



Universitetet
i Stavanger

HANDELSHØGSKOLEN VED UIS
MASTEROPPGAVE

STUDIEPROGRAM:

Master i regnskap og revisjon

ER OPPGAVEN KONFIDENSIELL?

(**NB!** Bruk rødt skjema ved konfidensiell oppgave)

TITTEL:

Bruk av kunstig intelligens i revisjonsbransjen

ENGELSK TITTEL:

The use of artificial intelligence in the audit industry

FORFATTER(E)

Kandidatnummer:

5013

.....

5000

.....

Navn:

Anders Skogseth

.....

Stig Herigstad

.....

VEILEDER:

Carmen Olsen

Forord

Denne oppgaven markerer avslutningen på vår mastergrad i regnskap og revisjon ved Universitetet i Stavanger.

Bakgrunnen for valget av temaet kunstig intelligens i revisjonsbransjen, er en forventning om at det vil endre revisors oppgaver i tiden fremover. Da temaet ikke har vært omtalt i pensum, ønsket vi å få en mer inngående kunnskap om det, for å være bedre forberedt til arbeidslivet. Arbeidet har vært meget interessant og lærerikt, men til tider svært utfordrende. Til tross for mye litteratur rundt kunstig intelligens, er det aller meste skrevet av dataingeniører og andre datakyndige personer, noe som ikke nødvendigvis er lettleselig for en revisor. Likevel gjorde de personene vi intervjuet temaet forståelig fra et økonomisk perspektiv.

Vi ønsker å rette en stor takk til veileder Carmen Olsen for gode og konstruktive innspill i prosessen. Til slutt vil vi også takke alle informantene som tok seg tid til å bli intervjuet i en hektisk periode midt i årsoppgjøret. Vi setter pris på respondentenes engasjement for oppgaven og ønske om å lese denne i etterkant.

Handelshøyskolen ved Universitetet i Stavanger, 17.06.2019

Sammendrag

Revisjonsprosessen er under stor endring, hvor forskning innenfor temaet viser at revisjonsbransjen er moden for å ta i bruk nye teknologiske løsninger. Målet med forskningen er å bidra til å økt forståelse for hva som skjer gjennom å undersøke hvilke verktøy som finnes i Norge, og hvilke konsekvenser økt bruk av kunstig intelligens har for revisjonsprosessen. De største revisjonsselskapene investerer mye tid og penger på å utvikle nye løsninger som skal ta bransjen inn i fremtiden. Ved bruk av kvalitativ metode har vi gjennomført dybdeintervju med seks personer fra de fire største revisjonsselskapene i Norge. Selskapene har utviklet tolv verktøy, hvorav fem er tilnærmet identiske. Disse øker kvaliteten og effektiviteten, men det er også utfordringer knyttet til manglende kompetanse, bias, tillit og revisjonsstandardene.

Løsningene som er utviklet blir i størst grad benyttet i planleggingsfasen, men det er også noen som kan benyttes som substanshandling, både i form av analytiske handlinger og detaljtesting. Revisjonshandlingene som utføres er foreløpig de samme, men måten disse utføres på er endret. Antall manuelle handlinger reduseres som følge av økt automatisering. Dette kan bidra til å oppnå formålet med revisjonen, altså å øke brukernes tillit til at regnskapet er uten vesentlig feil. Kunstig intelligens innenfor revisjonsbransjen er et relativt nytt tema, og vi vil til slutt adressere flere forslag til forskning.

Nøkkelord: Kunstig intelligens; Revisjonsprosessen; Konsekvenser; Endring; Verktøy

INNHOOLD

1	Introduksjon	1
2	Teori	4
2.1	<i>Revisjonsprosessen</i>	4
2.2	<i>Kunstig intelligens</i>	6
2.3	<i>AI innenfor revisjon</i>	9
3	Metode	13
3.1	<i>Forskningsdesign og -metode</i>	13
3.2	<i>Datainnsamling</i>	13
3.3	<i>Respondentene</i>	13
3.4	<i>Intervju</i>	15
3.5	<i>Personvern</i>	15
3.6	<i>Reliabilitet og validitet</i>	15
4	Resultater	17
4.1	<i>Verktøy</i>	17
4.2	<i>Konsekvenser for revisjonsprosessen</i>	25
5	Diskusjon	30
5.1	<i>Hvilke verktøy finnes i dag i revisjonsprosessen som er basert på kunstig intelligens (1)?</i>	30
5.2	<i>Hvilke konsekvenser gir økt bruk av kunstig intelligens for revisjonsprosessen? (2)</i>	34
6	Konklusjon og veien videre	38
	Referanser	39
	Vedlegg: Intervjuguide	42

FIGURER

Figur 1: Revisjonsprosessen	4
Figur 2: Tre kategorier maskinl�ring	9
Figur 3: Analyseverkt�y	19
Figur 4: Databasen med risikoomr�der for et selskap	20
Figur 5: Kartlegging av n�rst�ende parter	21
Figur 6: Oversikt over hvor i revisjonsprosessen AI-verkt�yene passer.....	24
Figur 7: Konsekvenser av AI.....	26
Figur 8: Kartlegging av verkt�y	31
Figur 9: Kategorisering av verkt�yene	32
Figur 10: Fordeler og ulemper knyttet til AI i revisjonsprosessen.....	34

TABELLER

Tabell 1: Oppsummering av planleggingsfasen	5
Tabell 2: Oversikt over respondenter	14
Tabell 3: Oversikt over hvordan revisjonshandlingene endres grunnet AI.....	23

FORKORTELSER

AI – Kunstig Intelligens

ISA – International Standards on Auditing

IAASB – International Auditing and Assurance Standards Board

RPA – Robotic Process Automation

1 INTRODUKSJON

Revisjonsbransjen har tidligere vært avventende når det gjelder implementering av nye teknologiske løsninger, men er moden for større grad av automatisering i tiden fremover. Dette skyldes arbeidsintensiteten til revisor, og at 80% av revisors oppgaver kan være gjenstand for automatisering (Abdolmohammadi, 1999, s. 54-58; Issa et al., 2016, s. 1; Kokina & Davenport, 2017, s. 116). Videre benytter kundene seg av avanserte systemer i større grad enn tidligere, noe som også legger press på at revisor må oppdatere sine verktøy (Alles, 2015, s. 2).

Vårt mål med denne masteroppgaven blir å samle kunnskap om bruk av kunstig intelligens innenfor revisjonspraksis. Det eksisterer store mengder av litteratur om kunstig intelligens generelt, men lite forskning om anvendelsen i revisjon (Issa et al., 2016, s. 4). Problemet er at kunstig intelligens er på et veldig tidlig stadium i revisjon. De fleste vet lite om temaet, og hva som vil skje fremover, noe som også inkluderer revisjonsselskapene og standardsetterne. Eksisterende forskning forsøker i stor grad å strukturere viktige problemstillinger og forskningsspørsmål, og det etterlyses forskning på sentrale områder. Flere av de spørsmålene som foreslås er imidlertid veldig spesifikke, og dermed vanskelige å besvare uten at kunnskap om temaet i større grad samles.

Et område vi opplever at eksisterende forskning har fokusert lite på, er hvor langt de største revisjonsselskapene har kommet i implementeringen av kunstig intelligens. Issa et al. (2016, s. 3) og Kokina og Davenport (2017, s. 119) nevner kort at de fire store – PwC, KPMG, Deloitte og EY - investerer store ressurser, og samarbeider med leverandører som IBM Watson og Kira. Revisjonsselskapene reklamerer i stor utstrekning om at kunstig intelligens kommer til å revolusjonere revisjonsbransjen, men det forklares derimot lite spesifikt hva man ønsker å oppnå, og hva de ulike verktøyene er i stand til. Vårt første forskningsspørsmål er derfor: **1) Hvilke verktøy finnes i revisjonsprosessen basert på kunstig intelligens i dag?**

Vi vil undersøke hvilke verktøy revisjonsselskapene i dag besitter, og bruksområdene for disse, og hvor i den tradisjonelle revisjonsprosessen verktøyene brukes. Her tas det utgangspunkt i en modell fra Arens, Beasley og Elder (2013) som deler prosessen i fire faser. For å se på hvorvidt verktøyene til selskapene baseres på kunstig intelligens (heretter kalt AI) vil vi trekke dem opp mot definisjonen av AI som anvendes i oppgaven.

Kokina og Davenport (2017, s. 120) peker på to utfordringer som krever forskning for å kunne gi veiledning til revisjonsselskapene og standardsetterne i tiden fremover. Den første utfordringen handler om bias i resultatet som produseres av AI. Dette kan forekomme grunnet mangler eller skjevhet i underliggende data, eller gjennom interaksjon mellom menneskene som trener eller benytter seg av maskinen. Den andre utfordringen handler om at AI-systemer som baserer seg på bruk av maskinlæring og dype nevralt nettverk, ofte opererer som «svarte bokser». Dette innebærer at det er vanskelig å forstå hvordan maskinen kommer frem til resultatet den produserer. Kokina og Davenport (2017, s. 120) mener at så lenge det er lite innsyn i arbeidet AI har utført, vil det være vanskelig for standardsettere og revisjonsselskapene å benytte teknologiens muligheter fullt ut. Det andre forskningsspørsmålet ser derfor på konsekvenser ved økt bruk av AI, og hvorvidt fordelene veier opp for utfordringene. Vi vil også se på hvordan revisjonsprosessen vil endres på henholdsvis kort og lang sikt. Det andre forskningsspørsmålet blir dermed: **2) Hvilke konsekvenser gir økt bruk av kunstig intelligens for revisjonsprosessen?**

For å få en dypere forståelse av temaet og for å kunne besvare våre to forskningsspørsmål har vi benyttet kvalitativ metode, med dybdeintervju som datainnsamlingsmetode. Vi tok kontakt med de fire største revisjonsselskapene i Norge, som alle ønsket å delta i undersøkelsen. Vi intervjuet seks personer, der to av selskapene stilte med to representanter. Alle respondentene hadde erfaring med bruk av AI i sitt selskap. Intervjuene ble gjennomført via Skype, da alle respondentene var lokalisert andre steder enn Stavanger. Det ble tatt utgangspunkt i en intervjuguide bestående av åpne spørsmål, for at respondentene skulle ha mulighet til å gi utfyllende svar. Det ble i tillegg benyttet oppfølgingsspørsmål ved behov.

Når det gjelder AI-verktøy som finnes i dag, viser funnene at det er variasjon i utviklingen hos revisjonsselskapene i Norge. Samlet identifiserte vi tolv verktøy. Da flere av selskapene har lignende verktøy med bare noen små ulikheter, ble de kategorisert i syv forskjellige typer basert på bruksområde. Verktøyene assisterer revisor i å 1) kartlegge risiko, 2) identifisere anomaliteter i hovedbok, 3) kategorisere regnskapslinjer, 4) identifisere relasjoner, 5) predikere regnskapstall, 6) skanne kontrakter og 7) automatisk estimere verdien på eiendeler. Dette utføres uten, eller med lav grad av menneskelig interaksjon. Satt opp mot revisjonsprosessen er det i planleggingsfasen AI benyttes i størst utstrekning i dag, særlig gjennom innledende analytiske handlinger, og vurdering av risiko. Til tross for dette er det tendenser til bruk også videre i revisjonsprosessen.

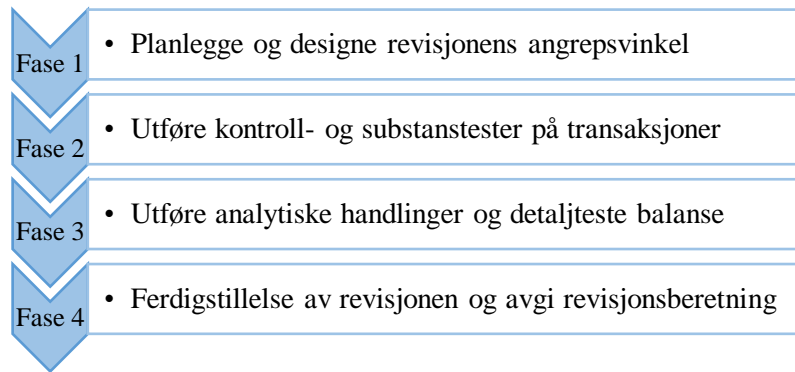
Våre funn viser at de positive konsekvensene ved økt bruk av AI i revisjonsprosessen, først og fremst er forhøyet kvalitet og effektivitet, der AI enten er i stand til å utføre en oppgave bedre enn et menneske, eller like godt, men mer effektivt. Dette oppnås gjennom at maskinen har mulighet til å analysere 100% av klientens transaksjoner på en slik måte at tradisjonelle menneskelige feil unngås. Likevel er det noen utfordringer knyttet til bruken som forsinker utviklingen. Det er manglende forståelse for AI blant revisjonsselskapene, kundene og tilsynsorganene. Dette skyldes at AI innenfor revisjon er på et tidlig stadium, og partene dermed har en begrenset kompetanse. Som en konsekvens av manglende forståelse, er det utfordringer rundt tilliten til verktøyene. Revisjonsmedarbeiderne kan ha for stor tillit, samtidig som ansvarlig partner har mindre tiltro til resultatet AI fremlegger. Manglende tillit vises også igjen i revisjonsstandardene, der funnene viser at standardene ikke nødvendigvis legger til rette for ny teknologi. Som en følge av disse utfordringene, tyder funnene på at det ikke vil bli noen store endringer for revisjonsprosessen på kort sikt. Flertallet mener likevel at revisjonsprosessen vil bli mer dynamisk på lengre sikt, der revisjonen mer eller mindre blir én kontinuerlig prosess. Dette skyldes at eksempelvis planlegging og testing vil utføres samtidig, noe som kan føre til at de fire fasene i revisjonsprosessen viskes ut.

Oppgaven starter med å presentere relevant teori sett opp mot forskningsspørsmålene. Dette inkluderer teori om revisjonsprosessen og AI, herunder bruk av AI i revisjon. Fremgangsmåte for datainnsamling etterfølges av en gjennomgående presentasjon av resultatene. Deretter diskuteres funnene i lys av forskningsspørsmålene og teori. Til slutt presenteres en konklusjon som besvarer oppgavens to forskningsspørsmål, og gir forslag til videre forskning.

2 TEORI

2.1 Revisjonsprosessen

I denne delen vil prosessen som revisor gjennomgår ved utførelsen av et revisjonsoppdrag forklares. Vi har valgt å ta utgangspunkt i den revisjonsprosessen som beskrives av Arens et al. (2013, s. 183). Her deles prosessen inn i fire faser, som vist i figur 1.



Figur 1: Revisjonsprosessen

Fase 1 – Planlegge og designe revisjonens angrepsvinkel

Målet med denne fasen av prosessen er å planlegge revisjonen på en slik måte, at revisor blir i stand til å trekke en rimelig konklusjon hvorvidt regnskapet er uten vesentlig feil. For at revisor skal klare å innhente tilstrekkelig og hensiktsmessig revisjonsbevis, og samtidig minimere kostnaden ved innhenting av bevis, er det nødvendig med en grundig planlegging. En kritisk del av planleggingsfasen er å opparbeide seg en forståelse av 1) klientens virksomhet og bransje, og 2) klientens interne kontroll (Arens et al., 2013, s. 183). En forståelse av klientens virksomhet og bransje vil bidra med å kartlegge selskapets iboende risiko. Dette gjøres ved å vurdere blant annet bransjerelaterte faktorer, nærstående parter, resultatet av tidligere revisjoner, mislighetsrisikofaktorer og i hvilken grad selskapet har kompliserte og/eller uvanlige transaksjoner (Arens et al., 2013, s. 284-286). For å få denne forståelsen gjennomfører revisor 1) forespørslar, 2) observasjon og inspeksjon, og 3) analytiske handlinger (ISA 315.6). Sistnevnte bidrar med å identifisere uvanlige mønstre i bokføringen sammenlignet med for eksempel tidligere år eller et bransjegjennomsnitt. Slike uvanligheter kan indikere at det foreligger forhold som har betydning for revisjonen (Arens et al., 2013, s. 240). Forståelsen av klientens virksomhet og bransje, og klientens interne kontroll brukes for å vurdere risikoen for vesentlig feil i regnskapet. Denne vurderingen brukes videre for å utforme en detaljert revisjonsplan og bestemme tidspunkt, type og omfang av revisjonshandlinger som utføres for å redusere revisjonsrisikoen til et akseptabelt lavt nivå (Arens et al., 2013, s. 184 og 279).

<p>Aksept av klient og utførelse av innledende planlegging (Arens et al., 2013, s. 229-233)</p> <p>Forstå selskapets virksomhet og omgivelser (Arens et al., 2013, s. 234-238)</p>	<ul style="list-style-type: none"> • Revisor kartlegger virksomheten til potensiell klient, finansiell stilling og erfaring med tidligere revisor. • Revisor vurderer egen kompetanse, risiko på oppdraget, nødvendige ressurser og uavhengighet. • Revisor opparbeider seg en bransjeforståelse. • Revisor opparbeider seg en forståelse av klientens viktigste inntektskilder, finansieringskilder, nærstående parter og eier- og ledelsesstruktur. • Revisor opparbeider seg en forståelse av selskapets strategi og mål samt hvordan selskapet måler og belønner ansatte og ledelsen.
<p>Vurder selskapets forretningsrisiko (Arens et al., 2013, s. 238-240)</p>	<ul style="list-style-type: none"> • Revisor anvender den samlede forståelsen fra foregående steg til å vurdere risikoen for at selskapet mislykkes i å oppnå sine mål. • Revisor vurderer hvorvidt klienten har etablert kontroller som kan redusere forretningsrisikoen.
<p>Utfør innledende analytiske handlinger (Arens et al., 2013, s. 240-242)</p>	<ul style="list-style-type: none"> • Revisor ser etter uvanligheter for å identifisere områder som er beheftet med særskilt risiko, og som dermed krever ytterligere oppmerksomhet.
<p>Bestem vesentlighetsgrenser og vurder akseptabel revisjonsrisiko og iboende risiko (Arens et al., 2013, s. 268-277)</p>	<ul style="list-style-type: none"> • Revisor bruker forståelsen som er opparbeidet til å fastsette totalvesentlighet og arbeidsvesentlighet ved hjelp av profesjonelt skjønn. • Revisor vurderer type og omfang av revisjonsbevis som er nødvendig ved anvendelse av revisjonsrisikomodellen.
<p>Vurder internkontroll og kontrollrisiko (Arens et al., 2013, s. 307-330)</p>	<ul style="list-style-type: none"> • Revisor opparbeider seg en forståelse av selskapets interne kontroller og gjør seg opp en forventning hvorvidt kontrollene er i stand til å forhindre vesentlig feil i regnskapet (vurderer kontrollrisiko).
<p>Vurder risikoen for misligheter (Arens et al., 2013, s. 354-359)</p>	<ul style="list-style-type: none"> • Revisor anvender profesjonelt skjønn for å identifisere risiko for misligheter som kan forårsake vesentlig feil i regnskapet enten som følge av uredelig regnskapsrapportering eller underslag av eiendeler.
<p>Utvikle overordnet revisjonsstrategi og revisjonsplan (Arens et al., 2013, s. 422-423) + ISA 300 punkt 7-9.</p>	<ul style="list-style-type: none"> • Revisor utvikler overordnet revisjonsstrategi som beskriver revisjonens innhold, når den skal utføres, samt angrepsvinkel. • Revisor utvikler revisjonsplan som fastsetter tidspunkt, type, og omfang av planlagte risikovurderingshandlingene samt videre revisjonshandlinger.

Tabell 1: Oppsummering av planleggingsfasen

Fase 2 – Utføre kontroll- og substanstester på transaksjoner

Revisor kan velge å bygge deler av revisjonen på test av kontroller, dersom kontrollrisikoen i planleggingen blir anslått for å være lav eller moderat. Dette vil redusere mengden substanshandlinger, som ofte vil være mer ressurskrevende (Arens et al., 2013, s. 308).

Uavhengig av om man bygger deler av revisjonen på test av kontroller, må det gjennomføres noe substanstesting. Dette skyldes blant annet at det er iboende begrensninger ved de interne kontrollene, for eksempel en mulighet for ledelsens overstyring. Med substanstesting menes handlinger som er utformet for å avdekke vesentlig feilinformasjon på påstandsnivå (ISA 330.4). Av effektivitetshensyn gjennomføres ofte kontroll- og substanstesting samtidig (Arens et al., 2013, s. 184).

Fase 3 – Utføre analytiske handlinger og detaljteste av balanse

Analytiske handlinger evaluerer finansiell informasjon ved å sammenligne den mot blant annet historiske tall og bransjegjennomsnitt. Et eksempel kan være bruttofortjenesteanalyse, hvor man sammenligner den månedlige bruttofortjenesten med tidligere perioder. For at analysen skal kunne benyttes som revisjonsbevis, er det avgjørende at revisor gjør seg opp en forventning av hva eksempelvis bruttofortjenesten bør være. Dette gjøres med bakgrunn i blant annet tidligere perioder, bransjetrender og kundens budsjett (Arens et al., 2013, s. 243). Detaljtesting av balanse utføres for å kunne konkludere om utgående balanse inneholder tallmessig feil av vesentlig karakter. Dette gjøres ofte ved å innhente eksterne forespørsler fra eksempelvis kunder og leverandører, noe som øker troverdigheten som revisjonsbevis (Arens et al., 2013, s. 184).

Fase 4 – Ferdigstillelse av revisjon og avgi revisjonsberetning

Revisor vil til slutt anvende alle innhentede revisjonsbevis for å konkludere hvorvidt regnskapet er uten vesentlig feil. Avgjørelsen om det er innhentet nok bevis er i høy grad en subjektiv vurdering, og avhenger av revisors profesjonelle skjønn. Dersom det konkluderes med at regnskapet er uten vesentlig feil, utstedes en ren beretning (Arens et al., 2013, s. 184).

2.2 Kunstig intelligens

Her vil det bli gitt en generell introduksjon til AI, før 2.4 omhandler mer spesifikt AI innenfor revisjon. For å forklare konseptet vil også noen av de teknologiene som muliggjør AI forklares, herunder maskinlæring og nevrale nettverk.

2.2.1 Definisjon

Til tross for økende popularitet og store mengder litteratur rundt AI, finnes det ingen entydig definisjon. I de mange forsøkene på å finne en definisjon, tas det normalt utgangspunkt i menneskelig intelligens (Kaplan & Haenlein, 2019, s. 17), som kan defineres som «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» ((Gottfredson, 1997, s. 13) egen oversettelse).

I 1950 forsøkte Alan Turing å definere en standard for hva som skal til for at en maskin skal kunne kalles intelligent. Denne standarden, som senere omtales som Turing-testen, går ut på at en datamaskin består, dersom et menneske, etter å ha kommunisert skriftlig med en samtalepartner, ikke greier å skille om partneren er en datamaskin eller et menneske. Etter at Turing definerte standarden, er det imidlertid få forskere som har forsøkt å skape systemer som kan passere testen. Dette skyldes at det oppleves som mer viktig å studere prinsippene bak intelligens, snarere enn å lage et system som imiterer menneskelig intelligens (Tørresen, 2013, s. 16-17). Det er dette som ligger til grunn når Kaplan og Haenlein (2019, s. 17) definerer kunstig intelligens som **«et systems evne til å tolke eksterne data riktig, å lære fra disse dataene, og å bruke lærdommen for å oppnå spesifikke mål og oppgaver gjennom fleksibel tilpasning»** (egen oversettelse).

I følge Burgess (2018, s. 30-31) er det tre målsettinger med AI: 1) innhente informasjon, 2) bestemme hva som skjer, og 3) forstå hvorfor det skjer. Målsettingen om å innhente informasjon handler for det første om at AI er i stand til å transformere ustrukturert informasjon (for eksempel et bilde av et ansikt) til strukturert informasjon (navnet til personen). For det andre greier AI å finne mønster i store datasett av strukturert informasjon. Til tross for at den menneskelige hjernen vanligvis er god til å finne mønster, greier den det ikke når datasettet består av hundrevis av variabler og millioner av datapunkter. Denne svakheten er det Simon (1979, s. 502) kaller for begrenset rasjonalitet, som handler om at mennesket er begrenset av den informasjonen man har tilgjengelig, hjernens kognitive kapabiliteter og tiden som er tilgjengelig. Den omtalte svakheten har ikke AI, og den greier derfor på en effektiv måte å gjøre den ustrukturerte informasjonen eller de store datasettene om til noe nyttig. Den neste målsettingen AI søker å oppnå er å bestemme hva som skjer. Informasjon har her vanligvis allerede blitt fanget av AI. Man har gjerne brukt talegjenkjenning for å gjøre om en lydfil til ord, men på dette stadiet vet man bare ordene ikke

betydningen. For å forstå innholdet i lydfilen kan språkbehandlingsverktøyer brukes. Den siste målsettingen med AI, nemlig å forstå hvorfor det skjer, er ikke mulig i dag. Dette betyr at selv om AI er i stand til å identifisere avvik så forstår ikke maskinen hvorfor akkurat dette utgjør et avvik.

2.2.2 Typer AI

I litteraturen skilles det gjerne mellom smal og generell AI. Vi har flere eksempler på AI i dag. Facebook er i stand til å gjenkjenne ansikter på bilder, Iphone's Siri er i stand til å forstå stemmer og handle deretter, og Tesla utvikler selvkjørende biler. Alt dette er eksempler på smal AI, ettersom maskinen kun er i stand til å utføre spesifikke oppgaver, like godt eller bedre enn mennesker. Målet til forskerne innenfor fagfeltet, er å utvikle generell AI.

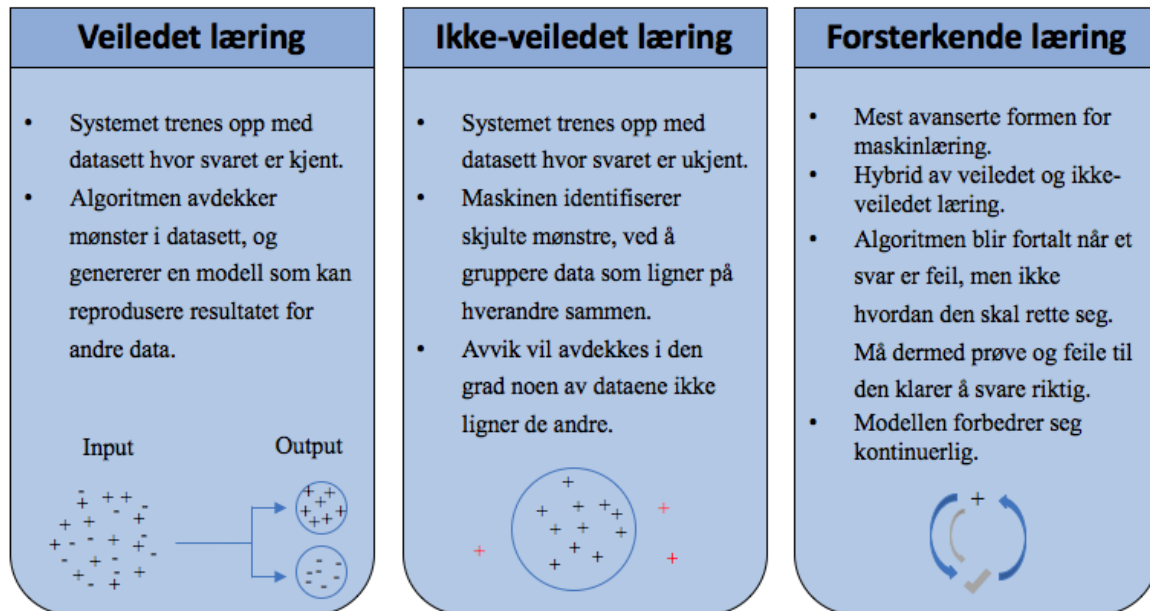
Maskinen vil da være i stand til å resonnerer, planlegge og finne løsninger på problemer den i utgangspunktet aldri var designet for å løse (Kaplan & Haenlein, 2019, s. 16). På grunn av manglende evne til å forstå hvorfor noe skjer, er dette foreløpig ikke mulig. Til tross for at det er sannsynlig at vi kommer til dette stadiet en dag, er det vanskelig å si noe om når (Burgess, 2018, s. 4).

Maskinlæring

Fra definisjonen av AI er det sentralt at maskinen er i stand til å lære. Dette gjøres ved hjelp av maskinlæring, som kan beskrives som metoder som gir datamaskiner evnen til å lære uten å bli spesifikt programmert (Samuel, 1959, s. 211). Den menneskelige utvikleren definerer algoritmen(e), før maskinen finner en løsning på problemet. For at maskinen skal bli lært opp til å utøve kunstig intelligens, trenger den normalt store mengder data. Dette skyldes at systemet trenger erfaring for å lære, noe som kommer ved at man tilfører informasjon. Ved bruk av statistisk modellering lærer maskinen å avdekke mønstre fra informasjonen den gis. Det genereres så en modell som er i stand til å gjenkjenne mønstrene når nye data blir prosessert av modellen. Dette betyr at når modellen blir tilført nye data, vil den basert på treningsdataene avgjøre hvilket mønster de nye dataene ligner mest på. På bakgrunn av dette kommer modellen med et estimert resultat (Datatilsynet, 2018, s. 6).

Det finnes i hovedsak tre kategorier maskinlæring, se figur 2 (Marsland, 2015, s. 6; Theobald, 2017, s. 15-19). Veiledet læring er den mest brukte formen for maskinlæring, og et eksempel på anvendelsesområde er prediksjon av markedspris for en bruktbil. En algoritme basert på veiledet læring vil ta utgangspunkt i den kjente prisen og jobbe bakover for å finne sammenhengen mellom pris og ulike variabler som årsmodell, merke og kilometerstand.

Mønsteret algoritmen finner i datasettet, brukes for å generere modellen som kan brukes for å predikere markedsprisen for en annen brukbil, ved å bruke de samme underliggende reglene som ble identifisert i testsettet (Burgess, 2018, s. 7; Theobald, 2017, s. 16).



Figur 2: Tre kategorier maskinlæring

Kunstige nevralt nettverk og dyp læring

Dyp læring er en underkategori av maskinlæring, som baserer seg på bruk av kunstige nevralt nettverk. Sistnevnte kan beskrives som maskinlæringsalgoritmer som forsøker å simulere beslutningsprosessen i det biologiske nervesystemet. Til forskjell fra maskinlæring, er det under dyp læring ikke nødvendig med menneskelig interaksjon, bortsett fra de fasitsvarene maskinen gis under treningen (om det benyttes veiledet læring). Dette fører til at hele læringsprosessen kan automatiseres (Telle, 2017, s. 194). Issa et al. (2016, s. 4-5) forklarer at dyp læring er spesielt relevant for å være i stand til å analysere ustrukturerte og semistrukturerte data, noe som er tilfellet ved eksempelvis gjennomgang av kontrakter og andre dokumenter. En utfordring ved bruk av nevralt nettverk er at datamengden som kreves for å trene opp systemet normalt er stor. En annen utfordring er at systemet i stor grad opererer som en «svart boks», i den forstand at man ikke nødvendigvis vet hvordan maskinen kommer frem til resultatet (Theobald, 2017, s. 80)

2.3 AI innenfor revisjon

I følge Kokina og Davenport (2017, s. 116) er revisjonsbransjen moden for bruk av Big Data Analytics og AI. Dette begrunnes med at det er utfordrende for revisor å benytte seg av all

den informasjonen som er tilgjengelig, for å få et innblikk i klientens finansielle og ikke-finansielle prestasjon. Selskaper genererer og samler inn store mengder data kontinuerlig. Når databasene består av så mye informasjon vil ikke tradisjonelle revisjonsmetoder lenger være effektive. Dette skyldes at mennesket ikke greier å utføre komplekse oppgaver som krever at man bruker mye informasjon fra ulike kilder, da man ofte vil overse relevant informasjon og mønstre (Issa et al., 2016, s. 9)

Oppgavene som søkes automatisert er først og fremst de mest arbeidsintensive, de som krever relativt lav kompetanse for å utføre (Rapport, 2016). Dette er oppgaver som er strukturerte og repeterbare, noe som gjør at de egner seg godt for bruk av AI (Kokina & Davenport, 2017, s. 116). I en studie av Abdolmohammadi (1999, s. 54-58) identifiseres 332 revisjonsoppgaver innenfor seks faser. Han kartlegger i hvilken grad hver av disse oppgavene er strukturerte, kompleksiteten til oppgavene, og om noen er egnet for beslutningsstøtte. Med strukturerte oppgaver menes de som har få alternative løsninger, og som dermed krever lite skjønn. Han finner at de fleste oppgavene i løpet av revisjonsprosessen er strukturerte eller semistrukturerte, henholdsvis 39 og 41%, mens kun 20% av oppgavene er ustrukturerte. Videre er det i substanstesting det er mest oppgaver som er strukturerte, med 67%. Det er også substanstestingsfasen som er mest egnet for beslutningsstøtte, til tross for at resultatene var for svake til noen generalisering. Basert på funnene fra Abdolmohammadi (1999) og påstanden fra Kokina og Davenport (2017) om at AI kan automatisere strukturerte oppgaver, kan det utledes at cirka 20% av de 332 revisjonsoppgavene ikke kan automatiseres. De resterende 80% kan dermed være gjenstand for automatisering og derfor effektiviseres ved bruk av AI.

Davenport og Kirby (2016, s. 21-22) har utviklet et rammeverk, hvor de kategoriserer ulike typer oppgaver og ser på i hvilken grad maskiner utøver intelligens for å løse oppgavene. Utgangspunktet har lenge vært at maskiner kun kan hjelpe mennesker med å ta beslutninger, men det er en relativt kort vei fra dette, til at maskiner selv kan ta beslutninger. Dette gjør seg særlig gjeldende i den grad det er snakk om strukturerte og repeterbare beslutninger. På dette nivået av intelligens vil maskinen støtte seg på regler, og klart definerte algoritmer, der prestasjonen ikke forbedres uten at et menneske justerer algoritmene. I dag har noen av de mest avanserte teknologiene evnen til å lære selv, og gjennom dette forbedre sin prestasjon. Basert på definisjonen som anvendes i denne oppgaven, er det ikke snakk om AI før maskinene er i stand til å utøve denne type intelligens.

Kokina og Davenport (2017, s. 117) bruker rammeverket i en regnskaps- og revisjonssammenheng, hvor de ser på hvorvidt oppgavetyperne numerisk analyse, analyse av ord og bilder, utførelse av henholdsvis digitale og fysiske oppgaver er aktuelle i denne sammenheng. Det er liten tvil om at førstnevnte er den mest relevante, som Davenport og Kirby (2016, s. 22) forklarer er maskiners evne til å analysere tall i strukturerte formater. Kokina og Davenport (2017, s. 117) påpeker at dette har blitt gjort over lengre tid, men det er først i senere år at de største selskapene har begynt å bruke maskinlæring for formålet. I revisjon er dette spesielt relevant for å identifisere anomalier i store datasett, som for eksempel at det bokføres uvanlig høy salgsinntekt mot slutten av en periode, eller store kostnader i begynnelsen av en periode. Funnene skaper grunnlag for revisors videre undersøkelser.

Neste oppgavetype er analyse av ord og bilder. En viktig del av menneskets kognisjon er å se ord og bilder, og forstå betydningen av disse (Davenport & Kirby, 2016, s. 22). Dette vil AI også kunne gjøre ved at maskinlæring og kunstige nevralt nettverk anvender det som gjerne kalles naturlig språkprosessering. I revisjonssammenheng vil maskinen kunne analysere og uthente den viktigste informasjonen fra blant annet salgsdokumenter, leasingavtaler, fakturaer og styremøtereferater (Agnew, 2016; Issa et al., 2016, s. 10). Issa et al. (2016, s. 10) argumenterer videre for at revisor kan benytte seg av slike verktøy, for å analysere poster på ulike forum samt kommentarer og samtaler på sosiale medier som kan tjene som supplerende revisjonsbevis. Med tiden kan det også bli mulig for maskinen å automatisk generere revisjonsberetningen. Når det gjelder analyse av bilder er dette lite utbredt i dag, men et eksempel som kan bli aktuelt er å erstatte revisors deltakelse på varetelling med droner (Kokina & Davenport, 2017, s. 118).

De siste to kategoriene er utførelse av henholdsvis digitale og fysiske oppgaver. Disse oppgavetyperne er mer avanserte enn de to foregående, da maskinen her må ha kunnskap til å ta informerte beslutninger, til forskjell fra å kun prosessere tall, ord eller bilder (Davenport & Kirby, 2016, s. 22). Når det gjelder utførelse av fysiske oppgaver er dette spesielt oppgaver som kan utføres av roboter. Foreløpig er ikke dette så relevant for revisjonsbransjen, men det kan være relevant ved eksempelvis varetelling, hvor det kan tenkes at revisor slipper å fysisk møte opp (Kokina & Davenport, 2017, s. 117). Utførelse av digitale oppgaver kan anses som en restkategori der maskinen skal ta beslutninger, hvor det ikke er behov for en robot. Eksempler på oppgaver i denne kategorien er klassifisering av selskapets kontoer på ulike

regnskapslinjer samt besvarelse av rutinehenvendelser fra kunder (Baldwin, Brown & Trinkle, 2006, s. 80; Davenport & Kirby, 2016, s. 23). Davenport og Kirby (2016, s. 21-23) viser til at intelligensen som utøves for denne kategorien på nåværende tidspunkt, er slik at systemet ikke lærer uten at et menneske justerer algoritmene. Dette gjør at definisjonen som presenteres i 2.3.1 ikke kan anses oppfylt. Likevel gjør nåværende intelligens det mulig med en automatisering av strukturerte oppgaver, ved hjelp av eksempelvis Robotic Process Automation (RPA) (Kokina & Davenport, 2017, s. 118-119).

Det er mange utfordringer med implementering av AI innenfor revisjonsbransjen. En utfordring er at implementeringen kan være kostnadskreven. Kinserdal (2017) påpeker at fremtidens revisor trenger en ny type kompetanse, samtidig som revisjonsmetodikken må oppdateres, for å sikre at de nye metodene gir like sikre konklusjoner og revisjonsbevis som tidligere. En stor andel av globale revisjonsselskaps verktøy utvikles dessuten internasjonalt. Det vil eksempelvis være problematisk å tilpasse AI-løsninger som kommer fra USA til norske forhold. Dette skyldes blant annet at revisjonsplikten i Norge strekker seg videre enn i USA, hvor revisjon er frivillig for private selskaper (Langli, 2015, s. 143), noe som gjør at verktøyene er tilpasset større og andre segmenter av kunder, enn hva som er tilfellet i Norge.

En annen utfordring er at bransjen er standarddreven. På den ene siden er det vanskelig for revisjonsselskapene å inkorporere ny teknologi når dette ikke kreves, eller har aksept fra de standardsettende organer (Issa et al., 2016, s. 9). På den annen side, er det vanskelig for standardsettere og tilsynsmyndigheter å godkjenne ny metodikk før denne er oppfunnet og testet (Kinserdal, 2017). Dette skaper en vanskelig situasjon, som kan hindre utviklingen av nye verktøy, og gi en mer målrettet og effektiv revisjon. IAASB (2018, s. 9), organet som utgir revisjonsstandardene (ISA-ene), startet i 2016 et prosjekt hvor de skulle undersøke hvorvidt revisjonsstandardene fortsatt er passende for den nye teknologien. Her kommer de frem til, at selv om standardene ikke oppfordrer til bruk av nye teknologiske løsninger, så hindrer de det heller ikke. Det er likevel flere som mener at standardene må oppdateres for å tillate og oppfordre bransjen til å utnytte de mulighetene som foreligger ved hjelp av AI (se blant annet (Bizarro & Dorian, 2017; Issa et al., 2016)). Dette begrunnes med at revisjonsstandardene er basert på tradisjonelle revisjonsmetoder, og de er skrevet på en tid der databasene var små, og utfordringene som finnes i dag ikke eksisterte (Issa et al., 2016, s. 9).

3 METODE

3.1 Forskningsdesign og -metode

Med forskningsdesign menes en plan for hvordan problemstillingen skal besvares. Designet vil variere med problemstilling, tiden og ressursene som er tilgjengelige for forskningsprosjektet. Forskningsdesignet kan enten være eksplorativt, deskriptivt eller kausalt. Da vi besitter lite kunnskap om AI innenfor revisjon på forhånd, og det eksisterer lite forskning rundt temaet, passer et eksplorativt design. Et slikt design brukes vanligvis når problemstillingen er uklar, man har begrensede forkunnskaper, og der det er vanskelig å formulere klare hypoteser (Grenness, 2001, s. 75-76). I følge Sekaran og Bougie (2016, s. 43) vil et eksplorativt forskningsdesign ofte medføre en kvalitativ forskningsmetode, noe som er tilfellet i denne masteroppgaven. Med en kvalitativ metode brukes data i form av ord som for eksempel intervjunotater og transkribering av lydopptak. Dette til forskjell fra kvantitativ metode som bygger på data i form av tall (Sekaran & Bougie, 2016, s. 2).

3.2 Datainnsamling

For å sikre en tilstrekkelig dyp forståelse av temaet, ble det utført semistrukturert dybdeintervju som datainnsamlingsmetode. Denne formen for intervju skiller seg ikke mye fra ustrukturerte intervju, hvor målet er å få respondentene til å uttrykke seg med egne ord om et tema. Forskjellen mellom de to intervjutypene er at man ved semistrukturerte intervju følger en intervjuguide som er en liste over spørsmål og temaer som ønskes dekket. Dette hindrer ikke forskerne fra å be respondentene utdype interessante eller uventede svar, men er med på å sikre en høyere grad av reliabilitet (Grenness, 2001, s. 144). Intervjuene ble gjennomført via Skype da respondentene befant seg på andre lokasjoner. Fordelen med et slikt intervju er at respondentene kan føle at det er enklere å utfolde seg, fremfor ved personlig intervju. Den største ulempen er at forskerne ikke får mulighet til å erfare kroppsspråket til respondentene ved besvarelse av spørsmålene (Sekaran & Bougie, 2016, s. 120).

3.3 Respondentene

Å velge ut riktige respondenter i kvalitative undersøkelser er viktig for å få mest mulig nyttig informasjon. Utfordringen ved bruk av intervju er å bestemme hvor mange, og hvem som skal intervjues. Med hensyn til mengden, så kan det hevdes at en bør intervju så mange at en til slutt ikke får ny informasjon (Johannessen, 2010, s. 104). I en masteroppgave hvor vi har begrenset med tid og ressurser, vil dette være vanskelig. Når det gjelder hvem som skal intervjues, så ønsker man ved kvalitative undersøkelser vanligvis å få mest mulig kunnskap

om fenomenet, fremfor statistisk generalisering, som er vanlig ved kvantitative undersøkelser. Dette gjør at respondentene velges strategisk ved kvalitative undersøkelser, fremfor tilfeldig som er vanlig ved kvantitative (Johannessen, 2010, s. 106).

Kandidater som har vært aktuelle i denne masteroppgaven er personer med IT-kompetanse innenfor revisjonsbransjen. En optimal kandidat for vår oppgave er en person med lang erfaring i revisjon, kjennskap til AI og som samtidig kjenner godt til sitt eget selskaps involvering i utviklingen. Utvalgsstrategien vår har derfor vært å finne personer som kunne tilfredsstillende disse kriteriene. Vi tok kontakt med de fire største revisjonsselskapene i Norge – EY, PwC, Deloitte og KPMG – med invitasjon til deltakelse i datainnsamlingen. Det ble etterspurt kandidater som kunne være i stand til å besvare spørsmål om AI innenfor revisjon. Disse fire selskapene ble valgt på bakgrunn av at alle er store aktører med gode forutsetninger for å investere i moderne teknologi. Selskapene opplyser om at det kun er omtrent fem personer i hver av de «fire store», som arbeider med AI-løsninger innenfor revisjon i Norge. Populasjonen av sakkyndige personer innenfor fagfeltet har derfor vært begrenset. Samtidig må det antas at de innhentede svar er representative i forhold til totalpopulasjonen når denne er liten, noe som styrker oppgaven.

Selskap	Respondent	Stilling	Erfaring i revisjon	Tidsbruk
A	A1	Direktør	0 år	60 minutter
	A2	Senior Manager	9 år	
B	B1	Direktør	0 år	70 minutter
	B2	Partner	15 år	
C	C1	Manager	5 år	65 minutter
D	D1	Associate	2 år	70 minutter

Tabell 2: Oversikt over respondenter

Innledningsvis i intervjuet stilte vi respondentene spørsmål om posisjon i selskapet samt erfaring innenfor henholdsvis revisjon og AI. Respondentene hadde ulik erfaring både når det gjaldt revisjon og AI, der to ikke har arbeidet innenfor revisjon. Disse to ble involvert på grunn av høy kompetanse innenfor digitalisering, der den ene leder sitt selskaps utvikling innen AI. De andre respondentene har mellom 2 og 15 års erfaring innenfor revisjon, og alle er involvert i utvikling av nye digitale løsninger, som dataanalyse og AI. Variasjonen i respondentenes erfaring mener vi er en styrke for oppgaven, da dette gir muligheter for ulike svar, alt etter hvilken bakgrunn de har.

3.4 Intervju

Intervjuene ble gjennomført ved å ta utgangspunkt i en intervjuguide bestående av elleve spesifikke spørsmål og fem generelle spørsmål. Målet med de elleve spesifikke spørsmålene var å innhente informasjon for å besvare oppgavens to forskningsspørsmål. Intervjuet ble strukturert i to temaer, der det første omhandlet syv spørsmål knyttet til selskapenes AI-verktøy. Det andre temaet omhandlet fire spørsmål knyttet til revisjonsprosessen, og hvordan den tradisjonelle revisjonsprosessen vil endres som følge av AI. De to temaene ble innledet med teori om hvordan vi definerer AI, og hvordan vi deler revisjonsprosessen inn i fire faser. Dette ble gjort for at respondentene skulle ha samme forståelse av begrepene og dermed samme utgangspunkt. Målet med de generelle spørsmålene var å kartlegge tre ting. For det første ønsket vi å få en forståelse av respondentenes bakgrunn og erfaring innenfor revisjon og AI. For det andre ønsket vi å høre deres tanker om hvordan utviklingen foreløpig har vært, og til slutt ønsket vi å få en forståelse av motivasjonen for utvikling av verktøy innenfor AI. I den grad det var behov for oppklaring av respondentenes svar, ble det stilt oppfølgingsspørsmål. Respondentene ble også kontaktet etter det opprinnelige intervjuet for å supplere med ytterligere informasjon der det var behov.

3.5 Personvern

Ved å bruke intervju som datainnsamlingsmetode behandles personopplysninger, noe som gjør det nødvendig å melde prosjektet til Norsk senter for forskningsdata (NSD). Prosjektet ble derfor meldt inn og godkjent av NSD. Intervjuene ble tatt opp ved hjelp av en elektronisk lydopptaker, noe respondentene ble informert om og godkjente. Det ble avtalt med respondentene at lydopptaket vil bli slettet når prosjektet er levert og evaluert. Intervjuobjektene og selskapene ble anonymisert, noe alle ble informert om på forhånd.

3.6 Reliabilitet og validitet

I forskningssammenheng opereres det med kriteriene, reliabilitet og validitet, for å vurdere undersøkelsesresultatene kvalitet (Grenness, 2001, s. 93). Med reliabilitet menes i hvilken grad man kan stole på at resultatene er pålitelige. Undersøkelsen vil være reliabel dersom man får de samme resultatene ved å gjenta den med samme eller ny metode (Gripsrud, 2004, s. 73).

Det er flere faktorer som er i stand til å true reliabiliteten i studien. For det første benyttes semistrukturert intervju, hvor samtalen til en viss grad varierer mellom hvert intervju. Dette

gjør det vanskelig å oppnå de samme resultatene om intervjuet eller hele studien skulle gjentas. For det andre holder selskapene vi har intervjuet fortsatt på med implementeringen av AI, hvor utviklingen skjer raskt. Dersom noen skulle undersøke det samme på et senere tidspunkt er det derfor sannsynlig at verktøyene og revisjonsprosessen allerede har endret seg. For det tredje kan vi som har utført studien påvirket respondentene under intervjuene. Trolig er det ingen med den samme kombinerte bakgrunnen og personlighetstrekkene, noe som kan gjøre det nesten umulig for andre å reprodusere resultatet. Selv om reliabiliteten må anses truet av disse faktorene, er den forsøkt styrket på flere måter. Bakgrunnen, formålet og den teoretiske forankringen til undersøkelsen er forsøkt beskrevet utførlig. Utvalgsmetode og metode for datainnsamling er beskrevet på en slik måte at nye forskere kan velge samme metode om ønskelig. Det at vi har vært to forskere gjør at vi har hatt en bedre kontroll over vilkårlighet og partisk subjektivitet i intervjuanalysen. Til slutt ligger intervjuguide vedlagt, noe som må anses å øke sannsynligheten for å reprodusere det samme resultatet.

Reliabilitet er en betingelse, men ikke tilstrekkelig for validitet. Med validitet menes det i hvilken grad resultatene 1) representerer dataene som er innsamlet (intern validitet) og 2) kan generaliseres (ekstern validitet) (Sekaran & Bougie, 2016, s. 349). Den interne validiteten kan være truet av at respondentene på noen av spørsmålene svarer ulikt, og det kan derfor være en risiko for at alle meninger ikke kommer frem. Vi har likevel forsøkt å få frem ulike meninger, slik at resultatet i størst mulig grad representerer de innhentede dataene. Dette viste seg å føre til at innhentet teori til en viss grad utfordres. Evnen til å inkludere slike resultater er med på å øke den interne validiteten (Sekaran & Bougie, 2016, s. 349). Når det gjelder ekstern validitet, så vil det være vanskelig å generalisere resultatet, med tanke på at vi har snakket med kun seks respondenter. Samtidig har vi snakket med det vi mener er relevante revisjonsselskaper, og vi har snakket med personer som har lang erfaring innenfor både AI og revisjon. Dette er med på å styrke evnen til generalisering.

Til tross for begrensningene nevnt ovenfor, egner intervjumetoden seg best for eksplorative spørsmål som stilles i denne oppgaven. Metoden gir mulighet for å innhente dybdeinformasjon om fenomenet AI som befinner seg i en implementeringsfase i revisjon. En stor styrke er å kunne kontakte respondentene på ny for å avklare områder som var uklare, og stille oppfølgingsspørsmål.

4 RESULTATER

Funnene fra intervjuene struktureres etter de to forskningsspørsmålene for oppgaven. I 4.1 presenteres resultatene fra spørsmålene som omhandler hvilke verktøy revisjonsselskapene har som er basert på AI. I 4.2 presenteres resultatene som omhandler hvilke konsekvenser økt bruk av AI har for revisjonsprosessen.

4.1 Verktøy

Som en innledning til de mer spesifikke spørsmålene spurte vi respondentene om hvordan veien frem til nå har vært med tanke på utvikling og bruk av AI. Respondentene er samstemte om at bruk av AI innenfor revisjon er på et tidlig stadium, og at det er innenfor analyse det foreløpig er størst progresjon. A2, C1 og D1 påpeker at dersom revisjonsselskapene skal kunne benytte seg av AI så må kunden være klar for dette. C1 mener at det har vært en betydelig endring de siste seks årene (2013-2019), hvor det vises til eksempel der revisor tidligere mottok papirbunker med dokumentasjon fra regnskapsfører. Dette er helt utenkelig i dag, hvor kunden i større grad ser nytten av ny teknologi, som eksempelvis kan gi bedre innsikt i ting som ikke fungerer.

A2 forteller at man i dag kan oppnå store effektivitetsgevinster gjennom å standardisere prosesser og bruke eksempelvis RPA fremfor selvlærende systemer som AI. C1 er enig at det er innen RPA næringslivet har kommet lengst, og at AI kun i begrenset grad blir benyttet både av revisor og av kundene. A2 mener imidlertid at AI vil komme mer i revisjon med tiden, da det i selskap A allerede blir benyttet i rådgivningssammenheng. B1 har derimot opplevd en større endring mellom 2014 og nå, som er tidsrommet vedkommende har jobbet med AI-løsninger i næringslivet. Respondenten forteller at tidligere var utviklingen av verktøy mye mer individuelt. Noen jobbet seg ferdig med sin del og leverte det, før noen andre fortsatte. I dag, ser man mer sammensatte lag, hvor man jobber mer sammen for å bygge AI inn i applikasjoner, maskiner og roboter. B1 konkluderer med at det har «vært en modning fra dette her har vært noe litt eksotisk som statistikere og avanserte analysefolk driver med, til at det nå er et verktøy i produksjon av en del tjenester og produkter».

Videre spurte vi hva som er motivasjonen for utvikling av verktøy innenfor AI. Her er alle respondentene enige om at kvalitet og effektivitet står i førersetet. B1 forteller at den overordnede motivasjonen er prosessforbedring. B2 utdyper med at AI vil endre måten kundene jobber på, og da bør også revisor endre måten arbeidet utføres på. B1 viser til

eksempel i forsikringsbransjen hvor bruk av AI og maskinlæring har vært brukt til prising og prediksjon av risiko i lengre tid. B2 forteller at en viktig del i utviklingen av AI-verktøy er å bygge kompetanse, for å kunne revidere kundenes bruk av AI. Respondenten forteller, at det handler om å ta i bruk noe som de tror kommer til å bli en «game changer» når det gjelder det å effektivisere og forbedre revisjonen gjennom å utnytte potensialet som ligger i store datamengder.

4.1.1 Eksisterende verktøy og funksjoner

Over på de spesifikke spørsmålene ønsket vi å kartlegge hvilke verktøy revisjonsselskapene har i dag, hvilke som er under utvikling, og hvordan disse passer opp mot definisjonen av AI som benyttes i denne oppgaven. I 4.1.2 vil vi se i hvilken fase i revisjonsprosessen verktøyene benyttes, og i hvilken grad tradisjonelle revisjonshandlinger endres som følge av AI.

Skanningverktøy

Felles for alle responderende selskaper er at det benyttes ulike skanningverktøy til å lese kontrakter, fakturaer og andre dokumenter. D1 forteller at slike verktøy leser dokumentene og henter ut kritisk informasjon til revisor, ved hjelp av naturlig språkprosessering. C1 forteller at «du får en first time skanning før du sitter med en oppsummering over ting som du bør forholde deg til, og se på». D1 supplerer med at verktøyet som er særlig passende for leasingavtaler, vil velge ut kontraktene som er relevant for å gi dekning for de ulike regnskapslinjene. Respondenten forteller at de har videreutviklet programmet, også til og automatisk kunne skanne og avstemme bankbrev. A2 forklarer at man med bildegjenkjenning, på faktura kan sammenligne med informasjon registrert i kundens ERP-system (Enterprise Resource planning)¹.

Alle respondentene mener at det benyttes og/eller utvikles AI i skanningverktøyet som er i tråd med definisjonen. D1 forklarer at revisor vil gjennomgå de kontraktene maskinen valgte ut, og konkludere hvorvidt AI treffer bra eller dårlig. Basert på konklusjonen vil den lære av dette, og gjøre optimaliseringer til neste gang. B1 forklarer at i det ene øyeblikket kan den benyttes til å hente ut problematiske formuleringer i kontrakten, og i neste øyeblikk kan den benyttes for å hente ut hvor i dokumentet det er en underskrift. Når det gjelder hvor i utviklingsstadiet de ulike selskapene er, så er samtlige i hovedsak enige om at det ikke er fullt

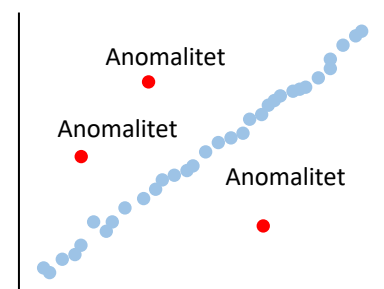
¹ ERP-systemer integrerer flere aspekter av et foretaks aktiviteter i et regnskapssystem (Arens et al., 2013, s. 409).

ut i bruk ennå, men at man befinner seg i en test- og utviklingsfase. D1 forklarer at verktøyet er utviklet i utlandet, og at det kan benyttes i Norge også, men på grunn av ulike årsaker er det ikke tatt i bruk ennå. Dette skyldes for det første at ansvarlig partner må stole på at verktøyet dekker risikoen på en god nok måte. For det andre må kunden godkjenne at en maskin skanner gjennom alle dokumenter, noe som kan være utfordrende med tanke på konfidensialitet og dataangrep.

Analyseverktøy for å identifisere anomaliteter i hovedbok

Selskap B, C og D beskriver alle et analyseverktøy som har likheter. Respondentene forklarer at verktøyet er i stand til å avdekke anomaliteter i bokføringen, og dermed varsle poster til revisor som er unormale ut fra transaksjonsflyten i hovedbok. B1 supplerer med at man kan identifisere blant annet: 1) uvanlige kontokombinasjoner, 2) uvanlige brukere og 3) uvanlig tidspunkt for bokføring.

For eksempel vil fire bokføringer klokken 0345 være emne for ekstra mistanke dersom selskapets resterende 500 000 transaksjoner er bokført mellom 0800-2300. Når det gjelder hvor i utviklingen av verktøyet, og hvordan det



Figur 3: Analyseverktøy

passer opp mot definisjonen av AI, er det noen ulikheter. B1 og C1 er enige når det gjelder progresjonen, der begge forteller at det for tiden er under testing. D1 forteller at programmet er i bruk per i dag, men siden det fremdeles ikke lærer av foregående år, tilfredsstilles ikke definisjonen. Respondenten utdyper at dersom verktøyet finner 100 anomalier, og revisor finner at 10 av disse er normale, så vil verktøyet neste år velge de samme 100. Koblingen som gjør at den faktisk lærer av foregående år mangler, men respondenten mener at dette vil utvikles innen relativt kort tid. C1 er enig, og forteller at det er nettopp verktøyets manglende evne til å lære som gjør det tvilsomt at den tilfredsstillende definisjonen. C1 mener at det ville nok blitt ansett som AI for noen år siden. Derimot mener B1 at deres verktøy tilfredsstillende definisjonen, nettopp grunnet evnen til å lære over tid. Av de tre selskapene som besitter dette verktøyet, er det derfor kun ett som oppfyller definisjonen, mens to foreløpig mangler evnen til å lære.

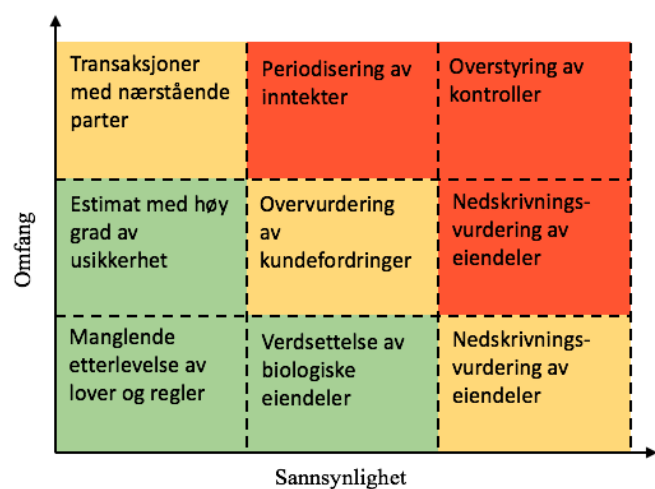
Verktøy for estimat

D1 forteller videre om et verktøy som er under utvikling, og som tilfredsstillende definisjonen av AI. Dette skal være i stand til å estimere markedsverdi av for eksempel eiendom, som så sammenlignes med kundens estimat. Prisen estimeres ved at man gir verktøyet tilgang til data

fra eksempelvis Finn og Google. Modellen finner mønster i dataene på bakgrunn av blant annet etterspørsel, beliggenhet og størrelse ved hjelp av lineær regresjon. Basert på disse faktorene vil verktøyet kunne estimere markedsprisen for kundens eiendom. Respondenten forteller at anvendelsesområdet er bredere enn kun eiendom, og at det også kan benyttes på for eksempel verdipapirer med lavt omsetningsvolum. D1 peker på at det er utfordringer rundt å bruke de vurderingene verktøyet gjør som revisjonsbevis. Dette skyldes kravene i revisjonsstandardene, særlig rundt etterprøvbarehet, i tillegg til en manglende innsikt og dermed en risikooversjon hos signerende partner.

Verktøy for kartlegging av risiko

D1 forteller om et verktøy som planlegges implementert i juni 2019. Dette skal hjelpe revisor med å kartlegge risiko på en mer nøyaktig og presis måte, som ofte vil være den vanskeligste delen av planleggingsfasen. Respondenten forteller at det ofte vil være vanskelig å vite hvor det er forhøyet risiko, noe som særlig gjelder ved nye klienter. Eksempelvis forteller D1 at revisjon av biologiske eiendeler kan være enkelt dersom man vet



Figur 4: Databasen med risikoområder for et selskap

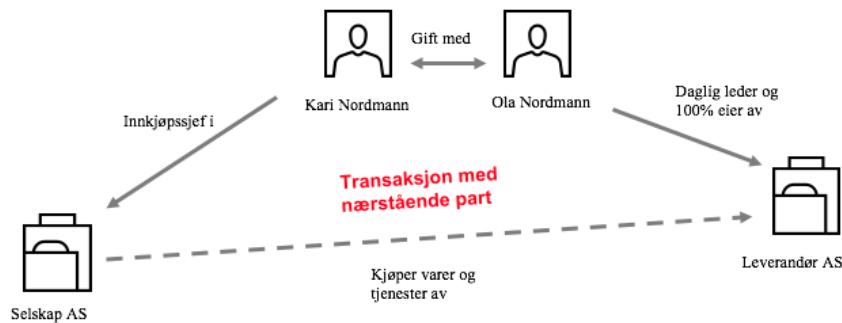
hvor risikoen ligger. Men, dersom man skulle bli introdusert for en stor kvegfarm i USA, så mister man fort oversikten over risikoen. Da vil dette verktøyet hjelpe, basert på en database bestående av tusenvis av gamle revisjoner hvor hver består av flere titalls ulike risikoer. Ved å anvende denne databasen finner verktøyet hvilke risikoer som eksisterer på klientens regnskapslinjer. Opp mot definisjonen så mener respondenten at denne tilfredsstilles. Det påpekes at verktøyet ikke treffer 100% ennå, men at ved input og tilbakemelding vil verktøyet lære, slik at det treffer bedre for hver gang det benyttes.

Verktøy som kategoriserer regnskapslinje

B1 forklarer at det for tiden testes ut en applikasjon som predikerer hvilken regnskapslinjekategori (FSLI-kode) hver konto i saldobalansen skal ha. Prediksjonen gjøres basert på blant annet kontoens nummer, navn, beskrivelse og beløpets fortegn. På bakgrunn av disse kriteriene vil kontoene automatisk bli plassert i eksempelvis lønnskostnad, driftskostnad eller skatt. Verktøyet har en 90% treffsikkerhet, som legger til rette for at revisor

kan kvalitetssikre arbeidet. Respondenten forklarer at verktøyet er «midt i smørøyet på definisjonen deres på kunstig intelligens». Dette skyldes at verktøyet er lært opp til å sortere og finne mønster basert på kombinasjonen av de ulike faktorene. Gjennom erfaring og informasjon vil verktøyet lære og utvikle seg til å sortere kontoene mer nøyaktig.

Relasjonsanalyseverktøy



Figur 5: Kartlegging av nærstående parter

B har utviklet og tatt i bruk et verktøy som i stor grad automatiserer en kartlegging av nærstående parter. Som forklart i 2.2 og i ISA 550.4 må revisor som en del av planleggingen identifisere klientens nærstående parter for å identifisere hvor det eksisterer forretningsrisiko. B1 eksemplifiserer bruken i revisjonen av anskaffelser, hvor verktøyet vil undersøke om det finnes noen koblinger mellom innkjøpere og leverandører. Verktøyet² leser ansatt- og leverandørlistene, og kartlegger automatisk relasjoner for hver enkelt ansatt og leverandør. Informasjon om de ansattes og leverandørens relasjoner hentes løpende ut fra proff.no og andre relevante kilder. Når kartleggingen er ferdig gjennomføres det et søk etter koblinger som kan være risikable. Disse presenteres for selskapets fagekspert, som kan utføre videre undersøkelser. B1 forklarer at verktøyet ligger i gråsonen av definisjonen av AI, som anvendes i denne oppgaven. Dette skyldes at det ligger en optimaliseringsalgoritme til grunn for å finne koblingene, og til tross for at slike algoritmer har mye til felles med veiledede maskinlæringsalgoritmer, så gjør den ingen fleksibel tilpasning av parameterne. Dette til forskjell fra de vanlige veiledede maskinlæringsalgoritmene. Verktøyet har en oppskrift hvordan den skal bygge et nettverk av relasjoner mellom selskap, og så har den en oppskrift på hvordan den skal finne potensielt risikable koblinger. Verktøyet har derimot ikke evnen til å lære, og det vil ikke endre mønster fra hvor den henter informasjon.

² Fremgår av tilsendt materiale fra selskap B.

Prediksjonsverktøy

B1 forteller om et verktøy som for tiden testes internasjonalt, men som ikke er tatt i bruk i Norge. Denne applikasjonen predikerer hva regnskapstallene burde være basert på tidligere års reviderte regnskap. Gapet mellom hva tallet burde være, og hva tallet faktisk er, undersøkes videre av revisor. Verktøyet bruker GBM³ (gradient boosting machine) for å lære sammenhengen mellom regnskapslinjer og drivende faktorer. Respondenten mener at definisjonen tilfredsstilles da det ligger maskinlæringsalgoritmer til grunn, som selvstendig bearbeider dataen den gis tilgang til. Verktøyet tilpasser prediksjonen basert på klientens aktuelle markedsområde, og lærer ved bruk.

Andre observasjoner

A2 forteller at revisjonsselskap A i Norge på dette tidspunktet ikke har noen spesifikke verktøy som benytter seg av AI i revisjon som er implementert. Likevel benytter selskapet seg av flere programmeringsverktøy som R, Python og Azure til å utvikle AI i rådgivningssammenheng. Felles for alle respondentene, er at det også utvikles verktøy i utlandet. B2 forklarer at til tross for at det er et stort samarbeid mellom de ulike selskapene i det globale nettverket, så eksisterer det utfordringer med å benytte løsninger som er produsert i eksempelvis USA. Dette skyldes i stor grad forskjeller når det gjelder størrelsen på kundene, hvor de fleste kundene i USA er langt større. Derfor vil en «analyseløsning som er perfekt for et stort børsnotert amerikansk konsern ... overhodet ikke passe, selv i den største norske virksomheten».

³ Forklares i tilsendt materiale fra selskap B. GBM er en maskinlæringsteknikk som produserer en prediksjonsmodell gjennom å samle flere svake modeller, typisk beslutningstrær (Friedman, 2002, s. 367).

4.1.2 Verktøyenes virkning på revisors handlinger og plass i revisjonsprosessen

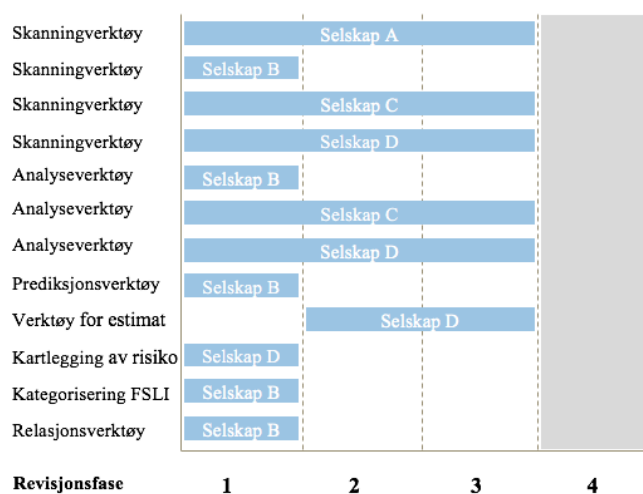
Verktøy	Revisjons-handling endres	Begrunnelse
Skanningverktøy	I noen grad	Tradisjonelt leses kontrakter og andre avtaler manuelt av revisor, for å hente ut kritisk informasjon. Gjennom skanningverktøy gjøres samme revisjonshandling, men nå automatisk. Arbeidet til revisor endres dermed fra å lese kontraktene manuelt og oppsummere kritisk informasjon, til å ta kontrakter gjennom et verktøy som oppsummerer den viktigste informasjonen. Revisor får dermed bedre tid til å bedømme hvordan dette vil påvirke regnskapet.
Analyseverktøy for identifisering av anomalier	I noen grad	Tidligere har revisor manuelt gjennomgått hovedbok for å avdekke anomaliteter. Analytiske handlinger er obligatorisk både i fase 1 og 3, men kan også benyttes som substanshandling i utførelse av test av transaksjoner (fase 2). Verktøyet vil automatisere denne prosessen, og utføre de samme handlingene på en mer effektiv måte samtidig som presisjonen av analysen øker.
Verktøy for estimat	I stor grad	Revisor har tidligere hentet markedsdata fra Finn, Google, finansielle data og andre relevante kilder, for å gi en indikator på verdien av ulike eiendeler. Ved bruk av AI vil verktøyet foreslå et estimat basert på tusenvis av scenarioer med ulik sannsynlighet og bygge opp dokumentasjon. Revisor vil fremdeles manuelt kvalitetssikre estimatet.
Verktøy for kartlegging av risiko	I noen grad	Tidligere har revisor kartlagt risiko per regnskapslinje manuelt gjennom erfaring fra tidligere års revisjon, og andre nærliggende klienter. Ved hjelp av AI vil verktøyet gjøre det samme, men med en høyere innsikt da databasen består av erfaringer fra langt flere klienter i hele verden. Revisor vil deretter fremdeles manuelt kvalitetssikre den kartlagte risikoen.
Verktøy som kategoriserer regnskapslinje	I noen grad	Revisor skal opparbeide seg betryggende sikkerhet for at hver regnskapslinje er uten vesentlig feil. For å gjøre dette må hver konto i saldobalansen kategoriseres under ulike regnskapslinjer. AI vil automatisere dette arbeidet, som tidligere har blitt gjort manuelt. Revisor vil deretter kvalitetssikre kategoriseringen.
Relasjonsanalyse-verktøy	I noen grad	Tradisjonelt kartlegges klientens nærstående parter ved at revisor henter informasjon fra kilder som proff.no, kontrakter og gjennom kommunikasjon med selskapet. Gjennom relasjonsanalyseverktøyet vil dette automatiseres ved at AI kartlegger relasjoner og risikable koblinger. Revisors manuelle arbeid reduseres til å kvalitetssikre.
Prediksjonsverktøy	I noen grad	Tradisjonelt sett har revisor utarbeidet forventninger i innledende analyse for hva regnskapstallene burde være, basert på historiske tall. Verktøyet utfører det samme ved å utnytte sammenhenger mellom regnskapstall og drivende faktorer. Revisor vil fortsatt kvalitetssikre arbeidet, men får mer tid til å granske områder med forhøyet risiko.

Tabell 3: Oversikt over hvordan revisjonshandlingene endres grunnet AI

Tabell 3 viser at revisjonsoppgavene er de samme selv om verktøyene i noen grad endrer hvordan revisjonshandlingene utføres. Respondentene påpeker at verktøyene først og fremst vil assistere og supplere revisors eksisterende oppgaver, fremfor å erstatte.

Etter å ha identifisert ulike verktøy ønsker vi å kartlegge hvor i revisjonsprosessen disse kan benyttes. I intervjuguiden lå revisjonsprosessens fire faser, som beskrevet i 2.2. Basert på tilbakemeldingene fra respondentene har vi i figur 6 kartlagt hvilke faser de ulike verktøyene passer under. Selskapene er i hovedsak enige om at verktøyene på nåværende tidspunkt først og fremst kan benyttes i planleggingsfasen (fase 1). A2 viser til at de ikke benytter kunstig intelligens i revisjonen, men rent hypotetisk, mener respondenten at det er planleggingsfasen for det meste. Dette skyldes at verktøyene er egnet til å kartlegge og avdekke områder hvor det kan være forhøyet risiko. B2 informerer om at «foreløpig er det ganske åpenbart i planleggingen, ettersom vi ikke får lov til å bruke det til noe annet». B2 begrunner dette med at det er krevende å oppfylle standardens krav til etterprøvbare og dokumentasjon når AI-løsningene etterhvert blir mer komplekse. Problematikken rundt revisjonsstandardene presenteres i delkapittel 5.2.

Til tross for at C1 og D1 er enige i bruksområdet mener respondentene at dette formålet er like relevant i fase 2 og 3. C1 legger til at verktøyene også kan benyttes i planleggingsfasen, men hvordan revisor skal håndtere revisjonsstrategien, vil i stor grad fortsatt gjøre og vurderes av mennesket. D1 forklarer at det er mest å hente ved bruk av AI i fase 2 og 3, samtidig som verktøyene kartlegger hvor revisor bør fokusere i fase 1. Respondenten avslutter med at selskapet foreløpig ikke har verktøy som knyttes til fase 4 i revisjonsprosessen, og sier «avslutningen av revisjonen, det vet jeg ikke helt om vi har så mye å hente på».



Figur 6: Oversikt over hvor i revisjonsprosessen AI-verktøyene passer

- Fase 1: Planlegge og designe revisjonens angrepsvinkel
- Fase 2: Utføre kontroll- og substanstester på transaksjoner
- Fase 3: Utføre analytiske handlinger og detaljteste balanse
- Fase 4: Ferdigstillelse av revisjon og avgi revisjonsberetning

4.2 Konsekvenser for revisjonsprosessen

4.2.1 Fordeler og ulemper knyttet til bruken

Når det gjelder fordeler er respondentene i all hovedsak enige om kvalitet og effektivitet. Alle er enige om at det å gå fra en stikkprøvetilnærming, til å se på 100% av populasjonen, øker kvaliteten ettersom man får en langt bedre indikasjon på hvor man bør gå i dybden. D1 mener at man ved AI greier å dekke risikoen på en bedre måte enn ved for eksempel detaljtesting. A2 forteller at man er i stand til å oppnå en effektivitetsgevinst fordi man kan allokere ressurser på en bedre måte. Man bruker ikke nødvendigvis mindre timer, men man kan fokusere på de konteringene som er gjort av daglig leder fremfor å bruke tid på å for eksempel manuelt bla gjennom uttrekk fra bankoverføringer. Da får kunden høyere kvalitet, og revisor får bedre revisjonssikkerhet. B1 gir ytterligere en fordel som handler om arbeidsmiljø og særlig trivsel til revisor. Respondenten forklarer at AI er best til å gjøre de oppgavene som er mest repetitive, og de som typisk krever lite bruk av skjønn. Når AI luker bort disse oppgavene, kan revisor heller konsentrere seg om de mer komplekse oppgavene, de som AI ikke klarte å løse. «Det er mer gøy å grave i hvorfor det er noe rart med tallene her, enn å sitte å regne gjennom det ene tallet etter det andre, bare for å se om svaret med to streker under, er noe rart. I 99 av 100 tilfeller, så var det ikke noe rart med det».

Respondentene forteller at en viktig del av implementering av ny teknologi er samspillet med kunden. Hvordan kunden opplever endringene er til en viss grad delte. A2 kan fortelle at kundene generelt opplever nye teknologiske løsninger som positive. C1 forteller «i stedet for å skrive en stil om prosessen, ta en stikkprøve på prosessen og se at det er bra, så ser du at vi bretter ut hele dataflyten ute hos kunden, og avslører hvordan prosessen egentlig har vært». Respondenten mener derfor at kunden får bedre innsikt i det som ikke fungerer, særlig opp mot internkontrollen. Likevel forteller B1 at kunden ikke nødvendigvis opplever en endring som følge av at revisjonsselskapene har begynt å ta i bruk AI. Dette skyldes at en del av verktøyene i stor grad effektiviserer interne prosesser, som ikke er merkbare for kunden.

Når vi ser på ulempene så er A2, B1 og D1 enige om at revisors manglende forståelse av teknologien kan være en utfordring. A2 påpeker viktigheten av at revisor forstår bruksområdet, og begrensningene til den anvendte teknologien. B1 utdyper og påpeker at man risikerer at brukerne stoler blindt på verktøyet, og sier at det vil være en fare for at man tenker «nå har AI skannet gjennom og fant ingen anomaliteter, og jeg har sett så mange flotte slides

hvor bra den er til å finne avvik, så nå er det helt sikkert greit». Poenget til B1 er at det kan fortsatt være noe som ikke blir plukket opp av de ulike applikasjonene, noe revisor må være klar over ved anvendelsen. D1 etterlyser en mer tverrfaglig kunnskap, da nåværende revisor ikke har den nødvendige kunnskapen om AI, samtidig som de med kunnskap om AI ikke forstår seg på revisjon. Dette reflekteres også i partnerskapet i revisjonsselskapene, hvor respondenten flere ganger peker på at det er en manglende forståelse for AI, og dermed lite villighet for å stole på de resultatene som AI gir. D1 mener at det er lettere for en partner å kritisere en medarbeider på hvorfor noe ikke ble gjort, enn å forstå hva som gikk galt i algoritmene. Dette skyldes at partnerne er av «den gamle skolen, som har saumfart x antall bilag».

C1 og D1 viser til at de resultatene man får av maskinen, er gitt den informasjonen den har tilgang til, og det den er programmert til å gi ut. D1 forklarer det med «garbage in, garbage out». Man er derfor avhengig av at informasjonen og algoritmene er så objektive og nøytrale som mulig. For å illustrere drar C1 paralleller mot tidsskrifter som hele tiden forsøker å gi leseren det han er interessert i, basert på søkehistorikk og tidligere preferanser. I den grad aviser kun gir deg sportsnyheter, så blir ikke leserens verdensbilde nødvendigvis større. Det samme mener C1 gjelder for AI, og sier at dersom man programmerer maskinen til å se etter enkelte ting, eller kun gir den deler av informasjonen som er riktig, så vil man få svar deretter. A2 forteller om utfordringer knyttet til konfidensialitet når revisjonen blir mer digital, og ting blir lastet opp i skyen. Det er derfor en del arbeid i utviklingen å finne løsninger som er enkle å bruke, som ikke bryter lover og regler, og som ikke har for mange svakheter.



Figur 7: Konsekvenser av AI

4.2.2 Målrettet og effektiv revisjon

Selskapene har samme syn på hvordan AI kan bidra til å skape en målrettet og effektiv revisjon. C1 forteller at til tross for at de er i startfasen når det gjelder bruk av AI, så får man en «mye mer effektiv gjennomgang av store datamengder, og dokumenter». B1 forteller at det i hovedsak er enkle og repetitive oppgaver, som for eksempel kategoriseringen av FSLI-koder, som kommer til å spare revisjonsbransjen for tid. B2 viser til at denne kategoriseringen kan gi en besparelse på 15 minutter på hvert revisjonsoppdrag. De mer utfordrende oppgavene, som for eksempel revisors bruk av profesjonelt skjønn, for å avgjøre om det er innhentet tilstrekkelig, og hensiktsmessig revisjonsbevis er det nok fremdeles mest effektivt at mennesket utfører. B1 forklarer at kartleggingen av selskapets nærstående parter som på enkelte selskaper kan ta opp til 37,5 – 75 timer å utføre, har blitt redusert til ca. 30 minutter ved hjelp av relasjonsanalyseverktøyet. D1 viser til at dersom det tar et menneske en time å lese en kontrakt, så vil skanningverktøyet bruke cirka ti sekunder. A2 forteller at AI vil automatisere en prosess ved å minimere antall manuelle steg fra eksempelvis 8 til 2, noe som fører til at man kan allokere ressursene mer effektivt.

Til tross for at respondentene er enige om at man ved hjelp av AI kan få en mer effektiv og målrettet revisjon, påpeker B1 at det tar en del tid å bygge applikasjonen også. A2 supplerer med at det normalt tar tre år før man er i stand til å oppnå en effektivitetsgevinst. Dette begrunnes med at i år 1 så vil man finne så mange avvik, og ha så mange spørsmål at man er heldig om man ikke bruker mer tid. I det andre året kan man grave litt dypere, og i år 3 er man på det tidspunktet hvor det foreligger en effektivitetsgevinst. B1 forteller videre at idealet er de ting som skal gjøres flere ganger, de prosessene som ikke endrer seg for mye fra år til år, og som ikke krever for mye spesialtilpasning på ulike kunder. Men, når man første har laget et verktøy som er i stand til dette, så vil det være en besparelse både i form av tid og penger, frem til ny og bedre teknologi utvikles.

4.2.3 Revisjonsstandardene

En utfordring som er nevnt i 2.4 er forholdet til revisjonsstandardene. Vi spurte derfor respondentene hvorvidt de mener at revisjonsstandardene åpner for bruk av AI. Generelt er respondentene enige om at forholdet til revisjonsstandardene er problematisk. A2 forteller at revisjonen er regelbasert, samtidig som AI har egenskapen av å utvikle og lage sine egne regler. Respondenten fortsetter med at per i dag er det kun adgang til å benytte seg av

regresjonsanalyse til risikovurdering. A2 sier «jeg vil ikke si at revisjonsstandardene per i dag stenger helt for bruk av kunstig intelligens, men det er vanskelig, de legger ikke til rette for det heller. Og i noen grad tenker jeg at det er berettiget». B2 utfyller, og mener at standardene definitivt åpner for bruk av AI som en del av risikovurderingen, men der kan man bruke alle former for informasjon. D1 forklarer at revisjonsstandardene muligens åpner for bruk av AI. Dette skyldes at standardene definerer målet og hva som dekker kravene, men «det er flere veier til Rom, og AI er blant dem».

I revisjonsstandardene er det oppgitt at «revisor skal» utføre en rekke oppgaver. Som et oppfølgingsspørsmål, spurte vi alle i den grad de ikke selv besvarte det på egenhånd, om det er problematisk at revisor i ISA 200 defineres som «den eller de personene som gjennomfører revisjonen ...» når man nå tar i bruk AI. Her var det noe uenighet blant respondentene. A2 og D1 mener at dette er uproblematisk på nåværende stadium. A2 sier at personer er tilstrekkelig nok involvert til å tilfredsstillere revisjonsstandardene, da «man ikke har kommet så langt at en eller annen form for automatisk prosess kjører gjennom hele eller vesentlige deler av revisjonen ...». B2 og C1 er mer enige om at dette er problematisk, og B2 sier at «når vurderingene gjøres av algoritmer, og ikke en person, så er egentlig ikke det i henhold til standarden». C1 tar et eksempel, og viser til to ytterpunkter dersom revisor skal bruke AI til å gjøre et estimat. På den ene siden kan man tenke seg at revisor får beskjed om at estimatet skal være dette beløpet, og dermed bare bruker dette tallet. På den andre siden kan man tenke at maskinen gir revisor alle beregninger som er gjort, alle forutsetninger som er brukt, alle inputparametere og det estimatet maskinen mener er riktig. Da legges det til grunn at et menneske tar den siste kontrollen, og tar en avgjørelse om det er fornuftig eller ikke, kontra å bare stole blindt på maskinen.

4.2.4 Bruk av AI for å forebygge og avdekke misligheter og feil

Som en konsekvens av at AI-verktøy har mulighet til å se over store datamengder, mener selskapene at det vil kunne øke sannsynligheten til å forebygge og avdekke misligheter og feil. B1 mener at maskinen kan fungere som en personlig assistent, og flagge ting for revisor, som virker mistenkelig. C1 påpeker at «revisor overser og har sine skjevheter i hvordan den ser på bildet, og har sine vesentlighetsgrenser når vi ser på ting». Respondenten mener at gevinsten ligger mye i at maskinen klarer å prosessere all datamengden på en annen måte enn hva revisor er i stand til. A2 forteller at algoritmer hjelper med å tolke og klassifisere transaksjoner, noe som øker sannsynligheten for å avdekke misligheter og feil. Respondenten

tilføyer at en maskin i utgangspunktet ikke har noen begrensninger. Dette er D1 enig i, og forteller at de tradisjonelle menneskelige feilene som oppstår ved sløvheter og uoppmerksomhet forsvinner. A2 sier at «den menneskelige hjerne er ikke skapt til å ta hensyn til flere dimensjoner samtidig, og at det kanskje er skjevhet på det man velger å se på, noe som en maskinlæringsalgoritme ikke vil være». Respondent A1 påpeker samtidig at selv om du kvitter deg med en skjevhet, så inviterer du en ny type skjevhet inn i form av AI.

4.2.5 Revisjonsprosess på kort og lang sikt

Til slutt spurte vi alle respondentene hvordan den tradisjonelle revisjonsprosessen vil bli endret som følge av AI på henholdsvis kort og lang sikt. Kort sikt ble definert til å være de neste fem årene, mens lang sikt som de neste 20 årene. På kort sikt er respondentene enige om AI ikke vil ha stor betydning på revisjonsprosessen. D1 begrunner dette med at det vil ta tid før revisorer og særlig ansvarlig partner opparbeider seg den forståelsen og kompetansen som er nødvendig «for å få den tradisjonelle revisjonsprosessen mer på AI baner». A2 viser til noen år tilbake, da Big Data var nytt og spennende. I etterkant mener respondenten at det ikke har vært så mange endringer som først antatt. C1 sier på sin side at AI vil være et hjelpemiddel som kan assistere revisor i de ulike delene av prosessen, men at utførelsen vil være mye på samme måte som tidligere. B1 er noe mer optimistisk og mener at på kort sikt, vil de mindre kompliserte og repetitive oppgavene forsvinne, slik at mennesker ikke utfører grovarbeidet lengre. Samtidig vil hyppigheten av slike oppgaver øke, fordi det ikke lenger er begrensninger på mengden.

B2 mener samtidig som revisjonsbransjen utvikler AI-verktøy, vil også regnskapsprodusenten utvikle flere verktøy. Verktøyene vil legge frem tall, hvor regnskapspåstandene ikke inneholder feil. Videre mener B2 at «på et eller annet tidspunkt, så møtes jo samme teknologi med to forskjellige perspektiv. Da går man kanskje fra å bruke AI til å revidere tallene, til at man heller bruker algoritmer til å revidere kundens algoritmer». A2 forholder seg noe mer konservativ, og mener AI ved verktøy som blant annet tekstgjenkjenning og vurdering av kontrakter først og fremst vil være i stand til å assistere revisor på lang sikt. D1 mener at den tradisjonelle revisjonsprosessen først vil endres når Finanstilsynet, revisor og ansvarlig partner får en større forståelse for AI. Respondenten tror ikke at Finanstilsynet, som revisor, har kunnskapen som er nødvendig for å kunne ettergå revisjonsbevisene. På lang sikt mener B2, C1 og D1 at AI åpner for en mer dynamisk revisjon. B2 ser for seg at fasene blir litt utvasket, og at revisjonen i større grad blir én kontinuerlig prosess.

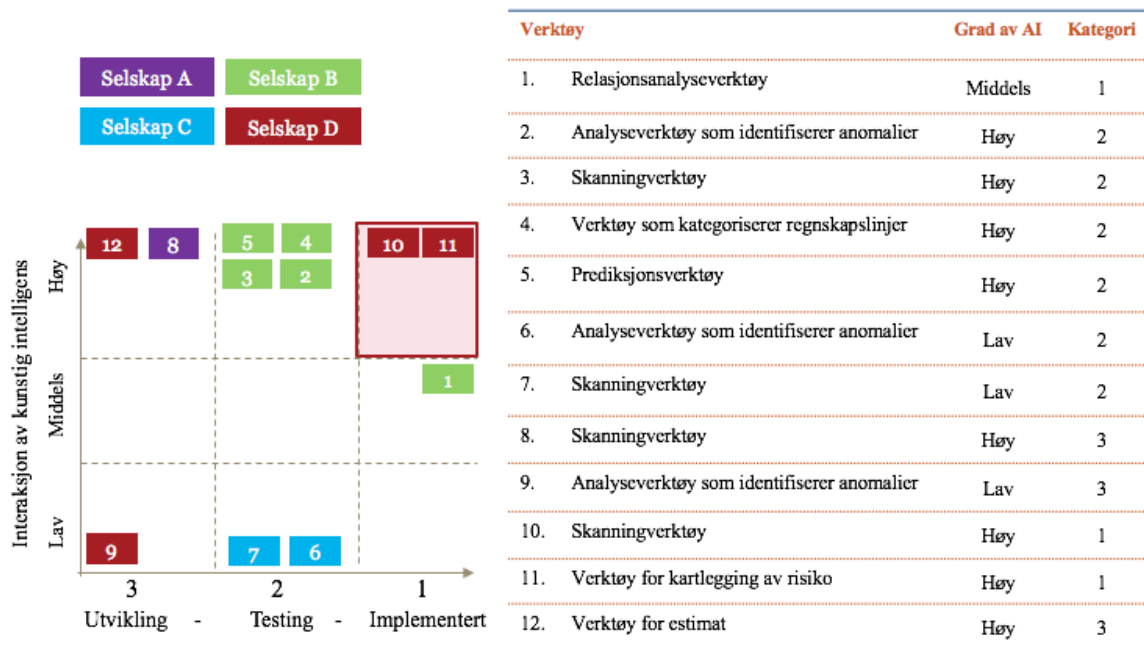
5 DISKUSJON

Vi vil i dette kapittelet drøfte de empiriske resultatene fra undersøkelsen opp mot våre to forskningsspørsmål. I 5.1 har vi samlet resultatene i en figur som viser hvilke verktøy som eksisterer, hvordan de forholder seg til den anvendte definisjonen, og hvorvidt verktøyet er tatt i bruk. Videre i drøftelsen knyttes verktøyene opp mot rammeverket til Kokina og Davenport (2017) som presenteres i 2.4. Vi vil i 5.2 drøfte hvilke konsekvenser økt bruk av AI har for revisjonsprosessen. Det er identifisert en rekke fordeler, men også noen ulemper, som kan by på utfordringer.

5.1 Hvilke verktøy finnes i dag i revisjonsprosessen som er basert på kunstig intelligens (1)?

AI har fått mye oppmerksomhet i samfunnet de siste årene. Det virker til å være en forventning om at AI skal revolusjonere flere industrier i årene fremover. Revisjonsbransjen er tungt regulert gjennom lover og regler, noe som kan sette begrensning på innovative og nye løsninger. Dette er noe svarene fra respondentene bærer preg av. Optimisme og realisme møttes når vi fikk et innblikk i etablerte verktøy og prosjekter under utvikling. I 2.4 påpekes det at Kokina og Davenport (2017, s. 116) mener revisjonsbransjen er klar for å ta i bruk AI i revisjonen. Etter å ha fullført intervjuene og analysert resultatene, opplever vi at dette er en sannhet med modifikasjoner.

Eksisterende forskning fokuserer i liten grad på hvilke verktøy selskapene benytter. Selskapenes offentlige informasjon bærer preg av å være konfidensiell og blir dermed i større utstrekning reklame enn å være informative om anvendelsesområde. Dette gjør det vanskelig å sammenligne våre resultater med andres funn, noe som nok kan forklares med at bruksområdet først i senere år har blitt identifisert. Basert på våre funn har vi i figur 8 kartlagt hvordan respondentene mener at verktøyet passer opp mot definisjonen (vertikal akse), og hvor langt verktøyet er utviklet (horisontal akse). Her defineres kategori 1 som ferdig utviklet og implementert, kategori 2 som testing og pilotering, og kategori 3 som under utvikling. I figuren presenteres alle de tolv identifiserte verktøyene. I resultatet ble disse presentert som syv unike verktøy, da verktøy for identifisering av anomalier og skanningverktøy var noe flere selskaper hadde.



Figur 8: Kartlegging av verktøy

Tidlig i kartleggingen, viste det seg å være en forskjell mellom selskapene. Selskap B og D har henholdsvis fem og fire verktøy av totalt tolv. Forskjellen viser seg også når det gjelder hvordan verktøyene passer opp mot definisjonen. De samme selskapene plasserer seg høyt oppe i grafen, noe som indikerer høy interaksjon av AI. Det kan være flere grunner til dette. Det foreligger ingen klar definisjon, noe som kom tydelig frem i både teori og resultat. Dette bygger opp om at det er bred uenighet om hva AI egentlig er, og at det som blir ansett som AI i dag, ikke nødvendigvis er det samme som for noen år tilbake. Hvordan samfunnet og en IT-ingeniør ser på teknologien er ikke nødvendigvis likt. Respondentene ble derfor introdusert med definisjonen vi har benyttet, som beskrevet i 2.3.1. Likevel kan respondentene ha tolket definisjonen ulikt, og ha forhåndsbestemt seg for hvorvidt de mener verktøyene deres bruker AI eller ikke.

Som figuren viser, er 2/4 av verktøyene under pilotering, og 1/4 under utvikling. Dette er i tråd med Kokina og Davenport (2017, s. 119) som forklarer at AI fremdeles er på et tidlig stadium i revisjon. Det at bare tre av tolv identifiserte verktøy er implementert, er noe lavere enn hva vi forventet, da selskapene er frempå med å reklamere for bruken av AI. Den ene respondenten trakk frem problematikken rundt at revisorer har liten forståelse for AI, mens IT-ingeniører ikke forstår revisjon. Det er derfor naturlig at det tar tid å implementere og

koble sammen to ulike fagfelt til ett verktøy. I delkapittel 5.2 kommer vi mer inn på utfordringene opp mot kompetanse og standarder, som selskapene mener setter en brems på farten på utviklingen.

Revisjonsbransjen er veldig forutsigbar, og det gjøres det samme mer eller mindre hvert år. Alle oppgavene verktøyene utfører er ikke nødvendigvis strukturerte, men alle er definitivt repeterbare. I 2.4 beskrives et rammeverk som fremstiller fire typer oppgaver forfatterne mener kan utføres ved hjelp av AI i revisjonsprosessen. Vi vil presentere hvorvidt de syv identifiserte verktøyene inngår i de fire oppgavetyperne, og i hvilken grad de passer inn i rammeverket.

Numerisk analyse	Analyse av ord og bilder	Utførelse av digitale oppgaver	Utførelse av fysiske oppgaver
<ul style="list-style-type: none"> Analyseverktøy som identifiserer anomaliteter Prediksjonsverktøy 	<ul style="list-style-type: none"> Skanning-verktøy 	<ul style="list-style-type: none"> Verktøy for estimat Verktøy for kartlegging av risiko Relasjonsanalyseverktøy Verktøy som kartlegger regnskapslinjer 	

Figur 9: Kategorisering av verktøyene

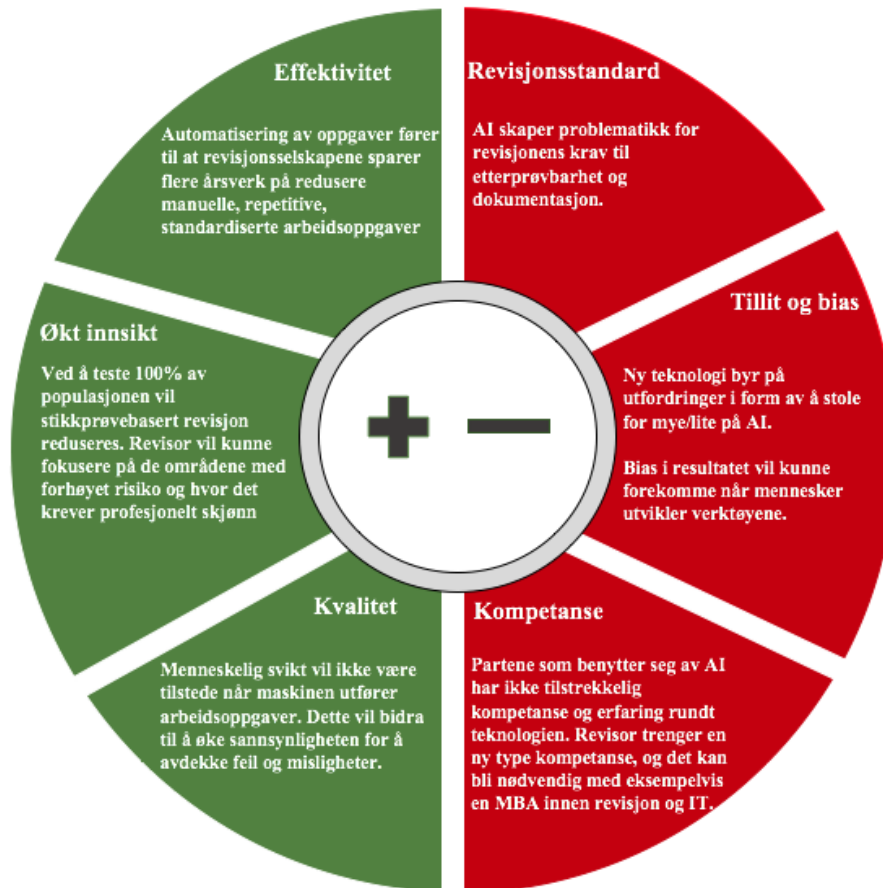
Numerisk analyse er i følge Davenport og Kirby (2016, s. 22) maskiners evne til å analysere tall i strukturerte formater. Gjennom våre intervju ble det avdekket to verktøy som passer under denne kategorien. Kokina og Davenport (2017, s. 117) beskriver at numerisk analyse er særlig nyttig for å identifisere anomalier i store datasett, noe som er sammenfallende med beskrivelsen for «analyseverktøy for å identifisere anomalier». Prediksjonsverktøyet estimerer hva regnskapstallene burde vært basert på tidligere års reviderte regnskapstall. Gapet mellom faktiske og estimerte tall utgjør en anomali som må forklares av revisor. Verktøyet passer derfor inn i kategorien for numerisk analyse. Den neste kategorien arbeidsoppgaver i rammeverket er analyse av ord og bilder. Alle fire selskapene utvikler eller har tatt i bruk skanningverktøy, som er i stand til å lese og innhente den viktigste informasjonen fra kontrakter og andre dokumenter. Dette er sammenfallende med hva Kokina og Davenport (2017, s. 118) beskriver som formålet med denne oppgavetyperne.

Den tredje kategorien arbeidsoppgaver i rammeverket er utførelsen av digitale oppgaver. I 2.4 ble dette omtalt som en restkategori for oppgaver som kreves at det tas en beslutning, men hvor det ikke behøves en robot. Skillet mellom denne kategorien og de to første er ikke helt

klar, noe som gjør kategoriseringen utfordrende. De fire verktøyene som er plassert i kategorien «utførelsen av digitale oppgaver», kunne også blitt kategorisert som numerisk analyse eller analyse av ord og bilder. Likevel er det tydelig at verktøyene gjør mer enn å kun prosessere tall eller ord. For eksempel vil «verktøy som kartlegger regnskapslinje» gjøre en prediksjon av FSLI-kode på bakgrunn av kontoens nummer, navn, beskrivelse og beløpets fortegn. Kokina og Davenport (2017, s. 118-119) viser til at det er typisk strukturerte digitale oppgaver som kan automatiseres på nåværende tidspunkt. «Verktøy som kartlegger regnskapslinje» vil nok falle innenfor denne beskrivelsen da kartleggingen er riktig eller feil, og dermed baseres lite på skjønn. Det er særlig for «verktøy for estimat» det gjør seg gjeldende at oppgaven er ustrukturert. Verktøyet estimerer markedsverdi av eiendom, noe som involverer høy grad av skjønn og alternative verdier. Videre beskriver selskapene at verktøyene er i stand til å lære av seg selv ved at revisor gir tilbakemelding hvorvidt resultatet er presist. Davenport og Kirby (2016, s. 21-23) mener at AI ikke er utviklet i denne grad enda for denne oppgavetypen, da verktøyene ikke lærer uten at et menneske omprogrammerer algoritmene. Basert på våre funn kan det virke som at revisjonsselskapene har kommet et steg lengre, enn hva som var tilfellet når rammeverket ble introdusert i 2016.

Den siste kategorien i rammeverket er fysiske oppgaver. Funnene viser ingen verktøy som utfører slike oppgaver foreløpig. Dette er i samsvar med Kokina og Davenport (2017, s. 117) som mener at dette er lite aktuelt i revisjon, med unntak av ved varetelling. Til tross for at respondentene ikke nevnte noe om dette kan det tenkes at bruk av droner kan bli aktuelt. Resterende fysiske oppgaver er eksempelvis kommunikasjon med klient og observasjon, noe som verken teori eller våre resultater tyder på vil automatiseres på verken kort eller lang sikt.

5.2 Hvilke konsekvenser gir økt bruk av kunstig intelligens for revisjonsprosessen? (2)



Figur 10: Fordeler og ulemper knyttet til AI i revisjonsprosessen

Effektivitet

ISA 200.3 beskriver at formålet med revisjon er å øke brukernes tillit til regnskapet, noe som oppnås ved at revisor gir uttrykk for en mening hvorvidt regnskapet er uten vesentlig feil. Respondentene forteller at verktøyene nevnt i 5.1, og AI generelt vil kunne øke effektiviteten i revisjonen. At en revisjon utføres mer effektivt, vil imidlertid ikke alene bidra til at formålet oppnås. Hva revisor bruker den frigjorte tiden til, vil derimot kunne påvirke brukernes tillit til regnskapet. For eksempel sparer selskap B mye tid på å benytte seg av relasjonsanalyseverktøy, noe som gir revisor mulighet til å fokusere mer på områder med forhøyet risiko som krever revisors profesjonelle skjønn. I 2.4 ble det utledet at kun 20% av 332 revisjonsoppgaver har mange alternative løsninger, og dermed krever høy grad av skjønn. De resterende 80% krever mindre skjønn og kan derfor lettere være gjenstand for automatisering gjennom AI. Frey og Osborne (2017, s. 277) mener som en konsekvens at revisoryrket vil forsvinne innen de neste 20 årene. Til tross for dette, mener de «fire store» at

AI vil assistere og ikke erstatte revisor foreløpig. På kort sikt, kan det virke som at revisjonsbransjen er en trygg bransje å gå inn i, men det er likevel en fare for at maskiner vil ta en større del i arbeidet. Foreløpig gjør den positive effekten at selskapene blir mer konkurransedyktige, samtidig som standardiserte og repetitive oppgaver faller bort.

Kvalitet og økt innsikt

Revisjonskvalitet kan defineres som sannsynligheten for at revisor både vil a) oppdage feil og misligheter i kundens regnskapssystem, og b) rapportere disse (Deangelo, 1981, s. 186). Selv om våre funn ikke indikerer at AI øker kvaliteten gjennom økt rapportering av funn (b), er alle respondentene enige om å gå fra en stikkprøvetilnærming til analyse av hele datasett øker kvaliteten. Dette skyldes en forhøyet sannsynlighet for at revisor vil oppdage feil og misligheter (a). Synet støttes også opp av Raphael (2017, s. 28) som viser til at risikoen for at en uvanlig transaksjon ikke oppdages reduseres. Fra et mislighetsperspektiv økes sannsynligheten for å både forhindre og avdekke misligheter. En frykt for at revisor kommer inn, og har muligheten til å se på alt, setter trolig terskelen for å bedrive misligheter høyere. Disse faktorene er med på å øke kvaliteten gjennom et kvantitetsaspekt. Dette er likevel bare en av flere faktorer som øker sannsynligheten for at revisor vil oppdage feil og misligheter, og dermed øke revisjonskvaliteten (Deangelo, 1981, s. 186).

Til tross for at maskinen ser på flere data, er det ingen automatikk i at den ser bedre på dem. På den ene siden har maskinen sin styrke gjennom at den er veldig god på det området den er utformet til å utføre (Burgess, 2018, s. 49). I den grad algoritmene er riktig programmert vil menneskelig svikt gjennom sløvheter og skjevhet i utvalg fjernes. I en undersøkelse⁴ kommer det frem at 94% av alvorlige bilulykker skyldes menneskelig feil. Eksempelet er ikke direkte relevant for revisjon, men illustrerer betydningen reduksjonen av menneskelige feil kan ha for kvalitet. På den annen side har den menneskelige hjernen sin styrke gjennom at den er bedre enn AI til å kombinere kunnskap, og bruke lærdom fra en situasjon, i en helt annen (Burgess, 2018, s. 49). Selv om et menneske ikke finner feil, vil det bruke denne kunnskapen videre til noe annet, på en måte maskinen ikke er i stand til. Maskinen mangler den kreative koblingen fra å forstå at selv om noe er riktig, kan det likevel bli galt en annen plass. På sikt er målet at AI skal ha mulighet til å forstå hvorfor noe skjer, og finne løsninger på ting maskinen aldri var designet for å utføre (Kaplan & Haenlein, 2019, s. 16). Vi er likevel ikke på det stadiet, hvor maskinen skal ta over for revisor, uten menneskelig interaksjon. Revisjonskvaliteten

⁴ Undersøkelsen er utført av National Highway Traffic Safety Administration (u.å.)

økes dermed på nåværende tidspunkt først og fremst gjennom at maskinen analyserer hele datasett og gir forslag til områder som har forhøyet risiko. På grunn av effektivitetsgevinster knyttet til AI får revisor videre mer tid til å granske disse, fremfor eksempelvis manuell lesing av kontrakter, noe som også bidrar til økt kvalitet.

Tillit og manglende kompetanse

En utfordring flere av respondentene påpeker er manglende forståelse for AI. Dette gjelder både revisor, ansvarlig partner og tilsynsmyndighetene. På den ene siden trekkes det frem at en manglende forståelse av bruksområdet, og verktøyenes begrensninger, kan føre til for stor tillit til verktøyene hos revisjonsmedarbeidere. På den annen side fører manglende forståelse til liten tillit hos ansvarlig partner som skal signere revisjonsberetningen, og «gå god» for alt arbeidet som er utført. Med en manglende forståelse av maskinens kapabiliteter og begrensninger, tør ikke partner å stole på arbeidet som er utført. Flere argumenterer for at manglende kompetanse hos revisor innen nye løsninger er en utfordring, og at fremtidens revisor trenger andre ferdigheter (se blant (Issa et al., 2016, s. 3; Kokina & Davenport, 2017, s. 120)). Tschakert, Kokina, Kozlowski og Vasarhelyi (2017) viser til at det er et økt fokus hos utdanningsinstitusjonene på digitalisering hvor det både utvikles ettårige mastergrader i analyse, og introduseres enkeltkurs i utdanninger innenfor regnskap. Våre funn viser at det er helt nødvendig med økt kompetanse og forståelse, for at revisjonsbransjen skal klare å dra nytte av de fordelene AI gir. Det kan derfor tenkes at det blir nødvendig med en MBA-utdanning innen revisjon og IT.

Bias

Til tross for at det menneskelige elementet av feil reduseres ved bruk av AI, er det likevel en utfordring at resultatet maskinen gir reflekterer bias fra menneskene som programmerer algoritmene, eller samhandler med maskinen (Kokina & Davenport, 2017, s. 120). Noen av respondentene forklarer problemstillingen med at maskinens resultater avhenger av treningsdataene. I den grad disse er skjeve blir også resultatet skjevt. Et eksempel er Google's bildeapplikasjon, som automatisk kategoriserer lignende bilder i kategorier i brukerens bildebibliotek. Her ble personer med mørkere hudfarge kategorisert som gorillaer. Dette skyldes antageligvis at treningsdataene inneholdt bilder av langt flere hvite mennesker, noe som gjør det vanskeligere for maskinen å gjenkjenne ikke-hvite ansikter (Crawford, 2016). Foreløpig vil revisor gjennomgå resultatene som AI produserer, noe som er med på å redusere sannsynligheten for feil. På lengre sikt er det derimot ønskelig at maskinen skal utføre oppgavene mer selvstendig, noe som gjør at bias kan bli mer problematisk. IBM (u.å.) jobber

med en løsning der algoritmer skal være i stand til å avsløre bias i resultatene. Dette gjøres ved å programmere maskinen slik at den anvender visse verdier og prinsipper som mennesker vanligvis bruker ved beslutninger.

Revisjonsstandardene

Innledningsvis i oppgaven løftes utfordringer som Issa et al. (2016) har påpekt ved å benytte seg av AI opp mot revisjonsstandardene. Vi observerer at alle revisjonsselskapene ser problematikken rundt dette, men vi ser samtidig at selskap B og D er de som ser størst på utfordringen. Dette er de samme to selskapene som i figur 8 anses for å ha høyest grad av AI involvert i verktøyene. Noen av selskapene løfter problematikken med at alt skal kunne etterprøves og dokumenteres. Revisor skal dokumentere fremgangsmåte på en slik måte at det er mulig å forstå arbeidet som er utført og de konklusjoner som er trukket.

For å forsøke å få svar på utfordringene nevnt i avsnittet over, tok vi kontakt med Finanstilsynet (FT) på mail. De påpekte tidlig at FT ikke har noen offisiell uttalelse rundt dette, men var likevel i stand til å gi oss et svar. I svaret fremgår det at FT ikke på generelt grunnlag, og på forhånd kan akseptere bruk av AI eller andre former for ny teknologi. Dette ville ha avskåret fra å kritisere en revisjonsutførelse der slik teknologi er benyttet, og hvor det i en konkret situasjon viser seg at det mangler tilstrekkelige, og hensiktsmessige revisjonsbevis. Utgangspunktet er derfor at i den grad revisjonsselskaper tar i bruk ny teknologi, må de selv vurdere forholdet til de bransjefastsatte standardene anvendt på det konkrete faktumet i saken. For at FT skal kunne ta stilling til om denne vurderingen er forsvarlig, må revisor dokumentere den. Det pekes videre på at det kan være knyttet vanskeligheter rundt det å etterprøve revisors arbeid, jf. Revisorloven § 5-3 første ledd. Denne etterprøvbarheten er helt sentral for at et offentlig tilsyn, og andre kontroller, skal fungere etter sin hensikt. FT har sagt seg villig til å gjennomgå oppdrag der revisjonsselskapet har tatt i bruk ny teknologi i revisjonsutførelsen, for å vurdere revisjonen opp mot kravene i revisorloven, herunder god revisjonsskikk. Slike kontroller er på nåværende tidspunkt under planlegging, men ikke gjennomført.

6 KONKLUSJON OG VEIEN VIDERE

Denne masteroppgaven har som mål å besvare problemstillingen; *Hvilke verktøy finnes i revisjonsprosessen og hvilke konsekvenser gir økt bruk av AI for revisjonsprosessen?*

Gjennom å forske på utviklingen de «fire store» har opplevd innen AI, og hvordan denne benyttes, skal vi svare på problemstillingen.

Forskningen viser en stor optimisme rundt revisjonsprosessen, og hvordan den vil endres i tiden fremover. På kort sikt vil det trolig ikke skje noen revolusjonerende endringer grunnet utfordringer i standard, bias og manglende kompetanse. Fordelene med AI-verktøyene, knyttet til økt innsikt, kvalitet og effektivitet, virker å være en gevinst allerede i dag. Som resten av samfunnet må også revisjonsbransjen utvikle seg og benytte seg av de teknologiske fremskrittene som foreligger. Kundene får i større grad økt teknologi som drives gjennom skyløsninger, noe som fører til at revisor kan få tilsendt data kontinuerlig. Med økt bruk av AI hos kunden vil regnskapet kunne presenteres uten vesentlig feil. Den tradisjonelle revisjonsprosessen vil som en konsekvens endres, hvor revisjonen blir en mer dynamisk prosess, der revisor reviderer kundens algoritmer fremfor selve substansen i tallene.

Da teknologien fremdeles er på et tidlig stadium, ville det vært interessant om forskning i fremtiden kvantitativt undersøker den tidsmessige besparelsen når det benyttes AI i revisjon. Samtidig opplyses det om at sannsynligheten for å avdekke feil og misligheter øker gjennom de identifiserte AI-verktøyene. Det ville derfor vært nyttig å forske på hvor mange færre feil som avdekkes, etter hvert som bruken av teknologien øker. For å imøtekomme utfordringene som medfølger bør utdanningsinstitusjoner innrette seg for at fremtidens revisor skal ha den nødvendige kompetansen og evnen til å omskolere seg, og tilpasse seg fremtidens revisjonsprosess.

REFERANSER

- Abdolmohammadi, M. J. (1999). A Comprehensive Taxonomy of Audit Task Structure, Professional Rank and Decision Aids for Behavioral Research. *Behavioral Research in Accounting*, 11, 51-92. Lastet ned fra <http://connection.ebscohost.com/c/articles/1706915/comprehensive-taxonomy-audit-task-structure-professional-rank-decision-aids-behavioral-research>
- Agnew, H. (2016, 9. mai). Auditing: Pitch battle, *Financial Times*. Lastet ned fra <https://www.ft.com/content/268637f6-15c8-11e6-9d98-00386a18e39d>
- Alles, M. G. (2015). Drivers of the use and facilitators and obstacles of the evolution of big data by the audit profession.(Report). 29(2), 439. doi: 10.2308/acch-51067
- Arens, A. A., Beasley, M. S. & Elder, R. J. (2013). *Auditing and Assurance Services, Global Edition* (Fifteenth edition. utg.): United Kingdom: Pearson Education M.U.A.
- Baldwin, A. A., Brown, C. E. & Trinkle, B. S. (2006). Opportunities for artificial intelligence development in the accounting domain: the case for auditing. *Intelligent Systems in Accounting, Finance & Management*, 14(3), 77-86. doi: 10.1002/isaf.277
- Bizarro, P. & Dorian, M. (2017). ARTIFICIAL INTELLIGENCE: THE FUTURE OF AUDITING. *Internal Auditing*, 32(5), 21-26.
- Burgess, A. (2018). The executive guide to artificial intelligence : how to identify and implement applications for AI in your organization
- Crawford, K. (2016, 25. juni). Artificial intelligence's white guy problem, *The New York Times*. Lastet ned fra <https://www.cs.dartmouth.edu/~ccpalmer/teaching/cs89/Resources/Papers/AIs White Guy Problem - NYT.pdf>
- Datatilsynet. (2018). *Kunstig intelligens og personvern*. (Rapport, januar 2018). Lastet ned fra <https://www.datatilsynet.no/globalassets/global/om-personvern/rapporter/rapport-om-ki-og-personvern.pdf>.
- Davenport, T. H. & Kirby, J. (2016). Just how smart are smart machines?(Intelligence)(Report). 57(3), 21.
- Deangelo, L. E. (1981). Auditor size and audit quality. *Journal of Accounting and Economics*, 3(3), 183-199. doi: 10.1016/0165-4101(81)90002-1
- Frey, C. B. & Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerisation? *Technological Forecasting & Social Change*, 114, 254-280. doi: 10.1016/j.techfore.2016.08.019

- Friedman, J. H. (2002). Stochastic gradient boosting. *Computational Statistics and Data Analysis*, 38(4), 367-378. doi: 10.1016/S0167-9473(01)00065-2
- Godfrey, J. M., Hodgson, A., Tarca, A., Holmes, S. & Hamilton, J. (2010). *Accounting theory* (7th ed. utg.). Milton: Wiley.
- Gottfredson, L. S. (1997). Mainstream science on intelligence: An editorial with 52 signatories, history, and bibliography. *Intelligence*, 24(1), 13-23. doi: 10.1016/S0160-2896(97)90011-8
- Grenness, T. (2001). *Innføring i vitenskapsteori og metode* (2. utg. utg.). Oslo: Universitetsforl.
- Gripsrud, G. (2004). *Metode og dataanalyse : med fokus på beslutninger i bedrifter*. Kristiansand: Høyskoleforl.
- Gulden, B. P. (2016). *Revisjon : teori og metode* (7. utg. utg.). Oslo: Cappelen Damm akademisk.
- Ibm. (u.å.). AI bias will explode. But only the unbiased AI will survive Hentet fra <https://www.research.ibm.com/5-in-5/ai-and-bias/>
- 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
- Iaasb. (2018). *Feedback statement - Exploring the growing use of technology in the audit, with a focus on data analytics*. Lastet ned fra <https://www.ifac.org/publications-resources/feedback-statement-exploring-growing-use-technology-audit-focus-data>.
- Johannessen, A. (2010). *Introduksjon til samfunnsvitenskapelig metode* (4. utg. utg.). Oslo: Abstrakt.
- Kaplan, A. & Haenlein, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, 62(1), 15-25. doi: 10.1016/j.bushor.2018.08.004
- Kinserdal, F. (2017). NHH skal forske på digitalisering i revisjonsbransjen. 79-86. Lastet ned fra <https://www.magma.no/nhh-skal-forske-pa-digitalisering-i-revisjonsbransjen>
- 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
- Langli, J. C. (2015). *Evaluering av unntak for revisjonsplikt i små aksjeselskaper*. Lastet ned fra

<https://www.regjeringen.no/contentassets/981a92dc8e474d6ea18e4be36d860255/revisjonsplikt.pdf>.

- Marsland, S. (2015). *Machine learning : an algorithmic perspective* (2nd edition. utg.). Boca Raton, FL: CRC Press.
- Nhtsa. (u.å.). Automated vehicles for safety Hentet fra <https://www.nhtsa.gov/technology-innovation/automated-vehicles-safety>
- Raphael, J. (2017). Rethinking the audit: innovation is transforming how audits are conducted - and even what it means to be an auditor. *Journal of Accountancy*, 223(4), 28.
- Rapport, M. (2016, 7. mars). Auditing firms count on technology for backup, *The Wall Street Journal*. Lastet ned fra <https://www.wsj.com/articles/auditing-firms-count-on-technology-for-backup-1457398380>
- Samuel, A. L. (1959). Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal of Research and Development*, 3(3), 210-229. Lastet ned fra <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5392560>
- Sekaran, U. & Bougie, R. (2016). *Research methods for business : a skill-building approach* (7th ed. utg.). Chichester: Wiley.
- Simon, H. A. (1979). Rational Decision Making in Business Organizations. *The American Economic Review*, 69(4), 493-513.
- Telle, J. A. (2017). Den nye maskinlæringen: Kunstig intelligens eller bare gode verktøy? *Nytt norsk tidsskrift*, 34(2), 192-204.
- Theobald, O. (2017). *Machine learning for absolute newbeginners* (2. utg. utg.): Scatterplot Press.
- Tschakert, N., Kokina, J., Kozlowski, S. & Vasarhelyi, M. (2017). How business schools can integrate data analytics into the accounting curriculum.(NEWS & VIEWS: Education). *The CPA Journal*, 87(9), 10.
- Tørresen, J. (2013). *Hva er kunstig intelligens* (vol. 49). Oslo: Universitetsforl.
- Varıcı, I. (2013). The Relationship between Information Asymmetry and the Quality of Audit. *Canadian Center of Science and Education*, 6(10), 132-140. Lastet ned fra doi:10.5539/ibr.v6n10p132

VEDLEGG: INTERVJUGUIDE

Informasjon om gjennomføring av intervju «kunstig intelligens i revisjon»

Vi er to studenter fra HHUiS som studerer master i regnskap og revisjon. Vi skriver masteroppgave om **Kunstig intelligens i revisjon**. Formålet med intervju spørsmålene er å undersøke følgende:

1. Hvilke verktøy finnes i dag i revisjonsprosessen som er basert på kunstig intelligens, og hvilke vil vi se i fremtiden?
2. Hvordan vil revisjonsprosessen endres som følge av kunstig intelligens?

Informasjon om gjennomføring av intervjuet

Intervjuet består av 12 intervju spørsmål som er forankret i revisjonsteori og forskning. Vi vil innledningsvis også stille noen spørsmål om din stilling og erfaring.

Deltakelse er frivillig og informasjonen vil anonymiseres.

Vi ønsker å ta lydopptak av intervjuet. Opptaket vil kun benyttes av undertegnede for å sikre korrekt informasjon. Opptaket slettes ved innlevering av masteroppgaven.

Estimert tidsbruk for intervju er ca. 60 minutter.

Studenter, HHUiS

Stig Herigstad og Anders Skogseth

Veileder: Carmen Olsen, NHH



Universitetet
i Stavanger

Generelle spørsmål

- Hvilken posisjon har du i ditt selskap?
- Hvor mange år har du jobbet med revisjon?
- Hvor mye erfaring har du innenfor kunstig intelligens?
- Hvordan har veien vært med bruk og utvikling av kunstig intelligens i ditt selskap?
- Hva er motivasjonen for utvikling av verktøy innenfor kunstig intelligens i ditt selskap?

Intervjuspørsmål

Tema 1: Hvilke verktøy finnes i revisjonsprosessen som er basert på kunstig intelligens og hvilke vil vi se i fremtiden?

Revisjonsbransjen har tradisjonelt sett vært treg i implementering av nye teknologiske verktøy. I de senere år har det imidlertid skjedd et skifte hvor særlig de største revisjonsselskapene investerer store ressurser i teknologier som kunstig intelligens (AI) og avansert dataanalyse. I åpenhetsrapporter og på selskapenes hjemmesider reklameres det i stor utstrekning for hvordan de selv ønsker å ta del i denne utviklingen som vi nå ser. Vi ønsker å undersøke og samle kunnskap om hvilke verktøy innenfor kunstig intelligens som finnes i markedet i dag, og hvilke vi med sannsynlighet vil se i fremtiden. På grunn av mange og ulike definisjoner vil vi definere kunstig intelligens som **«et systems evne til å tolke eksterne data riktig, å lære fra disse dataene, og å bruke lærdommen for å oppnå spesifikke mål og oppgaver gjennom fleksibel tilpasning»** ((Kaplan & Haenlein, 2019, s. 17) egen oversettelse)).

1. Hvilke verktøy innen revisjon har dere som eksisterer i dag, og hvilke funksjoner har disse?
2. Det er stor diskusjon og mange meninger om hva som defineres som AI. Hvordan mener du at ditt selskaps verktøy passer opp mot definisjonen beskrevet over av kunstig intelligens?
3. Hvilke fordeler og ulemper ser dere knyttet til disse verktøyene?
4. Hvordan presenterer dere disse verktøyene til kunden, og hvordan opplever kunden bruken av disse?

5. Mener dere at revisjonsstandardene åpner for bruk av AI, og hvor i revisjonsstandardene utfordres eventuelt utviklingen?
6. Har dere opplevd en besparelse i tid knyttet til revisjon, ved å implementere og ta i bruk verktøyene innenfor kunstig intelligens?
7. Er det noen prosjekt som er under utvikling per i dag?

Tema 2: Hvordan vil revisjonsprosessen endres som følge av kunstig intelligens?

Arens et al. (2013) deler revisjonsprosessen inn i de fire fasene: planlegging, utførelse av kontroll- og substanstester, utførelse av analytiske prosedyrer og detaljtesting av balanse, og avslutning av revisjonen med avgivelse av revisjonsberetning.

1. Med utgangspunkt i revisjonsprosessen som beskrevet over, kan du nevne de fasene i prosessen hvor deres verktøy knyttet til AI blir benyttet?
2. Hvordan mener dere at AI kan bidra til å skape en målrettet og effektiv revisjon?
3. Hvordan mener dere at AI kan bidra til å øke sannsynligheten for at revisor skal være i stand til å forebygge og avdekke misligheter og feil?
4. Hvordan mener dere AI vil endre den tradisjonelle revisjonsprosessen på henholdsvis kort og lang sikt?