



Kunstig intelligens og ansvar

En empirisk studie om hvordan norske organisasjoners ansvar blir annerledes ved innføring av kunstig intelligens

Ida Marie Barlaup Pedersen og Synne Spjøtvold Berg

Veileder: Katarine Kaarbøe

Masterutredning i Økonomi og Administrasjon

Hovedprofiler: Økonomisk styring og Business Analytics

NORGES HANDELSHØYSKOLE

Dette selvstendige arbeidet er gjennomført som ledd i masterstudiet i økonomi- og administrasjon ved Norges Handelshøyskole og godkjent som sådan. Godkjenningen innebærer ikke at Høyskolen eller sensorer inntår for de metoder som er anvendt, resultater som er fremkommet eller konklusjoner som er trukket i arbeidet.

Abstrakt

Januar 2020 la regjering frem en nasjonal strategi for kunstig intelligens. Regjeringen ønsker å opprettholde et bærekraftig velferdssamfunn ved å bruke kunstig intelligens til å utarbeide smarte og effektive løsninger for å opprettholde konkurransekraften (Kommunal- og moderniseringsdepartementet, 2020). I strategien påpekes det at ansvar er sentralt ved utvikling og bruk av kunstig intelligens. Derimot eksisterer det lite forskning på hvordan organisasjoner jobber med og forholder seg til ansvar når de tar i bruk kunstig intelligens. Derfor tar utredningen sikte på å undersøke hvordan norske organisasjoner diskuterer og jobber med ansvar relatert til kunstig intelligens for å sikre seg mulighetene som fremgår av teknologien.

Utredningen er bestått av en dokumentstudie og en casestudie. Gjennom en kvalitativ tilnærming tar utredningen sikte på å gi inngående kunnskap om relevante regler og anbefalinger, og hvordan organisasjoner forstår ansvar relatert til kunstig intelligens. Utredningen vil besvare følgende forskningsspørsmål:

«Hvordan blir ansvar annerledes ved innføring av kunstig intelligens i norske organisasjoner?»

Vi er nå i en fjerde industriell revolusjon. Til tross for at norske organisasjoner begynner å ta i bruk kunstig intelligens i sine prosesser, vet de kanskje ikke om organisasjonens ansvar endres ved å innføre kunstig intelligens. Vi mener et juridisk, moralsk og styringsansvar gir et helhetlig bilde av organisasjoners ansvar og at norske organisasjoners ansvar blir annerledes ved innføring av kunstig intelligens. Dette er basert på at kunstig intelligens skiller seg fra tradisjonell teknologi da kunstig intelligens har evne til å lære og oppfatte, samt bruke erfaringer til å utføre handlinger. Videre mener vi organisasjoners ansvar blir annerledes fordi det er en spenning mellom egenskapene som gjør at kunstig intelligens skiller seg fra tradisjonell teknologi og eksisterende lovverk.

Basert på våre funn, mener vi at norske organisasjoners ansvar blir annerledes ved innføring av kunstig intelligens ved at et juridisk ansvar ikke er tilstrekkelig og at moralsk og styringsansvar blir viktigere. Norske organisasjoner får derfor et økt ansvar ved innføring av kunstig intelligens.

Forord

Masterutredningen er gjennomført som en del av mastergradsstudiet i økonomi og administrasjon ved Norges Handelshøyskole (NHH), og er skrevet innenfor hovedprofilene Økonomisk Styring (BUS) og Business Analytics (BAN).

Temaet for utredningen er kunstig intelligens og ansvar. Kunstig intelligens er et dagsaktuelt tema med et stadig tilfang av ny litteratur og nye artikler. Videre er det stadig flere organisasjoner som ønsker å innføre kunstig intelligens. Utredningen tar opp viktige aspekter vedrørende ansvar som vi håper organisasjoner vil ha nytte av ved innføring av kunstig intelligens. I utredningen har vi blant annet gjennomført kvalitative intervjuer av ansatte i organisasjoner som har innført en form for kunstig intelligens. I tillegg gjennomførte vi innledningsvis to ustrukturerede intervjuer for å sette oss inn i nåværende situasjon i næringslivet. I den forbindelse ønsker vi å rette en stor takk til respondentene for at de tok seg tid til å bli intervjuet, til tross for en hektisk arbeidshverdag. Uten disse ville ikke utredningen blitt en realitet.

Avslutningsvis vil vi rette en spesiell stor takk til vår veileder, Katarina Kaarbøe, ved institutt for regnskap, revisjon og rettsvitenskap ved NHH. Hennes engasjement og gode humør har vært inspirerende helt fra første møte. Hennes støtte og råd har vært til god hjelp gjennom hele prosessen.

Bergen, 20. desember 2020

Ida Marie Barlaup Pedersen

Synne Spjøtvold Berg

Innholdsfortegnelse

1. INTRODUKSJON	1
1.1 BAKGRUNN	1
1.2 FORSKNINGSSPØRSMÅL	2
1.3 METODISK TILNÆRMING	3
1.4 VÅRT BIDRAG	3
1.5 DISPOSISJON	3
2. TEORI: KUNSTIG INTELLIGENS OG ANSVAR	5
2.1 KUNSTIG INTELLIGENS	5
2.1.1 <i>Historien til kunstig intelligens</i>	5
2.1.2 <i>Hva er kunstig intelligens?</i>	7
2.1.3 <i>Klassifisering av kunstig intelligens</i>	8
2.1.4 <i>Kritikk mot kunstig intelligens</i>	11
2.1.5 <i>Kunstig intelligens i vår utredning</i>	11
2.2 ANSVAR I TRE PERSPEKTIVER: JURIDISK, MORALSK OG STYRING	12
2.2.1 <i>Hvordan litteratur definerer ansvar</i>	12
2.2.2 <i>Juridisk ansvar</i>	14
2.2.3 <i>Moralsk ansvar</i>	15
2.2.4 <i>Styringsansvar</i>	16
2.2.5 <i>Hvordan vi definerer ansvar</i>	18
2.3 TEORETISK REFERANSERAMME	19
2.3.1 <i>Ansvar i tre deler</i>	20
2.3.2 <i>Vår tolkning av inndelingen</i>	23
3. METODE	25
3.1 FORSKNINGSTILNÆRMING	25
3.2 FORSKNINGSDESIGN	26
3.2.1 <i>Formål</i>	26
3.2.2 <i>Metode</i>	27
3.2.3 <i>Strategi</i>	27
3.2.4 <i>Tidshorisont</i>	29
3.3 DATAINNSAMLING	29
3.3.1 <i>Sekundærdata: dokumentstudie</i>	30
3.3.2 <i>Dokumentutvalget</i>	30
3.3.3 <i>Primærdata: semistrukturerte intervju</i>	32
3.3.4 <i>Intervjukandidater</i>	32
3.3.5 <i>Gjennomføring av intervju</i>	33
3.4 DATAANALYSE	33
3.4.1 <i>Dokumentanalyse</i>	34
3.4.2 <i>Caseanalyse</i>	34

3.5	FORSKNINGSKVALITET.....	35
3.5.1	<i>Begrepsvaliditet</i>	35
3.5.2	<i>Ekstern validitet</i>	36
3.5.3	<i>Reliabilitet</i>	36
3.5.4	<i>Dokumentkvalitet</i>	38
3.6	FORSKNINGSETIKK.....	39
3.7	OPPSUMMERING AV METODISKE VALG.....	39
4.	EMPIRI: DOKUMENT- OG CASESTUDIE	40
4.1	DEL EN: REGELVERK OG ANBEFALINGER FRA ET ANSVARSPERSPEKTIV	40
4.1.1	<i>Ansvar er sentralt ved utvikling og bruk av kunstig intelligens</i>	40
4.1.2	<i>Muligheter med kunstig intelligens</i>	42
4.1.3	<i>Utfordringer knyttet til bruk av kunstig intelligens</i>	44
4.1.4	<i>Konklusjon: Regelverk og anbefalinger fra et ansvarsperspektiv</i>	46
4.2	DEL TO: ORGANISASJONERS FORSTÅELSE AV KUNSTIG INTELLIGENS OG ANSVAR	48
4.2.1	<i>Kunstig intelligens</i>	49
4.2.2	<i>Ansvarsforståelse</i>	52
4.2.3	<i>Ansvar i organisasjoner</i>	57
4.2.4	<i>Lov og anbefalinger</i>	60
4.2.5	<i>Oppsummering: Organisasjoners forståelse av kunstig intelligens og ansvar</i>	63
5.	DISKUSJON.....	65
5.1	ANSVAR	65
5.2	ANSVAR PÅ FORHÅND OG ANSVAR ETTERPÅ.....	66
5.2.1	<i>Ansvar på forhånd</i>	66
5.2.2	<i>Ansvar etterpå</i>	68
5.3	DIREKTE ANSVAR OG MEDVIRKNING.....	69
5.3.1	<i>Direkte ansvar</i>	69
5.3.2	<i>Medvirkning</i>	70
5.4	SMALT OG BREDT ANSVAR	71
5.4.1	<i>Smalt ansvar</i>	71
5.4.2	<i>Bredt ansvar</i>	72
5.5	OPPSUMMERING: DISKUSJON	74
6.	KONKLUSJON.....	76
6.1	SVAR PÅ FORSKNINGSSPØRSMÅLET	76
6.2	IMPLIKASJONER FOR FORFATTERE AV DOKUMENTENE OG ORGANISASJONER	77
6.3	FORSLAG TIL VIDERE FORSKNING.....	78
6.4	UTREDNINGENS BEGRENSNINGER	79
	BIBLIOGRAFI.....	80
	VEDLEGG	85
	VEDLEGG 1: INTERVJUGUIDE.....	85

Tabelliste

TABELL 1: OVERSIKT OVER DOKUMENTENE I INNHOLDSANALYSEN	30
TABELL 2: OPPSUMMERING AV METODISKE VALG	39
TABELL 3: HOVEDFUNN – DEL EN: REGELVERK OG ANBEFALINGER.....	48
TABELL 4: OPPSUMMERING AV FUNN - DEL TO: ORGANISASJONERS FORSTÅELSE AV KI OG ANSVAR	64

Figurliste

FIGUR 1: ILLUSTRASJON AV SAMMENHENGEN MELLOM KI, ML OG DYP LÆRING	11
FIGUR 2: ILLUSTRASJON AV TEORETISK REFERANSERAMME FRA SYSE OG OLSEN (2013)	23
FIGUR 3: OPPSUMMERING GJENNOM TEORETISK REFERANSERAMME.....	75

1. Introduksjon

I dette kapitlet vil vi presentere utredningens bakgrunn i kapittel 1.1, samt forskningsspørsmål i kapittel 1.2. Deretter vil vi presentere metodiske valg og vårt bidrag i kapitlene 1.3 og 1.4. Avslutningsvis, i kapittel 1.5, er det en oversikt over resten av utredningens disposisjon.

1.1 Bakgrunn

Teknologi er i stadig utvikling og har de siste årene fått stor innvirkning på samfunnet og organisasjoner. Etter den første industrielle revolusjonen har utviklingen av teknologi økt, både i omfang og hastighet (Brynjolfsson & McAfee, 2016). Schwab (2014) argumenterer for at vi nå befinner oss i en fjerde industriell revolusjon da teknologien som blir utviklet er mer kompleks, altomfattende, samt endres raskere enn noen gang før. Vi står ovenfor et skifte, hvor nye kapabiliteter endrer måten vi lever og arbeider på. Etablerte organisasjoner utfordres blant annet av nye forretningsmodeller, trender og endringer i kundeatferd. En av disse trendene er teknologien kunstig intelligens, som blant annet er å finne i selvkjørende biler, personlig markedsføring og mønster- og talegjenkjenning (Adams, 2017).

Kunstig intelligens (KI) har fått mye oppmerksomhet den siste tiden. Ved å bruke søkeordet «kunstig intelligens» i søkefeltet til Dagens Næringsliv (DN) kommer det opp 944 treff i artikler og videoer. Flere store teknologiselskaper som Google, IBM og Facebook har gjort store investeringer i KI (Deloitte). Det å investere i teknologien er også blitt en viktig prioritet for norske selskaper (Ernst & Young, 2018). Ifølge Harry Shum, visepresident for Microsofts kunstige intelligens og forskningsgruppe, er Norge i en unik posisjon til å skape muligheter med kunstig intelligens. Dette på bakgrunn av digital modenhet og sterk økonomi (Ernst & Young, 2018). Januar 2020 la regjering frem en nasjonal strategi for kunstig intelligens. Her kommer det også frem at Norge har en rekke forutsetninger som gir Norge muligheter til å lykkes med KI. Noen av argumentene for dette er god økonomi, høy grad av tillit i samfunnet, godt utviklet digital forvaltning og digitalt kompetent næringsliv og befolkning. Regjeringen ønsker å opprettholde et bærekraftig velferdssamfunn ved å blant annet bruke KI til å utarbeide smarte og effektive løsninger for å opprettholde konkurransekraften (Kommunal- og moderniseringsdepartementet, 2020).

Til tross for at KI skaper muligheter er det også flere utfordringer med teknologien. I Global Risk Report (2017), gjennomført av World Economic Forum, fremgår det at KI er teknologien med flest muligheter, men også størst risiko for negative konsekvenser. Toppledere har flyttet oppmerksomheten fra å bekymre seg for om de skal innføre KI, til å bekymre seg om hvordan de skal anvende KI. Noen av de største barrierene ved innføring av KI i organisasjoner er utfordringer knyttet til lovverk, regler og tilgang på data (Brenna, Danesi, Finch, Goehring, & Goyal, 2018). I tillegg kan ansvar relatert til KI være en utfordring (Tennøe, 2017).

KI er integrert i store deler av hverdagen og har allerede ført til endringer (Stone, et al., 2016 som sitert i Dignum V., 2018). Ettersom utviklingen av KI går raskere enn noen gang, blir det viktig å vurdere ansvar ved bruk av teknologien (Dignum V. , 2017). Egenskapene til KI, spesielt når de brukes i kritiske situasjoner, utfordrer tradisjonelle tilnærminger til regulering og styring. Regelverk blir liggende bak da utviklingen av KI går raskere enn utviklingen av lover og regler (Burton, et al., 2020). Derfor blir ansvar spesielt viktig. Coeckelbergh (2019) sier temaet ansvar trenger mer diskusjon relatert til KI. Ettersom KI lærer og utfører handlinger, blir det viktig å diskutere hvem eller hva som har ansvar for KI (Coeckelbergh, 2019).

Vi vet at det eksisterer mye litteratur om KI (f.eks. Poole & Mackworth, 2017; Russel & Norvig, 2016; Sarangi & Sharma, 2018), og at det eksisterer mye litteratur om ansvar (f.eks. Fischer & Ravizza, 1998; Holdorf & Greenwald, 2018; Lindkvist & Llewellyn, 2003). I tillegg er det tidligere forskning om KI og ansvar som viser til at ansvar er et viktig aspekt ved utvikling og bruk av KI (f.eks. Coeckelbergh, 2019; Dignum V. , 2018). Så vidt vi vet fins det derimot lite litteratur som kobler ansvar og KI sammen fra et organisasjonsperspektiv. Vi ønsker dermed å bidra med økt innsikt om organisasjoners ansvar ved bruk av KI.

1.2 Forskningsspørsmål

Vi skal utforske norske organisasjoners ansvar ved bruk av KI. Vårt forskningsspørsmål er:

«Hvordan blir ansvar annerledes ved innføring av kunstig intelligens i norske organisasjoner?»

For å svare på forskningsspørsmålet tar vi utgangspunkt i tre offentlige dokumenter for å kartlegge regelverk og anbefalinger ved bruk av KI. Videre vil vi samle inn data fra norske

organisasjoner som har tatt i bruk en form for KI. For å besvare forskningsspørsmålet vil vi analysere og diskutere funn fra dokumentene og organisasjonene.

1.3 Metodisk tilnærming

For å svare på forskningsspørsmålet har vi benyttet en abduktiv forskningstilnærming. Ved å gjennomføre en eksplorerende studie med kvalitativ data hadde vi handlingsrom til å utforske ulike funn, som var hensiktsmessig da ansvar og KI er komplekse tema. Vi gjennomførte både en dokumentstudie og en casestudie. Vi startet med en gjennomgang av sekundærdata for å undersøke regelverk og anbefalinger som er relevante for norske organisasjoner ved innføring av KI. Videre gjennomførte vi to ustrukturerte intervju, med en KI-rådgiver og et tilsynsorgan, for å få økt kunnskap om viktige aspekter knyttet til KI og ansvar, samt anbefalinger til organisasjoner å intervju. Deretter samlet vi inn primærdata gjennom semistrukturerte intervju med ansatte som jobber direkte med KI i ulike norske organisasjoner.

1.4 Vårt bidrag

Målet er med utredningen er å bidra med innsikt i hvordan organisasjoners ansvar blir annerledes ved innføring av KI. For å få et helhetlig bilde har vi inkludert både offentlige og private organisasjoner. Ansvarsdiskusjoner kan være en avgjørende faktor for om organisasjoner ønsker å satse på og innføre teknologi som KI. Skal Norge bli ledende på området, slik som regjeringer ønsker, er det viktig at utfordringer rundt ansvar blir diskutert. Vårt mål er å bidra til en ansvarsdiskusjon slik at det kan bli enklere for organisasjoner å forholde seg til ansvar relatert til KI. Utredningen er dermed ikke en løsning på et komplekst og vanskelig problem, men forhåpentligvis et skritt i riktig retning.

1.5 Disposisjon

Utredningen er strukturert i seks kapitler hvor første kapittel er introduksjon. I kapittel to presenterer vi relevant teori og forskning, herunder hva kunstig intelligens er og litteratur om ansvar. Kapittel tre tar for seg metodiske valg og bakgrunn for disse valgene. I kapittel fire presenterer og analyserer vi relevante funn i to deler. Del en består av dokumentstudien og ser på relevant regelverk og anbefalinger. Del to består av casestudien og ser på organisasjoners

forståelse av KI og ansvar. Funnene blir diskutert i kapittel fem. I kapittel seks kommer vi med en konklusjon basert på vår diskusjon.

2. Teori: Kunstig intelligens og ansvar

I dette kapitlet presenteres teoretisk grunnlag for å svare på forskningsspørsmålet: «*Hvordan blir ansvar annerledes ved innføring av kunstig intelligens i norske organisasjoner?*». I Kapittel 2.1 presenterer vi hva kunstig intelligens er og hvilken KI vi legger til grunn i utredningen. Videre, i kapittel 2.2, presenterer vi hvordan litteraturen definerer begrepet ansvar før vi avslutter med hvordan vi forstår ansvar. Til slutt presenteres teoretisk referanseramme i kapittel 2.3.

2.1 Kunstig intelligens

For å forstå om ansvar blir annerledes ved bruk av KI er det viktig med en forståelse av KI. Ettersom KI er i stadig utvikling, starter vi med en gjennomgang av historien til teknologien og hvor langt utviklingen har kommet i dag. Denne kunnskapen gir et godt grunnlag til å definere hva KI er. Forståelse av KI bidrar i diskusjonen om hvordan ansvar blir annerledes ved innføring av KI.

2.1.1 Historien til kunstig intelligens

Tanker og ideer om at menneskelig intelligens kan sammenlignes med en maskin spores tilbake til filosofer som levde 400 f.Kr. (Russel & Norvig, 2016). Likevel har det ikke vært mulig å eksperimentere og bygge maskiner som tester hypotesene om intelligent atferd før de siste femti årene. Det er flere ulike fagdisipliner som har bidratt til utviklingen av KI. Deriblant matematikere, filosofer, psykologer, nevrologer, økonomer, biologer og dataingeniører (Buchanan, 2006). Utviklingen av KI har gjennomgått flere sykluser. Noen kjennetegnes av stor optimisme og gode resultater, mens andre kjennetegnes av vanskeligheter og varierende resultater (Russel & Norvig, 2016). Vi deler opp utviklingen i tre perioder med bakgrunn i litteratur.

Den første perioden (fra 1950)

Den første perioden kjennetegnes av entusiasme og store forventninger. Sommeren 1956 ble det første KI-studiet, Dartmouth-konferansen, arrangert (McCarthy, Minsky, Rochester, & Shannon, 1955). Dartmouth-konferansen har senere blitt omtalt som både fødestedet til KI og hvor teknologien fikk sitt navn (Russel & Norvig, 2016; Buchanan, 2006). Etter konferansen

ble det utviklet ulike teknologier som forøkte å imitere hvordan mennesker løser problemer (Russel & Norvig, 2016).

Til tross for stor entusiasme og lovende resultater er historien til KI også preget av perioder med desillusjonert stagnasjon (Kolbjørnsrud, 2017). Store løfter kombinert med mangel på resultater førte til en periode med mindre fremgang på slutten av 70-tallet. Dette ble omtalt som den første KI-vinteren (Sarangi & Sharma, 2018).

Den andre perioden (1980)

Etter den første KI-vinteren kom det en ny periode med fremgang tidlig på 80-tallet. Denne perioden kjennetegnes av stor optimisme rundt to underkategorier av KI; ekspertsystemer og nevrale nettverk (Poole & Mackworth, 2017; Russel & Norvig, 2016). I tillegg begynte organisasjoner i flere land å investere i forskning innenfor KI under denne perioden. KI-industrien fikk en stor fremgang og gikk fra å være en millionindustri i 1980 til å bli en milliardindustri i 1988. Dette førte til at flere hundre selskaper bidro til å bygge roboter, ekspertsystemer og programvarer for å utvikle KI (Russel & Norvig, 2016).

Den nye oppblomstringen førte igjen til store forventninger og optimisme, likevel klarte ikke forskningen å imøtekomme de ambisiøse målene. På slutten av 80-tallet begynte entusiasmen nok en gang å dale ettersom fremgangen ikke levde opp til forventningene (Russel & Norvig, 2016). Utviklingen av KI gikk inn i en ny periode preget av mindre fremgang, nærmere bestemt den andre KI-vinteren.

Den tredje og nåværende periode (fra 1990)

Til tross for den andre KI-vinteren opplevde teknologien ny fremgang på 90-tallet. Bedre bruk av vitenskapelig metodikk, ved eksperimentering og sammenligning av tilnærminger, har ført til at teknologien har utviklet seg raskere det siste tiåret (Russel & Norvig, 2016).

I løpet av denne perioden har det kommet flere eksempler på bruk av KI versus mennesker (Sarangi & Sharma, 2018). Tre milepæler for teknologien var da KI vant over de beste spillerne i spill som sjakk, Jeopardy og Go (Campbell, Hoane Jr, & Hsu, 2002; Ferrucci, 2012; Sautoy, 2019). Det ble også utviklet KI som skal assistere og samhandle med mennesker. Et område innen KI som vokste mye var interaktive kommunikasjonssystemer (Képuska & Bohouta, 2018). Flere store internasjonale teknologiselskaper som Microsoft, Apple, Amazon og Google

har brukt slike systemer til å utvikle virtuelle assistenter (Hoy, 2018). KI har også blitt brukt innen medisin for å blant annet hjelpe til med nøyaktig bildetolkning, som har ført til bedre arbeidsflyt og minimering av feil (Topol, 2019). I tillegg har selvkjørende biler allerede reist flere millioner mil ved hjelp av KI (Etzioni & Etzioni, 2017).

Dessuten økte interessen for KI i denne perioden. I 2017 ble det publisert mer enn åtte ganger så mye litteratur innenfor feltet som i 1996 (Shoham, et al., 2018). I tillegg har internett gjort informasjon lettere tilgjengelig. Forskning er ikke lenger bare for akademiske eksperter, men også for samfunnet generelt. Videre har bruken og mengden av stordata eksplodert. Dette i kombinasjon med økt lagrings- og beregningskapasitet, som er rimeligere enn noen gang, har ført til muligheter innen KI som tidligere ikke har vært til stede (Sarangi & Sharma, 2018).

Til tross for disse fremskrittene har KI fortsatt en lang vei å gå. KI som anvendes i dag, klassifiseres som smal KI. Smal KI er utviklet for å håndtere spesifikke områder og bestemte oppgaver. En KI som fungerer på større og mer avanserte problemer, nærmere bestemt generell KI, har uteblitt. Med de mulighetene som nå er tilgjengelig mener Sarangi og Sharma (2018) at utviklingen av KI vil foregå enda raskere enn tidligere.

2.1.2 Hva er kunstig intelligens?

Med en stadig økende interesse for KI og en utvikling som går raskere enn tidligere, er det sentralt med en forståelse for hva teknologien innebærer. En forståelse av KI er spesielt viktig for å forstå hvordan organisasjoners ansvar blir annerledes ved bruk av slik teknologi. Derfor vil vi definere begrepet KI og gå videre inn på klassifisering og underkategorier av KI.

Begrepet KI ble først definert under Dartmouth-konferansen. KI ble definert som problemet med å få en maskin til å oppføre seg på måter som ville blitt omtalt som intelligente dersom et menneske hadde oppført seg på samme måte (McCarthy, Minsky, Rochester, & Shannon, 1955). Siden den gang har det kommet flere definisjoner på hva KI er og definisjonene «*endrer seg gjerne i takt hva som er teknologisk mulig*» (Kommunal- og moderniseringsdepartementet, 2020, s. 9). En av årsakene til et mangfold av definisjoner er at begrepet «intelligens» er vanskelig å definere i seg selv (Kaplan & Haenlein, 2019). Ifølge Poole og Macworth (2017) har begrepet KI ført til forvirring da det kan tolkes som det motsatte av ekte intelligens. De

beskriver at KI handler om å skape intelligens gjennom noe kunstig, som for eksempel en datamaskin (Poole & Mackworth, 2017).

Russel og Norvig (2016) definerer KI som studien av agenter som mottar oppfatninger fra miljøet og som utfører handlinger. Noen definerer begrepet mer detaljert, slik som Kaplan og Haenlein (2019). De definerer KI som et system som har mulighet til å tolke ekstern data korrekt, og som lærer fra slik data for deretter å bruke disse erfaringene til å nå spesifikke mål og utføre oppgaver gjennom fleksibel adaptasjon. Andre forstår kunstig intelligens som «*IT-systemer som kan sanse, forstå, handle og lære*» (Kolbjørnsrud, 2017, s. 34). I disse definisjonene handler det mer om teknologiens evne til å lære gjennom å tolke data og forstå miljøet, fremfor å måtte ha intelligens.

I denne utredningen vil vi ta utgangspunkt i definisjonen gitt i regjeringens strategi for KI. De har definert begrepet på bakgrunn av definisjonen gitt av EUs ekspertgruppe slik:

«Kunstig intelligente systemer utfører handlinger, fysisk eller digitalt, basert på tolkning og behandling av strukturerte eller ustrukturerte data, i den hensikt å oppnå et gitt mål. Enkelte KI-systemer kan også tilpasse seg gjennom å analysere og ta hensyn til hvordan tidligere handlinger har påvirket omgivelsene» (Kommunal- og moderniseringsdepartementet, 2020, s. 9).

Det er dermed tydelig at det finnes et mangfold av definisjoner av KI. Det som er gjennomgående i litteraturen er at KI er skapt gjennom noe kunstig og som har evnen til å lære og oppfatte, samt bruker disse erfaringer til utføre handlinger. Vi mener det er nettopp dette som skiller KI fra annen teknologi.

2.1.3 Klassifisering av kunstig intelligens

Litteraturen deler KI i *generell kunstig intelligens* (artificial general intelligence) og *smal kunstig intelligens* (artificial narrow intelligence). Disse går også under navnene sterk (generell) KI og svak (smal) KI. Smal KI er trent for å utføre spesifikke oppgaver (Hanssen & Nichele, 2019). Deep Blue er et eksempel på smal KI. Den er svært godt trent på å spille sjakk, men kan ikke utføre andre intelligente handlinger. Generell KI kan derimot anvendes på en rekke uspesifiserte problemer, ta egne beslutninger og lære av seg selv. Med andre ord ha evnen til å

lære seg å lære i tillegg til å løse problemer, resonnere, planlegge og modellere (Hanssen & Nichele, 2019). En generell KI simulerer den menneskelige hjerne, men oppfører seg nødvendigvis ikke med menneskelig psyke eller moralske verdier (Meek, Barham, Beltaif, Kaadoor, & Akhter, 2016). Generell KI er utrolig avanserte systemer som foreløpig ikke eksisterer, men som det forskes mye på. Smal KI er derimot allerede godt implementert og brukes innen flere områder som for eksempel innen finans, medisinsk behandling, matematiske beregninger, værprediksjon og flere andre områder (Meek, Barham, Beltaif, Kaadoor, & Akhter, 2016). Ettersom generell KI foreløpig ikke eksisterer, vil vi ta utgangspunkt i smal KI i denne utredningen.

Underkategorier av kunstig intelligens

Videre kan KI deles i flere underkategorier. Dette er blant annet maskinlæring (ML), ekspertsystemer, naturlig språkgenerering (NLG) og språkprosessering (NLP), kunstige nevralt nettverk (KNN), datasyntese (Computer Vision), robotikk, planlegging og optimalisering (Russell & Norvig, 2016). Maskinlæring er en av de meste omtalte teknikkene i dagens litteratur. Når det er snakk om bruk av KI i dag er det som regel bruk av løsninger som baserer seg på ML (Kommunal- og moderniseringsdepartementet, 2020). Det er for eksempel ulike ML teknikker Facebook, Amazon og Netflix benytter seg av for å kunne tilby personalisert reklame (Müller & Guido, 2018). Med dette som bakgrunn ønsker vi å rette fokus på ML videre i utredningen. Vi vil dermed ikke gi en detaljert forklaring av de resterende underkategoriene.

Maskinlæring

Maskinlæring er et av dagens teknologiske felt som vokser hurtigst. Det er et vitenskapsfelt som ligger i krysningspunktet mellom informatikk og statistikk, og i kjernen til KI og datavitenskap. ML handler om å trekke kunnskap fra data, samt å bygge datamaskiner som forbedrer seg automatisk gjennom læring og erfaring (Jordan & Mitchell, 2015; Müller & Guido, 2018). Maskinene kan lære på ulike måter, som gjør at ML deles opp i ulike metoder etter hvilke problemer som løses. Dette er *veiledet læring* (supervised learning), *ikke-veiledet* (unsupervised learning), *læring med forsterkning* (reinforcement learning) og *dyp læring* (deep learning). Vi vil i det følgende presentere disse, men utelater de tekniske aspektene da det faller utenfor vårt omfang.

Veiledet læring er en metode som bruker et datasett med inngangsverdier og tilhørende utgangsverdier. Datasettet deles opp i trening- og testdata som muliggjør at modellen kan trenes

opp ved bruk av treningsdata før modellens grad av nøyaktighet måles ved bruk av testdata. Grad av nøyaktighet testes gjennom å sammenligne estimerte utgangsverdier modellen gir med kjente utgangsverdier fra testdata. *Ikke-veiledet læring* er en metode som bruker et datasett med inngangsverdier uten tilhørende utgangsverdier (Müller & Guido, 2018). Modellen forsøker å forstå sammenhengen mellom ulike variabler i data, da det ikke eksisterer utgangsverdier som kan veilede modellen (James, Witten, Hastie, & Tibshirani, 2013). *Forsterket læring* baserer seg på ikke-veiledet læring og er derfor ikke avhengig av utgangsverdier. Maskinen belønnes eller straffes etter hvorvidt iterasjoner fører nærmere en forhåndsdefinert ønsket løsning eller ikke. Metoden skiller seg derfor ut fra ikke-veiledet læring da maskinen får tilbakemelding som gir maskinen mulighet til å lære av egne feil og på den måten forbedre modellen (Russel & Norvig, 2016). I dag finner vi for eksempel bruk av veiledet læring i sortering av e-post, ikke-veiledet i identifisering av kundegrupper og læring med forsterkning i AlphaGo (Kommunal- og moderniseringsdepartementet, 2020).

Videre er *dyp læring* en maskinlæringsmetode som skiller seg ut fra de overnevnte metodene da den prøver å imitere den menneskelige hjerne. For å forstå dyp læring er det hensiktsmessig å forklare hva kunstig nevralt nettverk (KNN) er. KNN består av et inngangslag av nevroner, ett til tre skjulte lag av nevroner, og et siste lag med utgangsnevroner (Wang, 2003). Hver nevron produserer en egen verdi og er basert på prosessen i den menneskelige hjernen. Flere lag med nevroner gjør modellen mer kompleks. Nettverket er generelt ikke designet for å være en realistisk etterligning av den biologiske hjernen, men heller en enklere fremstilling av den (Russel & Norvig, 2016). Dyp læring kan sies å være en underkategori av KNN, men for at det skal regnes som dyp læring må det være mer enn ett skjult lag med nevroner. Dermed blir dyp læring en kompleks modell bestående av mange lag koblet sammen (Goodfellow, Bengio, & Courville, 2016). Dyp læring brukes blant annet i bildegjenkjenning (Kommunal- og moderniseringsdepartementet, 2020).



Figur 1: Illustrasjon av sammenhengen mellom KI, ML og dyp læring

2.1.4 Kritikk mot kunstig intelligens

Den historiske utviklingen til KI, samt senere års fremskritt, viser at KI har fått økt oppmerksomhet av akademikere, organisasjoner og generelt i samfunnet. Mye av oppmerksomheten har vært på bakgrunn av stor optimisme knyttet til mulighetene teknologien medfører. Imidlertid har det også kommet økt oppmerksomhet rundt utfordringer som følger av teknologien.

Når en teknologi utvikler seg i et raskt tempo, slik KI gjør nå, vil det oppstå nye utfordringer det må tas stilling til. Teknologien kan for eksempel enkelt og raskt utføre oppgaver som tidligere har vært vanskelige og tidkrevende for mennesker å løse (Mitchell, 2019). Dette er en stor fordel, men kan også føre til problemer hvor brukeren sitter igjen uten forståelse for hvorfor utfallet av løsningen ble slik den ble (Kommunal- og moderniseringsdepartementet, 2020). Videre finnes det eksempler hvor KI-systemer er lært opp av designers subjektive meninger eller av historiske data som inneholder skjevheter. Dette kan føre til at KI-systemer lærer og viderefører informasjon som i utgangspunktet ikke ble sett på som riktig (Danks & London, 2017). Kritikere tar også opp spørsmål som svikt i autonome maskiner og tap av personvern (Buchanan, 2006). Dette er bare et utvalg av utfordringer som følger av KI.

2.1.5 Kunstig intelligens i vår utredning

Basert på litteratur presentert i kapittel 2.1 er det tydelig at KI er en kompleks og omfattende teknologi. Som nevnt er KI skapt gjennom noe kunstig. Vi mener det som skiller KI fra annen teknologi er KIs evne til å lære og oppfatte, samt bruke disse erfaringene til å utføre handlinger.

Når det er snakk om KI i dag, er det som regel snakk om bruk av maskinlæringsmetoder (Kommunal- og moderniseringsdepartementet, 2020). Likevel finnes det flere teknologiske løsninger som kategoriseres som en underkategori av KI. For å svare på forskningsspørsmålet, om ansvar blir annerledes, har vi derfor ikke utelatt organisasjoner basert på hvilken KI de bruker. I utredningen har vi valgt å ha fokus på at organisasjonene har en eller annen form for KI, fremfor om organisasjonen har en spesiell type KI.

Uavhengig av underkategori vil ansvar være en del av å bruke KI. Ifølge Buchanan (2006) fører fremgangen og utviklingen av KI til et økt ansvar. Dette ansvaret gjelder spesielt overfor samfunnsmessige konsekvenser, å opplyse og bidra med informasjon til beslutningstakere og ovenfor samfunnet generelt (Buchanan, 2006). Det er dermed behov for diskusjoner om ansvar (Coeckelbergh, 2019). Ansvarsdiskusjoner er også et viktig aspekt for å forstå om ansvar endres ved å ta i bruk KI med tanke på at KI har evnen til å lære og oppfatte, og utføre handlinger.

2.2 Ansvar i tre perspektiver: juridisk, moralsk og styring

For å svare på forskningsspørsmålet «*Hvordan blir ansvar annerledes ved innføring av kunstig intelligens i norske organisasjoner?*» går vi nærmere inn på begrepet *ansvar*. Vi vil først gjennomgå hvordan litteratur definerer begrepet ansvar. Deretter vil vi presentere ansvar fra tre perspektiver. Vi har valgt å se på ansvar fra perspektivene juridisk, moralsk og styring da de gir et helhetlig bilde av organisasjoners ansvar. Til slutt vil vi presentere hva vi legger ned i begrepet ansvar i utredningen.

2.2.1 Hvordan litteratur definerer ansvar

Ansvar defineres ulikt i den eksisterende litteraturen. Dette kommer frem av Holdorf og Greenwald (2018), som skriver at den eksisterende litteraturen har ulike definisjoner og beskrivelser av ansvar. Fischer og Ravizza (1998) poengterer at ansvar blant annet kan deles inn i juridisk ansvar, bedriftsansvar og moralsk ansvar. Hvordan samfunnet ser på ansvar endrer seg over tid. I tillegg eksisterer ansvar i mange former og endres av hvilken sammenheng man befinner seg i (Sinclair, 1995).

Syse og Olsen (2013) forklarer at ansvar betyr å stå til svars og at det handler om en evne til å kunne svare for seg. Videre fremhever de poenget gjennom å bruke det engelske ordet for

ansvar: «*responsibility, det vil si respondere ('svar') + ability ('evne')*» (s. 18). Roberts og Scapens (1985) mener å holdes ansvarlig kan beskrives som å måtte gjøre rede for og rettfærdiggjøre dine handlinger overfor andre. Dette kommer også frem av sitatet til Tranøy (2018): «*forpliktelse til å stå til rette for, gjøre rede for eller bære utgiftene for noe*». Svensson (2011) mener også ansvar betyr å stille noen til svare og at ansvar er knyttet til enkeltpersoners handlinger. Ordboken Merriam-Webster (2020) beskriver ansvar gjennom flere punkter. Først beskrives begrepet som en tilstand ettersom man fikk noe til å skje. Ansvar innebærer dermed en handling. Deretter beskrives det som en plikt eller en oppgave, enten fordi det er forventet eller fordi det er et krav. Det tredje punktet omhandler at ansvar fører til noe man bør gjøre, enten fordi det er moralsk riktig eller fordi det er et lovlig krav. I tillegg beskrives ansvar som en egenskap (som sitert i Holdorf & Greenwald, 2018).

Til tross for ulike definisjoner og beskrivelser av begrepet ansvar, ser det ut til å være noen likheter. Det virker som det er enighet om at ansvar er relatert til hendelser og at ansvar handler om å stå til svare eller redegjøre for handlinger.

Handlefrihet

Ettersom ansvar er knyttet til handlinger, står handlefrihet sentralt. Vetlesen (2007) mener å tilskrive personer ansvar for sine handlinger er et grunnleggende trekk ved menneskets væremåte. Han skriver imidlertid at en generell forutsetning for å fordele ansvar, både til hverandre og oss selv, er at vi er frie vesener som har fullstendig kontroll over våre handlinger. Videre presiserer Vetlesen (2007) at det ikke fins noe bevis for at vi har en slik fullstendig kontroll, og at vi til tross for dette fordeler ansvar som om det skulle vært tilfellet. Svensson (2011) på sin side forklarer at handlefrihet eksisterer når vi har forskjellige alternativer å velge mellom. Han skriver at begrepet ansvar er «*rent allmänt knutet till individers handlingar. Frihet att handla föreligger när vi kan välja olika alternativ, och vi har då ansvar för handlingarna och deras konsekvenser för andra och oss själva*» (s. 307). Handlefrihet spiller dermed en sentral rolle for hvorvidt personer kan tilskrives ansvar for sine handlinger.

Konsekvenser

Det er imidlertid delte meninger rundt konsekvenser av handlinger. Fischer og Ravizza (1998) viser til to ulike syn på ansvar. De forklarer at et vanlig syn på ansvar er knyttet til negative konsekvenser. Videre beskriver de et bredere syn på ansvar, hvor ansvar assosieres med både positive og negative konsekvenser. Holdorf og Greenwald (2018) viser også til at det finnes

både positive og negative konsekvenser for utfall av handlinger. Positive konsekvenser kan for eksempel innebære takknemlighet, respekt og annerkjennelse, mens negative konsekvenser handler om straff og fordeling av skyld (Fischer & Ravizza, 1998). Hvorvidt konsekvensene er negative eller positive, avhenger av handlingens utfall.

Oppsummering

Den eksisterende litteraturen definerer ansvar ulikt. Dog, er det enighet om at ansvar handler om å stå til svars for sine handlinger, samt å være ansvarlig for konsekvensene som følger av dem. Det å tilskrive personer ansvar er et grunnleggende trekk ved menneskers væremåte. En forutsetning for å fordele ansvar er at vi som mennesker har handlefrihet. Handlefrihet eksisterer når vi kan fritt velge mellom ulike alternativer. Det er imidlertid uenighet om ansvar handler om kun negative konsekvenser eller både negative og positive konsekvenser. I tillegg vil synet på ansvar ofte endre seg over tid og avhenger av hvilken sammenheng man befinner seg i.

2.2.2 Juridisk ansvar

Juridisk ansvar handler om å ta konsekvensene av handlinger eller unnlaterelser som har ført til skade. Konsekvensene kan for eksempel komme i form av straff eller erstatningsplikt (Tranøy, 2018). Videre er et vilkår for å tildele juridisk ansvar at man etter lovverket kan tilregnes som skyldig, enten forsettlig eller uaktsomt. I noen tilfeller kan man også pålegges et erstatningsansvar for handlinger gjort av andre enn seg selv. For eksempel kan en arbeidsgiver tilskrives ansvar for feil gjort av sine ansatte når ansatte har vært på jobb (Falkanger, 2014). Dette kommer frem i skadeerstatningsloven § 2-1. Juridisk ansvar handler dermed om å ta konsekvensen av handlinger eller unnlaterelser som har ført til skade, gitt at man etter lovverket kan regnes som skyldig.

Ifølge norsk lov defineres selskap som et rettssubjekt, og kan dermed ilegges straff etter straffeloven av 1902 (SNL, u.d.). Videre kommer det frem av aksjelovens § 17-1 første ledd at det kan kreves erstatning av tap som skyldes handling eller unnlattelse av handling foretatt av enkeltpersoner med ulike roller i et aksjeselskap. En slik rolle er for eksempel et styremedlem hvor det i aksjeloven er definert at alle medlemmer av et styre kan bli erstatningsansvarlige. Videre viser aksjelovens § 17-1 andre ledd at skader som definert i § 17-1 første ledd også kan føre til erstatning for den som har medvirket til skaden (Østrem & Schwerdt, 2017). I tillegg til

lovene må offentlige organisasjoner forholde seg til hjemler. En hjemmel er et krav til rettsgrunnlag for en handling eller beslutning (Reusch, 2017). Dermed kan både organisasjoner og enkeltpersoner med ulike roller i organisasjoner bli straffet for lovstridige handlinger.

Organisasjoner er gjennom norsk lov pålagt å følge personopplysningsloven. Loven består av nasjonale regler og EUs personvernforordning (også kalt GDPR). Nærmere forklart ga EU ut et sett med regler som gjelder for alle EU og EØS-land som de kaller GDPR. Denne ble gjort til norsk rett gjennom personopplysningsloven som trådte i kraft 25. mai 2018. Loven handler om behandling, altså innsamling og bruk, av personopplysninger og gir enkeltpersoner en rekke rettigheter for sine personopplysninger (Datatilsynet, 2018; Datatilsynet, 2019). Organisasjoner er pålagt å følge personvernprinsippene i loven. I mange sammenhenger utvikles KI-løsninger av data som inneholder personopplysninger. Dermed er personopplysningsloven viktig når organisasjoner tar i bruk KI, da organisasjonene er avhengig av data for å utvikle KI-løsninger (Kommunal- og moderniseringsdepartementet, 2020).

Oppsummering

Juridisk ansvar oppsummeres som å ta konsekvenser for handlinger eller unnlaterelser som har ført til skade, gitt at det er hjemlet i lovverket. Ifølge norsk lov kan organisasjoner ilegges straff med hjemmel iblant annet straffeloven, aksjeloven eller personopplysningsloven. Det er dermed tydelig at organisasjoner har et juridisk ansvar og i visse tilfeller kan ansvaret også falle på personer med ulike roller i organisasjonen.

2.2.3 Moralsk ansvar

Holdorf og Greenwald (2018) forklarer at moralsk ansvar handler om å tilskrive handlinger til en person og holde personen ansvarlig for disse handlingene. På den måten blir den aktuelle personen ansvarlig for noe eller noen (Ricoeur, 1992; Robinson, 2009; Shoemaker, 2011; Twiss, 1977 som sitert i Holdorf & Greenwald, 2018). Moralsk ansvar defineres dermed relativt likt begrepet *ansvar* da moralsk ansvar også er relatert til hendelser utført av noen som blir ansvarlig.

Holdorf og Greenwald (2018) forklarer videre at tema som går igjen i litteratur om moralsk ansvar er en følelse av forpliktelse, være ansvarlig ovenfor andre når de oppfyller sine forpliktelser og å vise omtanke for andre (Ricoeur, 1992; Robinson, 2009; Shoemaker, 2011;

Twiss, 1977 som sitert i Holdorf & Greenwald, 2018). Tranøy (2018) beskriver moralsk ansvar på følgende måte: «*Moralsk ansvar innebærer forpliktelsen til å forsvare eller rettferdiggjøre handlinger under henvisning til en moralsk norm, regel eller autoritet, for eksempel samvittigheten*». Moralsk ansvar går dermed utover juridisk ansvar, ettersom moralsk ansvar også omhandler elementer som moralsk norm, samvittigheten, følelse av forpliktelse og omtanke ovenfor andre.

Det moralske aspektet er sterkt tilknyttet etikk da etikk er læren om moral (Sagdahl, 2020). Dette gjør at begrepene ofte brukes om hverandre. Moral betegner verdier og normer vi praktiserer og forholder oss til, mens etikk er å reflektere over eller diskutere disse verdiene og normene (Carson & Kosberg, 2011). Ifølge Dignum (2018) er det essensielt at forskning og utvikling av KI fokuserer på å forstå det etiske aspektet av KI-systemer. Både moral og etikk er relativt enkle tema i et stabilt samfunn, men hendelser som blant annet den industrielle revolusjonen har gjort at både sosial og politisk utvikling har gått fort de siste 300 årene. Dermed kommer den moralske kompetansen stadig mer på etterskudd. Det innebærer at teknologiske forandringer gjør at samfunnet kommer opp i nye situasjoner hvor gjeldene moralske normer blir utfordret (Hartvigsen, Johansen, & Måseide, 2000).

Oppsummering

Moralsk ansvar oppsummeres som en forpliktelse til å forsvare eller rettferdiggjøre handlinger i forhold til moralske normer, regler eller autoriteter. Moralsk ansvar handler om et ansvar utover juridisk ansvar ettersom moralske normer og samvittighet spiller inn. Når organisasjoner tar i bruk KI-løsninger kan det føre til at organisasjoner kommer opp i nye situasjoner hvor de etablerte moralske normene blir utfordret. Det moralske ansvaret har dermed en avgjørende rolle i diskusjoner om hvordan organisasjoners ansvar blir annerledes ved innføring av KI.

2.2.4 Styringsansvar

I styringslitteratur skilles det mellom begrepene *accountability* og *responsibility* når det er snakk om ansvar. Begge begrepene oversettes til ansvarlighet på norsk og brukes ofte om hverandre når man skal forklare hva ansvar er (Lindkvist & Llewellyn, 2003). Dette kan gjøre det vanskeligere å finne en entydig definisjon på begrepet. Med bakgrunn i dette ønsker vi innledningsvis å presentere hvordan litteratur skiller mellom begrepene *accountability* og *responsibility*. Deretter presenterer vi hvordan ansvar defineres i styringslitteratur.

Accountability og responsibility

Accountability omhandler instrumentalitet og eksterne kontroller (Bovens, 1998; Lindkvist & Llewellyn, 2003; Pedersen, 2013). Dette synet tilsvarende Bovens (1998) beskrivelse av *passivt ansvar*, som krever at den enkelte skal redegjøre for handlingen den har gjort, og med det svare på spørsmålet «hvorfør gjorde du det?». Dermed vil accountability også omfatte koblingen mellom autoritet og ansvar, slik som i en jobbsammenheng hvor ansatte blir holdt ansvarlig for ansvar delegert til dem. Accountability kan bli sett på som en del av responsibility, hvor sistnevnte er et bredere begrep som inkluderer personlig moral og verdier (Lindkvist & Llewellyn, 2003). Responsibility kan også beskrives som en dyd eller vilje til å tilpasse handlinger i samsvar med andre parters interesser og normer (Pedersen, 2013).

For å ta hensyn til de ulike begrepene har vi lagt til grunn at accountability relateres til eksterne kontroller, instrumentalitet, samt å redegjøre for handlinger man har gjort. Responsibility er et bredere begrep og handler om ansvar utover accountability. Responsibility inkluderer personlige verdier og moral, samt andre parters interesser. Til tross for at begrepene ofte brukes om hverandre vil vi i denne utredningen se på accountability som en del av responsibility, og omtaler begge under begrepet *ansvar*. Der det derimot er nyttig å klargjøre skillet mellom de engelske begrepene betegner vi accountability som *passivt ansvar* (Bovens, 1998).

Hvordan styringslitteratur definerer ansvar

Ettersom begrepet ansvar brukes i flere tilfeller og varierer mellom ulike fagområder vil nå presentere hvordan styringslitteratur ser på ansvar, utover det overnevnte skillet mellom ansvar og passivt ansvar.

Bovens (1998) forklarer ansvar ved hjelp av fem ulike kategorier: (1) *oppgave*, (2) *passivt ansvar*, (3) *kausalt ansvar*, (4) *kapasitet* og (5) *dyd*. Å forstå ansvar som en *oppgave* handler om ansvar som følger en bestemt rolle, posisjon eller autoritet og blir dermed ofte brukt synonymt med plikt. *Passivt ansvar* handler om legale eller politiske forpliktelser knyttet til en hendelse utført av en person. Ansvar knyttet til hendelser forårsaket av mennesker, ting eller andre omstendigheter går under kategorien *kausalt ansvar*. Videre forklarer Bovens (1998) at ansvar kan forstås som en *kapasitet* i tilfeller hvor ansvar er knyttet til kunnskap, makt eller mentale ferdigheter. Til slutt beskriver han ansvar som en *dyd* i form av egenskaper som at man

tar sine oppgaver og plikter seriøst, tenker før man handler og anser seg selv som ansvarlig for sine handlinger da de kan påvirke andre.

Oppsummering

Fra styringslitteratur kan ansvar forstås på flere måter. I engelsk litteratur skilles det ofte mellom *accountability* og *responsibility*. To begrep som er vanskelig å oversette til norsk. Imidlertid er begge begrepene viktige. Vi omtegner dem som *passivt ansvar* og *ansvar*. Passivt ansvar handler om eksterne kontroller og instrumentalitet. Ansvar omhandler et bredere syn på ansvar som også inkluderer personlige verdier og moral. Dermed vil passivt ansvar være en del av ansvar. I tillegg omfatter ansvar flere kategorier. Dette kan være ansvar som følger av en *oppgave*, *passivt ansvar*, *kausalt ansvar*, *kapasitet* eller *dyd*. Dette er alle viktige aspekter av begrepet som vi legger til grunn for å forstå begrepet *ansvar*.

2.2.5 Hvordan vi definerer ansvar

Vår utredning krever en helhetlig forståelse av organisasjoners ansvar. Vi har derfor valgt å se på ansvar fra tre perspektiver: juridisk, moralsk og styringsansvar. Perspektivene krever imidlertid en generell forståelse av begrepet ansvar. Ifølge Holdorf og Greenwald (2018) er det ulike definisjoner og beskrivelser av ansvar. Basert på gjennomgått litteratur i kapittel 2.2.1, definerer vi begrepet som å stå til svars for sine handlinger og for konsekvensene som følger av dem. Dette stemmer med Syse og Olsen (2013) og Svensson (2011) sine definisjoner av ansvar. Svensson (2011) utyper at ansvar er relatert til individers handlinger og at handlefrihet eksisterer når vi har flere alternativer å velge mellom. Som vist av Holdorf og Greenwald (2018) kan handlinger føre til både negative og positive konsekvenser. Vår definisjon av ansvar er viktig for å forstå juridisk, moralsk og styringsansvar. Videre forklarer vi de tre perspektivene for å forstå organisasjoner sitt ansvar i sin helhet.

Organisasjoner har et juridisk ansvar, da de må følge loven. Ifølge norsk lov defineres et selskap som et rettssubjekt (SNL, u.d.). Dette innebærer at organisasjoner må ta konsekvenser i form av straff eller erstatningsplikt for handlinger eller unnlatelser som har ført til skade. Dette gjelder imidlertid kun dersom organisasjoner etter lovverket regnes som skyldig (Tranøy, 2018). Loven omfatter også ansatte i organisasjoner (Østrem & Schwerdt, 2017). Enkeltpersoner i organisasjoner har dermed også juridisk ansvar. Dette viser at juridisk ansvar er sentralt da organisasjoner kan holdes ansvarlig ifølge loven.

Organisasjoner må ta ansvar utover det som defineres i lov. Dignum (2018) understreker at forskning og utvikling av ny teknologi bør fokusere på å forstå etiske aspekter. Det etiske aspektet handler om hva som er rett og galt uavhengig av lov. Å ta ansvar utover lov innebærer blant annet å ha omtanke for andre og en følelse av forpliktelse for sine handlinger. Forpliktelsen til å rettferdiggjøre handlinger er på bakgrunn av moralsk norm og samvittighet, i tillegg til regler. Denne formen for ansvar omtales som moralsk ansvar (Holdorf & Greenwald, 2018; Tranøy, 2018). Organisasjoner som utvikler og bruker KI-systemer må dermed vurdere ansvar utover loven.

Det er nødvendig å se på styringsansvar, da juridisk og moralsk ansvar ikke omfatter alt ansvar i en organisasjon. Det er nyttig å skille mellom begrepene passivt ansvar og ansvar da begge begrepene omtales i styringslitteratur. Passivt ansvar ses på som en del av ansvar, og viser til eksterne kontroller og instrumentalitet. Ansvar omhandler et bredere syn enn passivt ansvar og inkluderer personlige moraler og verdier (Bovens, 1998; Lindkvist & Llewellyn, 2003; Pedersen, 2013). For et mer nyansert bilde av begrepet ansvar deler Bovens (1998) ansvar i fem underkategorier. Underkategorien oppgave viser til ansvaret ansatte i organisasjoner eller organisasjoner selv får av deres bestemte rolle eller posisjon. Passivt ansvar viser til organisasjoners eller ansattes legale eller politiske forpliktelser. I tillegg spiller underkategorien kausalt ansvar en sentral rolle ettersom organisasjoner eller ansatte har ansvar for hendelser som er forårsaket av dem. De siste underkategoriene, dyd og kapasitet, viser til organisasjoners ansvar utover de førstnevnte underkategoriene, og referer til ansvar som en mental ferdighet og som en egenskap. Samlet sett gir styringsansvar nødvendige aspekter for hvordan organisasjoner kan forstå ansvar.

En helhetlig forståelse av organisasjoners ansvar tydeliggjøres ved hjelp av tre perspektiver, juridisk, moralsk og styring. Til tross for at perspektivene gir en omfattende forståelse på ansvar mener vi alle perspektivene er noe organisasjoner må forholde seg til hvis de skal ta i bruk KI.

2.3 Teoretisk referanseramme

I vår utredning anser vi både juridisk, moralsk og styringsansvar som sentrale for en helhetlig forståelse av organisasjoners ansvar. En slik forståelse er omfattende, og for å være i stand til å forstå respondentenes tolkning av ansvar er det nødvendig å operasjonalisere begrepet ansvar.

For å diskutere hvordan ansvar tolkes i regelverket og organisasjoner har vi valgt å operasjonalisere ansvar i tre deler: (1) *ansvar på forhånd og ansvar etterpå*, (2) *direkte ansvar og medvirkning* og (3) *smalt og bredt ansvar*. De tre perspektivene: juridisk, moralsk og styringsansvar kan være til stede i samtlige deler, men i varierende omfang. Inndelingen gir et mer nyansert bilde av hvordan organisasjoner ser på ansvar og er essensielt for å forstå hvordan organisasjoners ansvar blir annerledes ved innføring av KI.

2.3.1 Ansvar i tre deler

Vi har valgt å ta utgangspunkt i Syse og Olsens (2013) inndeling av ansvar for å diskutere hvordan organisasjoners ansvar blir annerledes ved innføring av KI. De har delt ansvar inn i følgende deler: (1) *ansvar på forhånd og ansvar etterpå*, (2) *direkte ansvar og medvirkning* og (3) *smalt og bredt ansvar*.

Ansvar på forhånd og ansvar etterpå

Den første inndelingen handler om et skille mellom ansvar på *forhånd* og ansvar *etterpå*. Dette skillet er også omtalt som *ex ante*- og *ex post*-ansvar, hvor den førstnevnte betyr på forhånd og den andre betyr etterpå.

Ansvar som er definert på forhånd kan for eksempel følge av en gitt rolle et vedkommende har. Dersom vedkommende er en bedriftsleder, vil det være knyttet ansvar til rollen. Ansvar som er definert på forhånd kan også være bestemt av samfunnets lover, eller følge av generelle etiske forventninger og spesifikke regler for den konkrete jobben (Syse & Olsen, 2013). Et konkret eksempel på ansvar på forhånd er styrets rolle i en organisasjon. Ansvar og oppgaver for styret er nedfelt i aksjeloven og omhandler blant annet organisering av organisasjonen og fastsetting av planer og budsjetter (Ness, 2002). Aksjeloven § 6-14 første og fjerde ledd viser at daglig leder står for den daglige ledelse av selskapet og skal sørge for at selskapets regnskap er i samsvar med lov og forskrifter. Utover det som følger av lovgivning kan organisasjoner ha eierstyring og selskapsledelse som klargjør rolledelingen mellom aksjeeiere, styre og daglig ledelse (NUES, 2018). Slike definerte roller vil være ansvar man er klar over på forhånd og kategoriseres dermed som ansvar på forhånd.

Ansvar etterpå handler om ansvar i ettertid av en hendelse. Ansvar etterpå kan til en viss grad være definert av ansvar på forhånd. En regnskapssjef kan for eksempel måtte påta seg ansvaret for at organisasjonen ikke har levert et regnskap til myndigheter og aksjonærer i tide. Dette er fordi det er mest sannsynlig regnskapssjefen som har hatt ansvar på forhånd for at regnskapet blir levert i henhold til tidsfrister (Syse & Olsen, 2013). I forhold til regnskapet har daglig leder et ansvar på forhånd for at regnskapet er i samsvar med lov og forskrifter, han vil derfor få et ansvar etterpå i tilfeller hvor det ikke er samsvar, jf. aksjeloven § 6-14 fjerde ledd. Ansvar etterpå kan også falle på noen som ikke hadde et ansvar som var definert på forhånd (Syse & Olsen, 2013). Et eksempel på dette kan forklares gjennom det som i norsk rett kalles objektivt ansvar. Det vil si ansvar uten skyld, altså at man på ulovfestet grunnlag kan få et objektivt erstatningsansvar når organisasjonen har gjort skade, opptrådte som et faremoment eller førte til økt risiko for omgivelsene (Lødrup & Gisle, 2010). Ansvar etterpå handler dermed om å finne ut hvem som hadde ansvaret i ettertid av en hendelse.

Direkte ansvar og medvirkning

Syse og Olsen (2013) skiller videre mellom *direkte ansvar* og *medvirkning*. «*Direkte ansvar er den typen ansvar som oppstår ved en direkte sammenheng mellom årsak og virkning*» (Syse & Olsen, 2013, s. 19). Direkte ansvar kan også gå under navnet kausalansvar hvor kausal betyr årsak eller grunn som igjen viser til at det må være et påvirkningsforhold mellom to fenomener (Dahlum, 2018). På en annen side finnes det kontekster hvor det er flere agenter som bidrar til et utfall og hvor det av den grunn ikke blir like tydelig hvem som har et direkte ansvar (Lagnado, Gerstenberg, & Zultan, 2013). Dette faller under kategorien medvirkning. Medvirkning omhandler også bistand, forledelse eller oppmuntring til en straffbarhandling som noen andre gjør, samt å unnlate å gripe inn hvis man har en forpliktelse til å gjøre det (Elden, 2020). Medvirker man til alvorlige lovbrudd vil man i de fleste tilfeller straffes på lik linje som en som er direkte ansvarlig (Lovdata, 2015).

Smalt og bredt ansvar

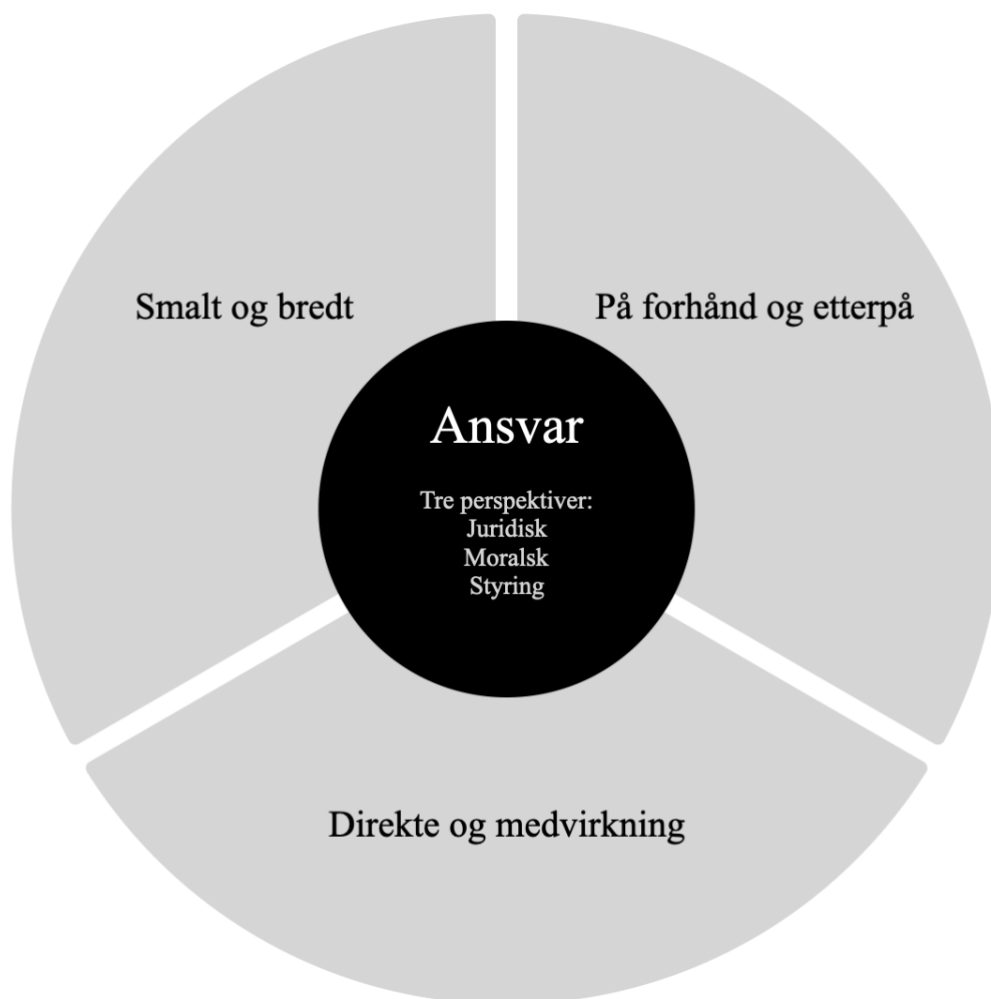
Den siste inndelingen omhandler et skille mellom *smalt* og *bredt ansvar*. Smalt og bredt ansvar blir ofte diskutert i forhold til hva en organisasjon er ansvarlig for. Mer bestemt går dette på hva som faller innenfor organisasjoners ansvarsområde og hva som faller utenfor og kan beskrives som andres ansvar (Syse & Olsen, 2013). I litteratur omtales smalt og bredt ansvar ofte som et shareholder view og et stakeholder view, noe vi videre vil omtale som *aksjonærteorien* og *interessentteorien*. Førstnevnte sier at lederens plikt er å maksimere

aksjonærenes avkastning (Smith, 2003). Martin (2010) støtter aksjonærteorien med bakgrunn i at organisasjoner må velge hvilke mål organisasjonen skal fokusere mest på, da det ikke fungerer å optimalisere flere mål samtidig. I slike tilfeller vil det ene målet alltid komme på bekostning av det andre. Gjennom å rette fokuset mot et konkret mål mener han at andre og mindre viktige mål vil komme av seg selv. Dermed faller aksjonærteorien under et smalt syn på ansvar.

Med et bredt syn på ansvar blir organisasjoner sett på som en aktør i et miljø. Dette kalles *interessentteorien* da den handler om å finne en balanse mellom aksjonærenes økonomiske interesser og interessene til andre interessenter i miljøet. Med dette synet tar organisasjoner hensyn til andre interessenter selv om det går på bekostning av aksjeeierens avkastning. Andre interessenter er blant annet ansatte, kunder og lokalsamfunnet (Smith, 2003). For mange organisasjoner vil det være naturlig å tenke på ansvar utover eierne da det er flere parter involvert i organisasjonen. Ofte kan det være parter organisasjonen er avhengig av for å lykkes. Ettersom bredt ansvar omfavner mye kan det være vanskelig å definere hvor grensen for eget ansvarsområde går, samt hva det er greit å gi andre ansvar for (Phillips, Freeman, & Wicks, 2015). Hvordan man løser dette vil avhenge av konteksten og innebærer som regel et delt ansvar for det omkringliggende miljøet (Syse & Olsen, 2013).

Det finnes dermed ulike inndelinger av ansvar. Videre presiserer Syse og Olsen (2013) at det er mange som kan ta ansvar. Poenget deres er at det ikke trenger å bare være lederposisjoner eller styret som tar ansvar, men også enkeltpersoner. Imidlertid påpeker de at det å ta ansvar krever at det er kultur for at det er mulig å påta seg ansvar. Ansvar handler dermed om mer enn å si ifra eller handle dersom noe er galt, men også om å skape en kultur som gjør det mulig å handle (Syse & Olsen, 2013). Når organisasjoner tar i bruk KI er det derfor viktig at det er en kultur som gjør det mulig å handle og påta seg ansvar, spesielt med tanke på at det ikke trenger å bare være lederposisjoner eller styret som tar ansvar.

Modellen under illustrerer sammenhengen mellom de tre perspektivene på ansvar og Syse og Olsens (2013) inndeling av ansvar:



Figur 2: Illustrasjon av teoretisk referanseramme fra Syse og Olsen (2013)

2.3.2 Vår tolkning av inndelingen

Basert på litteratur presentert i kapitlet 2 ønsker vi ta utgangspunkt inndelingen til Syse og Olsen (2013) for å få en god diskusjon rundt ansvar. Som nevnt deler de ansvar inn i følgende skiller: (1) *ansvar på forhånd og ansvar etterpå*, (2) *direkte ansvar og medvirkning* og (3) *smalt og bredt ansvar*. Vi skal bruke referanserammen som en konseptualisering av vårt forskningsspørsmål, i stede for en teoretisk referanseramme som vi ønsker å teste. Vi vil få et bilde av hvordan organisasjonene jobber med og diskuterer ansvar relatert til KI. Diskuterer organisasjoner ansvar både før og etter innføring av teknologien? Har de en rolle i organisasjonen som er direkte koblet til ansvar eller er det opp til hver av de ansatte? Forholder organisasjoner seg kun til ansvar definert av lovverket eller tenker de på ansvar utover dette?

Ved hjelp av Syse og Olsens (2013) første skille, ansvar på forhånd og ansvar etterpå, ønsker vi å se på hvordan organisasjoner forholder seg til ansvar før og etter de tok i bruk KI. Ansvar på forhånd er dermed relatert til hvordan organisasjonene diskuterte ansvar før de startet å bruke KI. Ansvar i ettertid er knyttet til hvordan organisasjonene forholder seg til ansvar etter KI-løsningen er tatt i bruk.

Det neste skillet går på direkte ansvar og medvirkning. Her vil vi koble direkte ansvar til en definert rolle som er direkte koblet til ansvar. Vi skal derfor å se om organisasjonen har en slik rolle når de har innført KI. Et eksempel på dette kan være om organisasjonen har en bestemt ansatt som har ansvaret for KI under utvikling og bruk. Det motsatte, altså det Syse og Olsen (2013) kaller medvirkning, er dermed tilfeller hvor det ikke er tydelig hvem i organisasjonen som har en slik rolle. I slike tilfeller vil det kanskje være opp til hver enkelt ansatt å finne ut av eventuelle tvister.

Ved hjelp av det siste skillet, smalt og bredt ansvar, ønsker vi koble smalt ansvar til juridisk ansvar og bredt ansvar til moralsk ansvar. Smalt ansvar er dermed knyttet til om organisasjonene utelukkende diskuterer ansvar i forhold til å holde seg innenfor lovverket. Et bredt ansvar er koblet til om organisasjoner tenker på ansvar utover lovverket og knytter inn aspekter som samvittighet, moral og omtanke for andre.

Samlet vil de tre delene av ansvar: før og etterpå, direkte og medvirkning og smalt og bredt, gi et helhetlig bilde av hvordan organisasjoner jobber med og diskuterer ansvar. Ved å operasjonalisere ansvar i tre deler vil vi få et nyansert bilde av organisasjoners syn på ansvar. Juridisk, moralsk og styringsansvar kan være til stede i hver inndeling, men i varierende omfang. Inndelingen vil dermed bidra til å forstå organisasjoners tolkning av ansvar, og hvordan organisasjoners ansvar blir annerledes ved innføring av KI.

3. Metode

I dette kapitlet beskriver vi metoden som er brukt for å svare på forskningsspørsmålet: «*Hvordan blir ansvar annerledes ved innføring av kunstig intelligens i norske organisasjoner?*». I det følgende presenterer og redegjør vi for valgene vi har tatt vedrørende forskningsdesign, datainnsamling og datanalyse. Kapitlet består av syv deler. I Kapittel 3.1 forklarer vi forskningstilnærming før vi i kapitlet 3.2 går inn på forskningsdesign. Metoden vi brukte ved innhenting av data presenteres i kapittel 3.3 før det i kapittel 3.4 utdypes hvordan vi analyserte dataene. Deretter evaluerer vi datakvaliteten i kapittel 3.5. Til slutt diskuterer vi forskningsetikk i kapittel 3.6. Siste del, kapittel 3.7, oppsummerer metodiske valg.

3.1 Forskningstilnærming

Forskningstilnærming handler om forskerens tilnærming til utvikling av teori (Saunders, Lewis, & Thornhill, 2016). Saunders et al. (2016) skriver at forskningstilnærming vanligvis deles opp i to tilnærminger; *deduktiv tilnærming* og *induktiv tilnærming*. Ved en deduktiv tilnærming går forskningen fra teori til empiri. Ved en induktiv tilnærming starter forskningen med datainnsamling for å utforske et fenomen eller å oppdage mønstre. Dataene analyseres og undersøkes for å formulere og utarbeide en teori. En tredje tilnærming, som kombinerer de overnevnte tilnærmingene, kalles *abduktiv tilnærming* (Suddaby, 2006 som sitert i Saunders, Lewis, & Thornhill, 2016).

I utredningen utforsker vi hvordan norske organisasjoners ansvar blir annerledes ved å ta i bruk KI. Vi startet utredningen med å gjennomføre en dokumentstudie. Gjennom dokumentstudien identifiserte vi et tema det var interessant å se nærmere på, mer spesifikt ansvar. På bakgrunn av dette var det hensiktsmessig med en teoretisk gjennomgang av temaet. Basert på teori utarbeidet vi en referanseramme som vi brukte for å konseptualisere forskningsspørsmålet. Den teoretiske referanserammen ble brukt til å utvikle intervjuguiden. Videre samlet vi inn primærdata som vi kunne relatere til referanserammen. Primærdata ble deretter analysert og diskutert for å identifisere tema og mønstre for å svare på hvordan ansvar blir annerledes i norske organisasjoner ved innføring av KI.

Vår prosess inneholder dermed både induktiv og deduktiv tilnærming. Vi hadde en iterativ prosess hvor vi gikk frem og tilbake mellom tilnærmingene. Derfor har vår utredning en abduktiv tilnærming (Saunders, Lewis, & Thornhill, 2016).

3.2 Forskningsdesign

Forskningsdesign handler om den generelle planen for hvordan studien går frem for å svare på forskningsspørsmålet (Saunders, Lewis, & Thornhill, 2016). Forskningsdesignet i utredningen består av fire komponenter: formål, metode, strategi og tidshorisont. Vi forklarer komponentene nærmere i kapitlene 3.2.1-3.2.4.

3.2.1 Formål

Saunders et al. (2016) forklarer at formål med studier kan kategoriseres som *eksplorative*, *deskriptive*, *forklarende* eller *evaluerende*. Formål kan også være en kombinasjon av disse. For å svare på forskningsspørsmålet har vi et eksplorativt formål. Eksplorative studier utforsker et fenomen eller tema av interesse for å få økt innsikt gjennom blant annet å stille åpne spørsmål (Saunders, Lewis, & Thornhill, 2016). Vår utredning kategoriseres som eksplorativ da vi utforsker hvordan norske organisasjoners ansvar blir annerledes ved innføring av KI (Johannessen, Tufte, & Christoffersen, 2016). I samsvar med teori undersøkte vi temaene nærmere gjennom å stille åpne spørsmål (Saunders, Lewis, & Thornhill, 2016).

Vi ønsket å få økt innsikt i hva regelverk og retningslinjer legger vekt på ved bruk av KI. Derfor analyserte vi tre offentlige dokumenter relatert til bruk av KI i Norge. Deretter utforsket vi hvordan organisasjoner forstår ansvar relatert til KI. Selv om KI og ansvar er temaer som er diskutert i eksisterende litteratur, er det lite utforsket hvordan ansvar blir annerledes ved innføring av KI i norske organisasjoner. Vi ønsker dermed å bidra til styringslitteraturen med å belyse fenomenet ytterligere.

Gjennom et eksplorativt formål fikk vi mulighet til å være fleksible og enkelt tilpasse oss endringer (Saunders, Lewis, & Thornhill, 2016). Vi startet med et bredt fokus som snevret seg inn etter hvert som forskningen utviklet seg, som korresponderer med eksplorative studier (Saunders, Lewis, & Thornhill, 2016). Innledningsvis i forskningsprosessen analyserte vi tre dokumenter, samt gjennomførte to ustrukturerte intervjuer med eksperter. Bidraget fra analysen

og intervjuene ga oss som forskere nødvendig informasjon over dagens situasjon, slik at vi fikk konkretisert våre tema ytterligere. Saunders et al. (2016) beskriver en slik prosess som en av flere måter å utføre eksplorativ forskning.

3.2.2 Metode

Metodevalg kategoriseres vanligvis som *kvantitativ* eller *kvalitativ* (Saunders, Lewis, & Thornhill, 2016). I utredningen valgte vi å benytte oss av en kvalitativ metode. Kvalitative studier utforsker ikke-numeriske data som deltakeres meninger og sammenhengen mellom ulike meninger (Saunders, Lewis, & Thornhill, 2016). En kvalitativ metode var hensiktsmessig da det lot oss gå i dybden av temaene KI og ansvar. Dette ga oss en bred forståelse av organisasjoners syn på ansvar ved bruk av KI, slik at vi kunne svare på forskningsspørsmålet (Johannessen, Tufte, & Christoffersen, 2016).

En kvalitativ metode var også hensiktsmessig i vår utredning da metoden ikke har en standardisert innsamlingsmetode (Saunders, Lewis, & Thornhill, 2016). Utredningen startet med en dokumentstudie som viste at ansvar er viktig ved bruk av KI. Imidlertid blir ikke ansvar utdypet, som kan gjøre det vanskelig for organisasjoner å forstå hva som legges ned i begrepet. Derfor valgte vi å gjennomføre semistrukturerte intervju slik at vi fikk belyst organisasjoners syn på ansvar gjennom innsikt i personlige meninger og refleksjoner rundt temaet. Ved å stille åpne spørsmål i intervjuene fikk vi innholdsrik data som ga dypere innsikt i hvordan organisasjoner forholder seg til ansvar. En kvalitativ metode ga oss mulighet til å gjennomføre en interaktiv og virkelighetsnær forskningsprosess (Saunders, Lewis, & Thornhill, 2016).

3.2.3 Strategi

Forskningsstrategi er en plan for hvordan forskeren skal gå frem for å svare på forskningsspørsmålet. Valg av strategi er derfor knyttet til forskningsspørsmålet, i tillegg til andre aspekter som blant annet forskningens tilnærming og design (Saunders, Lewis, & Thornhill, 2016). Kvalitative studier bruker som regel forskningsstrategier som *dokumentstudie*, *casestudie*, *etnografi*, *narrativ analyse* eller *grounded theory*. Strategiene er imidlertid ikke gjensidig utelukkende og det er mulig å kombinere ulike strategier i løpet av en studie (Saunders, Lewis, & Thornhill, 2016). For å svare på forskningsspørsmålet har vi kombinert to forskningsstrategier; *dokumentstudie* og *casestudie*.

Dokumentstudie

Dokumentstudie kan være en effektiv strategi, men avhenger av hvordan det passer sammen med forskningsspørsmålet og det å få tilgang til egnede dokumenter (Saunders, Lewis, & Thornhill, 2016). Vi gjennomførte en dokumentstudie for å få en bedre oversikt over hva offentlige dokumenter la vekt på ved bruk av KI. Dokumentstudien ga oss nyttig informasjon vi ikke hadde klart å samle inn selv. Gjennom denne informasjonen fant vi et tema som var interessant å utforske videre. Ettersom dokumentstudien tydeliggjorde at ansvar er viktig ved utvikling og bruk av KI uten å forklare hva ansvar innebærer, ønsket vi å belyse ansvar relatert til KI fra organisasjoners synspunkt. Dermed fant vi det nødvendig å kombinere dokumentstudien med en casestudie.

Casestudie

Ettersom vi utforsket hvordan ansvar blir annerledes ved å innføre KI i norske organisasjoner, var det hensiktsmessig med en casestudie. En casestudie gjør det mulig å få innsikt om temaer fra virkelige omgivelser som kan føre til innholdsrike empiriske beskrivelser og til utvikling av teori (Saunders, Lewis, & Thornhill, 2016). Ved å benytte oss av casestudie fikk vi gjennomført en detaljert og grundig undersøkelse av organisasjoners syn på ansvar ved innføring av KI (Yin, 2014 som sitert i Saunders, Lewis, & Thornhill, 2016). Yin (2014) påpeker at casestudier er relevante ved dagsaktuelle temaer, som samsvarer med at KI har fått økt oppmerksomhet. En casestudie vil dermed hjelpe med å forstå hvordan KI påvirker organisasjoners ansvar.

Yin (2014) skiller mellom to dimensjoner i design av casestudier. Det første skillet er *enkeltcasestudier* versus *flercasestudier* (Johannessen, Tufte, & Christoffersen, 2016). Vi benyttet en enkeltcasestudie. Enkeltcasestudie ga oss mulighet til å studere et fenomen som få har studert tidligere (Saunders, Lewis, & Thornhill, 2016). Vi valgte å se på Norge som en case, hvor KI og ansvar er fenomenet vi skal studere. I motsetning til flercasestudier er enkeltcasestudier ikke like robuste og overbevisende, som reduserer mulighet til generalisering (Johannessen, Tufte, & Christoffersen, 2016; Yin, 2014). Dog, er enkeltcasestudie mindre tid- og ressurskrevende, som er fordelaktig ved en masterutredning (Yin, 2014).

Det andre dimensjonen Yin (2014) skiller mellom er *én analyseenhet* (holistisk) versus *flere analyseenheter* («embedded»). Dersom casestudien studerer en organisasjon på et overordnet nivå og har et helhetlig perspektiv på organisasjonen, har studien med én analyseenhet. (Johannessen, Tufte, & Christoffersen, 2016). Vi benyttet oss av én analyseenhet ettersom vi

analyserte norske organisasjoner fra et helhetlig perspektiv. Selv om vi påpeker forskjeller mellom respondenter fra ulike organisasjoner, er vårt hovedfokus å undersøke hvordan ansvar blir annerledes ved innføring av KI i norske organisasjoner som en helhet.

3.2.4 Tidshorisont

I forskning kan tidshorisont deles inn i *longitudinell studie* eller *tverrsnittstudie*. Hvilken tidshorisont som velges vil blant annet påvirkes av forskningsspørsmålet. Vi har valgt en tverrsnittstudie da vi studerer et fenomen på et bestemt tidspunkt. Utredningen er derfor et øyeblikksbilde, som innebærer at resultater kun gjelder for det bestemte tidspunktet utredningen ble gjennomført (Saunders, Lewis, & Thornhill, 2016).

Dokumentene ble nedlastet i september og intervjuene ble gjennomført i oktober og november 2020. Tverrsnittstudier er vanlig å bruke når et fenomen kan ha ført til en endring. Dette samsvarer med forskningsspørsmålet da vi skulle se hvordan innføring av KI fører til at organisasjoners ansvar blir annerledes (Ghauri & Grønhaug, 2010). En svakhet ved tverrsnittstudier er at fenomenet som undersøkes ikke kan sammenlignes med tilstanden før eller etter studien. Denne svakheten ble mindre tellende ettersom vi så på flere organisasjoner, som sikret variasjoner som muliggjorde sammenligning av kritiske faktorer (Ghauri & Grønhaug, 2010). Den viktigste grunnen til at utredningen ble en tverrsnittstudie er imidlertid tid og ressurser, da masterutredningen utredes på et semester.

3.3 Datainnsamling

Innsamlet data kan kategoriseres som *primær-* eller *sekundærdata* (Saunders, Lewis, & Thornhill, 2016). For å svare på forskningsspørsmålet startet vi med en dokumentstudie. Gjennom dokumentstudien ble det tydelig at ansvar er et viktig aspekt ved innføring av KI, men det var lite informasjon om ansvar utover det. Derfor var det nødvendig med en ytterligere analyse for å innhente informasjon fra organisasjoners sitt syn på KI og ansvar. Derfor har vi både en analyse av sekundærdata og en analyse av primærdata. Ved å gjennomføre to delstudier fikk vi mulighet til å se på både regelverk og praksis.

3.3.1 Sekundærdata: dokumentstudie

Sekundærdata er data som i utgangspunktet ble samlet inn for et annet formål (Saunders, Lewis, & Thornhill, 2016). I vår utredning var det hensiktsmessig å benytte sekundærdata for å få en god forståelse av hva norske organisasjoner må forholde seg til ved innføring av KI. Imidlertid var vår sekundærdata utledet og publisert for et annet formål enn for å svare på vårt forskningsspørsmål. Derfor kan sekundærdataen inneholde informasjonen som er overdrevet eller skjevt i forhold til formålet med utredningen (Ghauri & Grønhaug, 2010). Vi mener sekundærdataen har relevant informasjon for å svare på forskningsspørsmålet ettersom den inneholder informasjon om regler og retningslinjer organisasjoner kan ha nytte av ved innføring og bruk av KI. Videre legger vi til grunn at sekundærdataen ikke er overdrevet da alle dokumentene trekker frem både positive og negative aspekter ved KI.

3.3.2 Dokumentutvalget

Datamaterialet i vårt dokumentutvalg består av tre dokumenter som alle er gratis og offentlig tilgjengelig på internett. Dokumentene ble valgt på bakgrunn av at vi mente de var mest relevante for organisasjoner og aktører i det norske samfunn da de er direkte koblet til bruk av KI i Norge.

Oversikt over dokumentene i innholdsanalysen				
Utgiver	Referanse	Tittel	År	Sider
Kommunal- og moderniseringsdepartementet	(Kommunal- og moderniseringsdepartementet, 2020)	Nasjonal strategi for kunstig intelligens	2020	66
European Commission	(European Commission, 2019)	Ethics Guidelines for Trustworthy AI	2019	39
Datatilsynet	(Datatilsynet, 2018)	Kunstig intelligens og personvern	2018	30

Tabell 1: Oversikt over dokumentene i innholdsanalysen

Nasjonal strategi for kunstig intelligens

Den nasjonale strategien for kunstig intelligens angir hva kunstig intelligens er og betyr. I strategien står det blant annet om Norge sitt grunnlag for å ta i bruk KI, samt hvordan regjeringen skal legge til rette for god infrastruktur for KI. Strategien er rettet både mot offentlig og privat sektor. Strategien legger EUs retningslinjer for pålitelige og tillitvekkende KI til grunn

og fokuserer på viktigheten av ansvarlig og pålitelig KI. Strategien presenterer utfordringer ved utvikling og bruk av KI og poengterer viktigheten av digital sikkerhet. Tidligere digitaliseringsminister, Nikolai Astrup, håper at strategien kan fungere som et slags rammeverk for organisasjoner som ønsker å utvikle og ta i bruk KI.

Ethics Guidelines for Trustworthy AI

Europakommisjonen nedsatte en gruppe med eksperter innenfor fagfeltet KI for å utarbeide etiske retningslinjer for pålitelig og tillitvekkende KI (*Trustworthy AI*). Rapporten er bygd opp i tre kapitler, hvor kapittel en presenterer grunnleggende rettigheter og etiske prinsipper som er sentrale ved utvikling og bruk av KI. Kapittel to lister opp syv krav med bakgrunn i de etiske prinsippene presentert i kapittel en. Disse kravene må være til stede for at KI-systemer skal være pålitelig og tillitvekkende. Videre presenterer ekspertgruppen tekniske og ikke-tekniske metoder som kan være til hjelp for å realisere pålitelig og tillitvekkende KI. I kapittel tre fremlegger de en vurderingsliste som skal hjelpe organisasjoner å oppfylle de syv kravene. Til slutt presenterer de konkrete eksempler om muligheter og risikoer som fremgår ved bruk av KI.

Kunstig intelligens og personvern

Datatilsynets rapport har som mål å beskrive og bidra til forståelse for hvordan personvernet blir berørt ved både bruk og utvikling av KI. I rapporten presenterer Datatilsynet hva KI er, herunder maskinlæring. Deretter trekker de frem utfordringer knyttet til personvernprinsipper og artiklene i personvernforordningen. Rapporten går inn på ulike prinsipper, som rettferdighet, formålsbegrensning, dataminimering og gjennomsiktighet, samt hvordan tilsyn med algoritmene kan gjennomføres. Avslutningsvis presenteres løsninger og anbefalinger for godt personvern i utvikling og bruk av KI.

Samtlige dokumenter tar for seg viktige diskusjoner rundt utvikling og bruk av KI. Det er gjort en innholdsanalyse av alle tre dokumentene. I analysen kom vi frem til at dokumentene peker på ansvar som et viktig aspekt, men at det likevel er begrenset informasjon om aspektet. Da sekundærkildene ikke ga godt nok svar på forskningsspørsmålet fant vi det hensiktsmessig å samle inn ny informasjon for å styrke vår analyse (Ghuri & Grønhaug, 2010). Derfor valgte vi å gjennomføre semistrukturerte intervju.

3.3.3 Primærdata: semistrukturerte intervju

Primærdata er ny data som innsamles spesielt for en spesifikk forskning. Hva forskeren leter etter, spør etter og samler inn avhenger av forskningsspørsmålet. Ved eksplorative studier er intervju en fordelaktig innsamlingsmetode (Saunders, Lewis, & Thornhill, 2016). I vår utredning har vi samlet inn primærdata gjennom syv semistrukturerte intervjuer. I forkant av intervjuene valgte vi å gjennomføre to ustrukturerte intervjuer for å få ytterligere innsikt om KI og ansvar. Vi hadde derfor samtale med et viktig tilsynsorgan og en AI-rådgiver. Gjennom disse samtalene fikk vi et helhetlig bilde av dagens situasjon fra to ulike synspunkt, samt en rekke anbefalinger til hvem det var relevant å intervjuer. Disse samtalene ble ikke transkribert og er ikke en del av datagrunnlaget.

Fordelen med at vi innhentet egne data var blant annet at dataen samsvarer med vårt forskningsspørsmål. På en annen side var det tidkrevende å gjennomføre og transkribere intervjuene (Ghuri & Grønhaug, 2010). Til tross for denne ulempen mener vi at intervju var en egnet metode for datainnsamling, da dataen ga nyttig informasjon for å svare på forskningsspørsmålet.

3.3.4 Intervjukandidater

Johannessen et al. (2016) presiserer at det er «*viktigere å skaffe et relevant utvalg av informanter enn å skaffe mange*» (s. 114). For å svare på forskningsspørsmålet benyttet vi oss av et ikke-sannsynlig utvalg, nærmere bestemt et hensiktsmessig («purposive») utvalg. Ettersom vi utforsket hvordan ansvar blir annerledes ved innføring av KI i norske organisasjoner, var det nødvendig for oss å finne intervjukandidater i norske organisasjoner som har tatt i bruk noen form for KI. For å finne kandidater, samt med tanke på at organisasjoner definerer sine roller ulikt, inneholder utvalget kandidater med ulike stillinger i sine organisasjoner. Hensiktsmessig utvalg gjorde at vi fikk innholdsrik data som gjorde det mulig å svare på forskningsspørsmålet (Saunders, Lewis, & Thornhill, 2016).

Vi mente det var nødvendig å inkludere både private og offentlige organisasjoner for å få en bred forståelse av om ansvar blir annerledes ved bruk av KI. I utgangspunktet ønsket vi like mange respondenter fra private og offentlige organisasjoner, men det viste seg å være vanskeligere å få tilgang hos private organisasjoner. Vi mener dessuten det virket fornuftig å

ha et ekstra intervju med en offentlig aktør da respondentene fra offentlige organisasjoner ga gode og utfyllende svar på hvordan de jobbet med ansvar.

Til tross for et begrenset antall respondenter argumenterer vi for en tilstrekkelig utvalgsstørrelse. Vi har i tillegg til semistrukturerte intervju også samlet informasjon fra dokumenter og to ustrukturerte intervju. Videre nådde vi et metningspunkt ved de to siste intervjuene da de fleste poenger allerede var nevnt i tidligere intervju. Vi anser dermed utvalgsstørrelsen som tilstrekkelig for å få innsikt i hvordan ansvar ble annerledes.

3.3.5 Gjennomføring av intervju

Før de semistrukturerte intervjuene ble gjennomført utformet vi en intervjuguide basert på teori presentert i kapittel 2. Intervjuguiden ligger i vedlegg 1. Ved å bruke intervjuguide fikk vi et godt grunnlag for å få svar på det vi ønsket, samtidig som vi hadde fleksibilitet til å stille oppfølgingsspørsmål ved behov. For å gjøre det mulig å sammenligne mellom intervjuene har vi i hovedsak delt intervjuguiden i to deler: KI og ansvar. Gjennom utviklingen av intervjuguide ble vi bedre forberedt på intervjuene.

Samtlige intervjuer ble gjennomført digitalt på grunn av den pågående pandemien Covid-19. Alle intervjuene var videosamtaler, med unntak av to intervjuer som var lyd samtaler. De semistrukturerte intervjuene ble gjennomført i perioden 26. oktober til 6. november og varte fra 45 til 60 minutter. Vi intervjuet en person i hver av organisasjonene i utvalget. I forkant av intervjuene har respondentene samtykket til et informasjonsskriv. Hvert intervju ble i tillegg startet med en kort innføring av oss selv og utredningen, samt en påminnelse av respondentenes rettighet til anonymitet. Vi la anonymitet til grunn slik at potensiell gjenkjenning ikke stod i veien for ærlige svar. Samtlige respondenter godkjente lydopptak slik at vi under intervjuene kunne holde fullt fokus på respondentene og svarene de ga.

3.4 Dataanalyse

Saunders et al. (2016) forklarer at prosessen med å analysere kvalitative data handler om å oppsummere, kode, kategorisere og gruppere innsamlede data på måter som gjør det mulig å strukturere og svare på forskningsspørsmålet. Vi brukte *tematisk analyse* i våre dataanalyser. Tematisk analyse blir ofte sett på en generisk metode å analysere kvalitative data på. Vi valgte denne metoden ettersom formålet vårt var å finne tema eller mønstre på tvers av de innsamlede

dataene (Saunders, Lewis, & Thornhill, 2016). I det følgende vil vi forklare nærmere hvordan vi gikk frem i dataanalysene.

3.4.1 Dokumentanalyse

Dokumentanalysen består av en *tematisk analyse* av innholdet i dokumentene. En tematisk analyse innebærer tre viktige steg (Saunders, Lewis, & Thornhill, 2016). Det første steget handler om å bli kjent med datamaterialet. Dette gjorde at vi fikk et overordnet perspektiv på relevant informasjonen i dokumentene. Neste steg handler om å kode dataen i kategorier. Her valgte vi å se på ansvar relatert til KI, muligheter med å bruke KI og utfordringer dokumentene mente var viktig ved innføring av KI i organisasjoner. Siste og tredje steg handler om å søke etter tema og sammenhenger i hver kategori. For oss handlet dette steget om å trekke frem de sentrale ansvarsaspekter, muligheter og utfordringer, samt utelate det som var mindre sentralt.

3.4.2 Caseanalyse

Caseanalysen består av en *tematisk analyse* av datamaterialet samlet inn gjennom intervju. Her blir dog de tre stegene gjennomgått noe ulikt. Denne analysen skulle i hovedsak se på organisasjoners forståelse av ansvar relatert til KI.

Vi startet med å transkribere lydopptakene etter gjennomført intervju. Transkriberingen gjorde vi samme dag eller senest en dag etter gjennomført intervju da det var ferskt i minnet. Vi skrev så ordrett som mulig, men utelot muntlige ord som «da» og «ikke sant» for å gjøre det enklere å lese. Dette førte til at vi fikk store mengder med data i form av tekst. For å bli kjent med datamaterialet gjennomgikk vi samtlige transkriberinger hver for oss. Da skrev vi ut hvert intervju slik at vi kunne streke under det vi mente var viktig og skrive notater hver for oss. Videre sammenlignet vi hva vi selv anså som relevant med hverandre, slik at vi fikk med datamateriale den andre overså.

I neste trinn kodet og organiserte vi datamaterialet nærmere. Denne gang handlet det om å finne sammenhenger og underkategorier i relasjon til utredningens forskningsspørsmål. Dette gjorde vi gjennom å klippe ut alle sitater vi mente var viktige. På denne måten ble vi kvitt irrelevant materialet. Videre samlet vi sitater i ulike kategorier, som for eksempel «ansvarsforståelse» og «ansvar i organisasjoner».

Under tredje og siste trinn gjennomgikk vi sitatene i hver kategori og forsøkte å finne mønstre og sammenhenger. Vi anså tema som gikk igjen, eller som ble påpekt som viktige under intervjuene, som relevant informasjon. I empirikapittelet har vi presentert og analysert relevant informasjon fra i intervjuene. I diskusjonskapittelet har vi så koblet hvert funn inn under inndelingen i vår teoretiske referanseramme.

3.5 Forskningskvalitet

Metodens kvalitet er et viktig aspekt ved forskning. Yin (2014) forklarer at det er fire kriterier for å vurdere kvaliteten på casestudier: begrepsvaliditet, intern validitet, ekstern validitet og reliabilitet. Intern validitet er ikke relevant for eksplorative studier, og utdypes derfor ikke (Yin, 2014). De tre resterende kriteriene for å evaluere forskningskvaliteten blir diskutert i kapittel 3.5.1 til 3.5.3. Avslutningsvis vil vi diskutere kvaliteten på dokumentene.

3.5.1 Begrepsvaliditet

Yin (2014) beskriver at begrepsvaliditet handler om å identifisere riktige operasjonelle mål for konseptet som studeres. Det vil si at vår operasjonalisering av begrepet samsvarer i størst mulig grad med det teoretiske begrepet. Denne typen validitet er spesielt utfordrende i casestudier (Yin, 2014). I vår utredning var det vanskelig å operasjonalisere mål for KI og ansvar da begge er komplekse begreper med flere definisjoner i litteraturen. Ved gjennomføring av intervju var det derfor viktig å stille tydelige og nøyaktige spørsmål slik at respondentene klarte å resonnerer rundt temaet (Saunders, Lewis, & Thornhill, 2016). Hvis respondentene ikke forstod hva vi spurte, kan vår data måle noe annet enn hva vi ønsket dataen skulle måle. Dermed kan utydelige spørsmål true begrepsvaliditeten i vår utredning.

For å motvirke trusselen intervjuet vi personer som er knyttet til bruk av KI i sin organisasjon. Dette fordi vi mente en grunnleggende forståelse for KI er viktig for å diskutere hvordan ansvar blir annerledes ved bruk av KI. Vi operasjonaliserte begrepet KI gjennom å spørre om organisasjonen har «noen form for KI» for å ikke utelate noen underkategorier av KI. Videre holdt vi spørsmålene under intervjuene åpne for å få frem respondentenes personlige forståelse og refleksjoner. Dermed fikk vi kartlagt hva respondentene selv la ned i begrepene KI og ansvar. Vi har også fortløpende tatt tak i respondentens utsagn når det oppstod tvil rundt hva de mente, som forhindret misoppfatninger rundt vanskelige begrep.

3.5.2 Ekstern validitet

Ekstern validitet handler om hvorvidt funnene i studien kan generaliseres (Yin, 2014). Yin (2014) kategoriserer ekstern validitet i to kategorier: analytisk generalisering og statistisk generalisering. Vi ser bort fra statistisk generalisering da det vanligvis ikke er relevant for casestudier. Analytisk generalisering handler om hvorvidt funnene fra en casestudie også kan gjelde for andre situasjoner og kontekster, basert på relevansen av teoretiske konsepter og prinsipper (Yin, 2014).

Utvalget består av syv respondenter fra ulike organisasjoner som bidro med sine synspunkter på KI og ansvar. Hvis vi finner sammenhenger på tvers av organisasjonene kan det tyde på at sammenhengene også kan gjelde for organisasjoner utover utvalget. I et slikt tilfelle kan våre funn bidra til styringslitteratur om ansvar relatert til KI. En svakhet med utredningen er imidlertid at utvalget består av et begrenset antall organisasjoner. Dermed kan det være vanskelig å påstå at våre funn gjelder for hele Norge. Selv om utvalget inneholder ulike organisasjoner, herunder både offentlige og private, er det ikke sikkert vi har fått med all relevant informasjon slik at funnene kan analytisk generaliseres til andre organisasjoner i lignende situasjoner.

3.5.3 Reliabilitet

Reliabilitet handler om hvorvidt utredningen og prosedyrene for datainnsamlingen fører til samme resultater dersom en annen ønsker å studere det samme igjen (Yin, 2014). Reliabilitet er dermed relatert til hvordan dataen er samlet inn, hvilke data som brukes og hvordan dataene håndteres. Saunders et al. (2016) beskriver fire utfordringer knyttet til reliabilitet. Utfordringene er *deltakerfeil*, *deltakerskjevheter*, *forskerfeil* og *forskerskjevheter* og utdypes nærmere i de følgende avsnittene.

Deltakerfeil

Deltakerfeil handler om intervjukandidatens ytelse, og kan for eksempel bli påvirket av om deltakeren intervjues på et ubeleilig tidspunkt eller om den har det travelt med å bli ferdig med intervjuet (Saunders, Lewis, & Thornhill, 2016). For å hindre deltakerfeil lot vi respondentene velge et tidspunkt som passet for dem. Samtidig holdt vi oss alltid innenfor avsatt tid. Dette gjorde at vi unngikk å komme opp i situasjoner hvor deltakerne ga ufullstendige svar på grunn

av dårlig tid. Vi avsluttet samtlige intervju med en avsluttende dialog som følte lett og naturlig, som styrker antakelsen om at deltakerne var avslappet og hadde tilstrekkelig med tid.

Imidlertid er det et intervju hvor det var økt risiko for deltakerfeil. På grunn av personlige årsaker gjennomførte vedkommende intervjuet i bilen utenfor legevakten. Det kan dermed kategoriseres som et ubeleilig tidspunkt. Likevel var denne deltakeren engasjert og ga gode og utfyllende svar. Svarene virket også gjennomtenkte, som førte til gode data til tross for ubeleiligheten.

Deltakerskjevhet

Deltakerskjevhet handler om faktorer som kan føre til uriktige svar. Dette kan skje hvis deltakeren sitter en plass hvor andre kan overhøre samtalen slik at det blir ubehagelig å gi ærlige svar, eller hvis deltakeren ønsker om å gi «riktige svar» isteden for ærlige svar (Saunders, Lewis, & Thornhill, 2016). Under intervju kan respondentene ha overdrevet hvor mye organisasjonen diskuterer ansvar da dette var utredningens hovedfokus. Respondentene kan også ha vært tilbakeholdne med å gi informasjon om utfordringer rundt KI og ansvar da de ikke ønsket å sette sin organisasjon i et dårlig lys. I tillegg kan lydopptak ha ført til tilbakeholdenhet da respondentene kan være forsiktige med å gi ærlige svar når de vet samtalen blir tatt opp.

For å redusere skjevheter understreket vi ønsket om ærlige svar og deres egne meninger rundt spørsmålene vi stilte. I tillegg valgte vi å holde organisasjonene og respondentene anonyme, slik at respondentene kunne si hva de mente uten å bekymre seg for ubehageligheter i etterkant. Valget om anonyme organisasjoner var nødvendig for å unngå at svar kobles til enkeltpersoner. Dette var med bakgrunn i at det kan være få personer i organisasjonen som oppfyller kravet om å være koblet til bruk av KI. Ikke minst kan respondentenes interesse og personlige meninger generelt vært over gjennomsnittet positive til ny teknologi da respon er koblet til KI, hvilket kan ha ført til fargede svar.

I tillegg gjorde digitale intervju det vanskeligere å etablere personlig kontakt med respondentene, som kan forhindre respondentene i å åpne opp om ansvar og KI. Digitale intervju gjorde også at vi ikke kunne forhindre faktorer som forstyrret respondentene, da vi ikke hadde kontroll over omgivelsene under intervjuene. Dette gjaldt spesielt i to intervju hvor vi blant annet grunnet manglende webkamera på hjemmekontor ikke kunne se respondentene.

Likevel var en fordel med hjemmekontor at respondenten var i kjente og trygge omgivelser, som reduserte sannsynligheten for at andre var til stede og overhørte samtalen.

Forskerfeil og forskerskjevheter

Forskerfeil omhandler alt som kan påvirke forskerens tolkning. Eksempelvis at forskeren er trøtt, uforberedt eller mistolker noen av deltakerens meninger. *Forskerskjevheter* omhandler faktorer som kan føre til subjektive påvirkninger i den innsamlede dataen, herunder at forskeren lar egne meninger påvirke tolkningen av respondentenes svar (Saunders, Lewis, & Thornhill, 2016). For å minimere forskerfeil og forskerskjevheter har vi forberedt oss til intervjuene ved å lese hvordan intervju bør gjennomføres og hvordan en intervjuguide bør utformes. Videre avklarte vi uklare momenter underveis i intervjuene. Lydopptak av intervjuene sikret også at vi fikk med alt respondentene snakket om. Vi mener også at det at vi har vært to forskere gjennom intervju- og analyseprosessen har bidratt til å redusere mistolkninger. Basert på de overnevnte argumentene mener vi derfor det er minimalt med forskerfeil og forskerskjevheter.

3.5.4 Dokumentkvalitet

Kvaliteten på sekundærkilder er like viktig som kvaliteten på primærkilder. Det er viktig at kildene faktisk gir den informasjonen forskeren trenger for å svare på forskningsspørsmålet. Videre er kvaliteten på sekundærdataen viktig og avhenger av metoden brukt for å samle de inn, samt hvilke kilder som ble brukt (Saunders, Lewis, & Thornhill, 2016). Vi går ut ifra at våre sekundærkilder har høy kvalitet da det som regel er tilfellet med data fra internasjonale organisasjoner og myndigheter (Ghauri & Grønhaug, 2010). Dokumentene er blant annet samlet inn av eksperter gjennom strengt regulerte metoder og bidrar positivt da de tilbyr informasjon som ville vært utfordrende for oss å samle inn i løpet av utredningens tidsaspekt.

Dokumentene er utarbeidet for et annet formål og kan derfor inneholde skjevheter i forhold til utredningens formål. Eksempelvis har regjeringens strategi et formål om å oppmuntre til bruk av KI slik at Norge kan bli et godt forbilde på området. Vi var kritiske til om strategien var skjev gjennom å fokusere mye på de positive aspektene ved KI, men mener det ikke var tilfellet da strategien også tar opp en rekke utfordringer og negative utfall. Videre kan dokumentene inneholde aggregert informasjon. Det er ikke gitt at aggregering som passer til dokumentene passer til utredningen. Dermed kan relevante detaljer for utredningen være utelatt eller nedtonet i dokumentene (Saunders, Lewis, & Thornhill, 2016).

Generelt har vi ved triangulering, å kombinere sekundær- og primærkilder for å innhente data, styrket utredningens reliabilitet og validitet (Saunders, Lewis, & Thornhill, 2016).

3.6 Forskningsetikk

I forskningssammenheng handler etikk om hvilken adferd forskeren skal ha for å ivareta rettighetene til de som blir berørt av forskningen (Saunders, Lewis, & Thornhill, 2016). Saunders et al. (2016) presenterer etiske prinsipper forskere bør tenke på når de gjennomfører en studie, som vi har brukt som en guide gjennom utredningen. I forkant av intervjuene har respondentene skriftlig samtykket til et informasjonsskriv. Informasjonsskrivet inkluderte informasjon om utredningen, deres frivillighet, rett til konfidensialitet og at vi tok lydopptak under videosamtalen. Disse punktene oppsummerte vi også i starten av hvert intervju.

Vi lagret direkte personidentifiserende opplysninger, som navn, og indirekte personidentifiserende opplysninger, som arbeidssted og stillingstittel, for en kort periode. Derfor var vi pliktige til å melde utredningen til Norsk senter for forskningsdata (NSD). Videre har vi vært bevisst på hvordan vi kunne påvirke resultatene vi fikk. For å unngå dette i størst mulig grad har vi vært bevisst vår integritet og objektivitet. Vi har gjort vårt ytterste for å analysere data og presentere funn på en ansvarlig måte ved å alltid ha respekt for andre. Dette for å ikke utsette respondentene for ydmykelse, skade eller andre materielle ulemper.

3.7 Oppsummering av metodiske valg

Dimensjon	Metodiske valg
Forskningstilnærming	Abduktiv
Formål	Eksplorativ
Metode	Kvalitativ
Strategi	Dokumentstudie og enkeltcasestudie
Tidshorisont	Tverrsnitt
Datainnsamling	Dokumenter og semistrukturerte intervju
Dataanalyse	Tematisk analyse

Tabell 2: Oppsummering av metodiske valg

4. Empiri: dokument- og casestudie

Empiri består av to deler. Vi har først gjennomført en dokumentstudie av EUs retningslinjer for pålitelig og tillitvekkende KI, regjeringens strategi for KI og Datatilsynets rapport om KI og personvern. Deretter gjennomførte vi en casestudie ved å intervjuer syv norske organisasjoner som har innført en form for KI. Empirikapitlet består av presentasjon av funn integrert med tematisk analyse fra begge studiene. Den første delen, i kapittel 4.1, handler om regelverk og anbefalinger som er relevante for organisasjonene som bruker KI. Den andre delen, i kapittel 4.2, er en casestudie som gir innsikt i organisasjoners forståelse av ansvar relatert til KI. Både første og andre del av utredningen vil bidra til å økt forståelse for hvordan ansvar blir annerledes ved innføring av KI i norske organisasjoner.

4.1 Del en: Regelverk og anbefalinger fra et ansvarsperspektiv

Før forskningsspørsmålet «*Hvordan blir ansvar annerledes ved innføring av kunstig intelligens i norske organisasjoner?*» ble utformet, gjennomførte vi en dokumentstudie for å se hvilke regelverk og anbefalinger som er relevante for norske organisasjoner som bruker KI. I den anledning valgte vi de tre nevnte dokumentene, utgitt av EU (European Commission, 2019), regjeringen (Kommunal- og moderniseringsdepartementet, 2020) og Datatilsynet (Datatilsynet, 2018). Dokumentstudien består av en innholdsanalyse som klargjorde at ansvar er et viktig aspekt ved innføring og bruk av KI. Hovedfunnene presenteres i kapittel 4.1.4 og oppsummeres i tabell 3.

4.1.1 Ansvar er sentralt ved utvikling og bruk av kunstig intelligens

En løpende ansvarsdiskusjon er nødvendig ved utvikling og bruk av KI. I samtlige dokumenter blir ordene *ansvar* og *responsibility* nevnt gjennomgående. EUs ekspertgruppe har ansvarlighet som et krav for å oppnå pålitelig og tillitsvekkende KI. Den nasjonale strategien nevner ansvarlighet som et etisk prinsipp for KI. Datatilsynet forklarer at personvernforordningen inneholder et prinsipp om ansvarlighet. Ansvar kommer dermed frem i anbefalinger og i lov.

EU

EUs ekspertgruppe har utviklet syv konkrete krav som KI-systemer burde følge, gjennom sin

livssyklus, for å være pålitelig og tillitsvekkende. Et av kravene er ansvarlighet og er et krav som utfyller de øvrige kravene. Kravet om ansvarlighet innebærer at det skal innføres mekanismer som sikrer både ansvar og passivt ansvar for KI og dens konsekvenser. Ekspertgruppen presiserer at utvikling og bruk av KI må være innenfor lovverket. Å sikre pålitelig og tillitsvekkende KI krever også en etisk kultur og tankesett, som kan skje gjennom offentlig debatt, utdanning og praktisk læring. Ifølge EUs ekspertgruppe er det kun gjennom pålitelig og tillitsvekkende KI europeere kan få fordelene fra KI-systemer, samtidig som de vet at det er gode tiltak som beskytter mot potensiell risiko. Derfor har ekspertgruppen valgt ansvarlighet som et krav for å oppnå pålitelig og tillitsvekkende.

Regjeringen

Regjeringen vil legge EUs syv krav til grunn for ansvarlig utvikling og bruk av KI i Norge. Den nasjonale strategien omtaler EUs krav som prinsipp og har derfor ansvarlighet som et etisk prinsipp for bruk av KI. Strategien forklarer at prinsippet om ansvarlighet utfyller de andre prinsippene og *«innebærer at det skal innføres mekanismer som sikrer ansvarlighet for løsninger basert på KI og deres resultater, både før og etter implementering av løsningene»* (s. 60). Til tross for at ansvarlighet utfyller de andre prinsippene, kan det oppstå spenninger mellom prinsippene. Da må organisasjonene avveie prinsippene på en rasjonell og metodisk måte. Generelt ønsker regjeringen en offentlig debatt om etisk bruk av KI i Norge. Dette er blant annet fordi KI-systemer selv tar beslutninger og utfører handlinger.

Datatilsynet

Utover anbefalingene som kommer frem i EUs rapport og den nasjonale strategien, fremkommer også ansvarlighet i regelverk. Datatilsynet skriver: *«KI forutsetter store mengder data for å kunne lære og ta intelligente avgjørelser»* (s. 3). Videre påpeker de at data i mange tilfeller er personopplysninger. Derfor er personopplysningsloven relevant ved utvikling og bruk av KI i Norge. Datatilsynet forklarer at reglene for behandling av personopplysninger bygger på grunnleggende lovfestede prinsipper. Disse står i personvernforordningen. Bruk av personopplysninger ved utvikling og bruk av KI, utfordrer flere av disse prinsippene. Et av prinsippene i personvernforordningen er *prinsippet om ansvarlighet* og handler om at behandlingsansvarlig har ansvar for, og skal kunne påvise, at de øvrige prinsippene i personvernforordningen overholdes. Det lovfestede personvernprinsippet om ansvarlighet ligger til grunn for all utvikling og bruk av KI i Norge.

Ansvar står dermed sentralt i både anbefalinger og regelverk. EUs ekspertgruppe har utviklet syv krav for å oppnå pålitelig og tillitsvekkende KI, hvor ansvarlighet er et av kravene som utfyller de øvrige kravene. Den nasjonale strategien har lagt EUs krav til grunn for å utvikle etiske prinsipper. Ansvarlighet kommer dermed frem i anbefalinger både i Europa og i Norge. Ansvarlighet kommer også frem i loven gjennom personvernforordningen. Personvernforordningen har et prinsipp om ansvarlighet som handler om at behandlingsansvarlige har et ansvar om å bevise at de øvrige prinsippene i personvernforordningen følges. Dermed kommer det frem både av anbefalinger og av lov at organisasjoner må ta ansvar for KI den utvikler og tar i bruk.

4.1.2 Muligheter med kunstig intelligens

I dokumentene kommer det frem at det eksisterer flere muligheter ved bruk av KI. Den nasjonale strategien påpeker at KI representerer store muligheter for enkeltmennesker og for samfunnet som helhet. EU identifiserer også positive aspekter ved bruk av KI, både for organisasjoner og for samfunnet generelt. At KI kan føre til muligheter for både samfunnet og enkeltindivider kommer også frem av Datatilsynets rapport. Samtlige dokumenter presiserer at mulighetene med KI krever at teknologien brukes på en ansvarlig måte. Dette er spesielt viktig da KI gjør det mulig å løse oppgaver på stadig bedre og helt nye måter. At KI kan løse oppgaver på helt nye måter tyder på at KI skiller seg fra annen teknologi. Ettersom det fremgår at KI fører til flere muligheter i samtlige dokumenter, vil vi gå nærmere inn på hvert av dokumentene.

EU

EUs ekspertgruppe peker på muligheter som effektivisering og forbedring av kundeopplevelser ved bruk av KI. Mulighetene ved KI kommer frem gjennom spesifikke eksempler på utnyttelse av pålitelig og tillitsvekkende KI. Ekspertgruppen mener det er gjennom pålitelig KI Europa vil få fordelene av teknologien, samtidig som KI utvikles i tråd med grunnleggende verdier med respekt for menneskerettigheter, demokrati og rettsstat. Ansvar er derfor et viktig aspekt for å realisere mulighetene. Videre gjør ekspertgruppen det klart at det er nå utviklerne og brukerne av KI kan legge til rette for å maksimere fordelene med KI, samtidig som risikoen minimeres.

Regjeringen

Regjeringen trekker frem at organisasjoner kan oppnå muligheter som å effektivisere prosesser og optimalisere ressursbruk ved å innføre KI. Videre påpeker de at kvaliteten i prosesser og

tjenester kan øke ved å automatisk oppdage sannsynlige avvik. KI vil ikke bare gjøre det mulig å løse oppgaver bedre, men også på helt nye måter. Regjeringen ønsker at Norge skal satse på KI for å utnytte fordelene med teknologien. Videre presiserer de et ønske om at Norge skal gå foran i utvikling og bruk av KI og at dette skal skje gjennom KI som har respekt for enkeltes rettigheter og friheter. I strategien kommer det frem at det er gjennom en ansvarlig bruk av KI konkurransefortrinnet opprettholdes. KI med respekt for enkeltes rettigheter og friheter kan i den globale konkurransen bli et viktig fortrinn. Dermed er det gjennom ansvarlig bruk av KI mulighetene med KI realiseres.

Datatilsynet

I Datatilsynets rapport kommer det også frem muligheter som et potensial for radikalt bedre tjenester og økonomisk gevinst. Videre fremgår det at KI kan bidra til samfunnsmessige fordeler som effektivitet, nye metoder for klima- og miljøvern og et sikrere samfunn. I rapporten kommer det frem at store datastrømmer kan utnyttes for å oppnå samfunnsgevinst ved å analysere data og finne mønstre og sammenhenger. Datatilsynet presiserer at «*KI kan gjøre en forskjell. Mens tradisjonelle analysemetoder er avhengige av å bli programmert til å finne sammenhenger og koblinger, lærer KI av alle dataene den ser. Datasystemene kan derfor hele tiden respondere på nye data og tilpasse analysen sin, uten menneskelig innblanding*» (s. 4). Videre fremgår det i rapporten at ansvarlig bruk av KI, gjennom å håndtere personopplysninger med personregelverket, kan føre til en konkurransefordel for norsk næringsliv. Derfor kommer det også frem av Datatilsynets rapport at det er gjennom en ansvarlig bruk av KI, mulighetene teknologien innebærer kan realiseres.

Det er dermed enighet om at KI medfører flere muligheter, til tross for at mulighetene fremkommer på ulike måter i dokumentene. EU har fokus på hvordan Europa kan realisere mulighetene KI gir. Strategien og Datatilsynet har fokusert på muligheter Norge kan realisere. Strategien har et stort fokus på muligheter, som ikke er overraskende da de ønsker å oppmuntre organisasjoner til å ta i bruk KI. Datatilsynet nevner muligheter, men i mindre grad. Det er heller ikke overraskende da Datatilsynets rapport har som mål å hjelpe å forstå hvordan personvernforordningen blir berørt av utvikling og bruk av KI. Til tross for at mulighetene fremkommer på ulike måter i dokumentene, er det enighet om at KI må brukes på en ansvarlig måte for å realisere mulighetene. Ansvar er spesielt viktig da KI kan løse oppgaver på nye måter. Både Datatilsynet og regjeringen forklarer at KI gjør det mulig å løse oppgaver på nye og bedre måter. Dette fordi KI kan lære av nye data og tilpasse analysen sin, uten menneskelig

innblanding. At KI kan løse oppgaver på helt nye måter og at KI kan lære av data tyder på at KI skiller seg fra annen teknologi.

4.1.3 utfordringer knyttet til bruk av kunstig intelligens

Som nevnt tidligere er ansvar sentralt i anbefalinger og lov. Imidlertid kommer det frem i samtlige dokumenter at lovverket kan føre til utfordringer. EUs ekspertgruppe skriver at en utfordring for brukere av KI kan være at lovverk ikke alltid er oppdaterte i forhold til teknologiske utviklinger. Den nasjonale strategien viser til en undersøkelse av organisasjoner i offentlig sektor som viste at: *«over halvparten av virksomhetene ser på organisasjonskultur, juridiske og regulatoriske forhold, og personvern og sikkerhet som de største utfordringene knyttet til KI»* (s. 53). Videre presiserer strategien at juridiske og regulatoriske utfordringer kan føre til at organisasjoner er unødvendig tilbakeholdne. Datatilsynet på sin side forklarer at personvernreglene har blitt strengere og presiserer at organisasjoner *«får større ansvar for at de behandler personopplysningene i tråd med regelverket, og kravene til åpenhet blir strengere»* (s. 4). Dermed er samtlige dokumenter enige i at det er utfordringer knyttet til lovverk. Utfordringer knyttet til lovverk kan hindre organisasjoner å ta i bruk KI.

I tillegg til nevnte utfordringer trekker samtlige dokumenter frem utfordringer på ulike måter. EUs ekspertgruppe trekker frem utfordringer knyttet til utvikling og bruk av KI. Utfordringene gjenspeiles i de syv kravene som må ligge til grunn for å oppnå tillitvekkende og pålitelig KI. Utfordringene kommer dermed frem gjennom beskrivelser av kravene og er for eksempel utfordringer knyttet til svart boks og dataskjevhet. Ekspertgruppen presiserer også at ulike situasjoner fører til ulike utfordringer. Eksempelvis reiser ikke et KI-anbefalingssystem for musikk de samme etiske bekymringene som KI som foreslår medisinske behandlinger. Den nasjonale strategien påpeker tidlig at KI byr på utfordringer og at utfordringene reiser vanskelige spørsmål. Strategien trekker frem fire sentrale utfordringer ved utvikling og bruk av KI: stordata og dataminimering, datakvalitet, mangel på transparens og autonomi. Datatilsynet presenterer personvernutfordringer ved utvikling og bruk av KI. De ser nærmere på fire KI-relevante utfordringer, herunder skjeve algoritmer møter prinsippet om rettferdighet, KI møter prinsippet om formålsbegrensning, KI møter dataminimering og den svarte boksen møter prinsippet om gjennomsiktig behandling. Dermed er det tydelig at samtlige dokumenter trekker frem utfordringer vedrørende KI.

Til tross for enighet om utfordringer, trekker dokumentene frem utfordringer i varierende grad og på ulike måter. For eksempel trekker Datatilsynet frem utfordringer særlig knyttet til personvernforordningen, mens ekspertgruppen og regjeringen trekker frem generelle utfordringer. Likevel er det fire utfordringer, *datakvalitet*, *mangel på transparens*, *autonomi* og to motstridende prinsipper – *stordata og dataminimering*, som går igjen i alle dokumentene. De er dermed relevante for norske organisasjoners ansvar og utdypes nærmere.

Stordata og dataminimering

KI krever tilgang på en viss mengde data når teknologien utvikles og brukes. Samtidig er det et grunnleggende personvernprinsipp om dataminimering i personvernforordningen. Strategien forklarer at prinsippet innebærer at mengden innsamlede personopplysninger begrenses «til det som er nødvendig for å realisere formålet med innsamlingen» (s. 57). Datatilsynet forklarer også denne utfordringen, og presiserer at prinsippet om dataminimering «bør ha en sentral rolle i utviklingen av kunstig intelligens slik at rettighetene til de registrerte og tilliten til modellene ivaretas» (s. 17). Videre forklarer Datatilsynet at prinsippet kan gjøre det vanskelig å ta i bruk KI for norske organisasjoner da «det kan være utfordrende å definere formålet med bruken fordi man ikke kan forutse hva algoritmen vil lære» (s. 17). I tillegg kan formålet endres etter hvert som maskinen lærer noe nytt. At KI krever en viss mengde data utfordrer prinsippet om dataminimering da det ikke er gitt hvilke data som er nødvendige for formålet.

Utfordringen med å definere formålet med bruk av data er også relatert til personvernprinsippet om formålsbegrensning. Datatilsynet forklarer at «prinsippet om formålsbegrensning innebærer at formålet for behandlingen av personopplysninger må være tydelig angitt og fastsatt når personopplysningene samles inn» (ss. 15-16). Prinsippet om formålsbegrensning kan bli utfordrende da KI ofte krever personopplysninger som organisasjoner i flere tilfeller har samlet inn til et formål og dermed ikke kan brukes til nye formål.

Datakvalitet

Store datamengder har ikke verdi dersom dataen ikke har god kvalitet og struktur. Dersom dataene inneholder feil, vil dette påvirke analysene og resultatene som fremkommer ved bruk av KI. Strategien forklarer at utvalgsskjevhet («bias») er en utfordring som er spesielt viktig ved KI. Videre påpeker de at utvalgsskjevhet kan oppstå enten ved at datasettet har mangler og ikke inneholder fullstendig informasjon, eller ved at data inneholder menneskelige feilvurderinger og historiske skjevheter. Dersom treningsdata inneholder slike skjevheter, vil

det videreføres til KI-modellen. Dårlig datakvalitet kan føre til dårlig utviklet KI-modeller som organisasjoner ikke ønsker å være ansvarlig for.

Mangel på transparens

Videre er mangel på transparens er en sentral utfordring med KI, og da spesielt KI-løsninger som er basert på dyp læring. Regjeringen forklarer utfordringen som at *«enkelte dyp lærings-algoritmer kan sammenlignes med en 'sort boks', der man ikke har innsyn i modellen som kan forklare hvorfor en gitt inndataverdi gir et gitt resultat»* (s. 58). Forklaringen viser at det kan være utfordrende å forstå resultater KI-modeller gir. Når det er vanskelig å vite hvordan resultat blir produsert er det også vanskelig å forklare ovenfor andre parter. Den sorte boksen gjør det utfordrende å oppfylle prinsippet i personvernforordningen om gjennomsiktighet. Datatilsynet er enige i at sort boks-problematikken er en utfordring, men legger til at det kan bli *«tilnærmet umulig å forklare hvordan opplysninger blir koblet og vektlagt i en spesifikk behandling»* (s. 18). Det kan dermed være spesielt utfordrende å ta i bruk KI-løsninger hvor organisasjonen ikke kan forklare resonneringer bak resultater løsningene gir. Dette fordi det kan være utfordrende for organisasjoner å være ansvarlig for noe de ikke kan forklare.

Autonomi

KI kjennetegnes av at maskinene lærer og utfører handlinger. Det kan oppstå flere utfordringer knyttet til autonome maskiner som tar avgjørelser uten at mennesker er involvert. Strategien fremhever at grad av autonomi hos KI-løsninger varierer, men det *«vil uansett reise spørsmål om ansvar for konsekvensene av slike beslutninger og hvordan man skal avgrense en slik autonomi»* (s. 58). En utfordring ved utvikling og bruk av KI handler dermed om maskinens grad av autonomi og hvem som blir ansvarlig for beslutningene KI-løsninger tar.

4.1.4 Konklusjon: Regelverk og anbefalinger fra et ansvarspektiv

Regelverk og anbefalinger viser først og fremst at ansvar er et viktig aspekt ved bruk av KI da ansvar legges vekt på både i anbefalinger og i lov. Ansvar er blant annet viktig fordi KI kan løse oppgaver på nye måter. KI gjør det mulig å løse oppgaver på nye måter da den kan finne sammenhenger og koblinger, og lære av data uten menneskelig innblanding. Mulighetene KI medbringer tyder dermed på at KI skiller seg fra tradisjonell teknologi. Dette underbygges også av utfordringene med autonomi og mangel på transparens da disse utfordringene er mindre gjeldene ved tradisjonelle analysemetoder. KI bryter med de tekniske begrensningene som

tradisjonelle metoder møter når stordata skal analyseres. KI skiller seg derfor fra tradisjonell teknologi. Da KI skiller seg ut fra tradisjonell teknologi og ansvar er et sentralt aspekt, er det viktig at ansvar relatert til KI er tydelig for organisasjoner som ønsker å ta i bruk KI. Dermed er det overraskende at ingen av dokumentene utdyper hva de legger i ansvar.

Den norske regjeringen oppfordrer organisasjoner til å ta i bruk KI, men utfordringer knyttet til KI og lovverk kan begrense bruken av KI. I den nasjonale strategien for KI er det skrevet at regjeringen ønsker at Norge skal gå foran i utvikling og bruk av KI, og at dette kan gi Norge et viktig fortrinn i den globale konkurransen. Til tross for oppfordringen er ikke alltid lovverk oppdatert i forhold til teknologiske utviklinger. KI finner løsninger og sammenhenger basert på store mengder data som mennesker ikke kan finne. Prinsippene om formålsbegrensning og dataminimering, som krever at formålet er fastsatt før data samles inn, og at innsamlede data begrenses til det som er nødvendig for å nå formålet, hindrer organisasjoner å bruke store mengder data. Prinsippet om gjennomsiktighet, som krever at organisasjonene kan forklare resoneringen bak resultater, hindrer KI-løsningen å finne løsninger og sammenhenger mennesker ikke kan finne. Dette kan tyde på at lovverket er bygd opp på en logikk som går ut på å vite hva man skal på forhånd, som er enkelt ved tradisjonelle metoder. Logikken bak KI er derimot at teknologien skal leke med data for å så finne løsninger. Lovverket er derfor bygd opp på en logikk som ikke samsvarer med logikken bak KI. Da logikkene bak KI og lovverket ikke samsvarer blir det utfordrende å følge oppfordringen til å innføre KI. Dermed er det en spenning mellom oppfordringen til å ta i bruk KI og lovverket.

Spenningen mellom oppfordringen og lovverket kan gjøre bruk av KI forvirrende for norske organisasjoner. Skal Norge gå foran i utvikling og bruk av KI er det nødvendig at norske organisasjoner innfører KI. Utfordringene knyttet til regler og anbefalinger kan gjøre dette vanskeligere. For å ta i bruk KI mener regjeringen det er nødvendig med en løpende diskusjon om hva som er en ansvarlig og ønsket utvikling. Til tross for at dokumentene kommer med retningslinjer og anbefalinger er det fortsatt uklart hva ansvar innebærer. Dette fører til at det kan være vanskelig for organisasjoner å forstå hvordan de skal forholde seg til ansvar når de utvikler og anvender KI. Spesielt da KI skiller seg fra tradisjonell teknologi og strukturen ikke samsvarer med logikken bak lovverket. Det er derfor et behov for mer konkret fremstilling av hva dokumentene legger i ansvar.

Basert på det overnevnte, mener vi det kan være vanskelig for organisasjoner å forstå hvordan de skal tolke ansvar ut fra både lov og anbefalinger ved utvikling og bruk av KI. Dette fordi ansvar ikke er definert. Datatilsynet fokuserer på utfordringer vedrørende lovverket, som tyder på at de legger vekt på et juridisk ansvar. På motsatt side har strategien og ekspertgruppen lite på fokus på loven da de fokuserer på etiske vurderinger. Dette tyder på at de legger vekt på moralsk og styringsansvar. Da ingen av dokumentene selv definerer eller presiserer hva de mener med begrepet ansvar kan det være uklart for organisasjoner å forstå hva regelverk og anbefalinger legger ned i begrepet ansvar ved utvikling og bruk av KI. Derfor har vi valgt å gjennomføre en casestudie for å forstå hvordan norske organisasjoner forholder seg til KI og ansvar. Ikke minst hvordan organisasjoner mener ansvar blir annerledes ved å innføre KI, da KI skiller seg fra annen teknologi.

De viktigste funne fra del en oppsummeres kort i følgende tabell:

Del en: regelverk og anbefalinger fra et ansvarsperspektiv	
Hovedfunn - dokumentstudie	
1	Ansvar er sentralt ved utvikling og bruk av KI
2	Spenning mellom oppfordring til bruk av KI og lovverket
3	KI skiller seg fra tradisjonell teknologi

Tabell 3: Hovedfunn – Del en: regelverk og anbefalinger

4.2 Del to: Organisasjoners forståelse av kunstig intelligens og ansvar

Dokumentstudien viste at ansvar er viktig ved utvikling og bruk av KI og at det er interessant å se på hvordan norske organisasjoner forstår ansvar relatert til KI. Basert på at KI skiller seg fra annen teknologi er det også interessant å se på om ansvar blir annerledes ved å ta i bruk KI. Slik kom vi frem til forskningsspørsmålet: «*Hvordan blir ansvar annerledes ved innføring av kunstig intelligens i norske organisasjoner?*». Med det ønsker vi å forstå hva organisasjoner tenker om KI og ansvar, og hvordan de diskuterer det i sin organisasjon. Vi har derfor gjennomført en casestudie for å samle informasjon fra et utvalg av norske organisasjoner.

Vi har valgt å presentere empirien uten innblanding av den teoretiske referanserammen. Temaene er basert på hva vi fant i datamaterialet fra de semistrukturerte intervjuene og som er i relasjon til utredningens forskningsspørsmål. Temaene er kunstig intelligens,

ansvarsforståelse, ansvar i organisasjoner og lov og anbefalinger som blir presentert i 4.2.1-4.2.4. Funnene oppsummeres i 4.2.5.

4.2.1 Kunstig intelligens

Funn 1: KI skaper muligheter for morgendagens organisasjoner

Gjennom intervju ble det tydelig at KI gir flere muligheter organisasjonene ønsker å benytte seg av. Samtlige respondenter påpekte effektivisering og forbedring av kundeopplevelser som to drivere bak innføring av KI. Da KI gir flere muligheter, kan innføring av KI i organisasjonen være viktig for å følge med på utviklingen i samfunnet. En av respondentene forklarte det slik:

«Så det ligger noen tanker om at maskinlæring er en type verktøy som vil kunne være med å bygge morgendagens selskap. (...) Det handler om å kunne bruke maskinlæring aktivt inn i de allerede eksisterende kjerneprosessene vi har.»

Mange av organisasjonene i utvalget har innført KI i interne prosesser som tidligere har blitt utført manuelt. Flere av respondentene uttrykte at de så på KI som et verktøy som kan bidra til å forbedre interne prosesser. Dette underbygges av sitater som: *«Å effektivisere og automatisere prosesser ved å bruke ulike maskinlæringsmodeller»*. I et digitalt samfunn kan et større formål bak innføring av KI bli essensielt for å henge med og forbli en konkurransedyktig organisasjon. Samtidig kan effektivisering være et krav for enkelte organisasjoner. En respondent forklarte at eierne satte krav om å effektivisere litt hvert år som førte til at organisasjonen måtte forbedre og effektivisere driften. Respondentene fortalte videre at kravet *«tvinger dem til å tenke nytt og ta i bruk ny teknologi»*. Effektivisering kan dermed forbedre organisasjonenes eksisterende prosesser.

Fordelene ved bruk av KI er imidlertid ikke bare relatert til organisasjonene selv, men også til kundene. Flere respondenter uttrykte at de ser mange muligheter med bruk av KI og at de i tillegg til effektivisering ønsket å forbedre kundeopplevelsene. I mange tilfeller henger mulighetene sammen. For eksempel kan bedre kundeopplevelser komme gjennom å effektivisere. En respondent forklarte hvordan de hadde gjort utplukk for kontroll bedre med KI:

«Der klarte vi å spare inn administrasjonskostnader fordi vi kontrollerte færre, også fikk vi samme utfall. Så det var vinn-vinn for alle parter. Våre eiere er veldig happy når vi sparer penger, og kundene er happy fordi vi ikke kontrollerte hele basen.»

Selv om flere respondenter så det å forbedre kundeopplevelse som en mulighet ved bruk av KI, var det imidlertid interessant at det var flest offentlige organisasjoner som la hovedvekt på denne muligheten. De private organisasjonene mente også at gode kundereiser og mer personaliserte produkt var en mulighet, men mer som et tilleggsgode. En offentlig organisasjon som var opptatt av å forbedre kundeopplevelser forklarte:

«Vi ønsker å bruke maskinlæring for å skape bedre kundeopplevelser og kundereiser. For eksempel, lage eller identifisere ulike kundegrupper eller kundesegmenter, slik at vi kan skape mer relevant og personifisert innhold til de ulike gruppene isteden for å sende samme innhold til alle.»

Sitatet underbygger at KI gjør det mulig å personalisere organisasjonenes produkter og tjenester, og forbedre kundeopplevelser på nye måter. KI kan oppdage mønstre og sammenhenger organisasjonen selv ikke kan se. En respondent forklarte at organisasjonen tok i bruk KI for å få *«hjelp til prioritering (...) se etter mønster som den enkelte ikke har mulighet til å se»*. Basert på at KI kan utføre visse oppgaver bedre enn mennesker, er det ikke overraskende at noen av organisasjonene brukte KI som et objektivt verktøy for blant annet å sette en mer rettferdig pris. KI gir dermed muligheter som å effektivisere og forbedre kundeopplevelser på måter mennesker ikke kan.

Oppsummering

Ved å innføre KI kan norske organisasjoner oppnå fordeler som økt effektivisering og bedre kundeopplevelser. Fordelene kommer blant annet av at KI ser mønstre mennesker ikke kan se. Dette tyder på at KI tilfører organisasjoner muligheter som tidligere ikke har vært til stede. Dessuten er effektivisering og gode kundeopplevelser aspekter som kan gjøre organisasjoner mer konkurransedyktige. Dermed kan en motivasjon for organisasjoner til å ta i bruk KI være å følge utviklingen i samfunnet og skape morgendagens selskap.

Funn 2: Mangel på transparens og datakvalitet er utfordrende for organisasjoner
I intervjuene ble det tydelig at KI ble ansett som et nytt verktøy for de fleste organisasjonene.

Dermed er det ikke overraskende at organisasjonene møtte utfordringer ved innføring av KI. En respondent påpekte: *«disse modellene er ganske firkantet og ustabile. Modellene kan finne på ting som er veldig rare og de må sees ganske godt etter»*. I utvalget var det enighet om at mangel på transparens og datakvaliteten er to utfordringer mange organisasjoner møter.

Flere respondenter forklarte at mangel på transparens var spesielt utfordrende da de etter loven må kunne forklare resultater KI kommer frem til. Ingen respondenter var imot personvern, men forklarte at det skapte utfordringer da det kan være vanskelig å forklare hvorfor KI-modellen kom frem til et gitt resultat. Dette gjelder spesielt bruk av persondata og da organisasjoner må forklare kunden hvordan persondataen ble brukt. En respondent utdypet hvorfor det kan være vanskelig å forklare:

«Å kunne forklare hva som skjer i en sort boks er til tider ikke bare vanskelig, men også umulig. For grunnen til at noen av disse systemene fungerer er fordi de ser mønster som vi ikke kan se. Og når de kan se mønster som vi ikke kan se, så er det ikke så lett for oss å skjønne forklaringen til systemet, fordi forklaringen ikke resonnerer inne i våres hoder.»

Videre kan utfordringen vedrørende transparens variere etter behov for forklarbarhet. Noen respondenter forklarte at ved å vurdere risiko tilhørende prosesser kunne man finne domener hvor forklarbarhet ble mindre relevant. *«Det kan være problematisk hvis det er et domene som har behov for forklaring, men det er ikke alltid man har et slikt behov.»*

Respondentene trakk også frem datakvalitet som en sentral utfordring ved bruk av KI. Flere respondenter påpekte at datakvalitet er et viktig område da data som brukes for å trene KI-modeller er den eneste sannheten KI kjenner til. *«Treningsdataene er det universet modellen kjenner til. Sånn sett kan du si at algoritmen bare synliggjorde problemene man hadde i utgangspunktet»*. Dersom organisasjoner sliter med dårlig data, som for eksempel at dataene inneholder feil eller mangler, kan KI-modellen videreføre feilen og gi uriktig resultater. En respondent forklarte:

«Man må være oppmerksomme på at dataene man sitter og modellerer med kan være bias i utgangspunktet. Man har kanskje studert noen typer mennesker ekstra nøye, og da kan det være farlig og videreføre den praksisen med å bygge modell på de dataene.»

I noen tilfeller hindrer utfordringer med datakvalitet organisasjonene å ta i bruk KI: «*Mange bedrifter har begrenset muligheter til å ta i bruk KI fordi de ikke har gode systemer for å samle inn og lagre data*». I andre tilfeller kan utfordringen med datakvalitet føre til en kostnad. En respondent forklarte at organisasjonen risikerte økte kostnader dersom algoritmen forurenses av dårlig data da dette «*kan føre til at du får en systematisk feil som har en kostnad. At du systematisk predikerer for høyt forbruk eller for høy produksjon som du blir nødt til å dekke inn*».

Utover utfordringen knyttet til datakvalitet og mangel på transparens, forklarte noen respondenter at de har lite problemer med teknologien som sådan. Et par respondenter påpekte at organisasjonen hadde lite utfordringer med å innføre KI da «*overgangen ifra å lage en matematisk modell til å lage en maskinlæringsmodell i praksis var veldig liten*». Imidlertid var dette gjeldende for organisasjoner som var vant til å jobbe med statistiske modeller.

Oppsummering

I intervjuene var det spesielt utfordringer vedrørende transparens og datakvaliteten som ble lagt vekt på. Transparens er knyttet til kunder da det er definert i lov at organisasjoner må kunne forklare sitt resonnement overfor sine kunder. Til tross for at personvernsprikkene er essensielle for å bevare personvernet, fører de også til utfordringer. Datakvalitet ble også trukket frem som en utfordring med KI. Utover utfordringen knyttet til datakvalitet og mangel på transparens forklarte noen respondenter at de har lite problemer med teknologien som sådan.

4.2.2 Ansvarsforståelse

Funn 3: Ansvar er et omfattende begrep med flere dimensjoner

Samtlige respondenter forklarte hvordan de forstår begrepet *ansvar*. Flere startet med å tydeliggjøre at det er et vanskelig begrep som omfattet mye. Videre ble det sagt at ansvar handlet om å gjøre det riktige, når det fremgår av lov eller retningslinjer. I tillegg spesifiserte flere av respondentene viktigheten av å ta ansvar utover lovverket, altså gjøre det riktige selv om det ikke fremkommer av lov. Dette kom frem gjennom følgende sitat:

«Jeg tenker at noen ganger skiller man mellom lov og moral. En ting er å følge lover og regler, en annen ting er å gjøre det som er riktig. Det er det jeg legger i å ta ansvar, det å gjøre det som er riktig også når det ikke omfattes av lov og krav. Videre tenker jeg at ansvar har mange

dimensjoner. Det er mye forretningsetikk i det, men det er også mye som går på samfunnsansvar, å se på i hvilken grad andre blir berørt og påvirket av virksomheten.»

Sitatet over viser at denne respondenten forstår ansvar som et todimensjonalt begrep, hvor ansvar har lov og krav på en side, og moral på en annen side. Det er flere av respondentene som forstår ansvar som et flerdimensjonalt begrep: «*Nå er det et veldig stort begrep (...) men jeg ser for meg i alle fall tre hoveddimensjoner*». En annen respondent trakk frem at det eksisterer ulike typer ansvar: «*et ansvar er det personlige ansvaret og et annet ansvar er organisasjonens ansvar*». Samme respondent presiserte at ansvar også handlet om å støtte hverandre i organisasjonen på tvers av det personlige og organisatoriske ansvaret. Det at ansvar beskrives på ulike måter og i form av ulike *dimensjoner* og *typer*, viser at ansvar er et omfattende begrep.

Oppsummering

Ettersom litteraturen har ulike definisjoner for begrepet *ansvar*, er det ikke overraskende at respondentene synes ansvar er vanskelig å forstå og forklare. Flere av respondentene forstår ansvar som et begrep med flere dimensjoner, som stemmer overens med litteraturen. Til tross for dette var det enighet om at ansvar omfattet å gjøre det riktige, uavhengig av om det kom frem av lover og regler.

Funn 4: Ansvar er viktig ved bruk av KI

Det er organisasjonen og menneskene som bruker KI som har ansvaret. Dette fremgikk av intervjuene hvor samtlige respondenter mente at det er menneskene som bruker KI som har et ansvar, ikke teknologien. En respondent forklarte det slik:

«Algoritmen i seg selv gjør hverken riktig eller feil, den finner bare mønstre og kommer med en anbefaling om hva utfallet skal være. Algoritmen er bare en matematisk teknikk, så ansvaret ligger hos de som bygger modellene og de som har tenkt å bruke teknologien. Det er opp til oss mennesker å undersøke om det er rettferdig eller ikke.»

Ansvaret ligger dermed hos organisasjonene som tar i bruk KI ettersom teknologien ikke kan tilskrives ansvar. Flere respondenter forklarte at organisasjoners ansvar er spesielt viktig ved bruk av KI ettersom ansvaret kan endres. Ansvaret kan endres da KI både er autonomt og mindre transparent i motsetning til tradisjonell teknologi. En respondent forklarte at KI skiller seg fra tradisjonell teknologi slik:

«I tradisjonell programmering gir man eksplisitte instruksjoner for å løse en bestemt oppgave, mens man i maskinlæring ikke gjør det. Ved maskinlæring trekker læringsalgoritmer ut sammenhenger fra data som gir eksempler på hvordan en oppgave blir løst. Sammenhengene brukes til å løse lignende oppgaver i fremtiden og generalisere for andre organisasjoner».

KI skiller seg fra tradisjonell teknologi da det ikke er like tydelig hva som foregår i en maskinlæringsmodell sammenlignet med eldre programmer, hvor instruksjonene kunne leses rett ut. Da maskinlæringsmodeller lærer og det er vanskelig å forklare hvordan modellen har løst en bestemt oppgave, blir ansvaret endret. Videre skiller KI seg fra tradisjonell teknologi da KI-løsninger tar autonome beslutninger, slik som denne respondenten forklarte:

«Enklere teknologi gjør ikke handlinger for deg, mens KI gjør handlinger som har konsekvenser i den fysiske verden. Hvis teknologien gjør noe galt, hvem er det som skal stå ansvarlig da? Hvordan skal vi løse opp i dette? I annen teknologi er det ofte slik at den virker når den skal og den er tilgjengelig når du trenger den, mens med KI er det noe som faktisk er autonomt og skal gjøre noe når ingen nødvendigvis sitter og overvåker.»

Til tross for at teknologien i større grad har en påvirkningskraft, påpekte to respondenter at ansvar ikke endres av hvilke typer teknologi eller verktøy organisasjonen bruker. Det kan ses i sammenheng med at organisasjonene respondentene jobbet for allerede brukte statistikk og matematiske modeller. Uansett var det enighet om at organisasjoner har ansvaret for teknologien den bruker og at ansvar er viktig ved bruk av KI.

Oppsummering

De fleste respondentene mente at ansvar endres ettersom KI er autonomt og mindre transparent. Videre var det enighet blant samtlige respondenter om at ansvar er viktig ved bruk av KI da organisasjonen har ansvar for teknologien den bruker.

Funn 5: Kontinuerlig ansvar fra start til slutt

Som nevnt i funn 4 er det tydelig at organisasjonene er ansvarlige for teknologien de benytter, inklusivt KI. En respondent presiserte det slik: *«ansvar betyr for meg at det vi bygger, er vi ansvarlig for»*. Det bunner ut fra konseptet om at KI er et verktøy som kan påvirke hvordan organisasjonen driver sin utadrettede aktivitet. En annen respondent trakk frem at

organisasjonen dermed har et ansvar for å forhindre at verktøy de bruker dreier innsatsen i feil retning: *«Man må være bevisst på at retningen faktisk er riktig».*

Videre er ansvar relatert til utvikling av teknologien. Organisasjoner har ansvar fra de starter å utvikle KI-løsninger. Dette kom frem gjennom intervju hvor respondentene forklarte at de utvikler algoritmer med data de kan stå inne for. En respondent forklarte at: *«Det handler om etisk ansvar. For eksempel, er det riktig å modellere med nasjonalitet?».* En annen respondent presiserte at organisasjoner må være ansvarsbevisste under utvikling av KI i forhold til kost og nytte ved innhenting av data. *«Fra et samfunnsperspektiv er det sannsynligvis ikke nyttig å rapportere mer for at vi skal ha gode datakilder».* Dermed har organisasjoner som utvikler KI-modeller også et ansvar under innhenting av data. Ansvar er relatert til utvikling av KI da organisasjonene må vurdere hva de ønsker å stå til svare for ved utvikling av KI.

I tillegg til utvikling og bruk av KI har organisasjoner også et ansvar for resultater og effekter som kommer av å bruke KI. Ansvar omfatter dermed konsekvenser ved utvikling og bruk av KI. Samtlige respondenter påpekte at organisasjonen har et ansvar for konsekvenser forårsaket av KI-løsninger. En respondent forklarte ansvar for konsekvenser på følgende måte: *«Hvis vi tenker innenfor KI så ser jeg ansvar for de resultatene som kommer av at noen integrerer med en algoritme vi har laget».*

De fleste respondentene beskrev ansvar for konsekvenser gitt at bruken av KI hadde ført til en negativ hendelse, med andre ord at noe gikk feil. En respondent sa: *«Hvem skal ta konsekvensene av eventuelle feil?».* En annen respondent poengterte imidlertid at ansvaret også omfatter positive hendelser, det vil si når bruk av KI gikk som planlagt:

«Alle var innforstått med hva som skjer dersom noe virker eller ikke virker. Hvem er det som skal ha både skryt og ansvar her, dersom det skjer noe både den ene eller den andre veien.»

Organisasjoner har ansvar fra de starter å utvikle KI, når de bruker KI og til de slutter å bruke KI. Organisasjonene har dermed et kontinuerlig ansvar fra start til slutt. En respondent forklarte det gjennom at det handler om å kontinuerlig *«være bevisst det ansvaret».*

Oppsummering

Når organisasjoner tar i bruk KI har de et ansvar fra de begynner å samle inn data, under

utvikling og til de slutter å bruke teknologien. Organisasjonene har dermed et kontinuerlig ansvar gjennom hele livssyklusen. Samtidig må organisasjonene også være bevisst et ansvar for resultater og effekter, uansett som de er negative eller positive.

Funn 6: Ønske om tydeligere rammeverk i fremtiden

Da våre funn baserer seg på et virkelighetsbilde slik organisasjonene forstår ansvar relatert til KI nå, fant vi det interessant å belyse respondentenes syn på hvordan ansvar for KI kan bli i fremtiden. En av respondentene påpekte at det har skjedd et generasjonsskifte: unge ansatte er innforstått med at ansvar går utover lovverk. En annen respondent innledet med: *«frem til nå har det skjedd ganske mye med hvordan man på ser på ting generelt i samfunnet, som går inn mot etikk og ansvar. Det har vært en stor utvikling»*. Flere respondenter legger tidligere utvikling til grunn for å svare på hvordan de tror ansvar vil utvikle seg i fremtiden. En respondent forklarte det på denne måten:

«Generelt tror jeg at grensen fort blir flyttet. Da tenker jeg på hvordan man behandler og deler informasjon litt sånn stort. Måten man gjør det på i dag kan ikke sammenlignes med hvordan man gjorde det for 10-20 år siden. Mulighetene for å samle inn og sammenstille data er på et enormt nivå sammenlignet med tidligere. Måten man jobber på i dag gjør at det er bare fantasien som setter grenser for hvordan man kan bruke informasjon og data.»

Da respondentene tenkte KI også vil gjennomgå utviklinger i fremtiden var det flere som uttrykte et ønske om et tydeligere rammeverk. En av respondentene så på dagens situasjon som en nybegynnerfase. Respondenten mente at KI er en gammel teknologi, men at det er først nå den er tatt i bruk. Gjennom bruken av KI vil organisasjonene få mer erfaring, som gjør at de skjønner hva konsekvensene av bruken er. Da kommer ansvarsdelen inn fordi ingen har lyst til å oppleve negative konsekvenser. Basert på det konkluderte respondenten at fremtiden vil se annerledes ut, og at *«når dette blir den nye normalen så bør det ligge et annet rammeverk til grunn for å håndtere ansvarsdelen»*.

En ytterligere respondent trakk også frem at et rammeverk er riktig vei å gå:

«Det vil nok endre seg slik at vi får en større industrialisering av denne type prosesser slik at det blir mer omfattende bruk av denne typen modeller. Da må man rigge et rammeverk for hvordan man skal reflektere rundt hva man holder på med. Hvis KI skal brukes i storskala så

kan man ikke holde på slik som i dag, hvor vi tar en grundig diskusjon for hver eneste modell og hver eneste implementering.»

Gjennom intervju ble det dermed klart at flere organisasjoner har et ønske om et tydeligere rammeverk i fremtiden. Dette kan bli spesielt viktig dersom hver organisasjon bruker flere KI-løsninger og ikke kan diskutere ansvar for hver enkel KI-løsning.

Oppsummering

Basert på utviklingen som har skjedd tidligere er mange av respondentene under den oppfatning at også fremtiden vil bringe utvikling. Da KI oppfattes å være i en nybegynnerfase og forventes å øke i skala ønsker organisasjonene et tydeligere rammeverk for å håndtere ansvarsdelen i fremtiden. Dette fordi det trolig blir for tidkrevende og omfattende å diskutere hver enkel KI-løsning ved bruk av mange løsninger.

4.2.3 Ansvar i organisasjoner

Funn 7: Varierende grad av behov for ansvarsdiskusjon før innføring av KI

Organisasjonene i utvalget diskuterte ansvar i varierende grad før organisasjonen innførte KI. To av respondentene jobber i organisasjoner som drev med statikk og matematiske modeller før de innførte KI. De forklarte at sine organisasjoner ikke hadde diskutert ansvar før innføring av KI. En av de nevnte respondentene forklarte:

«Det har noe å gjøre med at det ikke er noe nytt vi gjør med kunstig intelligens. Vi gjør det bare litt annerledes, så ansvaret er der fra før. Vi endrer ikke nødvendigvis på ansvaret med å sette på en algoritme.»

Andre respondenter forklarte at organisasjonen ikke hadde diskutert ansvar, de hadde kun diskutert hvilke prosesser som var egnet for en KI-løsning. En respondent presenterte et eksempel hvor de ansatte diskuterte forskjellige prosesser hvor de ville innføre KI. Videre forklarte respondenten at en prosess ble valgt *«fordi det var veldig lav risiko, hvor det er ingen som dør hvis vi gjør noe feil»*. Begrepet ansvar ble derfor ikke nevnt i diskusjonen før innføringen av KI, men å diskutere konsekvensene og risiko handler om å ta ansvar. Å diskutere og vurdere risiko er dermed en form for å diskutere ansvar.

Til slutt var det også noen respondenter som forklarte at ansvar ble diskutert før organisasjonen innførte KI. En respondent forklarte at diskusjon før innføringen kom naturlig for dem: «*Det kommer fra samfunnsoppdraget vi har. Så ansvar er den store tunge driveren bak det vi har gjort*». Det tyder på at nødvendigheten av ansvarsdiskusjon avhenger av organisasjonens formål.

Samlet sett viser utvalget at det er varierende grad av ansvarsdiskusjoner før organisasjoner innfører KI. Organisasjoner som før innføringen av KI brukte statistikk og matematikk virker ikke å ha et behov for en ytterligere ansvarsdiskusjon. Andre organisasjoner har derimot en større overgang og trenger å diskutere for eksempel risiko tilhørende ulike prosesser.

Oppsummering

Blant respondentenes organisasjoner var det varierende grad av ansvarsdiskusjoner før organisasjonene innførte KI. Hvorvidt det er nødvendig å diskutere ansvar før innføringen kommer blant annet an på organisasjonenes syn på KI. Ser organisasjonen på innføringen av KI som en stor overgang er det hensiktsmessig å ta en ytterligere ansvarsdiskusjon.

Funn 8: Organisasjoner har jevnlig ansvarsdiskusjoner etter innføring av KI

Ansvar blir stadig tatt opp og diskutert innad i organisasjonene etter organisasjonene tok i bruk KI, som ikke er overraskende da ansvar er et omfattende begrep. Respondentene forklarte at ansvarsdiskusjonen kommer opp som et kontinuerlig tema gjennom alt fra lovverk til etiske dilemma. Diskusjoner knyttet til ansvar og lovverk er spesielt knyttet til bruk og oppbevaring av data. En respondent presiserte: «*Vi har gode juridiske prosesser mellom myndighetene for utveksling av informasjon om databehandling*». Etiske dilemmaer var også jevnlig diskutert: «*Ansvar er på agendaen hele tiden. Etikk var for eksempel i fokus på en workshop med ledergruppen for litt siden (...). Slike typer diskusjoner er på radaren*».

Jevnlige diskusjoner rundt lovverk og etikk viser til et behov for kontinuerlige diskusjoner rundt ansvar etter innføring av KI. En respondent forklarte at ansvarsdiskusjoner er spesielt viktige når KI-løsninger skal settes i produksjon. «*En har ansvar med seg hele tiden. Det er spesielt jo nærmere en kommer produksjonen, jo lengere frem er ansvarsbevisstheten for de som eier det.*» Likevel var flest respondenter enige i at å bruke KI-løsninger i organisasjonen krevde en kontinuerlig bevissthet rundt ansvar. At ansvar jevnlig blir diskutert i organisasjoner som bruker KI underbygger at ansvar er et viktig aspekt.

Oppsummering

Respondentene forklarte at organisasjonene har jevnlig diskusjoner rundt lovverk og etiske dilemma etter innføring av KI. Tidligere funn viser at ansvar er et omfattende begrep og det er dermed ikke overraskende at ansvar krever jevnlig diskusjoner. Å innføre KI-løsninger i organisasjonen krever at organisasjonen kontinuerlig har et bevisst forhold til ansvar.

Funn 9: Ansvar fordelt på flere roller og ulike områder

Mange av respondentene mente organisasjonen de arbeidet i ikke hadde en bestemt rolle direkte knyttet til ansvar for KI. Til tross for at de ikke hadde en slik bestemt rolle, hadde organisasjonene likevel gode ordninger for fordeling av ansvar. Gjennom intervjuene ble det tydelig at ansvaret i organisasjonene var fordelt på flere. En respondent forklarte: *«Det er flere roller relatert til ansvar for den samme algoritmen eller systemet»*. Til tross for ansvaret var fordelt på flere var det likevel klart hvem ansatte skal forholde seg til dersom det kommer eventuelle tvister. Noen av respondentene mente lederen for det aktuelle området har ansvaret, men at denne personen får veiledning og hjelp til juridiske betraktninger fra konsernjuridisk avdeling. Blant offentlige organisasjoner er det vanlig å ha et personvernombud som har en viktig rolle vedrørende ansvar. Andre respondenter forklarte at de har prosesseiere som står som eier av datakilder. Disse har ansvar for å forvalte god sikkerhet og riktig bruk av dataene. En respondent forklarte ansvarsfordelingen på denne måten:

«Vi har ulike profiler eller fagansvarlige. Vi har for eksempel en leder for sikkerhet. Han sørger for at alle dataene ligger der de skal ligge uten at det skal være mulig å hacke eller stjele disse dataene ut av systemene våre. Så har vi en fagdirektør som er ansvarlig for kontroll. Han sørger for at kontrollaktivitetene vi gjør er i henhold til regelverket. Også har vi juristene. De sjekker hele tiden at det vi gjør er riktig. Også har vi domeneeksperter som sjekker om vi er på rett spor eller ikke.»

Oppsummering

Til tross for ingen rolle som er direkte koblet til ansvar i organisasjonene er det gode løsninger for ansvarsfordeling. At organisasjonene har gode løsninger for ansvarsfordelingen, er ikke overraskende da ansvarsfordeling generelt er viktig i organisasjoner. Imidlertid var det forskjeller mellom hvordan organisasjonene fordelte ansvar på ulike roller og områder for KI.

Dermed er det enighet om at ansvar fordeles på ulike roller i organisasjoner, til tross for at det ikke er en fasit på hvordan ansvar bør fordeles.

4.2.4 Lov og anbefalinger

Funn 10: Loven er nødvendig, men utfordrende

Samtlige respondenter påpekte viktigheten av å ha et lovverk. En respondent presiserte at det juridiske er *«egentlig den aller viktigste rammen man har i det daglige»*. Flere respondenter forklarte at lovverket er nødvendig og har bidratt i en positiv retning som fører til at organisasjonen blir mer ansvarsbevisste. En respondent forklarte at lovverket bidro positivt slik:

«Jeg opplever ikke at lovverket er et hinder i forhold til å ta i bruk KI. Det har bare gjort oss mer ryddig. Vi har fått mer struktur. Så alt i alt er det et positivt bidrag fra regelverket.»

Imidlertid uttrykte de fleste respondentene at lovverket var en utfordring som gjorde det vanskeligere å ta i bruk KI. Dette fordi de mente at lovverket er utydelig. Et utydelig lovverk fører til ansvarsutfordringer da det er vanskelig for organisasjonene å vurdere hvor grensene går. En respondent forklarte at utydelig lovverk sammen med få rettslige tolkninger på personvernforordningen gjør det spesielt utfordrende:

«Så mye av det er knyttet til GDPR, altså personverndirektivet fra EU-regelverket. Der avventer vi en rettslig tolkning på et høyt nok nivå til at det kan være en prinsipiell avklaring. (...) Derfor kan man trygt si at ansvarsrefleksjonen vi gjør begrenser hva vi kan gjøre med teknikkene ganske kraftig.»

Et lovverk som er utydelig og en manglende rettslig tolkning gjør det utfordrende å bruke KI. Da lovverket er vanskelig å tyde er det vanskelig å vite hva organisasjoner er ansvarlige for. Lovverket er dermed tett knyttet til diskusjoner rundt ansvar da lovverket ikke har tydelige grenser for hva organisasjoner er ansvarlige for.

Oppsummering

Samtlige respondenter var enige i at lovverket er en nødvendighet. Lovverket fører også til at organisasjoner blir mer ansvarsbevisste. Likevel mener flere respondenter at lovverket er

utydelig. Sett sammen med at det i tillegg er lite rettslig tolkning av loven er det være vanskelig for organisasjoner å vite hva de er ansvarlige for.

Funn 11: Utdatert lov fører til større avveininger

Gjennom intervjuene ble det tydelig at diskusjoner som omhandler lovverk og etisk riktighet står sentralt. Alle respondentene poengterte viktigheten av å gjøre det riktige. Respondentene forklarte at det er mye de har lov til å gjøre, men som de velger å ikke gjøre da de ikke kan stå inne for det. En av respondentene forklarte det slik: *«En ting er å ha lov til å gjøre det, en annen ting er om det er fornuftig og etisk holdbart».*

Til tross for at det juridiske blir omtalt som den viktigste rammen i det daglige må organisasjoner ta et større ansvar. En respondent forklarte at det er lurt å operere innenfor det man selv kan stå til svars for: *«Selv om man opererer innenfor regelverket så er det en god idé å operere innenfor det man synes er rimelig og det man selv kan gjøre rede for».* En annen respondent forklarte at organisasjonen tar et større ansvar gjennom et konkret eksempel hvor organisasjonen rettmessig kan bruke datakilder, men hvor organisasjonen selv ikke ønsker å stå til svars for datakildene. Dette fordi de mente: *«det virker litt sånn pussig. Ikke bare pussig, men kanskje over en etisk grense. (...) Dette er noe vi har i bakhodet hele tiden når vi etterspør nye datakilder».* Organisasjonene er dermed bevisste i sitt arbeid også når loven ikke strekker til. Dette underbygges av samtlige respondenter som påpekte at de har et utvidet syn på ansvar, slik som denne respondenten uttrykte:

«Noen konservative legger seg kanskje på listen at de plikter å følge regelverket. Punktum. Mens andre tenker kanskje de har lyst til å ta på seg noe mer, eller utvide hva de legger i ansvar. Vi ønsker å profilere oss som det siste, altså som et ansvarlig selskap som går lengre enn det som er strengt tatt nødvendig.»

Det er dermed enighet om at ansvar strekker seg utover lovverket. Samtidig kan avveiningen mellom lovverk og etisk riktighet tyde på at lovverket trenger en oppdatering. Sett opp mot utydelig lovverk og lite rettspraksis, som nevnt i funn 10, virker det som et utdatert lovverk gjør at organisasjoner må tenke på ansvar utover lovverket. Selv om organisasjonene etter loven gjør det som er rett, må de selv avveie om det er noe de ønsker å stå inne for. Loven er altså ikke tilstrekkelig slik den er i dag.

Når loven ikke er tilstrekkelig kreves det flere avveininger av organisasjonene. En spesifikk avveining er mellom effektivitetsgevinster for organisasjonen og samfunnsinteresser. Nærmere forklart situasjoner hvor organisasjonene må velge mellom fordeler for seg selv og fordeler for samfunnet generelt. Et konkret eksempel går på datadeling:

«Det er egentlig en frustrasjon i mye av arbeidet vi gjør at det fins mange barrierer for utveksling av data, som hindrer effektiv kontroll i et myndighetsperspektiv, men som selvfølgelig har en god del gevinster i et personvern- og borgerperspektiv.»

Avveingen mellom effektivitetsgevinster for organisasjonen og samfunnsinteresser kommer av at lovverket ikke er tilstrekkelig og trenger en oppdatering. Et utdatert lovverk fører også til at mange organisasjoner føler på en rastløshet. Rastløsheten kommer av at det i mange tilfeller tar endel tid å få gjennom endringer på grunn av et rigid lovverk. Likevel ble det presisert at lovverket er helt nødvendig:

«Jeg tror nok alle setter pris på at dette regelverket eksisterer som privatpersoner, men så merker jo vi at det dukker opp i veldig mange sammenhenger. Regelverket gjør mange prosesser tyngre for oss.»

Oppsummering

Ikke overraskende påpeker flere respondenter at de bryr seg om det etiske ansvaret. Dette blir eksemplifisert gjennom tilfeller hvor organisasjoner rettmessig kan bruke data og teknologi, men hvor de selv velger å ikke gå over en etisk grense. Når organisasjonene har et bevisst forhold til hvordan deres handlinger påvirker samfunnet blir det spesielt vanskelig å avveie mellom løsninger som potensielt kan gi store effektivitetsgevinster, men som på samme tid kan gå ut over klare samfunnsinteresser. Utdatert lov fører dermed til at organisasjoner må gjøre større avveiinger.

Funn 12: Behov for mer konkret strategi

Respondentene mente at anbefalinger og retningslinjer er vage. Gjennom intervju ble det tydelig at respondentene var positive til at regjeringen la frem en nasjonal strategi for KI. En respondent forklarte at det var bra at regjeringen oppfordrer til å ta i bruk ny teknologi: *«Jeg synes det er bra at de har laget en nasjonal KI-strategi hvor datadeling er viktig, personvern*

er viktig og om det å ta i bruk ny teknologi». En annen respondent forklarte at strategien var en guide organisasjonen forholdt seg til: «Det har vært en guide og noe vi har hatt i bakhodet».

Til tross for at en nasjonal strategi for KI er positivt, kom det frem gjennom intervjuene at den nåværende strategien ikke bidrar med så mye. En respondent påpekte at strategien ikke er imponerende da den er vag og for generell: «jeg leste gjennom veldig kjapt og jeg var ikke veldig begeistret eller imponert. (...) Jeg følte strategien var veldig lite modig og veldig mye fluff». Videre tyder respondentenes svar på at det er behov for en mer konkret strategi og for tydeligere grenser ved bruk av KI. En respondent presiserte at det er et behov for «å sette noen grenser for bruk av ny teknologi». Å konkretisere den nasjonale strategien og sette noen tydelige grenser kan gjøre at ansvar blir enklere å forholde seg til for organisasjonene.

Oppsummering

Organisasjonene mener det er bra med en nasjonal strategi for KI. Imidlertid er den nåværende strategien for vag og generell. Respondentens svar tyder på et behov for en mer konkret strategi og tydeligere grenser som organisasjonene enklere kan forholde seg til.

4.2.5 Oppsummering: Organisasjoners forståelse av kunstig intelligens og ansvar

I utredningen svarer vi på forskningsspørsmålet: «Hvordan blir ansvar annerledes ved innføring av kunstig intelligens i norske organisasjoner?». Derfor har vi gjennomført en casestudie for å videre utforske funnene i dokumentstudien. Gjennom semistrukturerte intervju ønsket vi å forstå respondentenes betraktninger om KI og ansvar og hvordan de diskuterer det i sin organisasjon, for så utforske om ansvar blir annerledes. Vi har presentert funnene tematisk basert på hva vi fant i datamaterialet fra intervjuene som er i relasjon til utredningens forskningsspørsmål. Temaene er: kunstig intelligens, ansvarsforståelse, ansvar i organisasjoner og lov og anbefalinger.

Våre funn tilsier blant annet at KI skaper flere muligheter for organisasjonene som gjør at organisasjonen kan følge utviklingen i samfunnet. Imidlertid møter organisasjonene nye utfordringer med KI, herunder datakvalitet og transparens. Som forventet oppfatter respondentene ansvar som et omfattende begrep som er vanskelig å forklare. At ansvar er omfattende styrker behovet for en mer konkret strategi. Til tross for at ansvar er vanskelig mener respondentene organisasjoner har et kontinuerlig ansvar fra de starter å utvikle KI til de

slutter å bruke teknologien. Generelt er ansvar viktig ved bruk av KI. Hvorvidt organisasjoner har behov for å diskutere ansvar før innføring av KI varierer, men etter innføring er det jevnliges ansvarsdiskusjoner i organisasjonene. Ansvar er også fordelt på flere roller og områder. Dersom KI blir brukt i større omfang i fremtiden, ønsker organisasjoner et tydeligere rammeverk for å håndtere ansvar for KI-løsninger. Dette underbygges av at lovverket er utydelig, men også nødvendig i forhold til ansvar og KI. Funnene oppsummeres i tabell 4.

Del to: organisasjoners forståelse av KI og ansvar	
Kunstig intelligens	
1	KI skaper muligheter for morgendagens organisasjoner
2	Mangel på transparens og datakvalitet er utfordrende for organisasjoner
Ansvarsforståelse	
3	Ansvar er et omfattende begrep med flere dimensjoner
4	Ansvar er viktig ved bruk av KI
5	Kontinuerlig ansvar fra start til slutt
6	Ønske om tydeligere rammeverk i fremtiden
Ansvar i organisasjoner	
7	Variierende grad av behov for ansvarsdiskusjon før innføring av KI
8	Organisasjoner har jevnliges ansvarsdiskusjoner etter innføring av KI
9	Ansvar fordelt på flere roller og ulike områder
Lovverk	
10	Loven er nødvendig, men utfordrende
11	Utdatert lov fører til større avveininger
12	Behov for mer konkret strategi

Tabell 4: Oppsummering av funn - Del to: Organisasjoners forståelse av KI og ansvar

5. Diskusjon

For å svare på forskningsspørsmålet: «*Hvordan blir ansvar annerledes ved innføring av kunstig intelligens i norske organisasjoner?*» skal vi nå diskutere ansvar med utgangspunkt i empiriske funn presentert i kapittel 4. Våre empiriske funn består av både en dokumentstudie og en casestudie. I dokumentstudien fant vi: (1) Ansvar er sentralt ved utvikling og bruk av KI, (2) Spenning mellom oppfordring til bruk av KI og lovverk og (3) KI skiller seg fra tradisjonell teknologi. I casestudien fikk vi innsikt i organisasjoners forståelse av ansvar relatert til KI. Diskusjonen av empiriske funn er strukturert etter den teoretiske referanserammen, som er basert på artikkelen av Syse og Olsen (2013). Diskusjonen viser at er ansvar etterpå, direkte ansvar, smalt ansvar og bredt ansvar blir annerledes, mens ansvar på forhånd og medvirkning kan bli annerledes. Denne diskusjonen vil ligge til grunn når vi svarer på forskningsspørsmålet i kapittel 6.

5.1 Ansvar

For å diskutere om ansvar blir annerledes ved innføring av KI er det nødvendig å diskutere hva ansvar er. Ansvar er et omfattende begrep som er vanskelig å forklare. Likevel mener vi organisasjoners ansvar innebærer juridisk, moralsk og styringsansvar, som forklart under vår definisjon av ansvar i kapittel 2.2.5. Ifølge Holdorf og Greenwald (2018) har den eksisterende litteraturen ulike definisjoner og beskrivelser av ansvar. Det er dermed vanskelig å finne en entydig definisjon av begrepet. Dette ble bekreftet gjennom dokumentstudien da ansvar ikke ble definert til tross for at det er et viktig aspekt ved utvikling og bruk av KI. Dermed er det ikke overraskende at begrepet også er vanskelig å forstå og forklare i praksis. Intervjuene bekreftet at ansvar er et omfattende begrep. Respondentene uttrykte at ansvar var vanskelig å forklare. Til tross for ulike beskrivelser av ansvar var flere respondenter enige i at ansvar handler om å gjøre det som er riktig, uavhengig av om det fremgår av lover eller regler. Fra et styringsperspektiv forstår Bovens (1998) ansvar som en dyd og kapasitet i tillegg til en oppgave, passivt og kausalt ansvar. Vi mener derfor for ansvar handler om mer enn å utføre en oppgave og en legal forpliktelse. Ansvar inkluderer personlige verdier og moraler som kapasitet og dyd.

Det er menneskene som bruker KI som har ansvaret for KI-systemers avgjørelser. Bovens (1998) beskriver at ansvar følger en bestemt oppgave. Kaplan og Haenlein (2019) forklarer at KI-systemer utfører oppgaver. Ifølge Russel og Norvig (2016) er KI agenter som mottar

oppfatninger fra miljøet og utfører handlinger. Dermed er det naturlig å tro at KI-systemer har ansvar for oppgaver, avgjørelser og hendelser forårsaket av KI. På en annen side har ikke KI evne til å vurdere konsekvenser av hendelser eller utføre moralske og etiske vurderinger. Derfor argumenterer vi for at KI ikke har ansvar. KI har hverken evne til å forstå ansvar som en kapasitet eller egenskap til å forstå ansvar som en dyd. At KI ikke har ansvar kom også frem gjennom intervjuene hvor respondentene forklarte at det er menneskene og organisasjonene som benytter seg av KI som har ansvaret. Organisasjoner og mennesker har dermed ansvaret ved bruk av KI.

Ansvar er spesielt viktig for organisasjoner da synet på ansvar vil potensielt endre seg over tid. Basert på litteratur om ansvar, vil synet på ansvar stadig endre seg, samtidig som ansvar eksisterer i mange former og ulike kontekster (Sinclair, 1995). Dermed er det naturlig å tro at ansvar for teknologi i en organisasjon også vil endre seg over tid. Ifølge Hartvigsen et al. (2000) fører ny teknologi til at samfunnet kommer opp i nye situasjoner hvor gjeldene moralske normer blir utfordret. Dette tyder på at organisasjoner må diskutere ansvar når de innfører KI. Ifølge Dignum (2018) er det viktig at utvikling av og forskning på KI fokuserer på etikk. Dette samsvarer med at samtlige respondenter mente ansvar er viktig når organisasjoner bruker KI. Videre var det flere respondenter som mente ansvar ble endret av å innføre KI i organisasjonen. For å kunne gi et helhetlig svar på om ansvar blir annerledes ved å ta i bruk KI vil vi gå inn på hvert av Syse og Olsens (2013) skiller. På denne måten kan vi gå mer konkret inn på hvordan ansvar blir annerledes.

5.2 Ansvar på forhånd og ansvar etterpå

5.2.1 Ansvar på forhånd

Basert på vår tolkning av rammeverket, i kapittel 2.3.2, er ansvar på forhånd relatert til hvordan organisasjoner forholder seg til ansvar før organisasjoner tar i bruk KI. Før en organisasjon skal innføre noen form for KI i sine prosesser er det naturlig å tro at diskusjoner rundt delegering av ansvar er viktig. Fra teori ser vi at ansvar definert på forhånd kan være bestemt av samfunnets lover, eller følge av generelle etiske forventninger og spesifikke regler for en konkret jobb (Syse & Olsen, 2013). Dermed er det klart at ansvar på forhånd må ses i sammenheng med begrepet passivt ansvar da det kobles til at ansatte blir holdt ansvarlig for ansvaret som er delegert til dem. En ytterligere kobling gjøres til juridisk ansvar i form av at ansvar er hjemlet i lovverk,

altså ansvar definert på forhånd. Basert på intervjuene mener vi det er variasjoner i behovet for å diskutere og delegerer ansvar for KI før teknologien innføres. I lys av variasjonene, kan det se ut til at innføring av KI ikke er en egen grunn til å ta opp ytterligere diskusjoner om ansvarsdelegering, utover ansvaret som allerede er innforstått i organisasjonen.

Blir ansvar på forhånd annerledes?

For å svare på om ansvar på forhånd vil endre seg ved innføring av KI er det viktig å se på organisasjonens eksisterende prosesser. Basert på intervju tyder det på at dersom organisasjonen ikke anser KI som noe nytt i forhold til eksisterende drift, vil ikke ansvaret som er delegert i organisasjonen endres i stor grad. Dette fordi innføring av KI bare førte til at organisasjonen gjorde eksisterende prosesser litt annerledes. En mulig forklaring til ansvar på forhånd ikke blir annerledes kan være at organisasjoner, som eksempelvis allerede driver med statistikk og matematiske modeller, vil ha en minimal overgang når de innfører KI. Organisasjoner som ikke anser KI som noe nytt i forhold til eksisterende drift anser dermed ikke ansvar som nødvendig å diskutere og delegerer før innføring av teknologien. Dette kan sees i sammenheng med Bovens (1998) forståelse av ansvar som en oppgave. Ettersom innføring av KI førte til at organisasjonen gjorde noe litt annerledes, endres ikke nødvendigvis oppgavene som er delegert på forhånd. Likevel kan det være hensiktsmessig å diskutere ansvar før innføring av KI. Det skaper økt bevissthet og forståelse rundt ansvar, som er viktig ettersom ansvar er et omfattende og vanskelig begrep.

I lys av Bovens (1998) forståelse av ansvar som en oppgave, kan det på en annen side tyde på at ansvar på forhånd blir annerledes hvis arbeidsoppgavene til ansatte blir annerledes. Gjennom intervju ble det klart at noen av respondentene diskuterte ansvar før de tok beslutningen av å innføre KI. Et par diskuterte ikke ansvar eksplisitt, men diskuterte heller hvilke områder som innebærer minst risiko. Å finne områder som har liten risiko mener vi innebærer å være ansvarlig da vurdering av risiko handler om å vurdere om organisasjonen ønsker å stå til svars for konsekvenser av handlinger. Videre kan noen organisasjoner ha ansvar som en sentral del av driften og dermed jevnlig oppdatere det delegerte ansvaret. Uansett vil en drastisk endring på eksisterende prosesser på grunn av innføring av KI gjøre det hensiktsmessig med nye diskusjoner og delegering av ansvar før KI innføres. Dette kommer av at oppgaver endrer seg, som gjør at ansatte får andre forpliktelser i form av andre oppgaver eller roller. Basert på teori kan fremgangen og utvikling av KI føre til et økt ansvar (Buchanan, 2006). Dermed kan en annen mulig forklaring for at ansvar på forhånd endres være at ansatte i organisasjoner får

annerledes ansvar i form av økt ansvar. Dermed kan en endring på eksisterende prosesser føre til annerledes ansvar på forhånd i form av endrede oppgaver eller økt ansvar.

5.2.2 Ansvar etterpå

Som presisert under vår tolkning av referanserammen, i kapittel 2.3.2, er ansvar etterpå relatert til hvordan organisasjoner forholder seg til ansvar etter organisasjonen innførte KI. Syse og Olsen (2013) forklarer at ansvar etterpå handler om hvem som har ansvar i ettertid av en hendelse. Ansvar etterpå sees dermed i sammenheng med konsekvenser. Ifølge Holdorf og Greenwald (2018) finnes det både positive og negative konsekvenser for utfallet av handlinger. Dette tyder på at organisasjonene som bruker KI har ansvar etterpå, både når KI har bidratt til en positiv hendelse og når KI førte til en negativ hendelse. Gjennom intervju ble det bekreftet at organisasjonene har et ansvar for resultater og effekter som kommer av bruken av KI, enten det er i form av å motta skryt eller stå for konsekvensene når noe ikke går som ønsket. Organisasjoner har derfor et ansvar for både positive og negative konsekvenser som følger bruken av KI.

Bli ansvar etterpå annerledes?

Ansvar etterpå blir annerledes for organisasjoner ved innføring av KI, da teknologien krever en kontinuerlig ansvarsdiskusjon. KI krever en kontinuerlig ansvarsdiskusjon på grunn av at KI skiller seg fra tradisjonell teknologi. I teori kommer det frem at KI enkelt og raskt kan utføre oppgaver som tidligere har vært både vanskelig og tidkrevende å løse for mennesker (Mitchell, 2019). Ifølge Kolbjørnsrud (2017) kan KI sanse, handle, forstå og lære. Dette kommer også frem av EUs ekspertgruppe sin definisjon som forklarer at KI-systemer utfører handlinger basert på tolkning av data, og tilpasning til hvordan tidligere handlinger har påvirket omgivelsene (Kommunal- og moderniseringsdepartementet, 2020). Disse definisjonene tyder på at KI skiller seg fra annen teknologi da KI kan lære, forstå og handle.

Gjennom dokumentstudien og intervju ble det bekreftet at KI skiller seg fra annen teknologi, blant annet fordi KI løser oppgaver på nye måter. Dermed er det samsvar mellom teori, dokumentstudien og casestudien om at KI skiller seg fra annen teknologi. Eksempelvis skiller KI seg fra annen teknologi da KI evner å tilpasse seg hvordan tidligere handlinger har påvirket omgivelsene, og handler og lærer selv. At KI bruker erfaringer og kan tilpasse seg, tyder på at KI krever kontinuerlig ansvarsdiskusjon. Kaplan og Haeinlein (2019) forklarer at KI bruker

erfaringer til å nå spesifikke mål og utføre oppgaver gjennom fleksibel adaptasjon. Dermed er det ikke overraskende at det kom frem av intervju at organisasjoner har jevnliges ansvarsdiskusjoner etter innføring av KI. Ansvar etterpå blir dermed annerledes da KI skiller seg fra annen teknologi og krever kontinuerlig ansvarsdiskusjon etter innføring i organisasjoner.

5.3 Direkte ansvar og medvirkning

5.3.1 Direkte ansvar

Ut ifra vår tolkning av rammeverket, i kapittel 2.3.2, er direkte ansvar relatert til om organisasjoner har en definert rolle som er direkte koblet til ansvar ved bruk av KI. Trolig er det nødvendig med en rolle som er direkte koblet til ansvar i organisasjoner ved bruk av KI. Dette fordi KI skiller seg fra annen teknologi og fordi ansvar er sentralt ved bruk av KI. Fra teori ser vi at direkte ansvar også går under navnet kausalansvar og handler om at det må være et påvirkningsforhold mellom to fenomener (Dahlum, 2018). Direkte ansvar oppstår ved en direkte sammenheng mellom årsak og virkning (Syse & Olsen, 2013). Med tanke på at KI utfører handlinger (Russel & Norvig, 2016), men likevel ikke er ansvarlig, må noen andre stå til svars for handlingene KI gjør. Som diskutert i 5.1 kom det frem at organisasjoner er ansvarlig for handlinger KI gjør. Til tross for at det ikke er en direkte sammenheng mellom årsak og virkning når organisasjonen er ansvarlig for KIs handlinger, vil det trolig være hensiktsmessig med en rolle i organisasjonen som er direkte koblet til ansvar. I lys av empiri, ser det ut til at det nødvendigvis ikke trenger å kun være én rolle. Organisasjonene i utvalget hadde gode ordninger for fordeling av ansvar til tross for at de ikke hadde en bestemt rolle som er direkte koblet til ansvar for KI. Det viktigste er at organisasjonen som helhet har et direkte ansvar for KI da KI utfører handlinger den ikke har ansvar for.

Bli direkte ansvar annerledes?

Det direkte ansvaret i organisasjoner blir annerledes i form av at organisasjoner får et ansvar for handlinger teknologien gjør. I intervjuene kom det frem at transparens mot kunden er en av to hovedutfordringer ved å bruke KI i organisasjonen. Mangel på transparens var også en av utfordringene som kom frem i dokumentstudien. Mangel på transparens fører til vanskelighet med å forstå og forklare resonneringen bak KI-løsningers handling og resultat, samtidig som

det er klart at det er organisasjonen som er ansvarlig for utfallet. Organisasjoner vil derfor mest sannsynlig oppleve nye situasjoner hvor de har et direkte ansvar for noe det er vanskelig å stå til svare for. Mangel på transparens vil også kunne være et argument for å være tilbakeholden til å innføre KI i sin organisasjon og er bakgrunn i en del av kritikken mot KI (Kommunal- og moderniseringsdepartementet, 2020). Det direkte ansvaret for organisasjoner som innfører KI blir annerledes, uansett om de innfører en definert rolle for ansvaret eller fordeler ansvaret på flere roller.

5.3.2 Medvirkning

Basert på vår tolkning av rammeverket, i kapittel 2.3.2, er medvirkning relatert til tilfeller hvor det ikke er tydelig hvem i organisasjonen som har direkte ansvar. Ifølge Lagnado et al. (2013) er det kontekster hvor det ikke alltid er like tydelig hvem som har direkte ansvar. Dette handler om det Syse og Olsen (2013) omtaler som medvirkning. Vi legger til grunn at KI ikke kan ha medvirkende ansvar, basert på argumentene for at KI ikke kan ha direkte ansvar. Igjen havner ansvaret hos menneskene og organisasjonen som bruker teknologien. Selv om det er klart at organisasjoner har ansvaret, kan det være mindre tydelig hvem av de ansatte i organisasjonen som har et direkte ansvar for KI-løsningen. I organisasjonene i utvalget var det som nevnt gode løsninger for ansvarsfordeling blant flere ansatte. Likevel kan organisasjonene komme i situasjoner hvor de har ansvar i form av medvirkning. Medvirkning kan være vanskelig da det er mindre tydelig fordi det omfatter blant annet bistand, oppmuntring eller unnlattelse til å gripe inn dersom en har forpliktelse (Elden, 2020). Etersom organisasjonene respondentene representerte hadde flere ulike roller som var direkte ansvarlig, er det naturlig at medvirkning ikke ble diskutert i like stor grad. Vi kan likevel ikke påstå at det er slik for alle organisasjoner i Norge da vi har et lite utvalg. Til tross for at medvirkning var lite diskutert blant respondentene, kan en som medvirker til alvorlige lovbrudd straffes på lik linje med en som er direkte ansvarlig (Lovdata, 2015). Vi argumenterer derfor for at medvirkning fremdeles er et sentralt aspekt, spesielt ettersom det kan oppstå nye situasjoner i fremtiden hvor medvirkning vil bli viktigere å diskutere.

Bli medvirkeransvar annerledes?

Hvorvidt medvirkning blir annerledes ved bruk av KI er mindre tydelig basert på empiriske funn. Dette er på bakgrunn av at vi fikk mange synspunkter vedrørende direkte ansvar og få synspunkter vedrørende medvirkning. Likevel kan vi ikke utelukke at medvirkeransvar ikke

blir annerledes. Dette er basert på at vi har et lite utvalg. I tillegg kan medvirkning bli straffet på lik linje med de som har direkte ansvar etter loven. Medvirkning kan dog bli mer aktuelt dersom det kommer flere eksempler på hendelser hvor KI har ført til negative konsekvenser. Som påpekt av respondentene er det lite rettspraksis på hendelser knyttet til KI. Lite rettspraksis er en utfordring da organisasjonene ikke har klare eksempler å forholde seg til. Trolig vil det komme mer rettspraksis etter hvert, som både kan tydeliggjøre spillereglene, samt endre dem. Eksempelvis kan det vise seg at organisasjoner eller ansatte får et medvirkeransvar for å dele data med andre. Avslutningsvis kan vi ikke konkludere med at medvirkning blir annerledes ved å ta i bruk KI ut fra vår empiri.

5.4 Smalt og bredt ansvar

5.4.1 Smalt ansvar

Med et smalt syn på ansvar mener vi organisasjoner som kun har fokus på å operere innenfor loven. Det vil si at organisasjonene begrenser ansvaret til det strengt nødvendige og bare forholder seg til det juridiske ansvaret. Et smalt syn på ansvar er koblet til aksjonærteorien da organisasjoner ifølge denne teorien tenker på ansvar i form av akkurat det de må gjøre og fokuserer på å tjene penger til eierne (Syse & Olsen, 2013). Ifølge Martin (2010) må organisasjoner velge hvilket mål de skal fokusere på. Han mener det ikke er mulig å optimalisere flere mål samtidig da det ene målet alltid vil gå på bekostning av de andre. Basert på empiri virket det ikke som organisasjonene synes det var problematisk å operasjonalisere flere mål samtidig da de hadde et bredt syn på ansvar. I intervju omtalte respondentene de som kun følger lovverket som konservative. Dette kan sees i sammenheng med hvordan respondentene forstod begrepet ansvar da samtlige respondenter uttrykte at ansvar handler om mer enn ansvar som fremgår av lov. I tillegg ble det gjennom empiri tydelig at det er et generasjonsskifte i forhold til hvordan ansatte ser på ansvar. Tidligere har det vært et smalere syn på ansvar, mens de som kommer ut i arbeidslivet nå ser på ansvar utover det som står i lover og regler. Dette viser at organisasjonen har et bredt syn på ansvar da det gjennom empiri ble tydelig at ansvar handlet om et ansvar utover lovverket.

Selv om organisasjonene hadde et bredt syn på ansvar ble det likevel klart gjennom intervjuene at loven er organisasjonens viktigste ramme i det daglige. Dette er ikke overraskende da organisasjoner og enkeltpersoner blir straffet for lovstridige handlinger. Imidlertid er det ikke

alltid gitt hvordan organisasjonene skal tolke lover og regler. Teori om den historiske utviklingen til KI viser at KI har utviklet seg i et raskt tempo det siste tiåret (Russel & Norvig, 2016). Videre har den teknologiske utviklingen gått raskere enn utviklingen av lover og regler (Burton, et al., 2020), som tyder på at lovverket henger bak den teknologiske utvikling. At lovverket henger bak den teknologiske utviklingen, kom frem i dokumentstudien og ble bekreftet i casestudien. Dette har ført til en spenning mellom loven og logikken bak KI da loven krever både dataminimering og forhåndsdefinerte formål, noe som er utfordrende med KI.

Vi mener spenningen fører til utydelighet som gjør det vanskeligere for organisasjoner å forstå hvordan de skal forholde seg til ansvar ved innføring av KI. Dermed kan det være vanskelig for organisasjoner å tolke ansvar ut ifra både lov og anbefalinger ved bruk av KI. Dette ble bekreftet gjennom intervju hvor flere respondenter uttrykte en frustrasjon over utydelig lovverk, herunder personvernforordningen og deling av data, som hindret en potensielt mer effektiv drift av organisasjonen. Imidlertid er det nødvendig med personvernforordningen for å beskytte enkeltindivider.

Bli r smalt ansvar annerledes?

Et smalt syn på ansvar vil trolig ikke endres ved innføring av KI. Med et smalt syn på ansvar vil lovverket være rammen for ansvar uavhengig av om organisasjonen innfører KI eller ikke. Derimot kan lover endres og det kan komme nye lover og regler. Da kan et smalt ansvar bli annerledes i form av at organisasjoner må forholde seg til lover og regler på nye måter. På en annen side kan smalt ansvar bli annerledes i form av at organisasjoner blir tvunget til å tenke på et bredere ansvar. Ettersom nåværende lovverk er utydelig, kan det bli vanskeligere å innføre KI for organisasjoner som kun fokuserer på å holde seg innenfor loven. Dermed kan organisasjonen bli tvunget til å tenke på ansvar utover lovverket. Dette ble tydelig gjennom empiri hvor respondentene uttrykte at organisasjonen måtte gjøre flere avveininger om hva som er ansvarlig da lovverket ikke er tydelig ved bruk av KI. Smalt ansvar blir dermed annerledes ved innføring av KI da organisasjonen må tenke på ansvar også utover det juridiske ansvaret.

5.4.2 Bredt ansvar

Med et bredt syn på ansvar forholder organisasjoner seg til ansvar utover det som står i loven. Syse og Olsen (2013) forklarer at et bredt syn på ansvar handler om å ta interessene til andre i betraktning, som lokalsamfunnet og kundene. Ifølge Tranøy (2018) skilles det ofte mellom

juridisk og moralsk ansvar. Vi mener begge perspektivene er viktige for å forstå et bredt syn på ansvar. Videre mener vi at Bovens' (1998) syn på ansvar, hvor ansvar også kan bli sett på som en dyd eller kapasitet, også viser til et bredt syn på ansvar.

Bredt ansvar handler om å ta et moralsk ansvar, sammen med det juridiske ansvaret. Gjennom intervjuene ble det tydelig at respondentene hadde et bredt syn på ansvar. Samtlige respondenter forklarte at ansvar innebærer å gjøre det som er riktig uavhengig av lov. I noen tilfeller kan organisasjonene etter lovverket rettmessig bruke data de selv mener ikke er moralsk riktig å bruke. Det tyder på at det ikke er tilstrekkelig å tenke på hva organisasjonen etter lovverket kan tilregnes som skyldig for (Falkanger, 2014). Faktorer som samvittighet og etiske vurderinger står dermed sentralt når organisasjoner skal ta i bruk KI. Til tross for at lovverket er nødvendig og fører til økt ansvarsbevissthet, er det fortsatt behov for at organisasjonene selv vurderer hva som er riktig når de bruker KI. Men er det virkelig slik at alle organisasjoner påtar seg et moralsk ansvar eller kan det være at organisasjoner ønsker å fremstå mer moralske enn de faktisk er? Basert på utsagn fra respondentene virker det som organisasjonene har et bevisst forhold til hvordan deres handlinger påvirker samfunnet, spesielt med tanke på hvordan de reflekterte i forhold til effektivitetsgevinster og samfunnsinteresser.

Det kan imidlertid være vanskelig å tydeliggjøre hvor grensen for eget og andres ansvarsområde går når organisasjoner har et bredt syn på ansvar (Phillips, Freeman, & Wicks, 2015). Dette kommer også frem gjennom intervjuene hvor flere respondenter påpekte at et utydelig lovverk og lite rettspraksis gjør det vanskelig å vite hva organisasjonen er ansvarlige for når de bruker KI. Organisasjonene må på grunn av utydelige regler i større grad se på ansvar som en følelse av forpliktelse, være ansvarlig ovenfor andre og vise omtanke for andre (Holdorf & Greenwald, 2018). I teori kom det frem at KI kan enkelt og raskt utføre oppgaver som tidligere har vært vanskelige og tidkrevende for mennesker å løse (Mitchell, 2019). Sett opp mot at respondentene etterlyser rettspraksis og tydelige regler, kan det se ut til at det vil være store variasjoner i tolkningen av grensene på tvers av organisasjoner. Dette underbygges av Syse og Olsen (2013) som mener at hvordan organisasjoner definerer egne og andres ansvarsområder avhenger av konteksten til organisasjonen. Et ytterligere aspekt ble trukket frem av respondentene i form av at organisasjonens formål med å innføre KI påvirker hva slags ansvar organisasjonen får. Eksempelvis vil noen formål medbringe større risiko og potensiale for negative konsekvenser. Dermed kan organisasjonens kontekst og KIs formål påvirke hvordan organisasjoner tolker ansvarsområder.

Blir bredt ansvar annerledes?

Når organisasjoner tar i bruk KI, kan et bredt syn på ansvar bli viktigere. Ved å ta i bruk KI-løsninger kan organisasjoner behøve å utvide det brede synet gjennom å ta hensyn til flere interessenter og flere nyanser enn tidligere. Basert på empiri, om at lovverket ikke er tilstrekkelig og trenger en oppdatering, kan det se ut til at det moralske ansvaret får økt betydning. Et mulig utfall kan være at det juridiske ansvaret får mindre betydning for organisasjoner fordi lovverket er utdatert og utilstrekkelig. Når et av perspektivene på ansvar får mindre betydning vil organisasjoner se mer på de andre to perspektivene, moralsk og styringsansvar. Dette er spesielt viktig i tilfeller hvor organisasjonen med rettmessighet kan utføre handlinger som kan føre til effektivitetsgevinster, men som til gjengjeld kan være imot organisasjonens moral og den allmenne interessen. Slik som ansvar er definert i kapittel 2.2.4, må organisasjonene stå inne for den teknologien de tar i bruk fordi de må stå til svars for handlingene sine. Dermed vil et bredt syn på ansvar bli viktigere når organisasjoner tar i bruk KI.

5.5 Oppsummering: Diskusjon

Ansvar på forhånd kan bli annerledes i form av endrede oppgaver. Dersom innføring av KI medfører store endringer av fordeling av oppgaver, er det behov for diskusjon av ansvar for KI-løsningen på forhånd. *Ansvar etterpå* blir annerledes da innføring av KI krever kontinuerlige ansvarsdiskusjoner i organisasjoner. *Direkte ansvar* blir annerledes da organisasjonen har ansvar for KI som tar egne beslutninger. Dermed vil det direkte ansvaret være uavhengig av om organisasjonen har en direkte rolle koblet til ansvar for KI eller om organisasjonen har fordelt ansvaret på flere roller. *Medvirkeransvar* kan bli annerledes, men kan ikke konkluderes basert på vår empiri. Imidlertid er medvirker ansvar viktig da en som medvirker til alvorlige hendelser kan bli straffet på lik linje som en som er direkte ansvarlig. *Smalt ansvar* blir annerledes da organisasjoner må gå over til et bredere syn på ansvar ved innføring av KI. Dette fordi lovverket er utydelig og fører til at organisasjonene må tenke på ansvar utover det juridiske ansvaret. *Bredt ansvar* blir annerledes i form av at det blir viktigere. Organisasjonene må i større grad vurdere selv hva de ønsker å stå til svars for, og ta hensyn til flere interessenter og flere nyanser enn tidligere.

Diskusjonen oppsummeres i følgende figur:



Figur 3: Oppsummering gjennom teoretisk referanseramme

6. Konklusjon

I dette kapitlet svarer vi på forskningsspørsmålet: «*Hvordan blir ansvar annerledes ved innføring av kunstig intelligens i norske organisasjoner?*». Svaret blir presentert i kapittel 6.1. Implikasjoner for organisasjoner og strategien blir presentert i kapittel 6.2. Kapittel 6.3 foreslår videre forskning og kapittel 6.4 tar opp utredningens begrensninger.

6.1 Svar på forskningsspørsmålet

Vi befinner oss nå i en fjerde industriell revolusjon hvor teknologien som utvikles endres raskere enn noen gang før (Schwab, 2014). Utviklingen har ført til at norske organisasjoner tar i bruk KI i sine prosesser og ifølge Dignum (2017) er det viktig å vurdere ansvar ved bruk av KI. Til tross for at litteratur og forskning har sett mye på KI, ansvar og ansvarlig KI, er det lite litteratur på ansvar relatert til KI fra et organisasjonsperspektiv. Utredningen ble gjennomført for å gjøre litteraturgapet mindre. Formålet var å forstå organisasjoners ansvar for KI fra et holistisk perspektiv. Vi har utforsket hva både regelverk og anbefalinger legger vekt på ved bruk av KI fra et ansvarsperspektiv, samt organisasjoners forståelse av ansvar relatert til KI. Med denne innsikten skal vi svare på forskningsspørsmålet:

«Hvordan blir ansvar annerledes ved innføring av kunstig intelligens i norske organisasjoner?»

Basert på diskusjon mener vi ansvar blir annerledes ved innføring av KI i norske organisasjoner. Vi argumenterer for at ansvar er et omfattende begrep som påvirkes av flere faktorer. Dette støtter forståelsen av ansvar, presentert i kapittel 2.2.5, hvor juridisk, moralsk og styringsansvar er sentralt for en helhetlig forståelse av organisasjoners ansvar. I diskusjonskapittelet så vi gjennom inndelingen til Syse og Olsen (2013) at ansvar etterpå, direkte ansvar, smalt ansvar og bredt ansvar blir annerledes, mens ansvar på forhånd og medvirkning kan bli annerledes. Basert på diskusjon mener vi de viktigste driverne av at delene blir annerledes er: (1) *KI skiller seg fra tradisjonell teknologi* og (2) *Spenning mellom lovverk og logikken bak KI*.

KI skiller seg fra tradisjonell teknologi da KI har evne til å lære og oppfatte, samt bruke disse erfaringene til å utføre handlinger. KI kan for eksempel å finne mønstre og løse oppgaver uten

menneskelig innblanding, samt løse enkelte oppgaver bedre enn mennesker. Da KI har evne til å tilpasse seg omgivelser og lære, blir organisasjoners ansvar annerledes fordi KI krever en kontinuerlig ansvarsdiskusjon. Organisasjoners ansvar blir annerledes ved at organisasjonene må tenke på ansvar utover det juridiske ansvaret i større grad enn før. Moralsk ansvar og styringsansvar, spesielt forståelse av ansvar som dyd og kapasitet, får en større betydning ved innføring av KI.

Det er *spenning mellom lovverk og logikken bak KI* fordi KI skiller seg fra tradisjonell teknologi. Logikken bak KI handler om at teknologien skal leke med data. Muligheter med KI, eksempelvis at KI kan se mønstre mennesker ikke kan se, er noen av årsakene til at organisasjoner ønsker å ta i bruk KI. Imidlertid finner KI løsninger mennesker ikke kan forklarte årsaken bak og krever store mengder data. Dette fører til en spenning mellom loven og logikken bak KI, da loven eksempelvis krever forhåndsdefinerte formål og dataminimering. Spenningen gjør det vanskelig å tolke forståelsen av ansvar ut ifra regler og anbefalinger. Organisasjoner må i større grad vurdere hva de ønsker å stå til svars for. Dette gjelder både når lovverket er utydelig og organisasjoner selv må tolke hva de synes er ansvarlig, og når organisasjoner etter lovverket rettmessig kan utføre handlinger de selv ikke vil stå til svars for. Derfor fører innføring av KI i norske organisasjoner til økt grad av moralsk og styringsansvar.

Ansvar blir annerledes ved innføring av KI i norske organisasjoner fordi KI krever at organisasjoner legger mer vekt på moral og styringsansvar da juridisk ansvar ikke er tilstrekkelig.

6.2 Implikasjoner for forfattere av dokumentene og organisasjoner

Denne masterutredningen kom frem til at norske organisasjoners ansvar blir annerledes ved innføring av KI. For organisasjoner som ønsker å innføre en form for KI i sine prosesser kan det være lurt å tenke gjennom ansvar relatert til KI før innføringen. Flere organisasjoner opplever KI som en teknologi som skiller seg fra tradisjonell teknologi, dermed kan ansvar for KI-løsninger være vanskeligere å forholde seg til enn ansvar for tradisjonell teknologi. Ved bruk av KI er det et behov for å tenke utover det juridiske ansvaret da KI gjør handlinger som kan få konsekvenser for mange av organisasjonens interesser. Basert på diskusjonen vil

innføring av KI kreve et økt ansvar fra organisasjoner. Med et økt ansvar tenker vi på kontinuerlige ansvarsdiskusjoner, og at moralsk og styringsansvar blir viktigere.

Det er påbegynt arbeid for å håndtere utfordringene med bruk av KI og personvern. Den 1. desember 2020 presenterte Datatilsynet en regulatorisk sandkasse. Datatilsynet presiserer at dårlig ivaretatt personvern i utviklingen av KI kan hindre grad av innovasjon (Datatilsynet, 2020a). Sandkassen skal fungere som et kontrollert testmiljø for organisasjoner som ønsker å bruke KI og hjelpe med utvikling av KI samtidig som prinsippene i personvernforordningen møtes. Målet med sandkassen er å stimulere innovasjon og etisk og ansvarlig KI samtidig som personvernet ivaretas (Datatilsynet, 2020b). Imidlertid er det behov for å tydeliggjøre hva regler og anbefalinger legger ned i begrepet ansvar. Ansvar er et omfattende begrep og for å hjelpe organisasjoner som ønsker å innføre KI vil det være nyttig å tydeliggjøre begrepet ansvar. For å hjelpe organisasjoner kan for eksempel regjeringen oppdatere den nasjonale strategien for KI. Her tenker vi en oppdatering gjennom å tilføre hva strategien legger i ansvar relatert til KI. Etersom samtlige dokumenter i dokumentstudien presiserer at ansvar er viktig, er det naturlig at dokumentene også presiserer hva de mener organisasjoner som bruker KI er ansvarlige for. Gjennom en tydelig definisjon av ansvar i den nasjonale strategien for KI mener vi det blir lettere for norske organisasjoner å følge regjeringens oppfordring til å ta i bruk KI.

6.3 Forslag til videre forskning

KI og ansvar får i dag økt oppmerksomhet i Norge, likevel er det vanskelig å finne hva organisasjoner blir ansvarlig for når de bruker KI. Derfor vil vi oppmuntre akademikere til å forske videre på organisasjoners ansvar for KI. Videre vil vi oppmuntre til å videre studere personvernforordning for å utforske hvordan personvernet kan opprettholdes samtidig som utfordringene med bruk av KI reduseres.

Mulig forskning ved et større tidsperspektiv er å gjennomføre en forklarende studie med samme forskningsspørsmål. Ved å kontrollere for andre variabler, samt sammenligne organisasjoner før og etter innføringen av KI får man et nytt perspektiv på saken.

På grunn av vår abduktive tilnærming og utforskende design har vi hatt en bred tilnærming til organisasjoners forståelse av KI og ansvar. Dermed har vi fått noen aspekter som faller utenfor vårt område. Et viktig aspekt respondentene påpekte handler om organisatoriske utfordringer

ved å innføre KI. Smal KI, som brukes i dag, trenger definerte prosesser de kan ta over. Organisasjoner må derfor finne aktuelle prosesser det er nyttig å innføre KI. Ikke minst er det helt essensielt å «selge inn» KI-løsningen til andre ansatte i organisasjonen for at de skal ønske å ta teknologien i bruk. Vi vil derfor oppfordre forskere til å se videre på implementeringsprosessen av KI i norske organisasjoner.

6.4 Utredningens begrensninger

Det er flere begrensninger ved utredningen. For det første er KI et omfattende og flertydige begrep. I utredningen hadde vi krav om at organisasjonene i utvalget hadde innført en form for KI. Dette på bakgrunn av at KI har flere underkategorier som går under betegnelsen KI. Dermed er det ikke gitt at organisasjonene definerer KI likt, som kan føre til en svakhet med utredningen da respondentenes svar er farget av deres forståelse av KI. Gjennom intervju ble det imidlertid klart hvilke KI som var innført i organisasjonene og det tydet på at respondentene definerte KI relativt likt.

Videre er KI et tema hvor det stadig kommer ny innsikt og litteratur. Gjennom utredningen kom det derfor stadig ny innsikt og informasjon å forholde seg til. Dermed var det vanskelig å kontinuerlig passe på at innholdet i utredningen var oppdatert. Til tross for denne utfordringen er utredningen en tverrsnittstudie som studerer KI og ansvar på et bestemt tidspunkt.

En annen utfordring er at temaet ansvar kan oppfattes som uklart og tvetydig. For å best fange respondentenes forståelse av ansvarsbegrepet, stilte vi spørsmålet «Hvordan forstår du begrepet ansvar?». På denne måten hadde vi noe å ta utgangspunkt i og forsikret oss om at vi og respondenten ikke snakket om to forskjellige ting. Likevel er det en svakhet med utredningen at ansvar er et vanskelig begrep. I tillegg er det naturlig å tro at ingen av respondentene ønsket å sette organisasjonen i et dårlig lys. En ytterligere svakhet med utredning er derfor at svarene kan være farget av et ønske om å opprettholde et godt omdømme.

Avslutningsvis er en annen svakhet er at vi har et begrenset utvalg med syv respondenter i vår casestudie. Dermed er det ikke gitt at svarene i utredning kan generaliseres til norske organisasjoner som en helhet.

Bibliografi

- Adams, R. (2017, januar 10). *Forbes.com*. Hentet fra 10 Powerful Examples of Artificial Intelligence In Use Today: <https://www.forbes.com/sites/robertadams/2017/01/10/10-powerful-examples-of-artificial-intelligence-in-use-today/#5368480c420d>
- Bovens, M. (1998). *The Quest for Responsibility. Accountability and Citizenship in Complex Organisations*. Cambridge University Press.
- Brenna, F., Danesi, G., Finch, G., Goehring, B., & Goyal, M. (2018). *IBM*. Hentet fra Shifting toward Enterprise-grade AI: <https://www.ibm.com/downloads/cas/QQ5KZLEL>
- Brynjolfsson, E., & McAfee, A. (2016). *The Second Machine Age: work, progress, and prosperity in a time of brilliant technologies*. W. W. Norton & Company.
- Buchanan, B. G. (2006). A (Very) Brief History of Artificial Intelligence. *AI Magazine*, 26(4).
- Burton, S., Habli, I., Lawton, T., McDermid, J., Morgan, P., & Porter, Z. (2020). Mind the gaps: Assuring the safety of autonomous systems from an engineering, ethical, and legal perspective. *Artificial Intelligence*(279).
- Campbell, M., Hoane Jr, A., & Hsu, F. (2002). Deep Blue. *Artificial Intelligence*, 134(1-2), ss. 57-83.
- Carson, S., & Kosberg, N. (2011). *Etikk. Teori og praksis*. Cappelen Damm Akademisk.
- Coeckelbergh, M. (2019). Artificial Intelligence, Responsibility Attribution, and a Relational Justification of Explainability. *Science and Engineering Ethics*, 26, ss. 2051–2068 .
- Dahlum, S. (2018). *Store Norske Leksikon*. Hentet november 2020 fra Kausalitet: <https://snl.no/kausalitet>
- Danks, D., & London, A. (2017). Algorithmic Bias in Autonomous Systems. *International Joint Conference on Artificial Intelligence*. IJCAI.
- Datatilsynet. (2018, 11. januar). *Rapporter og utredninger*. Hentet 16. september 2020 fra Kunstig intelligens og personvern: <https://www.datatilsynet.no/globalassets/global/dokumenter-pdf-er-skjema-ol/rettigheter-og-plikter/rapporter/rapport-om-ki-og-personvern.pdf?fbclid=IwAR0FTMIIAEYqbC8eRo9UEdxN5HTrE3G8IVe1P8fKeo9-yjD8lHwebJiGT0M>
- Datatilsynet. (2019). *Datatilsynet*. Hentet november 2020 fra Regelverk og verktøy: <https://www.datatilsynet.no/regelverk-og-verktoy/lover-og-regler/om-personopplysningsloven-og-nar-den-gjelder/>
- Datatilsynet. (2020a). *Datatilsynet*. Hentet desember 2020 fra Aktuelle nyheter 2020: regulatorisk sandkasse: <https://www.datatilsynet.no/aktuelt/aktuelle-nyheter-2020/regulatorisk-sandkasse-for-utvikling-av-ansvarlig-kunstig-intelligens/>
- Datatilsynet. (2020b). *Datatilsynet*. Hentet desember 2020 fra regelverk og verktøy: <https://www.datatilsynet.no/regelverk-og-verktoy/sandkasse-for-kunstig-intelligens/>

-
- Deloitte. (u.d.). *www2.deloitte.com*. Hentet fra Avmystifiser AI:
<https://www2.deloitte.com/no/no/pages/technology-media-and-telecommunications/articles/avmystifiser-artificial-intelligence.html>
- Dignum, V. (2017). Responsible Autonomy. *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence* (ss. 4698-4704). IJCAI.
- Dignum, V. (2018). Ethics in artificial intelligence: introduction to the special issue.
- Elden, J. (2020). *Store Norske Leksikon*. Hentet november 2020 fra Medvirkning:
<https://snl.no/medvirkning>
- Ernst & Young. (2018). *Artificial Intelligence in Europe: Norway*. EY LLP.
- Etzioni, A., & Etzioni, O. (2017). Incorporating Ethics into Artificial Intelligence. *The Journal of Ethics*(21), ss. 403-418.
- European Commission. (2019, 8. april). *Reports and studies*. Hentet 16. september 2020 fra Independent High-Level Expert Group on Artificial Intelligence set up by the European Commission: Ethics Guidelines for Trustworthy AI:
<https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai?fbclid=IwAR3YgoXsvFPqPjIEyMjvNNpKAwbn27gt4pzsN4oCmDlyd7IadzKWE2iq7jo>
- Falkanger, T. (2014). *SNL*. Hentet november 5, 2020 fra Ansvarlighet:
<https://snl.no/ansvarlighet>
- Ferrucci, D. A. (2012). Introduction to “This is Watson”. *IBM Journal of Research and Development*, 56(3.4), ss. 1:1-1:15.
- Fischer, J. M., & Ravizza, M. (1998). *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge University Press.
- Ghuri, P., & Grønhaug, K. (2010). *Research Methods in Business Studies: Fourth edition*. Pearson.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT press.
- Hanssen, A. B., & Nichele, S. (2019). Ethics of Artificial Intelligence Demarcations. I K. Bach, & M. Ruocco (eds), *Nordic Artificial Intelligence Research and Development. NAIS 2019. Communications in Computer and Information Science* (Vol. 1056). Springer, Cham.
- Hartvigsen, G., Johansen, D., & Måseide, A. (2000). *Etikk og informasjonsteknologi*. Fagbokforlaget.
- Holdorf, W. E., & Greenwald, J. M. (2018). Toward a taxonomy and unified construct of responsibility. *Personality and Individual Differences*, 132, ss. 115-135.
- Hoy, M. (2018). Alexa, Siri, Cortana, and More: An Introduction to Voice Assistants. *Medical Reference Services Quarterly*, 37(1), ss. 81-88.

- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to Statistical Learning: with Applications in R*. New York: Springer.
- Johannessen, A., Tuftes, P., & Christoffersen, L. (2016). *Introduksjon til samfunnsvitenskapelig metode*. Abstrakt forlag.
- Jordan, M., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245), ss. 255-260.
- 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), ss. 15-25.
- Képuska, V., & Bohouta, G. (2018). "Next-generation of virtual personal assistants (Microsoft Cortana, Apple Siri, Amazon Alexa and Google Home). *2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC)* (ss. 99-103). IEEE.
- Kolbjørnsrud, V. (2017). Kunstig intelligens og lederens nye jobb. *Magma*, 20(6), ss. 33-42.
- Kommunal- og moderniseringsdepartementet. (2020, 14. januar). *Regjeringen.no*. Hentet 16. september 2020 fra Rapporter og planer: Strategi for kunstig intelligens: <https://www.regjeringen.no/no/dokumenter/nasjonal-strategi-for-kunstig-intelligens/id2685594/?ch=1>
- Lagnado, D., Gerstenberg, T., & Zultan, R. (2013). Causal Responsibility and Counterfactuals. *Cognitive Science*, 37(6), ss. 1036-1073.
- Lødrup, P., & Gisle, J. (2010). *Store Norske Leksikon*. Hentet november 2020 fra Objektivt ansvar: https://snl.no/objektivt_ansvar
- Lindkvist, L., & Llewellyn, S. (2003). Accountability, responsibility and organization. *Scandinavian Journal of Management*, 19(2), ss. 251-273.
- Lovdata. (2015). *Ny straffelov*. Hentet november 2020 fra Medvirkning: https://lovdata.no/artikkel/ny_straffelov_-_medvirkning/1585
- Martin, R. (2010). The Age of Costumer Capitalism. *Harvard Business Review*.
- Müller, A., & Guido, S. (2018). *Introduction to Machine Learning with Python: A Guide for Data Scientists*. O'Reilly Media, Inc.
- McCarthy, J., Minsky, M., Rochester, N., & Shannon, C. E. (1955, August 31). *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*. Hentet fra <http://jmc.stanford.edu/articles/dartmouth.html>
- Meek, T., Barham, H., Beltaif, N., Kaadoor, A., & Akhter, T. (2016). Managing the Ethical and Risk Implications of Rapid Advances in Artificial Intelligence: A Literatur Review. *Engineering and Technology Management Faculty Publications and Presentations*(108).

-
- Merriam-Webster. (2020). *Dictionary*. Hentet november 2020 fra Responsibility: <https://www.merriam-webster.com/dictionary/responsibility>
- Mitchell, M. (2019). Artificial Intelligence Hits the Barrier of Meaning. *Information*, 10(2).
- Ness, R. (2002). Styrets rolle, ansvar og innflytelse - ti bud for godt styrearbeid. *MAGMA*, 3. Hentet fra <https://www.magma.no/styrets-rolle-ansvar-og-innflytelse-ti-bud-for-godt-styrearbeid>
- NUES. (2018). *Norsk utvalg for eierstyring og selskapsledelse*. Hentet november 2020 fra Eierstyring og selskapsledelse: <https://nues.no/eierstyring-og-selskapsledelse/>
- Pedersen, L. T. (2013). *Managing in dynamic business environments: between control and autonomy*. (K. Kaarbøe, P. N. Gooderham, & H. Nørreklit, Red.) Edward Elgar Publishing Limited.
- Phillips, R., Freeman, E., & Wicks, A. (2015). What Stakeholder Theory is Not. *Business Ethics Quarterly*, 13(4), ss. 479-502.
- Poole, D. L., & Mackworth, A. K. (2017). *Artificial Intelligence: Foundations of Computational Agents, 2nd Edition*. Cambridge University Press.
- Reusch, M. (2017). *Store Norske Leksikon*. Hentet november 2020 fra Hjemmel: <https://snl.no/hjemmel>
- Ricoeur, P. (1992). *The concept of responsibility: An essay in semantic analysis in Ricoeur The Just*. Chicago: University of Chicago Press .
- Roberts, J., & Scapens, R. (1985). Accounting systems and systems of accountability - understanding accounting practices in their organisational contexts. *Accounting, Organizations and Society*, 10(4), ss. 443-456.
- Robinson, S. (2009). The nature of responsibility in a professional setting. *Journal of Business Ethics*, 88(1), ss. 11-19.
- Russel, S., & Norvig, P. (2016). *Artificial Intelligence: A Modern Approach Third Edition*. Boston: Pearson Education Ltd.
- Sagdahl, M. (2020, juli). *Store Norske Leksikon*. Hentet november 2020 fra Etikk: <https://snl.no/etikk>
- Sarangi, S., & Sharma, P. (2018). *Artificial Intelligence: Evolution, Ethics and Public Policy*. Routledge India.
- Saunders, M., Lewis, P., & Thornhill, A. (2016). *Research Methods for Business Students: Seventh edition*. Harlow, England: Pearson Education Limited.
- Sautoy, M. (2019). *The creativity code: how AI is learning to write, paint and think*. HarperCollins Publishers Limited.
- Schwab, K. (2014). The Fourth Industrial Revolution. *World Economic Forum, First edition*.

- Shoemaker, D. (2011). Attributability, answerability, and accountability: Toward a wider theory of moral responsibility. *Ethics*, 121(3), ss. 602–632.
- Shoham, Y., Perrault, R., Brynjolfsson, E., Clark, J., Manyika, J., Niebles, J. C., . . . Bauer, Z. (2018). *Artificial Intelligence Index 2018 annual report*.
- Sinclair, A. (1995). The Chameleon of Accountability: Forms and Discourses. *Accounting, Organizations and Society*, 20(2/3), ss. 219-237.
- Smith, H. J. (2003). The Shareholders vs. Stakeholders Debate. *MIT Sloan Management Review*, 44(4), ss. 85-90.
- SNL. (u.d.). *Store Norske Leksikon*. Hentet november 2020 fra Juridisk person: https://snl.no/juridisk_person
- Stone, P., Brooks, R., Brynjolfsson, E., Calo, R., Etzioni, O., Hager, G., . . . Teller, A. (2016). *Artificial Intelligence and Life in 2030: One Hundred Year Study on Artificial Intelligence: Report of the 2015–2016 Study Panel*.
- Suddaby, R. (2006). What grounded theory is not. *Academy of Management Journal*, 49(No 4), ss. 633-642.
- Svensson, L. G. (2011). Profession, organisation, kollegialitet och ansvar. *Socialvetenskaplig tidskrift*, Vol 18(4).
- Syse, H., & Olsen, J. B. (2013). Hva er ansvar? *Magma*, ss. 17-22.
- Tennøe, T. (2017). *Kunstig intelligens: smart eller skremmende?* Hentet desember 2020 fra Norge 2030: <https://teknologiradet.no/kunstig-intelligens-smart-eller-skremmende/>
- Topol, E. (2019). High-performance medicine: the convergence of human and artificial intelligence. *Nature Medicine*, 25, ss. 45-56.
- Tranøy, K. E. (2018). *Store Norske Leksikon*. Hentet november 5, 2020 fra Ansvar: <https://snl.no/ansvar>
- Twiss, S. B. (1977). The problem of moral responsibility in medicine. *Journal of Medicine and Philosophy*, 2(4), ss. 330–375.
- Vetlesen, A. (2007). *Hva er etikk*. Universitetsforlaget.
- Wang, S. (2003). *Artificial Neural Network*. Boston, MA: Springer.
- World Economic Forum. (2017). *The Global Risks Report 2017, 12th Edition*. Geneva: World Economic Forum.
- Yin, R. K. (2014). *Case Study Research: Design and Method (5th edn)*. London: Sage.
- Østrem, K., & Schwerdt, T.-A. (2017). Personlig ansvar som styremedlem i aksjeselskap. *Selskapsrett*(4), ss. 30-34.

Vedlegg

Vedlegg 1: Intervjuguide

Del 1: Introduksjon

1.1 Om oss og prosjektet

- a) Takke for at respondentene er villig til å delta
- b) Kort presentasjon om oss
- c) Kort presentasjon om masterutredningen
 - i) «Vi skal skrive en master om ansvar relatert til kunstig intelligens i norske organisasjoner.»

1.2 Generell informasjon

- a) Oppsummer deltakeres rett til konfidensialitet og anonymitet. Er det greit at vi tar opp denne videosamtalen?
- b) Informer om selve intervjuprosessen
 - i) Vi skal stille noen spørsmål – svar ærlig og hva du mener

1.3 Om intervjuobjekt

- a) Hva jobber du med?
- b) Bruker dere noen form for kunstig intelligens eller maskinlæring i organisasjonen?
 - i) Hva er formålet med bruken?

Del 2: Respondentenes personlige refleksjoner

2.1 Kunstig intelligens

- a) Hvilke muligheter ser du ved bruk av kunstig intelligens for din organisasjon?
- b) Har dere hatt problemer ved bruk av kunstig intelligens?
 - i) Hvordan har dere håndtert problemene?
 - ii) Har problemene påvirket videre bruk av teknologien?
- c) Har dere sett på regjeringens strategi for bruk av kunstig intelligens?
 - i) Hva med andre rapporter?

2.2 Ansvar

- a) Hvordan forstår du begrepet ansvar?

- b) Hva er organisasjonens ansvar i forhold til bruk av kunstig intelligens?
 - i) Er ansvaret annerledes enn ansvar ovenfor tradisjonell teknologi?
- c) Ble ansvar diskutert før innføring av kunstig intelligens?
 - i) Har ansvar blitt diskutert i ettertid av innføringen?
- d) Har dere en rolle i organisasjonen som er direkte koblet til ansvar relatert til kunstig intelligens?
 - i) Hvis ikke, hvordan forholder ansatte seg til ansvarsproblematikk?
- e) Hva mener du er viktig i forhold til eksisterende lovverk ved bruk av deres teknologi?
 - i) Tenker dere på ansvar utover dette?
- f) Har organisasjons ansvar påvirket deres mulighet til å ta i bruk kunstig intelligens?
- g) Tror du organisasjoners ansvar vil endre seg i fremtiden?

Del 3: Avsluttende bemerkninger

- a) Er det noe du vil legge til nå på slutten, eller føler du at du ikke fikk svart godt nok på noe vi har vært inne på?
- b) Har du noen spørsmål til oss?
- c) Tusen takk for at du deltok!