. Det er på den andre siden vanskelig å si hvilke spesifikke standarder som begrenser bruk av KI ettersom vår empiri og litteratur ikke diskuterer spesifikke standarder konkret. Det er også beheftet delte meninger rundt dette. Samtidig mener de fire informantene som mener standardene må oppdateres at KI kan benyttes innen enkelte deler av dagens standarder. Våre funn viser også at de som er i praksis vil ha mer konkret og spesifikk veiledning gjennom en viss trygghet rundt bruken av dette, mens standardsetterne er åpne for det meste og mener standardene er «gode» nok. Standardenes begrensning på KI avhenger derfor av øynene som ser og hva man ønsker. Til tross for dette vil det for mer prinsipielle holde med tilstrekkelig og hensiktsmessig bevis, uavhengig innhentingemetoden. Standardsetterne vil se at det fungerer effektivt og hensiktsmessig der de samtidig slipper behovet for å kontinuerlig måtte oppdatere standardene for å tillate og reflektere nye verktøy, teknikker, metoder og teknologier. Standardenes begrensning på KI er derfor styrende for egne tolkninger og insentiver partene har. Vi har ikke grunnlag for å si hvordan tilsynsmyndighetene ser på dette. Øvrige revisjonsstandarder har vi heller ikke grunnlag for å uttale oss om.

6.17 Forskningsspørsmål 21

«Hvilke krav bør stilles til KI sin etterlevelse av standarder, dokumentasjon av den interne strukturen og etterprøvbarehet?»

Samtlige informanter understøtter dette forskningsspørsmålet som en praktisk utfordring. Flere påpeker mangelfull dokumentasjon, forståelse eller etterprøvbarehet som kritiske faktorer. Samtidig mener de fleste at dette i noen grad vil løses gjennom en verifisering fra de riktige instansene.

A mener en kombinasjon av programmerere, utviklere og revisorer må skrive koden bak KI sammen for å sikre at revisor forstår hva som ligger til grunn i systemet. Videre uttrykker A at de som har utviklet systemet også må vite formål og hensikt, hvor den ene kompetansen ikke kan utelukke den andre. C mener noen

må verifisere algoritmene og at de er gode nok. Eksempelvis at noen eksperter verifiserer og sertifiserer den type verktøy, ettersom revisor ikke forstår alt eller har den type utdannelse. Dersom verktøyet eller maskinen er sertifisert, kan en annen revisor gå inn å vurdere om dette er godt nok mener C, men påpeker vi ikke er der i dag. Verifisering av KI-verktøyet støttes av informant E og F. Informant E utdyper at dette er en stor utfordring, men er usikker på om verifisering er godt nok. Informant B mener det alltid vil være noen som kontrollerer og verifiserer maskinen og dermed ikke lar maskinen ta en beslutning helt alene.

Informant A sier man må dokumentere den logikken som er programmert i systemet, vedrørende innhenting og behandling av data, og hvilke kriterier som er lagt til grunn for vurderingene. B supplerer med at koder må gjennomgås og testes, hvor man viser til testresultater og dokumentasjon på at analysene og verktøyene fungerer. Videre mener B at systemene må kunne åpnes, men ikke nødvendigvis av en revisor. D sier det på et eller annet tidspunkt kommer til å komme et behov for å sikre algoritmene og vil sette en begrensning på utviklingen, men samtidig medføre en større trygghet. Dersom det ikke er etterprøvbart, dokumenterbart eller mulig å åpne, vil det ikke holde i retten. Informant D mener en mulighet på denne problematikken vil være å benytte en kontrollerende algoritme (motalgoritme), dokumentere utviklingsprosessen i logikken bak teknologien som er brukt eller en kombinasjon av alt. C legger til en beskrivelse av hele prosessen hvor man forsøker å beskrive hva som skjer internt i verktøyet. Informant C supplerer med at man må avstemme og verifisere fullstendigheten og nøyaktigheten av data når den type verktøy gjør dette på egenhånd, og dokumentere hvordan man har sikret fullstendig og nøyaktig informasjon som er vurdert av KI. En eller annen form for digital signatur må være på plass for å hindre muligheten for å mate maskinen med feil data eller manipulere input hvor C ytterligere påpeker behovet for å ha på plass sikkerhet rundt dette. Dokumentasjon støttes av informantene E og F.

Informant A viser til et scenario hvor maskinen begynner å tenke selv og gir forslag til ting, hvor den underbygger med hvilke data og kilder informasjonen kommer fra og sammenstiller etter beste evne. Det vanskelige for oss blir å forklare og dokumentere hva maskinen har gjort, ettersom den har blitt så mye smartere enn oss forteller A.

E mener en vanlig revisor vil slite med å etterprøve alle de kildene KI har innhentet informasjon fra. E hevder videre en løsning kan være at de som skal etterprøve vil komme med sin egen «robot», og etterprøve, å se om de får tilnærmet likt svar. Arbeidspapirene forklarer hva som er gjort og tenkt, hvor en annen ser om dette er fornuftig. Tilsvarende kan en annen sertifisert «robot» benyttes for å teste systemet og få tilnærmet likt svar, som en mulighet. E foreslår også en IT-ekspert til å se på algoritmene. F forteller en mulighet kan være å snakke med KI slik som med et menneske å se hva den svarer. Både informant C og F bekrefter problematikken rundt etterprøvbarehet. F lurer på om tankesettet ikke står seg i lengden, fordi det er så mye skjønnsutøvelse slik det er i dag. Mens C tror dette kan bli lettere dersom verktøyet er verifisert.

Informant D påpeker også behovet for å forstå hva som skjer inni den svarte boksen og sier revisor må kunne ansvarliggjøres, hvor ansvarlig revisor vet dette er gjort på en hensiktsmessig måte. “Akkurat dette er en stor jobb og tar litt tid å finne ut av hva som er godt nok”. E mener hvis en annen får det samme resultatet, får man en viss sikkerhet for at den gjør det riktig, for vi klarer ikke å plukke opp om det er en liten feil i algoritmen. Man kan forstå til en viss grad forklarer E, og understreker at man må være trygge på det man lanserer og tar i bruk i revisjonen. Dette fordrer at man stoler på det man bruker.

Vi ba de utvalgte kandidatene ta stilling til påstanden:

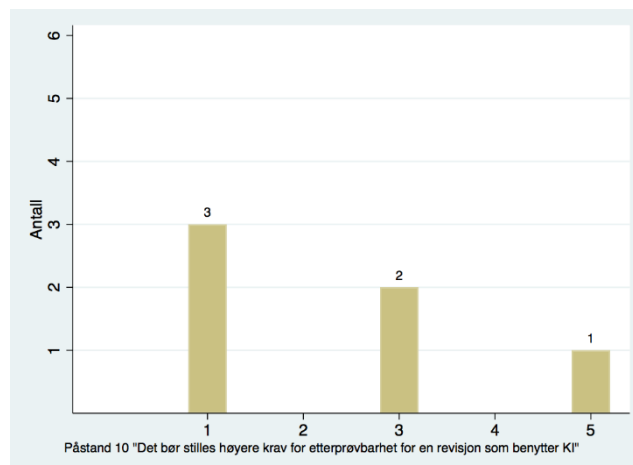
Påstand 10: «Det bør stilles høyere krav for etterprøvbarehet for en revisjon som benytter KI»

| n | Gj.snitt | Std.avvik | Median | Min | Max | Var.bredde |
|---|----------|-----------|--------|-----|-----|------------|
| 6 | 2,33 | 1,632 | 2 | 1 | 5 | 4 |

Tabell 15: Deskriptiv statistikk om påstand 10

Resultatene fra tabell 15 og figur 21 viser stor variasjon i påstand 10. Det er uenighet vedrørende krav til etterprøvbarehet ved en revisjon med KI. En gjennomsnittlig respondent hos oss er uenig til påstanden om at kravet skal være strengere. Respondent D tilføyer at kravet til etterprøvbarehet vil forbli den samme, men at måten det gjøres på vil endre seg.

Gjennom litteraturgjennomgangen i delkapitlene 3.2 og 4.2.5 avdekket vi problemer med å forstå dynamikken og innholdet i komplekse datamaskiner, omtalt som den svarte boksen. Manglende transparens og åpenhet av begrunnelsen til sine prediksjoner ble trukket frem.



Figur 20: Frekvenstabell, påstand 10

Dette understøtter uttalelsen til Skatteetaten i delkapittel 3.3. Samtidig har vi påpekt at dette ikke kun gjelder KI. Mennesker og utøvelse av skjønn og skepsis, samt forbedring av disse, er forbundet med en svart boks, hvor en detaljert forklaring av handlinger kan være problematisk. I delkapittel 2.6 presenterte vi relevant teori rundt viktige krav til dokumentasjon for revisorer og hvor man hele tiden skal forstå hva som foregår i en revisjon. Mangelfull evne til å forstå og tolke den interne strukturen, samt faktorene avgjørende for maskinens resultat utgjør derfor en utfordring for å hensiktsmessig ta dette i bruk i revisjon i dag. Videre i delkapittel 4.2.5 presenterte vi ulike metoder som kan løse denne problematikken selv om disse teknikkene har sine begrensninger. Etterlevelse av relevante krav må i første omgang reflekteres gjennom koding og oppbygging av den intelligente maskinen. Det er derimot visse utfordringer, eksempelvis at maskinen kan reflektere menneskelige bias gjennom programmering eller i treningsdata. Vi presenterte også hva som kunne forventes å måtte dokumenteres i en slik prosess.

Flere av momentene trukket frem i delkapittel 4.2.5 reflekteres også i våre data fra informantene. Det ser tilsynelatende ut til at informantenes syn på dokumentasjon av den intelligente maskinen er sammenfallende med påkrevde krav til en vanlig revisor vedrørende godkjenning, dokumentasjon og etterprøvbarehet. Derimot påpekes det at KI-verktøyet, enten det gjelder et program, verktøy eller maskin bør verifiseres, og at dette trolig skjer internasjonalt. Kravet til etterprøvbarehet vil forbli det samme, men informantene påpeker visse utfordringer når KI har utført en oppgave. Her vil en annen revisor se om det virker fornuftig eller benytte sitt avanserte program til å etterprøve.

Vi vil oppsummere med at KI må programmeres hensiktsmessig til å reflektere formål, hensikt og etterlevelse av relevante krav. Videre stilles det strenge, men fornuftige krav til dokumentasjon av programmet, maskinen eller verktøyet selskapet benytter. Vi har ikke funnet støtte for at mangelfull dokumentasjon av KI vil stå seg. Empirien og litteraturen viser at krav til etterprøvbarehet vil forbli uendret. Derfor er dokumentasjon av det revisor bygger sine antakelser og videre handlinger på, vesentlig å dokumentere ved en etterprøvbarehet.

6.18 Forskningsspørsmål 22

«Hva må en revisjonslogg inneholde for å tilfredsstille ISA 230?»

Informant E sier det skal eksistere en rød tråd fra start til slutt. D ser en utfordring i en lærende algoritme som vil forandre seg hele tiden. Dette er noe man er nødt til å ha kontroll på som også ble poengtert av informant A. Videre forteller E at du må forklare hva du har tenkt, resultatet og hvordan dette er fulgt opp. Dersom en annen revisor skal forstå hva som har skjedd når KI blir benyttet i revisjon må det komme en standardtekst for revisjonsloggen, for eksempel med versjonsnummer og annen dokumentasjon, sier E. Det er imidlertid knyttet usikkerhet for hva som blir praksisen rundt dette. B mener dette kommer til å bli drevet internasjonalt, at disse systemene blir testet og godkjent der, og at regulatorer videre vil standardisere dette. C svarer at konkrete input-data må beskrives og hvordan man har sikret at disse er fullstendige og nøyaktige. A supplerer med viktigheten av å ha kontroll på kontrollmiljøet rundt datakilden, for å sikre datakvaliteten. Deretter må man ha en beskrivelse av KI-verktøyet og dokumentere hva algoritmene gjør. Resultater må vurderes, utfordres og beskrives. Det samme svarer F med at man må ha dokumentasjon på hvordan maskinen har tenkt. A legger vekt på at man må sikre dokumentasjon av logikken i systemet, altså hvordan data blir innsamlet og behandlet, samt hvilke vurderinger som ligger bak.

Litteraturen mener man må inkludere dokumentasjon av all feilinformasjon, som ikke er klart uvesentlig. Det må også forklares og dokumenteres hvorfor avvik er avvik. I tillegg bør benyttet metode inkl. scripts til å uthente data være tilgjengelig i revisjonsloggen. Det oppsummeres i litteraturen at det bør etableres en logg av et

revisjonsspor som følger datainnhenting, datavask, transformering, analyse, avvik til visualisering, evaluering og tolking av informasjonen og mot handling.

Det er noe usikkerhet og uenighet blant våre informanter. Flere kilder mener at dokumentasjonen må inneholde hvordan data har blitt innhentet og behandlet. Avvik og vurderinger må ytterligere dokumenteres. Vi kan ikke gi en utfyllende liste over konkret dokumentasjon revisjonsloggen må inneholde, fordi det er bred uenighet. Antakeligvis vil dette sannsynligvis bli standardisert. Likevel har vi kartlagt potensielle viktige dokumenteringskrav i revisjonen, i delkapittel 4.2.5. Utover dette har vi ikke data til å ytterligere se dette opp mot ISA 230.

7. Avslutning og konklusjon

Formålet med denne masteravhandlingen har vært å identifisere muligheter og utfordringer rundt bruk av KI i revisjon. Vi har hatt som mål å skape en bedre forståelse av KI og hvilke muligheter og utfordringer som ligger ved å ta i bruk KI. Formålet og problemstillingen til studien reflekteres gjennom 22 forskningsspørsmål. Først ble en robust litteraturgjennomgang presentert. Deretter innhentet vi empiri gjennom intervju for å analysere og drøfte våre funn opp mot presentert litteratur. Intervjuene og spørreskjemaene tilfører oppgaven synspunkter rundt KI i revisjon fra revisjonsbransjen selv. Til slutt oppsummerte vi betydningen av dataene våre under hvert enkelt forskningsspørsmål. Resultatene tyder på at våre funn er i samsvar med litteraturen på en rekke områder, mens noe sprikende på enkelte forskningsspørsmål. Nedenfor vil vi oppsummere og konkludere på problemstillingen for deretter å presentere praktiske implikasjoner. Oppgavens begrensninger og forslag til videre forskning blir deretter belyst.

7.1 Oppsummering og konklusjon

Gjennom studien har vi kartlagt en rekke muligheter og utfordringer. I det kommende oppsummeres og poengteres våre viktigste funn fra studien. Litteratur, intervju og spørreskjema har dannet grunnlaget for våre funn og konklusjoner.

Våre viktigste funn innenfor muligheter av KI i revisjon er følgende:

KI er i første omgang spesielt effektivt i de handlingene som bruker data. KI kan i denne sammenheng benyttes til å innhente, analysere og vurdere data. En mulighet er at KI evner å innhente informasjon fra mange kilder samtidig, hvilket vi mennesker har begrenset evne til. KI har stor prosesseringskraft og kan raskere gjennomgå strukturerte og ustrukturerte data og hele populasjoner. Fordelen ved dette er oppnåelse av en lavere risiko, fordi man sitter på mer informasjon. KI kan videre bistå i utarbeidelse av revisjonsplan og utføre risikovurderings- og risikohåndteringshandlinger. Dette medfører at revisjonen effektiviseres og hever samtidig revisjonskvaliteten. Betydningen av dette er at revisor oppnår en dypere innsikt om klienten og styrker tilliten i markedet. KI gir også revisor mulighet til å bruke mer av sin tid på oppgaver som krever vurderinger og bruk av profesjonelt

skjønn og skepsis. Våre data indikerer at KI ikke kan gjøre det sistnevnte per i dag, men bevege oss mer mot etterhvert.

Våre funn representerer de viktigste identifiserte mulighetene og fordelene med KI i revisjon. Informantene uttrykte at ingen vet hvor teknologien vil stoppe og teknologien som delvis umoden som gjør det vanskelig å identifisere konkrete muligheter utover dette.

Våre viktigste funn innenfor utfordringer av KI i revisjon er følgende:

En av de største utfordringene vi har identifisert er problematikken med den svarte boksen og mangelfull evne til å forstå, dokumentere og etterprøve den interne strukturen i komplekse maskiner. Dette i kombinasjon med revisors bekymring for eventuelle brudd på standarder, lovkrav og tilsynsmyndighetenes betraktninger utgjør en særlig utfordring. Mange er skeptiske til å ta KI i bruk fordi de ikke vet om dette er godt nok eller om det står seg. Våre funn tilsier at den største utfordringen vil bransjen og selskapene oppleve dersom de ikke blir med på den teknologiske utviklingen. Som en konsekvens er det andre relevante utfordringer rundt dette. Utfordringer trukket frem er videreutvikling av ansatte, investeringskostnader, juridiske- og andre regulatoriske hensyn. Integritet, sikkerhet, konfidensialitet, sikring av revisjonsspor i KI-verktøyet som sikrer fullstendig og nøyaktig input, pålitelighet i output og tillit til systemet viktige hensyn å ta høyde for. Dette gjelder også regulering og verifisering av den intelligente maskinen. Som en konsekvens av disse utfordringene er det større usikkerhet knyttet til å benytte KI til å utøve profesjonelt skjønn og skepsis.

Flere av faktorene trukket frem ovenfor er i overensstemmelse mellom litteratur og empiri. Dette er viktige funn som understøtter utfordringer rundt KI i revisjon. Det er enkelte utfordringer som intervju kandidatene i større grad enn litteraturen fremhever økt fokus rundt. Dette gjelder tilsvarende motsatt. Selv om litteraturen og empirien fokuserer på enkelte ulike utfordringer, er dette utfordringer som må hensyntas ved implementering for praktisk anvendelse. Når man tar i bruk KI fordrer dette en nevnt utfordring – at man har tillit til systemet man bruker.

Vi avdekket samsvar mellom litteratur og empiri på forskningsspørsmål 1, 2, 3, 5, 7, 8, 11, 13, 14, 18 og 21. Vi finner ikke empirisk støtte med litteraturen for 6, 9, 17 og 19. På forskningsspørsmål 2, 10, 12 og 22 utfyller litteraturen og empirien hverandre godt. Oppsummeringsvis ser vi at KI i første omgang vil implementeres innen risikovurdering og planleggingsfasen, samt misligheter. Det neste steget vil være implementering av KI i gjennomføringsfasen hvor KI implementeres i revisjonshandlinger for å innhente revisjonsbevis. Deretter vil man til slutt se om teknologien kan benyttes til å utøve profesjonelt skjønn og skepsis, og dermed foreta mye av revisjonen på eget grunnlag. Vi vil avslutningsvis si at det foreligger mange muligheter og fordeler i revisjon med bruk av KI. Tilsvarende er bruk av KI beheftet med utfordringer som må hensyntas.

7.2 Teoretiske og praktiske implikasjoner ved KI i revisjon

I en studie om KI i revisjon er det behov for å kartlegge de teoretiske implikasjonene gjennom empiri. Det er ytterligere behov og ønskelig å presisere praktiske implikasjoner av litteraturen og empirien. Denne delen tar for seg hvordan ny teori og praktiske hensyn kan utledes gjennom studiens forskningsprosess.

Enkelte deler av vår forskning er beheftet med svak eller mangelfull teori og litteratur. Teoretiske implikasjoner fra våre funn er at multi-modale revisjonsbevis kan benyttes i revisjonen til en viss grad og under visse forutsetninger. Samtidig understrekes det at KI kan innhente eksterne data for å bekrefte interne data. Videre er det gjort rede for at man kan benytte droner i revisjon og KI til å bistå og vurdere bevisets relevans. Enten gjennom direkte oppkobling mellom revisjonshandling til påstand eller KI til å gjøre en vurdering av dette. Dette kan sikres gjennom en verifisering av maskinen eller verktøyet. Det er beheftet delte meninger rundt KI og revisjonsstandarder, samt i skjønn og skepsis. Våre funn viser at standardsetterne er positive, mens utøverne er skeptiske og «problematiserer» bruken. I tillegg innebærer funnene fra forskningsspørsmål 2 teoretiske implikasjoner for utfordringer ved bruk av KI, utover det som er dekket i litteratur. Hvorvidt forskningsspørsmålene blir bekreftet eller avkreftet av empiri sammenstilt mot intervju eller at de utfyller hverandre, vil være med å avgjøre de teoretiske implikasjonene.

Resultatene fra studien fremlegger praktiske faktorer revisjonsbransjen må hensynta. Litteratur og empiri viser at mulighetene og fordelene ved å ta i bruk KI har en oppside for revisjonsprofesjonen. De viktigste ble trukket frem i delkapittel 7.1. Vi vil rette fokus mot implikasjonene rundt utfordringene.

KI innebærer bruk av teknologi som kan programmere komplekse og sofistikerte systemer. Vår empiri samsvarer med Datatilsynet (2018) og Nugent og Cunningham (2005) som viser at denne kompleksiteten gir utfordringer rundt transparensen og begrunnelsen til maskinens resultater. Den svarte boksen utgjør en barriere sett i lys av en profesjon som krever forståelse, dokumentasjon og etterprøvbarhet av resultater. Forståelse, håndtering og løsning på denne problematikken, etter beste evne, kan derfor føre til mer utstrakt bruk av KI i revisjon.

Revisjonsprofesjonen er strengt regulert og følger internasjonale standarder i sin yrkesutførelse. Resultatene våre viser at dagens standarder henger etter fremskrittene i samfunnet og teknologien. Dette innebærer at man må kunne standardene for å vite om man er innenfor, når de utfordres. Videre synes det at standardene ikke forbyr, men samtidig ikke fremmer bruk av ny teknologi (IAASB 2018a). Det tyder på KI kan benyttes, men fordrer utøvelse av varsomhet og trygghet på det man tar i bruk. Videre fremmer IAASB (2016) at bransjen må være proaktive og utforske uten støtte i litteratur og imøtekomme fremtidige utviklinger, uten at standardene kontinuerlig oppdateres. Vi mener derfor bransjen i større grad bør utfordre fremfor å satse på den klassiske revisjonen, hvor Kinserdal (2017) påpeker at det er vanskelig for tilsynsmyndighetene og standardsetterne å godkjenne ny revisjonsmetodikk før den er oppfunnet og testet. Vi mener ytterligere at utvalg og stikkprøver som praksis og «sedvanerett» kan selskapene endre på. Dette kan skje gjennom trygg og sikker bruk gjennom aktiv bruk av slike verktøy til analyse og testing av hele populasjoner. Samtidig må det føres trygg og etisk styring av systemet. Trusler for erstatningsansvar og andre sanksjoner er styrende for at revisjonsselskaper begrenser bruken i praksis eller gjør dette i tillegg til den klassiske revisjonen for å være på sikker grunn. Selv om standardsetterne er positive til utviklingen med KI og økt revisjonskvalitet, er det revisjonsselskapene som sitter med risikoen dersom noe går galt.

Samtidig som utviklingen av teknologi skjer raskt, tar det tid å bygge KI. Dette gjelder eksempelvis Google Duplex. Kompetanse rundt oppbyggingen av systemet er kun en side. Resultatene våre viser at videreutvikling av ansatte og studiefakultetenes ansvar og relevans av nyutdannede for bruk av slike verktøy og forståelse av resultatene er praktiske utfordringer som må imøtekommes. På denne måten kan fremtidens revisor opprettholde relevansen og eksistensen i markedet. Resultatene fra studien viser at bransjen må utvikle seg og hvor det er de som forbedrer og forstår bruken av kognitiv teknologi er de som vil lykkes. Samtidig ser vi utøvelse av myke verdier, skjønn og skepsis blir viktigere fremover. Etter hvert som teknologien modnes mer og systemene avanseres ytterligere, viser studiens litteratur og funn at KI kan benyttes i skjønn- og skepsis-vurderinger. Det er derimot beheftet med delte meninger rundt dette. I første omgang vil KI bli utøvet som et beslutningsstøtteverktøy i prediksjoner og konklusjoner til å støtte oppunder revisors egne vurderinger.

7.3 Oppgavens begrensninger

Studiens tema og problemstilling utgjør en begrensning i seg selv. Samtidig er problemstillingen forholdsvis vid og åpen. Denne begrenses ytterligere gjennom konkrete forskningsspørsmål forankret i presentert litteratur. En begrensning for oppgaven er at litteraturen har vært styrende for forskningsspørsmålene. Dermed er det mulig at vi ikke har avdekket andre potensielle muligheter og utfordringer. Vi har så langt vi kunne, forsøkt å imøtekomme dette gjennom grundig research av relevant litteratur, men også sett i lys av empirien. Dette har medført en forholdsvis bred, men samtidig en robust studie. Vi har forsøkt å forklare og utdype de ulike delene tilstrekkelig og hensiktsmessig for å sette ting i perspektiv.

Intervju og spørreskjema med seks kandidater gjør det vanskelig å generalisere funn. Vi har av den grunn påvist varsomhet i våre konklusjoner. Selv om intervju-kandidatene er dyktige, utgjør de et lite utvalg samtidig som de vet mer om temaet enn en vanlig revisor. På en annen side har vi intervjuet «nøkkelpersonell» med innsikt i temaet. Dette bidrar til innsikt og forståelse rundt de problemstillingene vi drøfter. Likevel ser vi i etterkant av innhentet empiri at det ville vært hensiktsmessig å inkludere et intervju med en akademiker innen revisjon, «KI-ekspert» og Finanstilsynet, selv om sistnevnte ble forsøkt. Spørreskjemaet ble

rettet mot de samme kildene vi intervjuet som begrenser dens funn ved at svarene allerede kan reflekteres i intervjuene. Som påpekt i kapittel 5 er det i en studie om KI i revisjon vanskelig å få tak i de rette kildene. Av praktiske hensyn ble derfor spørreskjemaet rettet mot de vi intervjuet.

Gjennom MRR-studiet har vi fått tilstrekkelig og hensiktsmessig forståelse for revisjon. En potensiell begrensning i en studie om Revisjon med KI er at vi ikke besitter teknologisk kompetanse. Vi har forsøkt å kompensere dette med en grundig og intensiv studie hvor vi har satt oss inn i dette temaet. Av den grunn føler vi at vi har oppnådd god generell kompetanse om KI, hvor vi har forsøkt å se KI og revisjon i sammenheng gjennom teori, litteratur og empiri.

Tildeling av PwC sitt Scholarship kan medføre begrensninger ved studien som påpekt i kapittel 5. På en annen side føler vi til gjengjeld et ansvar for å levere en avhandling med innhold av god kvalitet. Dette underbygges videre ved at vi ønsker at oppgaven skal komme flere andre interessenter til nytte. Et slikt fokus føler vi styrker denne masteravhandlingen.

7.4 Videre forskning

Noen forslag til videre forskning i forbindelse med muligheter og utfordringer til bruk av KI i revisjon ble gitt i delkapittel 5.1. Dette er forskningsspørsmål som vi anser som relevante for vår studie, men ikke hadde kapasitet eller empiri til å svare på. Vi må presisere hva som er lagt i begrepet videre forskning i vår studie. Gjennom kartlegging av relevant litteratur kan særskilt forskning eller litteratur falt utenfor våre søkekriterier, til tross for søkekriterier som var rettet mot KI med underliggende teknologier og tilknyttet revisjon. Videre forskning er derfor brukt som en samlebetegnelse for utført forskning som vi ikke har avdekket eller som det er behov for å utføre. Det var også artikler som var vanskelige å få tilgang til. I den sammenheng er det stor sannsynlighet for at vi ikke har fått tilgang til eller funnet andre gode relevante artikler. Dette utgjør en begrensning for oppgavens del, med å foreslå forskning som allerede kan være utført eller igangsatt.

En robust litteraturgjennomgang og innhentet data gir oss anledning til å supplere videre forskning på andre områder enn de vi har presentert i kapittel 5.1. Ettersom

vi innledningsvis i oppgaven introduserte KI til å favne bredt og mer enn bare ekstern finansiell revisjon, vil også videre forskning presentere forslag til forskning innen KI og andre typer revisjon.

Ettersom vår oppgave har et begrenset utvalg på seks personer, vil det være av interesse med en annen studie der utvalget i større grad kan generaliseres gjennom kvantitativ forskning. Case, eksperiment og særlig gruppeintervju av KI i revisjon ville også vært studier med høy relevans. Representerte avgrensninger i oppgaven utgjør også gode forslag til videre forskning. Vi vil avslutningsvis påpeke at mye av gjennomgått litteratur også presenterer relevante forslag til videre forskning.

Innspill til videre forskning:

- F4: Hvilke muligheter eller utfordringer med avansert analyse (KI og dataanalyse kombinert) eksisterer i stegene data tilgjengelighet, -innhenting, -vask, -transformasjon, -analyse og tolkning.
- F15: Hvordan sikre datakvaliteten i data som revisor innhenter og tolker. I delkapittel 4.2.1 introduserte vi egenskapene til data hvor disse blir viktigere i en mer datarik revisjon.
- F16: Kan standardenes henvisning til CAAT benyttes med revisjonsselskapets KI-verktøy, og må revisor forstå klientens IT-miljø og tilhørende kontroller i større eller mindre grad enn før. Vi har i delkapittel 4.2.1 og gjennom empiri funnet at det strides hvorvidt detaljtesting eller IK-testing vil bli den dominerende fremover med mer data og avanserte verktøy. Samtidig er et tredje syn å benytte eksterne data for å si noe om kvaliteten på den interne dataen.
- F20: Vil tilstrekkelig og hensiktsmessig revisjonsbevis endres som følge av KI. Relevant litteratur om dette ble presentert i delkapittel 4.2.4, der vi kartla at ny teknologi kan påvirke revisors skjønn i vurderingen av tilstrekkelig og hensiktsmessig revisjonsbevis. Samtidig antyder Brown-Liburd og Vasarhelyi (2015) at målet på tilstrekkelighet vil skifte mot tidsriktig tilgjengelighet og revisors evne til å innhente, analysere, tolke og evaluere informasjonen effektivt og meningsfullt.
- En utfyllende studie om KI og misligheter i en mer moderne revisjon.
- En mer detaljert beskrivelse av anvendelsen av droner i revisjon

- Dagens revisjonsstandarder mot nye teknologiske løsninger med fokus på konkrete standarder og punkter som utfordrer automatiseringsteknologier.
- En dypere studie om KI til å utøve skjønn og skepsis
- KI betydning og anvendelse for andre revisorer og i andre typer revisjoner, bør undersøkes ettersom anvendelsesområdet for KI er bredt.

Referanseliste

- Abdolmohammadi, M. (1999). A comprehensive taxonomy of audit task structure, professional rank and decision aids for behavioral research. *Behavioral Research in Accounting*, 11, 51-92.
- ACCA. (2015). The Data Revolution. http://www.accaglobal.com/content/dam/ACCA_Global/Technical/tech/pi/ACCA-IMA-data-revolution-report.pdf
- ACCA. (2017). *Exploring the Growing Use of Technology in the Audit, with a focus on Data Analytics A public consultation issued by IAASB* Hentet fra <https://www.ifac.org/system/files/publications/exposure-drafts/comments/ACCAresponsetoIAASBDataAnalyticsCP.pdf> doi
- Agrawal, A., Gans, J. S., & Goldfarb, A. (2016). *Exploring the Impact of Artificial Intelligence: Prediction versus Judgment* Hentet fra <https://webcache.googleusercontent.com/search?q=cache:OH0hdJXxZXAJ:https://www.aeaweb.org/conference/2017/preliminary/paper/BHKkz394+&cd=1&hl=no&ct=clnk&gl=no> doi
- AICPA. (2015). *Audit Analytics and Continuous Audit: Looking Toward The Future* Hentet fra https://competency.aicpa.org/media_resources/208500-audit-analytics-and-continuous-audit-looking-toward doi
- AICPA. (2017). *Guide to audit data analytics fact sheet* Hentet fra <https://www.aicpa.org/content/dam/aicpa/interestareas/frc/assuranceadvisorieservices/downloadabledocuments/guide-to-audit-data-analytics-fact-sheet-web.pdf> doi
- Alain, G., & Bengio, Y. (2016). *Understanding intermediate layers using linear classifier probes*.
- Allen, C., Smit, I., & Wallach, W. (2005). Artificial Morality: Top-down, Bottom-up, and Hybrid Approaches. *Ethics and Information Technology*, 7(3), 149-155. doi:10.1007/s10676-006-0004-4
- Alles, M., & Gray, G. L. (2016). Incorporating big data in audits: Identifying inhibitors and a research agenda to address those inhibitors. *International Journal of Accounting Information Systems*, 22, 44-59. doi:<https://doi.org/10.1016/j.accinf.2016.07.004>
- Andersen, L. M., & Bakkeli, M. (2015). *Big Data: Hva er Big Data, og hva betyr Big Data for deg?* Hentet fra <https://www.pwc.no/no/publikasjoner/information-management/big-data.pdf> doi
- Ang, S. H. (2014). *Research design for business & management*. Thousand Oaks, CA: Sage.
- Appelbaum, D. (2016). Securing Big Data Provenance for Auditors: The Big Data Provenance Black Box as Reliable Evidence. *Journal of Emerging Technologies in Accounting*, 13(1), 17-36. doi:10.2308/jeta-51473
- Appelbaum, D., Kogan, A., & Vasarhelyi, M. (2017a). Introduction to Data Analysis for Auditors and Accountants. *The CPA Journal*, 7-15.
- Appelbaum, D., Kogan, A., & Vasarhelyi, M. A. (2017b). Big Data and Analytics in the Modern Audit Engagement: Research Needs. *AUDITING: A Journal of Practice & Theory*, 36(4), 1-27. doi:10.2308/ajpt-51684
- Appelbaum, D., & Nehmer, R. A. (2017a). The Coming Disruption of Drones, Robots, and Bots: How Will It Affect CPAs and Accounting Practice? *CPA Journal*, 87(6), 40-44.

- Appelbaum, D., & Nehmer, R. A. (2017b). Using Drones in Internal and External Audits: An Exploratory Framework. *Journal of Emerging Technologies in Accounting*, 14(1), 99-113. doi:10.2308/jeta-51704
- Appelbaum, D. A., Kogan, A., & Vasarhelyi, M. A. (2018). Analytical procedures in external auditing: A comprehensive literature survey and framework for external audit analytics. *Journal of Accounting Literature*, 40, 83-101. doi:<https://doi.org/10.1016/j.acclit.2018.01.001>
- Arens, A. A., & Beasley, M. S. (2013). *Auditing and Assurance Services, Global Edition* (Fifteenth edition. utg.): United Kingdom: Pearson Education.
- Arens, A. A., Elder, R. J., & Beasley, M. S. (2016). *Auditing and Assurance Services, Global Edition* (16th ed. utg.). Harlow: Harlow : Pearson Education Limited.
- Asklund, A. (2016). Digger digitaliseringen. *Revisjon og Regnskap*, 7, 12-14.
- Asklund, A. (2017). Spiller på lag med teknologien. *Revisjon og Regnskap*, 7, 20-21.
- AUASB. (2017). AUASB Submission on IAASB Request for Input, Exploring the Growing Use of Technology in the Audit, with a Focus on Data Analytics. https://www.ifac.org/system/files/publications/exposure-drafts/comments/AUASBIAASBDAWGRFI_2March2017_final.pdf
- Aurstad, T. (2017). Revisjonsbevis i en digital hverdag. *Revisjon og Regnskap*(7).
- Banavar, D. G. (2016). *Learning to trust artificial intelligence systems* Hentet fra https://www.research.ibm.com/software/IBMResearch/multimedia/AIEthics_Whitepaper.pdf doi
- Barghini, T. (2017). *Auditing Gets Smart* Hentet fra <https://www.gfmag.com/magazine/march-2017/corporate-finance-future-audits> doi
- BBC. (2016). *Artificial intelligence: Go master Lee Se-dol wins against AlphaGo program* Hentet fra <http://www.bbc.com/news/technology-35797102> doi
- Birkeland, K. (2017). *Ekstern revisors handleplikter*. Oslo: Cappelen Damm akademisk.
- Bizarro, P., & Dorian, M. (2017). ARTIFICIAL INTELLIGENCE: THE FUTURE OF AUDITING. *Internal Auditing*, 32(5), 21-26.
- Boer, M. d., Eimers, P., & Elsas, P. (2014). *Reengineering the audit in a digitized environment - developments in practice, challenges for auditing standards and opportunities for further research* Hentet fra http://www.isarhq.org/2014_downloads/papers/ISAR2014_Boer_Eimers_Elsas.pdf doi
- Brazel, J. F., & Schaefer, T. J. (2015). Executive Summary: State of the Art Research Related to Auditor Professional Skepticism (2013 – 2015). https://www.iaasb.org/system/files/meetings/files/20151207-IAASB-Agenda_Item_7B-Executive-Summary-State_of_Art_Research_Related_to_Auditor_Professional_Skepticism-final_0.pdf
- Brennan, B., Baccala, M., & Flynn, M. (2017). Artificial Intelligence Comes to Financial Statement Audits. <http://ww2.cfo.com/auditing/2017/02/artificial-intelligence-audits/>
- Brown-Liburud, 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. doi:10.2308/acch-51023

- Brown-Libur, H., & Vasarhelyi, M. A. (2015). Big Data and Audit Evidence. *Journal of Emerging Technologies in Accounting*, 12(1), 1-16. doi:10.2308/jeta-10468
- Brown, C. E., Wong, J. A., & Baldwin, A. A. (2007). A Review and Analysis of the Existing Research Streams in Continuous Auditing. *Journal of Emerging Technologies in Accounting*, 4(1), 1-28. doi:10.2308/jeta.2007.4.1.1
- Brørs, T., & Sellæg, F. E. (2015). Automatisering av regnskapsfunksjonen. *Praktisk økonomi & finans*, 31(04), 307-318.
- Burgess, A. (2018). *The executive guide to artificial intelligence : how to identify and implement applications for AI in your organization*: Springer International Publishing : Imprint: Palgrave Macmillan.
- Byrnes, P. E. (2015). Developing automated applications for clustering and outlier detection. doi:<https://doi.org/doi:10.7282/T3R78H7Q>
- Bøhren, Ø. (2011). *Eierne, styret og ledelsen : corporate governance i Norge*. Bergen: Fagbokforl.
- Chabert, J. L., Weeks, C., Barbin, E., Borowczyk, J., Chabert, J. L., Guillemot, M., . . . Martzloff, J. C. (2012). *A History of Algorithms: From the Pebble to the Microchip*: Springer Berlin Heidelberg.
- Chan, D. Y., & Vasarhelyi, M. A. (2011). Innovation and practice of continuous auditing. *International Journal of Accounting Information Systems*, 12(2), 152-160. doi:<https://doi.org/10.1016/j.accinf.2011.01.001>
- Chiu, T., & Jans, M. (2017). *Process Mining of Event Logs: A Case Study Evaluating Internal Control Effectiveness*, Working paper Hentet fra <https://ssrn.com/abstract=3136043> doi:10.2139/ssrn.3136043
- Chiu, T., Vasarhelyi, M., Alrefai, A., & Yan, Z. (2018). *Validating Process Mining: A Framework Integrating Auditor's Risk Assessment*, Working paper Hentet fra <https://ssrn.com/abstract=3136007> doi:10.2139/ssrn.3136007
- Clark, A. (2017). Machine Learning Audits in the 'Big Data Age'. *CIO Insight*, 1-1.
- Copeland, M. (2016). What's the Difference Between Artificial Intelligence, Machine Learning, and Deep Learning? , <https://blogs.nvidia.com/blog/2016/07/29/whats-difference-artificial-intelligence-machine-learning-deep-learning-ai/>
- CPA-Canada. (2017a). Audit Data Analytics Alert. *Survey On Use of Audit Data Analytics in Canada*, <https://www.cpacanada.ca/-/media/site/business-and-accounting-resources/docs/01453-rg-audit-data-analytics-alert-survey-using-ada-in-canada-october-2017.pdf>
- CPA-Canada. (2017b). Re: Exploring the growing use of technology in the audit, with a focus on data analytics. <https://www.ifac.org/system/files/publications/exposure-drafts/comments/CPACanadaADAResponsetoIAASB-Final-Feb152017.pdf>
- CPA-Canada. (2018). Audit Data Analytics Alert. *Update on the Use of Audit Data Analytics and the Evolution of the Financial Statement Audit*, <https://www.cpacanada.ca/-/media/site/business-and-accounting-resources/docs/g10492-rg-audit-data-analytics-alert-uwcisa-symposium-oct2017-january-2018.pdf?la=en&hash=A37B1F326924C49DAC8C161EDE80A951B056CBE3>

- CPA. (2016). Audit Data Analytics Alert. <https://www.cpacanada.ca/en/business-and-accounting-resources/audit-and-assurance/canadian-auditing-standards-cas/publications/audit-data-analytics-alert-pace-of-change>
- Dai, J., & Vasarhelyi, M. A. (2016). Imagineering Audit 4.0. *Journal of Emerging Technologies in Accounting*, 13(1), 1-15. doi:10.2308/jeta-10494
- Datatilsynet. (2018). *Kunstig intelligens og personvern (Rapport, januar 2018)* Hentet fra <https://www.datatilsynet.no/globalassets/global/om-personvern/rapporter/rapport-om-ki-og-personvern.pdf> doi
- Davenport, T. (2016). *Innovation in audit takes the analytics, AI route* Hentet fra <https://www2.deloitte.com/insights/us/en/focus/cognitive-technologies/audit-analytics-artificial-intelligence.html> doi
- Deloitte. (2016). *Deloitte forms alliance with Kira Systems to drive the adoption of artificial intelligence in the workplace* Hentet fra <https://www2.deloitte.com/us/en/pages/about-deloitte/articles/press-releases/deloitte-forms-alliance-with-kira-systems-to-drive-the-adoption-of-artificial-intelligence-in-the-workplace.html> doi
- Deloitte. (2017). *Auditing the risks of disruptive technologies* Hentet fra <https://www2.deloitte.com/content/dam/Deloitte/us/Documents/risk/us-risk-digitization-banking.pdf> doi
- Earley, C. E. (2015). Data analytics in auditing: Opportunities and challenges. (58), 493-500 doi:<https://doi.org/10.1016/j.bushor.2015.05.002>
- Economist, T. (2016). Automation and anxiety: Will smarter machines cause mass unemployment?
- EY. (2017a.). EY scaling the use of drones in the audit process. Hentet fra <http://www.ey.com/gl/en/newsroom/news-releases/news-ey-scaling-the-use-of-drones-in-the-audit-process>
- EY. (2017b). Request for Input – Exploring the Growing Use of Technology in the Audit, with a Focus on Data Analytics. <https://www.ifac.org/system/files/publications/exposure-drafts/comments/EYGCommentLetterDataAnalyticsFINAL.pdf>
- EY. (2017c). *Welcome to the machines* Hentet fra [http://www.ey.com/Publication/vwLUAssets/ey-reporting-ai-welcome-to-the-machines/\\$FILE/ey-reporting-ai-welcome-to-the-machines.pdf](http://www.ey.com/Publication/vwLUAssets/ey-reporting-ai-welcome-to-the-machines/$FILE/ey-reporting-ai-welcome-to-the-machines.pdf) doi
- FASB. (2006). Conceptual Framework for Financial Reporting: *Objective of Financial Reporting and Qualitative Characteristics of Decision-Useful Financial Reporting Information*, http://www.fasb.org/jsp/FASB/Document_C/DocumentPage?cid=1218220340119&acceptedDisclaimer=true
- Fenstad, J. E. (2017). Struktur og algoritmer: hva kunnskap er. *Nytt Norsk Tidsskrift*, 34(02), 182-191. doi:10.18261/issn.1504-3053-2017-02-07
- Fenwick, N. (2014). *The Future Of Business Is Digital* Hentet fra <https://www.forbes.com/sites/forrester/2014/03/10/the-future-of-business-is-digital/#4227083f7db8> doi
- Ferreira, D. R. (2017). *A Primer on Process Mining : Practical Skills with Python and Graphviz* SpringerBriefs in Information Systems, Hentet fra doi
- Fisher, I., Garnsey, M., & Hughes, M. (2016). Natural Language Processing in Accounting, Auditing and Finance: A Synthesis of the Literature with a Roadmap for Future Research. *Intelligent Systems in Accounting, Finance and Management*, 23(3), 157. doi:10.1002/isaf.1386
- Fjørtoft, L. E. (2018). Digitalisering og disrupsjon i revisjonsbransjen. *Revisjon og Regnskap*(1), 24-26.

- Forum, W. E. (2016). *The Future of Jobs* Hentet fra http://www3.weforum.org/docs/WEF_Future_of_Jobs.pdf doi
- Foy, H. (2016). PwC to deploy drone army. <https://www.irishtimes.com/business/technology/pwc-to-deploy-drone-army-1.2640692>
- FRC. (2017). Exploring the Growing Use of Technology in the Audit, With a Focus on Data Analytics. <https://www.ifac.org/system/files/publications/exposure-drafts/comments/IAASBDAWGConsultationFRCResponse.pdf>
- Gardner, H. (1999). *Intelligence reframed : multiple intelligences for the 21st century*. New York: Basic Books.
- Gartner. (2009). *Digitalization* Hentet fra <https://gartner.com/it-glossary/digitalization> doi
- Gee, S. (2014). *Fraud and Fraud Detection: A Data Analytics Approach*: United States: John Wiley & Sons Inc.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*: MIT Press.
- Gottfredson, L. S. (1994). *Mainstream Science on Intelligence: An Editorial With 52 Signatories, History, and Bibliography* Hentet fra <https://www-sciencedirect-com.ezproxy.library.bi.no/science/article/pii/S0160289697900118> doi:[https://doi.org/10.1016/S0160-2896\(97\)90011-8](https://doi.org/10.1016/S0160-2896(97)90011-8)
- Grandal, B. H. (2018). Regulatoriske krav stopper revisjons-revolusjon. *Finansavisen*(5. april).
- Gray, G. L., & Debreceeny, R. S. (2014). A taxonomy to guide research on the application of data mining to fraud detection in financial statement audits. *International Journal of Accounting Information Systems*, 15(4), 357-380. doi:<https://doi.org/10.1016/j.accinf.2014.05.006>
- Griggs, A. (2017). *Artificial intelligence and machine learning – the inevitable changes in professional service firms* Hentet fra <https://www.accountancyage.com/2017/08/14/artificial-intelligence-and-machine-learning-the-inevitable-changes-in-professional-service-firms/> doi
- Gulden, B. P. (2016). *Revisjon : teori og metode* (7. utg. utg.). Oslo: Cappelen Damm akademisk.
- Heggernes, T. A. (2017). *Digital forretningsforståelse : fra store data til små biter* (2. utg. utg.). Bergen: Fagbokforl.
- High, J., Harriott, N., Sporrer, T., Davies, S., & Poh, C. (2016). Comments to the International Auditing and Assurance Standards Board on exploring the growing use of technology in the audit, with a focus on data analytics. https://www.ifac.org/system/files/publications/exposure-drafts/comments/DataAnalyticsCommentLetter_.pdf
- Hindberg, T. (2015). Big Data og revisjon. *Revisjon og Regnskap*(3).
- IAASB. (2016). *Exploring the Growing Use of Technology in the Audit, with a Focus on Data Analytics* Hentet fra <https://www.ifac.org/system/files/publications/files/IAASB-Data-Analytics-WG-Publication-Aug-25-2016-for-comms-9.1.16.pdf> doi
- IAASB. (2018a). Feedback statement: Exploring the Growing Use of Technology in the Audit, with a Focus on Data Analytics.
- IAASB. (2018b). *ISA 315 (Revised), Identifying and Assessing the Risks of Material Misstatement* Hentet fra <http://www.iaasb.org/system/files/meetings/files/20180312->

- [IAASB_CAG_Agenda_Item_G3_ISA-315-Revised-Application-Material-Marked-to-extant-Final.pdf](#) doi
- IAIS. (2017). RE: IAASB's Request for Input - Exploring the Growing Use of Technology in the Audit, with a Focus on Data Analytics.
- IBM. (2018.). Deep Blue. Hentet fra <http://www-03.ibm.com/ibm/history/ibm100/us/en/icons/deepblue/>
- ICAEW. (2016). Data analytics for external auditors. <https://www.icaew.com/-/media/corporate/files/technical/iaa/tecpln14726-iaae-data-analytics---web-version.ashx>
- ICAEW. (2017a). *Artificial intelligence and the future of accountancy* Hentet fra <https://www.icaew.com/-/media/corporate/files/technical/information-technology/technology/ai-report-web.ashx?la=en> doi
- ICAEW. (2017b). *The Future of Audit: Technology* Hentet fra <https://www.icaew.com/en/technical/audit-and-assurance/faculty/the-future-of-audit/the-future-of-audit-technology> doi
- ICAS. (2016). Auditor Skill in a Changing Business World. https://www.icas.com/_data/assets/pdf_file/0010/259516/Auditor-skills-ICAS-FRC-20.9.16-Final.pdf
- IDW. (2017). Re.: IAASB Data Analytics Working Group 'Request for Input "Exploring the Growing Use of Technology in the Audit, with a Focus on Data Analytics"'. <https://www.ifac.org/system/files/publications/exposure-drafts/comments/DataAnalytics31Jan2017.pdf>
- ISCA. (2017). Response to the IAASB request for input, exploring the growing use of technology in the audit, with a focus on data analytics. <https://www.ifac.org/system/files/publications/exposure-drafts/comments/IAASBRFIDataAnalytics-CommentLetterFINAL.pdf>
- Issa, H., Sun, T., & Vasarhelyi, M. A. (2016). Research Ideas for Artificial Intelligence in Auditing: The Formalization of Audit and Workforce Supplementation. *Journal of Emerging Technologies in Accounting*, 13(2), 1-20. doi:10.2308/jeta-10511
- Jacobsen, D. I. (2005). *Hvordan gjennomføre undersøkelser? : innføring i samfunnsvitenskapelig metode* (2. utg. utg.). Kristiansand: Høyskoleforl.
- Jans, M., Alles, M. G., & Vasarhelyi, M. A. (2014). A Field Study on the Use of Process Mining of Event Logs as an Analytical Procedure in Auditing. *The Accounting Review*, 89(5), 1751-1773. doi:10.2308/accr-50807
- JICPA. (2017). JICPA Comments on the Request for Input, Exploring the Growing Use of Technology in the Audit, with a Focus on Data Analytics. <https://www.ifac.org/system/files/publications/exposure-drafts/comments/20170215-JICPACommentsontheRequestforInputExploringtheGrowingUseofTechnologyintheAuditwithaFocusonDataAnalytics.pdf>
- Johansen, Ø. K. (2015). *Verdien av tillit : revisjonens historie*. Oslo: Kagge.
- Jonker, J. (2007). *The essence of research methodology : a concise guide for master and Phd students in management science*. Heidelberg: Springer.
- Kassambara, A. (2018). *Machine Learning Essentials: Practical Guide in R*: CreateSpace Independent Publishing Platform.
- Kinserdal, F. (2017). NHH skal forske på digitalisering i revisjonsbransjen. <https://www.magma.no/nhh-skal-forske-pa-digitalisering-i-revisjonsbransjen>
- Kinserdal, F. (2018). *DIGITAL REVISJON: INTRO; FORELESNING 1 OG 2*. NHH.

- Knudsen, E., Siebke, H. A., Cordt-Hansen, H., & Revisorforeningen. (2010). *Revisorloven med kommentarer : lov om revisjon og revisorer av 15. januar 1999 nr. 2* (4. utg. [oppdatert og supplert ved] Henning Alme Siebke og Espen Knudsen. utg.). Oslo: Den norske revisorforening.
- Kogan, A., Alles, M. G., & Vasarhelyi, M. A. (2010). Analytical Procedures for Continuous Data Level Auditing: Continuity Equations. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.174.240&rep=rep1&type=pdf>
- Kokina, J., & Davenport, T. H. (2017). The Emergence of Artificial Intelligence: How Automation is Changing Auditing. *Journal of Emerging Technologies in Accounting*, 14(1), 115-122. doi:10.2308/jeta-51730
- Kolbjørnsrud, V. (2017). *Kunstig intelligens og lederens nye jobb* Hentet fra <https://www.magma.no/kunstig-intelligens-og-lederens-nye-jobb> doi
- Koskivaara, E. (2004). Artificial neural networks in analytical review procedures. *Managerial Auditing Journal*, 19(2), 191-223. doi:10.1108/02686900410517821
- KPMG. (2016a). *Game Changer The Impact of Cognitive Technology on Business and Financial Reporting* Hentet fra <https://assets.kpmg.com/content/dam/kpmg/pdf/2016/05/game-changer-impact-of-cognitive-technology.pdf> doi
- KPMG. (2016b). *Revisjonens digitale reise* Hentet fra <https://home.kpmg.com/no/nb/home/nyheter-og-innsikt/2016/03/revisjonens-digitale-reise.html> doi
- KPMG. (2017a). *KPMG Clara A smart audit platform* Hentet fra <https://assets.kpmg.com/content/dam/kpmg/xx/pdf/2017/05/kpmg-clara-a-smart-audit-platform.pdf> doi
- KPMG. (2017b). Request for Input: Exploring the Growing Use of Technology in the Audit, with a Focus on Data Analytics. <https://www.ifac.org/system/files/publications/exposure-drafts/comments/KPMGresponsetoDataAnalyticsRFI.pdf>
- Kraheil, J. P., & Titera, W. R. (2015). Consequences of Big Data and Formalization on Accounting and Auditing Standards. *Accounting Horizons*, 29(2), 409-422. doi:10.2308/acch-51065
- Krishnaswami, D. O. R., & Satyaprasad, D. B. G. (2010). *Business Research Methods*: Himalaya Publishing House.
- Kristoffersen, T. (2014). *Virksomhetsstyring og regnskapsorganisering : lærebok*. Bergen: Fagbokforl.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). *Deep Learning* Vol. 521. Hentet fra http://www.cnd.mcgill.ca/~ivan/Papers/deep_learning_review.pdf doi:doi:10.1038/nature14539
- Li, Q., & Vasarhelyi, M. (u.å.). Developing A Cognitive Assistant For Audit Plan Brainstorming Sessions. <http://raw.rutgers.edu/docs/wcars/41wcars2/Li.pdf>
- Lindberg, T. M. (2017). Fremtidens revisor. *Revisjon og Regnskap*(6), 5.
- Liu, B. (2011). *Web data mining : exploring hyperlinks, contents, and usage data*. Berlin, Heidelberg: Springer.
- Macaulay, M. (2016a). *How Cognitive Tech Is Revolutionizing the Audit* Hentet fra <https://www.financialexecutives.org/Topics/Strategy/How-Cognitive-Tech-Is-Revolutionizing-the-Audit.aspx> doi

- Macaulay, M. (2016b). *How Technology Is Transforming the Audit* Hentet fra <https://www.forbes.com/sites/kpmg/2016/11/23/how-technology-is-transforming-the-audit/#21fa15fa716c> doi
- Marsland, S. (2015). *Machine learning : an algorithmic perspective* Chapman & Hall/CRC machine learning & pattern recognition series, Hentet fra doi
- Mayes, J. (2017). Machine Learning 101. https://docs.google.com/presentation/d/1kSuQyW5DTnkVaZEjGYCkfOxvzCqGEFzWBy4e9Uedd9k/present?ueb=true&slide=id.g168a3288f7_0_58
- Mazzei, C., & Duffy, N. (2018). *Putting artificial intelligence (AI) to work* Hentet fra <http://www.ey.com/gl/en/issues/business-environment/ey-innovation-matters-putting-artificial-intelligence-to-work> doi
- McKee, T. E. (2018a). *Audit Data Analytics Documentation requirements*. NHH.
- McKee, T. E. (2018b). *Cluster analysis*. NHH.
- McKee, T. E. (2018c). *Cluster analysis for anomaly detection*. NHH.
- McKinsey. (2017). *What the future of work will mean for jobs, skills, and wages* Hentet fra <https://www.mckinsey.com/global-themes/future-of-organizations-and-work/what-the-future-of-work-will-mean-for-jobs-skills-and-wages> doi
- Meek, T. (2017). *How Humans And AI Will Share The Auditing Function Of The Future* Hentet fra <https://www.forbes.com/sites/workday/2017/07/10/how-humans-and-ai-will-share-the-auditing-function-of-the-future/#597346a84fa1> doi
- Milne, R. (2017). *Nordea to axe 6,000 jobs in drive to digital future* Hentet fra <https://www.ft.com/content/fa627912-ba50-11e7-8c12-5661783e5589> doi
- Milner, C., & Berg, B. (2017). Tax Analytics: Artificial Intelligence and Machine Learning–Level 5 <https://www.pwc.no/no/publikasjoner/Digitalisering/artificial-intelligence-and-machine-learning-finall.pdf>
- Miralles-Pechuán, L., Rosso, D., Jiménez, F., & García, J. (2017). A methodology based on Deep Learning for advert value calculation in CPM, CPC and CPA networks. *A Fusion of Foundations, Methodologies and Applications*, 21(3), 651-665. doi:10.1007/s00500-016-2468-4
- Moffitt, K. C., & Vasarhelyi, M. A. (2013). AIS in an Age of Big Data. *Journal of Information Systems*, 27(2), 1-19. doi:10.2308/isys-10372
- Montgomery, D. C., Peck, E. A., & Vining, G. G. (2015). *Introduction to Linear Regression Analysis*: Wiley.
- Nagar, Y., & Malone, T. W. (2011). *Making Business Predictions by Combining Human and Machine Intelligence in Prediction Markets* Hentet fra <https://www.mckinsey.com/global-themes/future-of-organizations-and-work/what-the-future-of-work-will-mean-for-jobs-skills-and-wages> doi
- Nearon, B. (2005). Foundations in Auditing and Digital Evidence. *The CPA Journal*, 75(1), 32-34.
- Nilsson, N. J. (2009). *The quest for artificial intelligence : a history of ideas and achievements*: Cambridge University Press.
- NTNU Kveld: Kunstig intelligens: fantastisk eller farlig? [videoklipp]. (2018.). Hentet fra <https://youtu.be/Mxy7xxIxWmA>
- Nugent, C., & Cunningham, P. (2005). A Case-Based Explanation System for Black-Box Systems. *Artificial Intelligence Review*, 24(2), 163-178. doi:10.1007/s10462-005-4609-5
- Nygaard, N. (2007). *Skade og ansvar* (6. utg. utg.). Bergen: Universitetsforl.

- Omoteso, K. (2012). Review: The application of artificial intelligence in auditing: Looking back to the future. *Expert Syst. Appl.*, 39(9), 8490-8495. doi:10.1016/j.eswa.2012.01.098
- Ot.prp.nr.75. (1997-1998). *Om lov om revisjon og revisorer (revisorloven)* Hentet fra <https://lovdata.no/pro/PROP/forarbeid/otprp-75-199798> doi
- Ot.prp.nr.78. (2008-2009). *Om lov om endringer i revisorloven og enkelte andre lover (gjennomføring av revisjonsdirektivet)* Hentet fra <https://www.regjeringen.no/contentassets/d5f4e330fcde4c42a3d29ed92c8fe3f1/no/pdfs/otp200820090078000dddpdfs.pdf> doi
- Oxford. (2018). *Algorithm* Hentet fra <https://en.oxforddictionaries.com/definition/algorithm> doi
- Partington, A. (1996). *The Oxford Dictionary of Quotations*: Oxford University Press.
- PCAOB. (2017). THE USE OF DATA AND TECHNOLOGY IN AUDITS. <https://pcaobus.org/News/Events/Documents/05242017-SAG-meeting/DTA-Briefing%20Paper%20May-2017.pdf>
- Pedersen, J. S. (2016). Dataanalyse i revisjon. *Revisjon og Regnskap*(7).
- Persico, F., & Boillet, J. (2017.). What Impact Will AI Have on the Audit? Hentet fra <https://daily.financialexecutives.org/impact-will-ai-audit/>
- Persico, F., & Sidhu, H. (2017). *Voices How AI will turn auditors into analysts* Hentet fra <https://www.accountingtoday.com/opinion/how-ai-will-turn-auditors-into-analysts> doi
- Postholm, M. B. (2010). *Kvalitativ metode : en innføring med fokus på fenomenologi, etnografi og kasusstudier* (2. utg. utg.). Oslo: Universitetsforl.
- PwC. (2016a). Global Market for Commercial Applications of Drone Technology Valued at over \$127 bn. <https://press.pwc.com/News-releases/global-market-for-commercial-applications-of-drone-technology-valued-at-over%E2%80%93%20127-bn/s/AC04349E-C40D-4767-9F92-A4D219860CD2>
- PwC. (2016b). *PwC wins 'Audit Innovation of the Year' at the Accountant & International Accounting Bulletin awards 2016* Hentet fra <https://press.pwc.com/News-releases/pwc-wins--audit-innovation-of-the-year--at-the-accountant---international-accounting-bulletin-awards/s/9fed3272-3c5a-4fef-aebe-a91a4105c548> doi
- PwC. (2017a). Harnessing the power of AI to transform the detection of fraud and error. <https://www.pwc.com/gx/en/about/stories-from-across-the-world/harnessing-the-power-of-ai-to-transform-the-detection-of-fraud-and-error.html>
- PwC. (2017b). Investing in America's data science and analytics talent. http://www.bhef.com/sites/default/files/bhef_2017_investing_in_dsa.pdf
- PwC. (2017c). *PwC wins Audit Innovation of the Year award* Hentet fra <https://www.pwchk.com/en/services/audit-and-assurance/publications/accounting-listing-rule-updates/pwc-wins-prestigious-international-accounting-bulletins-2016-audit-innovation-of-the-year-award.html> doi
- Rammen, K. (2017). *Nettbaserte regnskapsprogrammer* Hentet fra <https://finanssans.no/regnskapsprogram> doi
- Raphael, J. (2015). *How Artificial Intelligence Can Boost Audit Quality* Hentet fra <http://bit.ly/2q301ZM> doi

- Raphael, J. (2017). *Rethinking the audit* Hentet fra <https://www.journalofaccountancy.com/issues/2017/apr/rethinking-the-audit.html> doi
- Rapoport, M. (2016). Auditing Firms Count on Technology for Backup. *Wall Street Journal* (8. mars).
- Reuters, T. (2017). 4 Keys to the Future of Audit. <https://tax.thomsonreuters.com/wp-content/private/pdf/checkpoint/whitepapers/Checkpoint-AI-Anderson-Whitepaper.pdf>
- Revisorforeningen. (2009a). ISA 200 *Overordnede mål for den uavhengige revisor og gjennomføringen av en revisjon i samsvar med de internasjonale revisjonsstandardene*.
- Revisorforeningen. (2009b). ISA 230 *Revisjonsdokumentasjon*.
- Revisorforeningen. (2009c). ISA 315 *Identifisering og vurdering av risikoene for vesentlig feilinformasjon gjennom forståelse av enheten og dens omgivelser*.
- Revisorforeningen. (2009d). ISA 330 *Revisors håndtering av anslåtte risikoer*.
- Revisorforeningen. (2009e). ISA 500 *Revisjonsbevis*.
- Revisorforeningen. (2009f). ISA 620 *Bruk av arbeidet til en ekspert engasjert av revisor*.
- Revisorutvalget. (1997). *NOU 1997: 9 Om revisjon og revisorer* Hentet fra <https://www.regjeringen.no/contentassets/670c887a3d5f49e983b100775576446a/no/pdfa/nou199719970009000dddpdfa.pdf> doi
- Rezaee, Z., Sharbatoghlie, A., Elam, R., & McMickle, P. L. (2002). Continuous Auditing: Building Automated Auditing Capability. *AUDITING: A Journal of Practice & Theory*, 21(1), 147.
- Rolstadås, A., Krokan, A., Dyrhaug, L. T., & Norges tekniske, v. (2017). *Teknologien endrer samfunnet*. Bergen: Fagbokforl.
- Russell, S. J., Norvig, P., & Davis, E. (2016). *Artificial Intelligence: A Modern Approach, Global Edition*: Pearson Education M.U.A.
- Rutgers. (2017). The Rutgers CarLab is pleased to provide comments on the IAASB Data Analytics Working Group Consultation Paper, Exploring the Growing Use of Technology in the Audit, with a Focus on Data Analytics. https://www.ifac.org/system/files/publications/exposure-drafts/comments/IAASBComment_Fina12-15-17.pdf
- Rutgers, & AICPA. (u.å.-a). Multidimensional Audit Data Selection (MADS) Project. <https://www.aicpa.org/content/dam/aicpa/interestareas/frc/assuranceadvisoryservices/downloadabledocuments/radar/radar-multidimensional-audit-data-selection.pdf>
- Rutgers, & AICPA. (u.å.-b). Process Mining Project. <https://www.aicpa.org/content/dam/aicpa/interestareas/frc/assuranceadvisoryservices/downloadabledocuments/radar/radar-mining-project-summary.pdf>
- Rutgers, & AICPA. (u.å.-c). Visualization as Audit Evidence. <http://raw.rutgers.edu/docs/radar/Visualization%20Project%20Summary.pdf>
- Ruud, T. F. (2000). Assurance services : nye muligheter for revisors yrkesutøvelse (del 1). *Revisjon og Regnskap*(7), 20-26.
- Ruud, T. F. (2001). Assurance services : nye muligheter for revisors yrkesutøvelse (del 2). *Revisjon og Regnskap*(4), 52-57.

- Sandberg, B. T. (2017). *Kunstig intelligens er en utfordring for Norge* Aftenposten Hentet fra <https://www.aftenposten.no/meninger/kronikk/i/QvLp4/Kunstig-intelligens-er-en-utfordring-for-Norge--Bjorn-T-Sandberg> doi
- Saunders, M. N. K., Lewis, P., & Thornhill, A. (2015). *Research Methods for Business Students*: Pearson Education Limited.
- Schmee, J., & Oppenlander, J. E. (2010). *JMP Means Business: Statistical Models for Management*: SAS Institute.
- Schwab, K. (2017). *The Fourth Industrial Revolution*: Crown Publishing Group.
- sentralbyrå, S. (2006). Retningslinjer for visuell utforming av spørreskjema. <https://www.ssb.no/a/publikasjoner/pdf/DP/handbok-88.pdf>
- Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., . . . Hassabis, D. (2017). *Mastering the game of Go without human knowledge* Vol. 550. Hentet fra <https://drive.google.com/file/d/1pjhZ1OzM0e8TUttVpK7E2mfqxWScpxDR/view> doi:doi:10.1038/nature24270
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). *Dropout: A Simple Way to Prevent Neural Networks from Overfitting* (Vol. 15).
- Sundet, J. M. (2015). *Hva er intelligens* (Vol. 56). Oslo: Universitetsforl.
- Telle, J. A. (2017). Den nye maskinlæringen: Kunstig intelligens eller bare gode verktøy? *Nytt Norsk Tidsskrift*, 34(2), 192-204.
- Thagaard, T. (2013). *Systematikk og innlevelse : en innføring i kvalitativ metode* (4. utg. utg.). Bergen: Fagbokforl.
- Tjønndal, A. (2018). *Statistisk analyse i Stata*. Oslo: Cappelen Damm akademisk.
- Tysiac, K. (2017). *How to enable audit innovation* Hentet fra <https://www.journalofaccountancy.com/issues/2017/apr/audit-innovation.html> doi
- Tørresen, J. (2013). *Hva er kunstig intelligens* (Vol. 49). Oslo: Universitetsforl.
- van Der Aalst, W. (2012). Process Mining. *Association for Computing Machinery. Communications of the ACM*, 55(8), 76. doi:10.1145/2240236.2240257
- van der Aalst, W. (2016). *Process Mining : Data Science in Action* Hentet fra doi
- Vasarhelyi, M. A. (2017). Auditing, the Technological Revolution, and Public Good. <http://raw.rutgers.edu/MiklosVasarhelyi/PIOBAuditing.TheTechnologicalRevolutionandPublicGoodslides.pdf>
- Vasarhelyi, M. A., & Sun, T. (2017). *Deep Learning and the Future of Auditing* Hentet fra <https://www.cpajournal.com/2017/06/19/deep-learning-future-auditing/> doi
- Wei Koh, P., & Liang, P. (2017). *Understanding Black-box Predictions via Influence Functions*.
- Yoon, K. (2016). *Three Essays On Unorthodox Audit Evidence* Hentet fra <https://rucore.libraries.rutgers.edu/rutgers-lib/50579/PDF/1/play/> doi
- Yoon, K., Hoogduin, L., & Zhang, L. (2015). Big Data as Complementary Audit Evidence. *Accounting Horizons*, 29(2), 431-438. doi:10.2308/acch-51076
- Zhang, L., Pawlicki, A. R., McQuilken, D., & Titera, W. R. (2012). The AICPA Assurance Services Executive Committee Emerging Assurance Technologies Task Force: The Audit Data Standards (ADS) Initiative. *Journal of Information Systems*, 26(1), 199-205. doi:10.2308/isys-10277

- Zikmund, W. G., Babin, B. J., Carr, J. C., & Griffin, M. (2010). *Business research methods* (8th ed. [rev. by] Barry J. Babin, Jon C. Carr, Mitch Griffin. utg.). Mason, Ohio: South Western Cengage Learning.
- Zohrabi, M. (2013). Mixed Method Research: Instruments, Validity, Reliability and Reporting Findings.
<http://www.academypublication.com/issues/past/tpls/vol03/02/tpls0302.pdf#page=56> doi:10.4304/tpls.3.2.254-262