



Høgskulen på Vestlandet

MOØ300 Masteroppgave

MOØ300

Predefinert informasjon

Startdato:	07-05-2021 09:00	Termin:	2021 VÅR
Sluttdato:	21-05-2021 14:00	Vurderingsform:	Norsk 6-trinns skala (A-F)
Eksamensform:	Masteroppgave		
Flowkode:	203 MOØ300 1 O 2021 VÅR		
Intern sensor:	Kristin Lofthus Hope		

Deltaker

Navn:	Frida Nord Arnesen
Kandidatnr.:	206
HVL-id:	587821@hvl.no

Informasjon fra deltaker

Egenerklæring *: Ja
Jeg bekrefter at jeg har Ja
registrert
oppgavetittelen på
norsk og engelsk i
StudentWeb og vet at
denne vil stå på
vitnemålet mitt *:

Gruppe

Gruppenavn: 11
Gruppenummer: Maiken Johannesen
Andre medlemmer i gruppen:

Jeg godkjenner avtalen om publisering av masteroppgaven min *

Ja

Er masteroppgaven skrevet som del av et større forskningsprosjekt ved HVL? *

Nei

Er masteroppgaven skrevet ved bedrift/virksomhet i næringsliv eller offentlig sektor? *

Nei



Høgskulen
på Vestlandet

MASTEROPPGAVE

Kunstig intelligens og Big Data i
forsikringsbransjen

Artificial intelligence and Big Data in the
insurance industry

Frida Nord Arnesen & Maiken Johannesen

Master i Innovasjon og ledelse
Institutt for økonomi og administrasjon
Veileder: Kjersti Berg Danilova
21.05.2021

Vi bekrefter at arbeidet er selvstendig utarbeidet, og at referanser/kildehenvisninger til alle kilder som er brukt i arbeidet er oppgitt, jf. Forskrift om studium og eksamen ved Høgskulen på Vestlandet, § 12-1.

Masteroppgavens tittel:	Kunstig intelligens og Big Data i forsikringsbransjen
Levert dato:	21.05.2021
Masteroppgavens omfang:	30 studiepoeng
Forfattere:	Frida Nord Arnesen og Maiken Johannesen
Mastergrad:	Innovasjon og ledelse
Antall sider uten vedlegg:	73
Antall sider med vedlegg:	101
Studieobjekt:	Innovasjonsprosjekt: Fraud Detection Platform
Metodevalg:	Kvalitativ studie
Sammendrag:	<p>Kunstig intelligens og Big Data er digitale teknologier som blir stadig mer relevant i flere bransjer, også i forsikringsbransjen som vi studerer i denne oppgaven. Studien belyser hvordan forsikringsbransjen på tvers av konkurrerende selskaper kan ta i bruk KI og Big Data for å bekjempe forsikringssvindel, og hvilke utfordringer som kan oppstå ved en slik løsning. Vi har derfor valgt å studere et innovasjonsprosjekt som jobber med å validere et konsept for trygg datadeling mellom konkurrerende selskaper for å bekjempe et felles problem. Vi fokuserer på de teknologiske, forretningsmessige og juridiske dimensjonene ved prosjektet. Det teoretiske rammeverket bygger derfor på flere begrep; innovasjon, åpen innovasjon, samkonkurranse, kunstig intelligens og personvern/GDPR. Oppgaven er en kvalitativ single case studie hvor vi studerer ett spesifikt innovasjonsprosjekt. Vi har i den forbindelse samlet inn data ved hjelp av semi-strukturerte dybdeintervjuer og gruppeintervjuer med deltakere i prosjektet. Kvalitativ metode ble ansett som mest hensiktsmessig ettersom konseptet vi studerer er innenfor et relativt nytt felt hvor det eksisterer lite forskning fra før. Ved å snakke med alle involverte parter i prosjektet har vi innhentet ulike synspunkt, men også fått avdekket at deltakerne i prosjektet har en felles forståelse om hva som er deres største</p>

utfordringer. Studien viser at det eksisterer et behov for å innovere og samarbeide i denne bransjen ettersom norske forsikringsselskaper alene kan bli for små i møte med økte kundeforventninger og internasjonal konkurranse. Norge er i dag preget av en relativt streng personvern- og konkurranselovgivning som potensielt kan hemme forsikringsbransjens muligheter og evne til å innovere. Videre avdekkes det et behov for retningslinjer og lovverk som tilrettelegger for innovasjon som gjør at forsikringsselskapene enklere kan etterleve kundenes forventninger i lys av digitalisering, økt internasjonal konkurranse og redusere økonomiske tap i forbindelse med forsikringssvindel.

Summary:

Artificial intelligence (AI) and Big Data are digital technologies that is becoming increasingly more relevant in several industries, including the insurance industry which we are studying in this master thesis. The study illuminates how the insurance can use AI and Big Data across competing companies to combat insurance fraud, and what challenges that may arise with such a solution. We have therefore chosen to study an innovation project that works to validate a concept for secure data sharing between competing companies in order to combat a common problem We focus on the technological, business and legal dimensions of the project. The theoretical framework is therefore based on several concepts: innovation, open innovation, coopetition, AI and privacy/GDPR. The thesis is a qualitative single case study where we look closer into a specific innovation project. In that context, we have collected data using semi-structured in-depth interviews and group interviews with all participants in the project. A qualitative method was considered most appropriate as the concept we are studying is within a relatively new field where little research has been

conducted in the past. By talking to all parties involved in the project, we have obtained different points of view, but also revealed that the participants have a common understanding of what their biggest challenges are. The study shows that there is a need to innovate and collaborate in this industry, as Norwegian insurance companies alone may become too small in the face of increased customer expectations and international competition. Norway is characterized by a relatively strict privacy and competition legislation that potentially can hamper the insurance industry's opportunities and ability to innovate. Furthermore, there is revealed a need for guidelines and legislation that facilitates innovation and makes it easier for insurance companies to live up to customers' increased expectations in light of digitalization, increased international competition and reducing financial losses due to insurance fraud.

Stikkord for biblioteket: Innovasjon, åpen innovasjon, samkonkurrans, kunstig intelligens, maskinl ring, Big Data, personvern, GDPR

Forord

Denne masteroppgaven er en avslutning på det toårige masterprogrammet Innovasjon og ledelse ved Høgskulen på Vestlandet. Oppgaven er skrevet av Frida Nord Arnesen og Maiken Johannesen ved institutt for Økonomi og administrasjon.

Vi ønsket å bruke denne oppgaven for å studere et innovasjonsprosjekt; Fraud Detection Platform, som handler om hvordan forsikringsbransjen kan dra nytte av kunstig intelligens og Big Data i kampen mot forsikringssvindel. Finansiell kriminalitet som hvitvasking har lenge vært et aktuelt tema i finansbransjen, men vi har sett lite til arbeid som er knyttet direkte mot å bekjempe forsikringssvindel. Det er i Norge avdekket forsikringssvindel for flere hundre millioner kroner hvert år, og det anses derfor som et reelt samfunnsproblem.

Innovasjonsprosjektet vekket interesse ettersom de benytter ny teknologi, digital innovasjon og samkonkurranse for å finne en bærekraftig modell for svindeldeteksjon. Dette er også temaer som har vært aktuelle gjennom flere av kursene i vårt masterprogram.

Uten bidrag fra deltakerne i innovasjonsprosjektet hadde det ikke vært mulig å skrive denne masteroppgaven, og vi vil derfor rette en stor takk til alle våre informanter. Vi vil også takke Høgskulen på Vestlandet for å tilby et svært interessant og relevant masterprogram som legger et godt grunnlag for vår vei videre. Videre vil vi takke vår veileder Kjersti Berg Danilova for verdifulle råd, tilbakemeldinger og ikke minst gode diskusjoner.

Til slutt vil vi takke hverandre for et godt samarbeid gjennom fem år med bachelor- og masterstudier.

Bergen, mai 2021

Frida Nord Arnesen & Maiken Johannesen

Innholdsfortegnelse

1.	INNLEDNING	1
1.1.	Forskningsspørsmål	2
1.2.	Studiens bidrag	2
1.3.	Oppgavens oppbygning	3
2.	TEORETISK RAMMEVERK	3
2.1.	Forsikring.....	4
2.1.1.	Finansiell svindel.....	4
2.1.2.	Digitalisering og digital innovasjon i forsikringsbransjen	5
2.2.	Innovasjon.....	7
2.2.1.	Åpen innovasjon.....	9
2.2.2.	Samkonkurransen	12
2.2.3.	Næringsklynger	13
2.3.	Teknologi.....	14
2.3.1.	Big Data.....	14
2.3.2.	Kunstig intelligens.....	16
2.3.3.	Typer intelligenser.....	17
2.3.4.	Felt innen kunstig intelligens – maskinlæring og dyplæring	19
2.3.5.	Black Box-problematikken.....	21
2.3.6.	Bias.....	22
2.4.	Personvern	23
2.4.1.	GDPR	26
2.5.	Teoretisk relevans	26
3.	METODE	27
3.1.	Forskningsdesign og tilnærming.....	27
3.2.	Forskningsmetode.....	28
3.3.	Casestudieforskning.....	29
3.4.	Valg av informanter	29

3.5.	Datagrunnlag.....	30
3.5.1.	Kvalitative sekundærdata og litteratur	31
3.5.2.	Kvalitative primærdata	31
3.6.	Analyse av data.....	33
3.7.	Datamaterialets kvalitet	34
3.7.1.	Validitet	34
3.7.2.	Reliabilitet	35
3.7.3.	Etiske betraktninger og personvern	36
3.8.	Studiens begrensninger	36
4.	EMPIRISK KONTEKST	36
4.1.	Presentasjon av innovasjonsprosjektet	37
4.2.	NCE Finance Innovation	37
4.3.	Tryg Forsikring	37
4.4.	Frende Forsikring.....	38
4.5.	Fremtind Forsikring	38
4.6.	Webstep	38
4.7.	PwC.....	39
5.	ANALYSE	39
5.1.	Fra problem til MVP.....	39
5.2.	Teknologisk løsning og utfordringer knyttet til denne	44
5.3.	Samkonkurransen	47
5.4.	Personvern	53
5.5.	Oppsummering av funn	59
6.	DISKUSJON	61
6.1.	Innovasjon.....	61
6.2.	Teknologi – KI og Big Data	62
6.3.	Samkonkurransen	63

6.4.	Personvern og GDPR.....	65
6.5.	Utfordringer knyttet til skalering	67
7.	KONKLUSJON.....	69
7.1.	Faglige implikasjoner	71
7.2.	Praktiske implikasjoner	72
7.3.	Metodiske begrensninger og forslag til videre forskning	72
8.	REFERANSELISTE	74
9.	VEDLEGG	82
9.1.	Vedlegg 1 – Samtykkeskjema NSD	82
9.2.	Vedlegg 2 – Godkjenning fra NSD	85
9.3.	Vedlegg 3 - Intervjuguide	87

Figurer:

- Figur 1: A map of innovation search space
- Figur 2: Åpen innovasjonsmodell
- Figur 3: The four intelligences
- Figur 4: Categories of machine learning algorithms according to training data nature

Tabeller

- Tabell 1: Prinsipielle forskjeller mellom åpen og lukket innovasjon
- Tabell 2: Intervjuobjekter
- Tabell 3: Fremstilling av kvalitative primærdata
- Tabell 4: Oppsummering av funn

1. INNLEDNING

Finansbransjen i Norge, herunder forsikring, har de siste årene opplevd store endringer på bakgrunn av digitalisering og nye digitale innovasjoner. Noen av disse endringene kommer som følge av en holdningsendring hvor bransjen er blitt mer kundeorientert, men også som følge av konkurranse fra store internasjonale selskaper. Begreper som FinTech (Financial Technology) og InsurTech (Insurance Technology) er blitt stadig mer vante begreper og preger bransjen i stor grad. Den tradisjonelle bank- og forsikringsbransjen har sett seg nødt til å være innovative for å kunne opprettholde konkurransefortrinn i et marked under endring. Som følge av FinTech og InsurTech har begreper som Big Data og Kunstig Intelligens (KI) blitt noe flere kjenner til, og som nå viser seg å være høyst relevant innen effektivisering og forbedring av tilbud og tjenester i forsikringsbransjen. I en rapport fra McKinsey Global Institute (2011), gjengitt i Datatilsynet (2013, s. 9), blir det fremhevet at Big Data vil øke konkurransekraften og innovasjonstakten i både offentlig og privat næringsliv. Datatilsynet henter opp denne tråden i en rapport fra 2019 (Regjeringen 2019, s. 1) hvor de fremhever at KI ligger høyt på agendaen i de aller fleste sektorer på bakgrunn av det store potensialet for bedre tjenester, forskningsmessige gjennombrudd og økonomiske gevinster. Selv om Big Data og KI er blitt mer relevant, er arbeidet med denne type teknologier så vidt startet i Norge, og Datatilsynet legger vekt på at vi befinner oss i startgroppen av noe som kan ha en betydelig effekt på samfunnet (Regjeringen, 2019, s. 1).

Forsikringsbransjen har imidlertid fått opp øynene for KI og Big Data, og ser at de kan benytte disse teknologiene i kampen mot forsikringssvindlere. Forsikringssvindel regnes som et samfunnsproblem og kriminell aktivitet, og forsikringsselskapene i Norge har derfor både plikt og interesse av å bekjempe dette problemet. I og med at forsikringsselskapene i Norge er relativt små sammenlignet med internasjonale konkurrenter, har det vokst fram en idé om at selskapene kan dele data seg imellom. Ved å dele data kan selskapene få tilgang til et større datasett (Big Data) hvor de kan trene sine respektive prediksjonsmodeller for svindel (KI). Basert på denne ideen har et innovasjonsprosjekt i regi av næringsklyngen NCE Finance Innovation vokst fram. Innovasjonsprosjektet har ønske om å validere et konsept for trygg datadeling mellom konkurrerende selskaper for å bekjempe et felles problem; forsikringssvindel. Vi ønsker gjennom denne studien å undersøke dette prosjektet nærmere, med fokus på de teknologiske, forretningsmessige og juridiske dimensjonene ved prosjektet.

1.1. Forskningsspørsmål

Kunstig intelligens og Big Data er teknologier som har fått mye oppmerksomhet de siste årene. Vi mennesker blir stadig mer digitale, og vi legger fra oss større mengder data enn noensinne. Store datamengder legger også til rette for å ta i bruk kunstig intelligens for å enten effektivisere prosesser eller bedre kunne være i stand til å løse utfordringer. Basert på dette ønsker vi i denne kvalitative casestudien å avdekke hvordan forsikringsselskaper kan dele data på tvers av egne organisatoriske grenser, for å sammen bekjempe forsikringssvindel. På bakgrunn av at vi studerer et spesifikt innovasjonsprosjekt som jobber med å validere en slik løsning ved bruk av kunstig intelligens og Big Data, vil vi også se på hvilke utfordringer som kan oppstå i et slikt prosjekt. Tema for oppgaven er derfor kunstig intelligens og Big Data i forsikringsbransjen, og vi ønsker å besvare følgende forskningsspørsmål:

Hvordan kan forsikringsbransjen i Norge skape en felles stordataplattform for å bekjempe forsikringssvindel?

Hva er de største utfordringene i utviklingen av en felles stordataplattform som potensielt kan forbedre arbeidet med å detektere og predikere forsikringssvindel?

For at vi best mulig skal være i stand til å svare på oppgavens forskningsspørsmål har vi valgt å ta i bruk kvalitativ tilnærming for å få en bedre forståelse av temaet vi studerer. Vi ønsker derfor å gjennomføre dybdeintervjuer med deltakerne i det aktuelle innovasjonsprosjektet, hvor målet er å tilegne oss mest mulig informasjon som kan bidra til forståelse og refleksjon rundt oppgavens tema og forskningsspørsmål.

1.2. Studiens bidrag

Studien vil kunne bidra til å belyse et stadig mer relevant område innenfor utviklingen av teknologi i forsikringsbransjen. Dette spesielt med tanke på utviklingen og bruken av kunstig intelligens og Big Data. Studien vil også bidra til å poengtere viktigheten av å bekjempe forsikringssvindel som et samfunnsproblem på lik linje med annen finansiell kriminalitet. Masteroppgaven vil gi økt innsikt i innovasjonsprosjektet som sådan. Studien fremmer viktigheten av at konkurrerende selskaper går sammen for å bekjempe et felles samfunnsproblem, samt hvilke utfordringer som kan oppstå ved arbeidet med et slikt komplekst prosjekt. Oppgaven belyser også at gjennomføring av teknologiske

innovasjonsprosjekter på tvers av ulike selskaper krever at man løser utfordringer knyttet til både teknologi, samkonkurranse og personvern.

1.3. Oppgavens oppbygning

Kapittel to tar for seg oppgavens teoretiske rammeverk, og har som hensikt å gi oss som forskere et godt teoretisk grunnlag for å kunne svare på våre forskningsspørsmål, samt gi leseren en forståelse av studiens tema. Videre presenteres redegjørelse for oppgavens metodiske valg i kapittel tre. Oppgavens fjerde kapittel tar for seg presentasjon av caset som studeres, samt involverte aktører. I kapittel fem presenteres funnene innhentet i vår datainnsamling. Videre tar kapittel seks for seg diskusjon hvor vi analyserer funnene våre opp mot det teoretiske rammeverket for oppgaven. Avslutningsvis i kapittel syv vil vi presentere oppgavens konklusjon med svar på studiens forskningsspørsmål, implikasjoner og forslag til videre forskning.

2. TEORETISK RAMMEVERK

I dette kapittelet presenterer vi teori og litteratur som er relevant for oppgavens tema og forskningsspørsmål. Innovasjonsprosjektet har ikke vært forsket på tidligere, og vi har derfor sett det nødvendig med en grundig litteraturgjennomgang for å finne et teoretisk grunnlag som passer med oppgavens tema. Litteratur og teori som er presentert i dette kapittelet har bidratt til å skape forståelse for innovasjonsprosjektet og dets utfordringer.

På bakgrunn av at innovasjonsprosjektet finner sted i en spesifikk bransje ønsker vi å innlede vårt teoretiske rammeverk med en beskrivelse av forsikringsbransjen, hva forsikringssvindler innebærer, samt digitalisering og digital innovasjon i bransjen. Innovasjonsprosjektet bygger på digital innovasjon, og det blir dermed naturlig å gjøre rede for dette og tilhørende begreper. Dernest legger vi fram begrepet innovasjon, med tilhørende underemner; åpen innovasjon, samkonkurranse og næringsklynger. Videre presenteres bakgrunnen for den teknologiske løsningen som er kunstig intelligens og Big Data. Vi avslutter med å presentere personvernlovgivningen. Underveis i det teoretiske rammeverket blir det også lagt fram utfordringer knyttet til begrepene samkonkurranse, teknologi og personvern, som danner grunnlaget for videre analyse.

2.1. Forsikring

Begrepet forsikring har eksistert like lenge som samfunnslivet, og kan spores langt tilbake i tid. Det vokste fram av man som medlem av gilder eller broderskap, ga hverandre gjensidig hjelp ved ulike tilfeller som blant annet dødsfall, brann og tyveri (Døving, 2018). Dette gjorde at man som enkeltperson ikke stod med hele risikoen selv dersom man ble utsatt for en av nevnte tilfeller, men man tok støytten sammen. Basert på denne tankegangen ble forsikringsselskaper etablert. Forsikring er en gjensidig ordning, og defineres i Store Norske Leksikon av Døving & Loen (2018, u.s.) som; «en ordning som har til oppgave å erstatte økonomisk tap som forårsakes av tilfeldige, uforutsette hendelser, som for eksempel tyveri, brann, bil- og flyulykker, forlis, naturskader, ulykkeshendelser, dødsfall og skade påført annen person eller annen persons eiendom».

Som individ har man ofte store verdier knyttet til eiendeler som blant annet bil, båt og hus, og man vil gjerne på bakgrunn av dette ikke sitte alene med risikoen om det skulle oppstå et skadetilfelle. Det er derfor mange velger å inngå en forsikringsavtale. Denne avtalen mellom forsikringstaker og forsikringsgiver er kontraktmessig regulert i forsikringsbeviset og forsikringsvilkårene (Døving & Loen, 2018). Ved å inngå en slik forsikringsavtale vil man som forsikringstaker betale en gitt sum i året for å forsikre ønskede produkter, som ofte kalles en forsikringspremie. Dersom man da skulle være så uheldig å oppleve et skadetilfelle, vil man som forsikringstaker i de fleste tilfeller kun måtte betale en egenandel hvor forsikringsselskapet vil dekke resten av kostnadene i forbindelse med skaden.

Avtaleforholdet mellom forsikringsgiver og forsikringstaker er basert på gjensidig tillit. Forsikringsgiver krever at forsikringstaker ikke sender inn uriktige skadekrav, er ærlig ved utfylling av helseskjema og at all informasjon som gis er korrekt. På den andre siden krever forsikringstaker at forsikringsgiver gir utbetalinger i de tilfeller forsikringstaker har krav på dette, og at denne prosessen skjer sømløst. Det er likevel ikke alle som bryr seg like mye om dette tillitsforholdet, og som ønsker å utnytte systemet.

2.1.1. Finansiell svindel

Finansiell svindel blir sett på som et stort samfunnsproblem og defineres av Association of Certified Fraud Examiners gjengitt i Zhou og Kapoor (2011, s. 570) som; «et bevisst bruk av ulovlige metoder eller praksis for å oppnå økonomisk gevinst». Denne definisjonen er også blitt spesifisert inn mot forsikringsbransjen. Folkehelseinstituttet beskriver forsikringssvindel

som «et bevisst bedrag begått mot eller av et forsikringsselskap med et formål om økonomisk gevinst» (FHI, 2018, s. 9). Forsikringssvindel kan finne sted når som helst i forsikringsprosessen, og kan også utføres av selskapets ansatte, men som oftest handler forsikringssvindel om at kunden prøver å tilegne seg midler den ikke har krav på (West & Bhattacharya, 2016, s. 51).

Begge parter i et forsikringsforhold, forsikringsselskap og forsikringstaker, er juridisk pålagt å handle i god tro mot hverandre som forplikter dem til å gjensidig oppgi all den informasjonen som er nødvendig (Viaene & Dedene, 2004, s. 314). Å handle i mangel på god tro vil ikke automatisk regnes som svindel, men uriktige handlinger som med viten og vilje er gjennomført for å oppnå fordeler man ikke har krav på, vil regnes som svindel (Viaene & Dedene, 2004, s. 314-315). Det finnes utallige måter å utføre forsikringssvindel på, og svindlere finner stadig nye måter å lure selskapene. Folkehelseinstituttet (2018, s. 9) presenterer de mest vanlige måtene å svindle på som for eksempel «oppblåsing av faktiske krav, feilaktig fremstilling av fakta på en forsikringsøknad, innsending av skader som aldri har skjedd, og planlagte ulykker». Spesielt innenfor bilskade ser man at flere forsøker å svindle; falske krav blir meldt inn, mens noen kollisjoner er også planlagte (West & Bhattacharya, 2016, s. 51).

2.1.2. Digitalisering og digital innovasjon i forsikringsbransjen

Forsikringsbransjen har over tid gjennomgått store forandringer – fra oppstart i fysiske lokaler med personlige relasjoner, til å gå over til en mer digital tilnærming og ofte uten en personlig relasjon. Endringen fra den fysiske tilstedeværelsen til at alt nå foregår digitalt, kan ses i sammenheng med *digitalisering*. Digitalisering er blitt et moteord (Andersen & Sannes, 2017, s. 18) og begrepet brukes i stadig flere sammenhenger; alt fra hverdagslige samtaler, til diskusjoner blant politikere og ledere i både små og store organisasjoner (Osmundsen et al., 2018, s. 1). Det finnes mange ulike definisjoner av digitalisering, og ofte ser man at de fleste definisjonene bygger på definisjonen som er utarbeidet av Yoo og kollegaer (Osmundsen et al., 2018, s. 3);

Med digitalisering mener vi transformasjon av sosiotekniske strukturer som tidligere ble formidlet av ikke-digitale gjenstander eller relasjoner, til strukturer som nå blir formidlet gjennom digitale gjenstander eller relasjoner. Digitalisering handler om mer enn bare en teknisk prosess for koding av analog informasjon til et digitalt format (dvs. digitisering)

og innebærer å organisere nye sosiotekniske strukturer med digitaliserte gjenstander, samt endringer i selve gjenstandene (Hentet og oversatt fra; Yoo et al., 2010, s. 6).

Mange er av den oppfatning av at digitalisering kun innebærer å gå vekk fra å sende dokumenter i papirformat, og heller behandle papirene i digitale verktøy. Å konvertere noe fra analogt til digitalt, eksempelvis oversendelse av dokumenter, kan ifølge Andersen og Sannes (2018, s. 197) ikke beskrives som digitalisering, men som *digitisering*. Til sammenligning vil digitalisering være når man knytter sammen systemer som gjør at man ikke trenger å sende dokumenter i det hele tatt (Andersen & Sannes, 2018, s. 197). Digitalisering i et selskap handler i stor grad om å tilpasse funksjoner og prosesser for å være konkurransedyktige og produsere mer effektivt, noe som også er svært relevant for forsikringsbransjen.

Flere selskaper i bransjen benytter seg i dag av løsninger som er vokst fram ved hjelp av ny teknologi. Eksempler på dette kan være automatisert skadebehandling, stemmeassistenter, chatbots og maskinlæring, noe som kan kalles for *digital innovasjon*. Digital innovasjon ses ofte i sammenheng med digitalisering. Det finnes heller ikke her en klar og forenlig definisjon av begrepet, og det eksisterer dermed flere tolkninger. Noen forskere bygger definisjonen på at innovasjon er en prosess, mens andre heller fokuserer på sluttresultatet av det som skapes (Osmundsen et al., 2018, s. 5). Ved innovasjon som bygger på prosess legges det ofte vekt på selve arbeidet med å utvikle, produsere, skape og transformere nye produkter, tjenester eller ideer (Osmundsen et al., 2018, s. 5). Mens innovasjoner som bygger på sluttresultatet i hovedsak fokuserer på utfallet av en innovasjon (Fichman et al., 2014, s. 333). Forskning viser likevel en gjensidig avhengighet mellom digital innovasjon som prosess og resultat, og at det uansett handler om noe som er helt nytt, i tillegg til at innovasjonen er basert på digital teknologi (Osmundsen et al., 2018, s. 6).

Digitale innovasjoner og endringer i kundeatferd regnes som store pådrivere for de endringene forsikringsbransjen til nå har vært gjennom, men også med hensyn til endringstakten de opplever i dag. I finansbransjen generelt, kan man se at disruptive Finansielle Teknologiske (FinTech) selskaper har etablert seg og forstyrrer det eksisterende markedet. Derfor har også tradisjonelle finansielle selskaper sett seg nødt til å tenke nytt for å ikke bli utkonkurrert av FinTech-selskaper som leverer produkter og tjenester med større fokus på teknologi, innovasjon, effektivitet og brukervennlighet. Basert på FinTech har begrepet InsurTech vokst frem. Dette begrepet tar for seg bruken av teknologi for å kunne

tilby mer effektive forretningsmodeller spesifikt for forsikringsbransjen (Cortis et al., 2019, s. 72).

Ny teknologi som forsikringsbransjen eksperimenterer med i dag, er for eksempel kunstig intelligens og maskinlæring. Disse teknologiene kan bidra til mer nøyaktighet, og hjelpe til med å predikere komplekse sammenhenger (Holtan, 2019) på tilbudssiden, men også i selve skadebehandlingen. Gode prediksjonsmodeller for svindel basert på maskinlæring og Big Data er et dagsaktuelt tema og denne type teknologi for svindeldetektering blir stadig mer relevant ettersom digitaliseringen i forsikringsbransjen har bidratt til større økonomiske kostnader i forbindelse med svindel (Viaene & Dedene, 2004, s. 313). Vi kommer nærmere inn på aktuelle teknologier i kapittel 2.3.

Når nye teknologier og digitale innovasjoner påvirker forretningsmodeller og arbeidsprosesser i selskapet kan det kalles *digital transformasjon*. Digital transformasjon kan ses på som et resultat av digitisering, digitalisering og digital innovasjon, og innebærer den endringen som forekommer i økonomi, institusjoner og samfunn (Andersen & Sannes, 2018, s. 197). Dette blir ofte utløst ved endringer i kundeatferd og forventninger, og sammen med økt bruk av ny teknologi har dette endret måten selskaper drives på, samt hvordan de konkurrerer og samhandler (Osmundsen et al., 2018, s. 5).

2.2. Innovasjon

Innovasjon er et begrep som er blitt stadig mer populært som følge av digitalisering, og ansett som avgjørende for økonomisk vekst, både for selskaper så vel som regioner og nasjoner (Njøs og Sjøtun, 2016, s. 4). Joseph Schumpeter, en østerisk økonom på 1900-tallet, blir i mange tilfeller omtalt som innovasjonens far, og lyktes i stor grad med å utvikle en forståelse for at innovasjon var et sosialt fenomen som bidro til å forme økonomisk utvikling (Fagerberg, 2003, s. 135). Det er i all hovedsak Schumpeters forståelse og definisjon av innovasjon denne oppgaven baserer seg på. Definisjonen er godt etablert i Norge, og brukes blant annet aktivt av Regjeringen og Innovasjon Norge; «En ny vare, en ny tjeneste, en ny produksjonsprosess, anvendelse eller organisasjonsform som er lansert i markedet eller tatt i bruk i produksjonen for å skape økonomiske verdier» (Regjeringen, 2008-2009, s. 13).

Definisjonen som er utarbeidet av Schumpeter, baserer seg på 5 ulike dimensjoner innenfor innovasjon og presenteres i boken «*The Theory of Economic Development*»:

1. Introduksjon av et nytt produkt eller tjeneste, eller en betydelig forbedring av et produkt eller en tjeneste som allerede eksisterer.
2. Introduksjon av en ny produksjonsmetode.
3. Introduksjon eller etablering i et nytt marked – uavhengig om dette markedet eksisterte fra før, eller ikke.
4. Fremskaffelse av nye leveranskilder eller råvarekilder for halvfabrikata og/eller produktkomponenter.
5. Innføring av nye metoder for organisering av både organisasjon og bransje, utvikle en monopolposisjon, eller oppløsning av et eksisterende monopol.

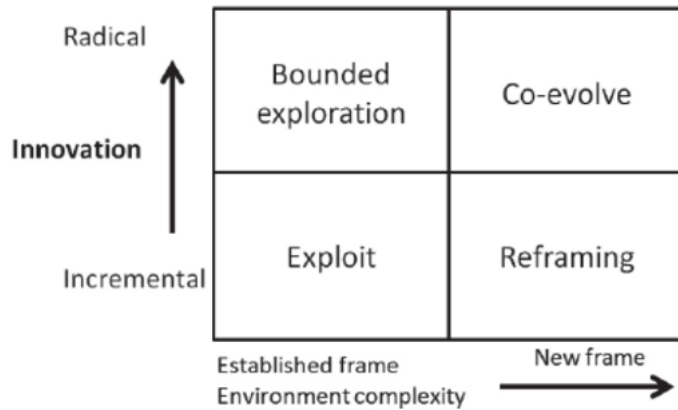
(Hentet og oversatt fra Schumpeter, 1934, s. 66).

Selv om begrepet innovasjon, og definisjonen av Schumpeter er godt etablert, finnes det mange tolkninger av både begrepets definisjon og dimensjoner. Likevel ser vi at de aller fleste bygger på de elementære byggsteinene i definisjonen til Schumpeter. Blant annet kan vi se på en mer moderne definisjon av begrepet som er gjort av Tidd og Bessant; «Proessen med å gjøre ideer om til virkelighet, og fange verdi av disse ideene. Det innebærer også at en idé, eller en oppfinnelse, ikke kan regnes som en innovasjon før den er praktisk anvendt» (Oversatt fra Tidd & Bessant, 2018, s. 16).

Tidd og Bessant bygger også sin definisjon på flere dimensjoner, men av samme karakter som Schumpeter. Dimensjonene har til felles at de bygger på produktinnovasjon, prosessinnovasjon, markedsinnovasjon og organisatorisk innovasjon (Tidd & Bessant, 2018, s. 21). Felles for definisjonene er at de begge legger vekt på at ideen ikke kan regnes som en innovasjon før den faktisk er anvendt, eller satt i funksjon. Et annet viktig trekk ved innovasjon er at det ofte blir omtalt som en pådriver for utvikling og økonomisk vekst (Njøs og Sjøtun, 2016, s. 4), men at man også må drive med kontinuerlig innovasjon for å kunne lykkes på lang sikt (Tidd og Bessant, 2018, s. 16). Det underbygges av Marx og Schumpeter som sier at økonomisk belønning med vellykket innovasjon er forbigående, og forsvinner så snart en tilstrekkelig mengde etterlignere har kommet på banen (Fagerberg, 2003, s. 130).

Tidd og Bessant (2018, s. 228) skiller også mellom inkrementelle og radikale innovasjoner. Inkrementell innovasjon innebærer at vi forbedrer det vi allerede gjør og blir omtalt som

utnyttende (exploitation), mens radikal innovasjon innebærer at vi gjør noe helt nytt og blir omtalt som utforskende (exploration). Tidd og Bessant har utviklet en modell som viser de ulike gradene av innovasjon, og denne består av fire soner;



Figur 1: A map of innovation search space (Tidd og Bessant, 2018, s. 235).

Sone 1 (exploit) forutsetter at det ligger en stabil og delt ramme, hvor adaptiv og inkrementell utvikling finner sted. Sone 2 (bounded exploration) innebærer at man utforsker et nytt territorium, men det foregår fortsatt innenfor et etablert rammeverk. Sone 3 (reframing) innebærer å utforske i et rom der alternative arkitekturer blir generert, og forskjellige elementer i dette miljøet. Til slutt har vi sone 4 (co-evolve); denne sonen representerer kaotiske og komplekse omgivelser hvor innovasjon fremstår som et produkt av en prosess med stor grad av samutvikling (Tidd og Bessant, 2018, s. 236-237).

Innovasjon kan ses i sammenheng med konkurranse, og for mange selskaper er det innovasjon som bidrar til at de holder seg relevante i markedet. Innovasjon trenger nødvendigvis ikke være en materiell idé eller en oppfinnelse, men vi kan også se innovasjon i sammenheng med blant annet strategi, prosesser og forretningsmodeller.

2.2.1. Åpen innovasjon

Det kommer stadig nye disruptive aktører som forstyrrer dagens marked og bidrar til en hurtigere endringstakt – desto viktigere blir det å ha en endringsdyktig forretningsmodell som kan møte disse endringene på en god måte. Forretningsmodeller blir mindre fleksible og mer endringsresistente over tid, ettersom de ikke er designet for å endres. Desto lengre en forretningsenhet eksisterer, jo mer vil forretningspraksis og dens tilnærming til å løse utfordringer, bli mer inngrodd. Inngrodd tilnærming kan ses i sammenheng med selskapets

kultur (Christensen et al., 2016, s. 34). For å møte endringer i markedet er det gjerne nødvendig å gjøre organisatoriske endringer for å kunne møte nye utfordringer på best mulig måte, her aktuelt; innovasjon i selskapets forretningsmodell. Forretningsmodellinnovasjon kan hjelpe selskaper med å være i forkant av endringer i markedet. Amit og Zott (2012, s. 2) refererer til en bedriftsleder i sin studie som har uttalt følgende; «Du er alltid én innovasjon unna å bli utkonkurrert, noe som eliminerer behovet for ditt produkt».

Innenfor innovasjon og innovasjon i forretningsmodeller, finner vi begrepet *åpen innovasjon*. Henry Chesbrough, en amerikansk organisasjonsteoretiker, har utarbeidet en definisjon av begrepet som vi ønsker å legge til grunn i oppgaven; «innstrømming og utstrømming av kunnskap som skal bidra til å øke hastigheten av intern innovasjon, og henholdsvis utvide markedene for ekstern bruk av innovasjon» (Egen oversettelse; Chesbrough, 2005, s. 2). Med andre ord kan det åpne innovasjonsparadigmet forstås som motsatsen til tradisjonelle, interne FoU-aktiviteter (Chesbrough, 2005, s. 2), som ofte blir omtalt som *lukket innovasjon*. Ved lukket innovasjon vil selskaper generere, utvikle og bygge ideer internt, samt markedsføre, distribuere, vedlikeholde og finansiere disse internt (Chesbrough, 2003, s. xx), noe som setter stor lit til egne ressurser. Det lukkede innovasjonsparadigmet har i senere tid blitt utfordret av kompetente medarbeideres økende mobilitet blant arbeidsgivere, men også av den økende innovasjonstakten (Chesbrough, 2003, s. 34) som setter høyere krav til selskapene. På bakgrunn av dette er lukket innovasjon ikke lenger regnet som bærekraftig på lang sikt (Chesbrough, 2003, s. xxiv).

Chesbrough (2003, s. xxvi) illustrerer en tabell med ulike prinsipper som skiller mellom åpen og lukket innovasjon;

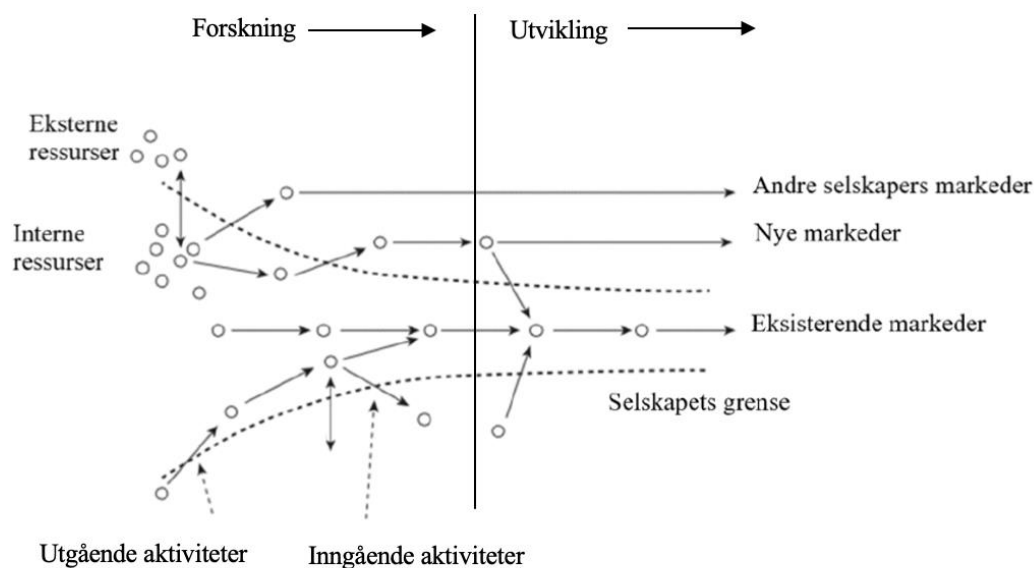
Lukket innovasjon	Åpen innovasjon
De mest kompetente på feltet jobber for oss.	Ikke alle de mest kompetente jobber for oss. Vi må jobbe med kompetente ressurser både innenfor, men også utenfor selskapet.
For å skape fortjeneste av FoU må vi oppdage, utvikle og distribuere det selv.	Ekstern FoU kan skape signifikant verdi; intern FoU er nødvendig for å kreve en del av den verdien.
Har en idé opphav hos oss, får vi det på markedet først.	Ideer trenger ikke å ha opphav hos oss for at vi skal ha fortjeneste på den.

Selskapet som kommer med en innovasjon på markedet først, vinner.	Å bygge en bedre forretningsmodell er bedre enn å komme på markedet først.
Genererer vi mange av de beste ideene på markedet, vinner vi.	Dersom vi optimaliserer bruken av interne og eksterne ressurser, vinner vi.
Vi må ha kontroll på våre immaterielle rettigheter (IP), slik at ikke konkurrenter tjener på våre ideer.	Vi burde dele vår IP og tjene på det, samt kjøpe andre sine IP når det gagnar vår forretningsmodell.

Tabell 1: Prinsipielle forskjeller mellom åpen og lukket innovasjon (Hentet og oversatt fra Chesbrough, 2003, s. xxvi).

Selv om man skiller mellom åpen og lukket innovasjon er det ikke nødvendigvis slik at selskaper skiller mellom disse i praksis. Som oftest ser vi at selskaper investerer i både lukkede og åpne innovasjonsaktiviteter samtidig (Enkel et al., 2009, s. 312). Enkel et al. (2009, s. 312) hevder også at for mye åpenhet kan ha negativ påvirkning på selskapenes innovasjonssuksess på lang sikt fordi det kan bidra til mangel på kontroll og tap av viktig kompetanse. Samtidig vil selskaper som opererer mer lukket, ikke ha mulighet til å tjene kravet om kortere innovasjonssykluser og tid til markedet. I samme artikkel legges det også vekt på at en passende balanse mellom åpne og lukkede innovasjonsaktiviteter er fremtiden; der selskaper bruker alle tilgjengelige verktøy for å skape vellykkede produkter og tjenester raskere enn konkurrentene, samtidig som de bygger på intern kjernekompetanse og beskytter deres intellektuelle ressurser. En kan derfor si at selskaper som har evne til å kombinere åpne og lukkede innovasjonsaktiviteter, gjerne også har større konkurransekraft.

Blant åpne innovasjonsaktiviteter skiller man mellom *inbound* og *outbound* aktiviteter, heretter kalt inngående og utgående. Ved inngående åpen innovasjon vil eksterne ideer eller teknologi strømme inn i selskapet, mens ved utgående åpen innovasjon vil interne ideer eller teknologi brukes av et annet selskap som er bedre skikket til å videreutvikle og kommersialisere dem (Usman & Vanhaverbeke, 2017, s. 174). Åpen innovasjon og aktivitetsprosessene i denne modellen kan illustreres som en trakt; Chesbrough (2003, s. xxv; 2012, s. 23) har utarbeidet en åpen innovasjonsmodell som illustrerer denne trakten. Vi har gjenkonstruert denne modellen i en egendefinert modell hvor vi også illustrerer selve innovasjonsaktivitetene;



Figur 2: Åpen innovasjonsmodell (Gjenkonstruert og oversatt fra Chesbrough, 2003, s. xxv; 2012, s. 23).

2.2.2. Samkonkurransen

De fleste selskaper står i dag overfor omfattende og raske endringer som følge av digitalisering, ny teknologi og økt internasjonal konkurranse. Dette gjør at det stilles høyere krav til innovasjon. Samkonkurransen, samarbeid og konkurranse på samme tid, er et konsept som er antatt å lede til innovasjon (Nesse, 2018, s. 61). Bengtsson og Kock definerer samkonkurransen, på engelsk kalt *coopetition*, som: «det dyadiske og paradoksale forholdet som oppstår når to eller flere selskaper samarbeider om visse aktiviteter, eksempelvis en strategisk allianse, og samtidig konkurrerer i andre aktiviteter» (Egen oversettelse fra Bengtsson & Kock, 2000, s. 412).

Tidligere har man tradisjonelt sett at forhold mellom selskaper enten fokuserer på konkurranse (*competition*) eller samarbeid (*cooperation*). Konkurranse handler om at selskaper fokuserer på å handle etter egne mål ved å prøve å utkonkurrere andre selskaper i markedet, mens samarbeid handler om at flere selskaper går sammen for å nå et felles mål. Begrepet *coopetition* er vokst ut av disse to begrepene, og tar for seg at konkurranse og samarbeid kan anses som gjensidig uforenelige, men likevel ikke helt uavhengige av hverandre (Bengtsson & Kock, 2002, s. 414). Samkonkurransen er derfor når konkurranse og samarbeid foregår på samme tid, mellom to eller flere rivaler som konkurrerer i samme

marked (Luo, 2007, s. 129). Bakgrunnen for at selskaper finner dette attraktivt er på grunn av økt kompleksitet, kort produktlivssyklus og økende globalisering. Dette vil gjøre at selskaper kan holde seg relevante, og samarbeidet blir ofte ansett som verdifullt med hensyn til fordeling av kostnader og arbeidskapasitet, samt deling av kompetanse.

Samkonkurransen kan bli sett på som en styrke innen utviklingen av innovative løsninger. Chris McCarthy illustrerer viktigheten av å drive med samkonkurransen, gjennom sport. Profesjonelle syklister koordinerer bevegelsene sine slik at de som gruppe kan gjøre det bedre enn det de ville gjort alene (McCarthy, 2018, s. 26). På lik linje kan selskaper benytte seg av samme strategi for å styrke deres innovative kraft, og på denne måten komme ut av alliansen som mer konkurransedyktig enn de hadde vært alene.

Ved å inngå samarbeid med andre selskaper vil likevel dette kunne tilrettelegge for bedre flyt av ressurser, kapasitet og kunnskap, som igjen kan styrke deres innovative kraft på et høyere nivå enn man ville gjort alene (Bouncken et al., 2015, s. 579). Selv om selskaper inngår slike samarbeid, er det viktig å ikke glemme at de er konkurrenter. I Norge er slike samarbeid regulert av konkurranselovgivningen. Lovens formål er å fremme konkurranse for derigjennom å bidra til effektiv bruk av samfunnets ressurser, og lyder som følger; «Ved anvendelse av denne lov skal det tas særlig hensyn til forbrukernes interesser» (Lovdata, 2004).

For å kunne håndtere kompleksiteten og gjøre det mulig å dra nytte av et slikt forhold, er det viktig å skille de to delene av samkonkurransen (Bengtsson & Kock, 2000, s. 412). Mistillit og konflikter er to sentrale utfordringer som kan spole samarbeidet og true deres samarbeidsevne, ved at man eksempelvis ikke har like strategiske mål (Hamel et al., 1989, s. 135). På den andre siden kan felles interesser, tillit og god kommunikasjon gi partene verdifullt utbytte av samarbeidet. For at samspillet mellom samarbeid og konkurranse skal kunne lykkes, er det nødvendig å oppnå en balanse mellom disse begrepene (McCarthy, 2018, s. 28).

2.2.3. Næringsklynger

Et tiltak for å tilrettelegge for samarbeid på tvers av selskaper og bransjer er dannelsen av næringsklynger. Innovasjon Norge beskriver en næringsklynge som et samarbeidsgrunnlag som omfatter etablert næringsliv, gründere, investorer, kunnskapsmiljøer og offentlige

utviklingsaktører (Innovasjon Norge, u.å.). I Norge har vi et klyngeprogram i regi av Innovasjon Norge med navn Norwegian Innovation Cluster. Michael Porter definerer en klynge som; «geographic concentrations of interconnected companies and institutions in a particular field» (1998, s. 78), altså en geografisk konsentrasjon av selskaper og organisasjoner innenfor et bestemt felt. Norwegian Innovation Cluster består av flere klynger innenfor ulike felt; blant annet finans, sjømat og forskning. Formålet med slike klynger er å forsterke samarbeidsbaserte utviklingsaktiviteter, øke klyngens dynamikk og attraktivitet, samt øke det enkelte selskaps innovasjonsevne (Innovasjon Norge, u.å.). Porter (1998, s. 80) underbygger dette ved å presisere at klyngesamarbeid bidrar til å øke produktivitet, forsterker innovasjonstakten og stimulerer dannelse av nye selskaper, som igjen utvider og styrker selve klyngen.

Ved å være medlem i en klynge åpner det seg muligheter for økt tilgang til ekstern informasjon, kompetanse og teknologi, samt muligheter for samarbeid med relaterte selskaper, som kan øke motivasjonen for forbedring av produkter og tjenester (Porter, 1998, s. 81). Klyngesamarbeid er også en kjent driver for økt innovasjonsevne; for det første har selskapene i klyngen gjerne bedre innsikt i markedet enn de ville hatt dersom de opererte alene; for det andre, bidrar det til at selskapene kan lære av hverandre; og for det tredje, klyngen bidrar til økt fleksibilitet og kapasitet til å reagere raskt på endringer i markedet. Samtidig er klyngen et sted hvor selskapene kan eksperimentere på nye løsninger for lavere kostnader, og unngå store forpliktelser inntil man er sikker på at en gitt innovasjon vil være en suksess (Porter, 1998, s. 83).

2.3. Teknologi

2.3.1. Big Data

Data refererer til fakta og statistikk som er samlet for å vurderes i forbindelse med beslutningstaking, eller informasjon i elektronisk form som kan lagres og brukes av en datamaskin for diverse analyser (Cambridge Dictionary, 2021). Data i store mengder blir omtalt som Big Data; på norsk brukes også oversettelsen stordata. Big Data er verken et nytt konsept, eller en ny idé (Zaslavsky et al., 2013, s. 2), og begrepet refereres ofte til store datamengder og aktivitetene knyttet til å både samle inn, lagre og analysere disse dataene (Datatilsynet, 2013, s. 7). Vi har valgt å legge til grunn EU-kommisjonens rådgivende organ i personvernspørsmål, Artikkel 29-gruppen, sin definisjon av Big Data. Definisjonen er gjengitt i en Big Data-rapport utarbeidet av Datatilsynet;

Big Data refererer til den enorme økningen i tilgang til, og automatiserte bruk av, opplysninger: det refererer til gigantiske mengder digitale data som er kontrollert av selskap, myndigheter og andre store organisasjoner, og som gjøres til gjenstand for omfattende analyse ved bruk av algoritmer. Big Data kan bli brukt til å identifisere generelle trender og sammenhenger, men kan også bli benyttet slik at det berører enkeltindivider direkte (Datatilsynet, 2013, s. 7).

Datatilsynet (2013, s. 4) legger vekt på at Big Data kan brukes til mange gode og samfunnsnyttige formål ved å blant annet identifisere og forutse trender og sammenhenger. Minelli et al. (2013, s.1) snakker også godt om Big Data, hvor de blant annet nevner at Big Data vil bidra til både kostnadseffektivitet og topplinjevekst. Zaslavsky et al. (2013, s. 4) underbygger denne påstanden ved å legge vekt på at Big Data potensielt kan ha store gevinster gjennom et forretningsperspektiv, som for eksempel økte inntekter, redusert risiko, samt enklere predikere fremtidige utfall. Ved hjelp av Big Data kan selskaper hente ut data som er mer spesifikke, og dermed av større nytte. Zaslavsky et al. (2013, s. 2) trekker fram tre karakteristikk som alle bidrar til å gi en god beskrivelse av hva Big Data innebærer;

- *Volum* → Volum knyttes til størrelsen, eller måleenheten, av dataene og datalagringskapasiteten.
- *Variasjon* → Variasjon ses i sammenheng med alle ulike typer data som samles.
- *Hastighet* → Dette viser til hvor ofte data blir generert. For eksempel om det gjøres hvert millisekund, sekund, minutt, time, daglig, ukentlig og årlig.

Prosesseringsfrekvensen avhenger av hvilken type data; noen data må prosesseres i sann tid, mens andre prosesseres kun ved nødvendighet.

Ved at data er blitt digitale har også informasjonsflyten og kommunikasjonen i moderne selskaper blitt mer effektiv – digitale data lager et grunnlag for standardisering, behandling, overføring og lagring (Andersen & Sannes, 2018, s. 198). Disse digitale dataene kan samles i store mengder og deretter analyseres. Ved å analysere store mengder data legges det til rette for nye beslutningsmodeller, der selskaper kan ta beslutninger basert på helt eller delvis automatisert analyse av omgivelsene innenfor produksjon, kundegrensesnitt og systemvedlikehold (Andersen & Sannes, 2018, s. 200). Denne type automatisert analyse kan ses i sammenheng med bruk av kunstig intelligente systemer.

2.3.2. Kunstig intelligens

Utviklingen av kunstig intelligens har hatt store fremskritt de siste årene og mulighetene virker lovende: et bedre og mer effektivt næringsliv, nye metoder for klima- og miljøvern, et sikrere samfunn og kanskje også løsningen på kreftgåten (Datatilsynet, 2018, s. 4). Kunstig intelligens er en teknologi som kan få stor betydning for samfunnsutviklingen, og som kan bidra til å gi oss nye verktøy og metoder for å løse samfunnsutfordringer, forbedre offentlige tjenester og bidra til økt verdiskaping i næringslivet (Regjeringen, 2020, s. 8). Vi er med andre ord i startfasen på noe potensielt som vil ha en betydelig effekt på samfunnet (Datatilsynet, 2018, s. 4).

«Mennesker anses å være de mest intelligente skapningene som noen gang har eksistert på jorden» (Goodwin, 2020, s. 11). Våre evner til å løse komplekse problemer er så unike at det kan være vanskelig å forestille seg at andre også kan ha slike evner – hvert fall ikke maskiner og dataprogrammer. Hos mennesker finnes intelligens naturlig, men denne intelligensen kan også gjenskapes i maskiner og dataprogrammer. Det er først når slike maskiner og dataprogrammer begynner å oppføre seg intelligente uten å få instruksjoner fra et menneske, at man kan omtale oppførselen som kunstig intelligens (Goodwin, 2020, s. 11; Tidemann, 2020). Begrepet kunstig intelligens stammer fra det engelske ordet artificial intelligence, og blir forkortet til henholdsvis KI og AI. Vi har i denne oppgaven valgt å benytte det norske begrepet og dets forkortelse KI.

Det finnes mange forslag til definisjoner av kunstig intelligens, og disse definisjonene endrer seg også gjerne i takt med hva som er teknologisk mulig (Regjeringen, 2020, s. 9). Vi har i denne oppgaven valgt å ta utgangspunkt i definisjonen som blir presentert i Regjeringens nasjonale strategi for kunstig intelligens;

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.

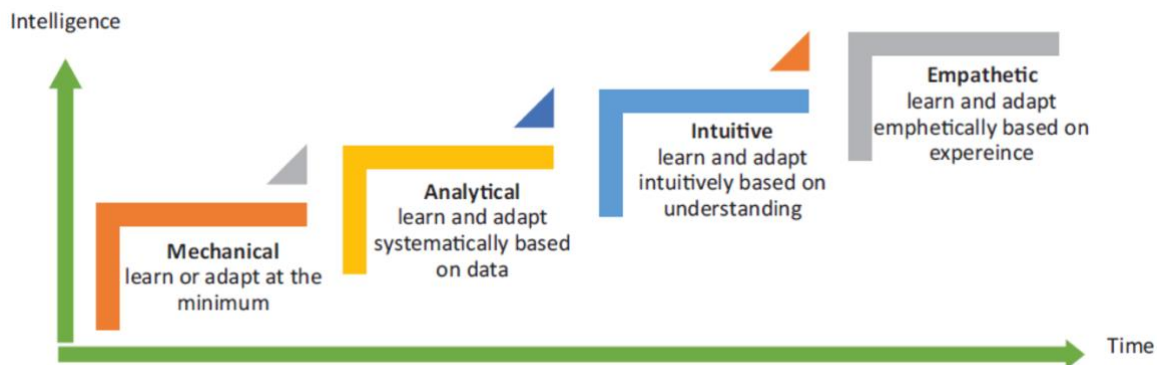
(Regjeringen, 2020, s. 9)

Kunstig intelligens ble allerede på 1950-tallet kjent som et begrep og en teknikk det ble knyttet store forhåpninger til (Datatilsynet, 2018, s. 4). Dartmouth-professor John McCarthy

er kjent som en av grunnleggerne av kunstig intelligens (Sabanovic et al., 2012, s. 99) gjennom et prosjekt hvor man arbeidet med å lage algoritmer som får dataprogrammer og maskiner til å forstå språk, være kreativ og ha intelligens (Goodwin, 2020, s. 55). Dario Gils, visepresident for KI og Quantum Computing i IBM, hevder at KI ikke fungerte på en god måte før 2012. Årsaken til dette var at man da fikk stadig billigere og større tilgang til data, og tilgangen til store mengder data (Big Data) har vært helt essensielt for utviklingen av kunstig intelligens de siste årene (Canals & Heukamp, 2018). Den økte tilgangen på data kan ses i sammenheng med den digitale utviklingen; data som tidligere kun har vært tilgjengelig på papir er blitt digitale slik at de kan behandles i datasystemer, i tillegg til at de kan automatisere enkle rutineoppgaver som tidligere har blitt behandlet manuelt av mennesker (Andersen & Sannes, 2018, s. 207). Videre ønsker vi å gå inn på de ulike typene av kunstig intelligens som har vokst fram som følge av digitaliseringen.

2.3.3. Typer intelligenser

Avhengig av hvilken tjeneste som skal utføres, kreves det ulike typer intelligens. Huang og Rust (2018, s. 156) sier at det handler om å utvikle maskinbasert intelligens (KI) for å tilegne seg menneskelig intelligens (HI) i form av kunnskap, resonnement, problemløsning, læring, formidling, oppfatning og handling. De skiller mellom følgende fire nivåer av kunstig intelligens; mekanisk, analytisk, intuitiv og empatisk;



(Figur 3: *The Four Intelligences*. Huang & Rust, 2018, s. 158).

Mekanisk intelligens

Mekanisk intelligens er det laveste nivået av de fire intelligenstypene, og omhandler evnen til å utføre rutinemessige oppgaver. For mennesker krever ikke mekaniske prosesser mye kreativitet da prosessene har blitt utført så mange ganger, at de til slutt blir utført med liten eller ingen ekstra tanke. For å utføre mekaniske oppgaver er det ikke nødvendig med

menneskelig faglært arbeidskraft, opplæring eller høyere utdanning. Mekanisk KI er designet for å ha begrenset læring – aktivitetene som kan gjennomføres er derfor konsistente og ikke tilpasningsdyktige. Roboter er et eksempel på bruk av mekanisk intelligens. Dette er en type teknologi som kan utføre enkle oppgaver, operere autonomt uten behov for instruksjon, og styres av datamaskiner uten menneskelig hjelp. Robotene forstår ikke miljøet rundt seg, og kan derfor ikke tilpasse seg situasjonen automatisk; i stedet vil robotens kunnskap oppdateres i nødvendige tilfeller (Huang & Rust, 2018, s. 158).

Analytisk intelligens

Analytisk intelligens handler om å behandle informasjon for å løse problemer, og deretter lære. På dette nivået lærer og tilpasser man seg systematisk basert på tilgjengelig data, og teknologien må behandle informasjon, foreta logiske resonneringer og ha matematiske ferdigheter. Disse ferdighetene er tilegnet ved hjelp av trening, ekspertise og spesialisering innen kognitiv tenking. Maskinlæring og dataanalyse er de viktigste KI-applikasjonene på dette nivået. Analytisk intelligens bruker i all hovedsak algoritmer for gjentakende læring, og finner på denne måten innsiktsfull informasjon uten å være programmert for hvor den skal lete etter informasjonen. Analytisk KI er derfor tilsynelatende intelligent, men kan ikke simulere intuisjon. Dens systematiske natur gjør den derimot egnet for massetilpasning basert på Big Data fra kunder (Huang & Rust, 2018, s. 158).

Intuitiv intelligens

Den intuitive intelligensen har evnen til å tenke kreativt, og effektivt kunne tilpasse seg nye situasjoner. Intuitiv intelligens krever faglige ferdigheter som bidrar til en kreativ og innsiktsfull problemløsning, og er dermed designet for å fungere mer fleksibelt i likhet med et menneske. Den viktigste egenskapen som skiller intuitiv KI fra analytisk KI, er forståelse. Intuitiv KI sees på som sterk KI, og er designet for å være fleksibel. Denne intelligensen er bygget for å etterligne et bredt spekter av menneskelig kognisjon og lære på samme måte som et menneskebarn. På bakgrunn av sin datakraft og tilkobling vil den intuitive intelligensen ikke begå samme feilen flere ganger på bakgrunn av dens erfaring (Huang & Rust, 2018, s. 159).

Empatisk intelligens

Empatisk intelligens omhandler evnen til å gjenkjenne og forstå følelser, reagere følelsesmessig passende, samtidig som den skal kunne påvirke andres følelser. På dette

nivået av intelligens inkluderes sosiale ferdigheter som hjelper mennesker å være følsom og samarbeide godt med andre. Empatisk KI kan være en maskin som kan føle, eller i det minste oppføre seg som at den har følelser (Huang & Rust, 2018, s. 159).

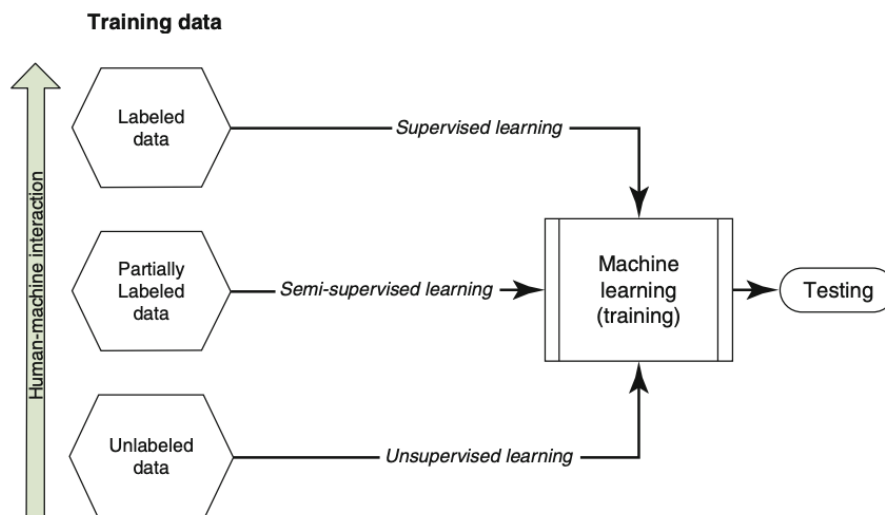
2.3.4. Felt innen kunstig intelligens – maskinlæring og dyplæring

Det finnes flere felt innenfor kunstig intelligens ettersom det omfatter alle intelligente systemer, men denne oppgaven legger vekt på maskinlæring og dyplæring i det teoretiske grunnlaget. Feltene er komplekse og kan inngå i hverandre.

Maskinlæring

Maskinlæring er en gren innenfor kunstig intelligens, og innebærer å få maskinen til å optimalisere seg selv (Mendling et al., 2018, s. 4; Bergsjø & Bergsjø, 2019, s. 55). Det finnes flere definisjoner på maskinlæring. Vi har valgt å ta i bruk Arthur Samuel sin definisjon «maskinlæring er en studieretning som gir datamaskiner muligheten til å lære uten å være eksplisitt programmert», og Ethem Alpaydin sin definisjon som sier at; «maskinlæring er programmering av datamaskiner for å optimalisere en gitt ytelse ved bruk av testdata eller tidligere erfaring» (Hentet og oversatt fra; Naqa & Murphy, 2015, s. 6). Det disse to definisjonene har til felles er at maskiner veiledes til å ta intelligente beslutninger ved å lære av gjentakende tilfeller (Naqa & Murphy, 2015, s. 6).

Ved bruk av maskinlæring prøver man å lære noe om dataenes underliggende struktur og regelmessigheter, og på bakgrunn av dette utvikle modeller, ofte kalt prediksjonsmodeller. Maskinene prøver å finne mønstre i data for å bedre kunne predikere fremtidige tilfeller (Iriondo, 2021; Theodoridis, 2015, s. 1). Den økte tilgjengeligheten av data, har også økt potensialet for bruken av maskinlæring. Det kan brukes både store og små mengder med data for å trene maskiner, men større mengder data vil i de fleste tilfeller kunne gi mer nøyaktige resultater. Det er i tillegg viktig å huske på at resultatet avhenger av kvaliteten på dataene og hvilke maskinlæringsalgoritmer som blir brukt (Mendling et al., 2018, s. 4; Iriondo, 2021). Maskinlæringsalgoritmer kan lære på forskjellige måter. Vi tar videre for oss de tre kategoriene av læring; veiledet læring, ikke-veiledet læring og semi-strukturert læring.



Figur 4: Categories of machine learning algorithms according to training data nature (Naqa & Murphy, 2015, s. 7).

Veiledet læring – Supervised Learning

«Ved veiledet læring blir algoritmen trent med et datasett der både inndata og resultatet er gitt» (Regjeringen, 2020, s. 11). Basert på dette vet maskinen på forhånd hvilken oppgave den står ovenfor, og hva som er den riktige løsningen på oppgaven. Maskinen bruker denne informasjonen videre til å bygge modellen slik at den senere vil kunne ta en beslutning kun basert på inndata (Regjeringen, 2020, s. 11; Bergsjø & Bergsjø, 2019, s. 55). For at denne metoden for læring skal kunne fungere optimalt, er det nødvendig med et stort datasett å trene på (Bergsjø & Bergsjø, 2019, s. 55). Ved veiledet læring blir det benyttet kategoriserte data, hvor hver data har sin egen merkelapp som kategoriserer den (Datatilsynet, 2018, s. 7).

Ikke-veiledet læring – Unsupervised Learning

Ved ikke-veiledet læring får algoritmen et datasett hvor det ikke eksisterer en fasit. Algoritmen må selv finne mønstre i det gitte datasettet, og ta beslutninger basert på dette (Regjeringen, 2020, s. 11). Forskjellen på veiledet og ikke-veiledet læring er at veiledet læring både får input og output data, mens ikke-veiledet læring får kun input data (Naqa & Murphy, 2015, s. 7). Data som blir benyttet er ikke forhåndskategoriserte, og ved denne typen av læring ønsker man at maskinen selv skal finne strukturene i dataene og klare å gruppere sammen data som ligner på hverandre. En utfordring ved denne type læring er at maskinen ikke vil kunne plassere data i andre grupper enn det som ble oppdaget i læringsprosessen. For at denne typen av læring skal fungere optimalt, er det derfor viktig at grunnlaget av treningsdataene er representativt (Datatilsynet, 2018, s. 8).

Semi-strukturert læring – Semi-Supervised Learning

Semi-strukturert læring, også kalt forsterket læring, er en form for læring som bygger på en kombinasjon av veiledet læring og ikke-veiledet læring. Her er noen av dataene til dels kategorisert. Disse til dels kategoriserte dataene brukes til å prøve å forstå den ikke-kategoriserte delen (Naqa & Murphy, 2015, s. 7). Denne formen for læring bygger derfor algoritmen sin basert på ikke-veiledet læring. Videre får algoritmen tilbakemeldinger fra bruker og/eller operatør om beslutningen som blir tatt kan kategoriseres som god eller dårlig. Denne tilbakemeldingen tas så med tilbake i modellen, og bidrar til forbedring (Regjeringen, 2020, s. 11). Ved semi-strukturert læring gir maskinen flere poeng til algoritmene som fungerer bra, og færre til de som ikke fungerer like bra. På bakgrunn av dette optimaliseres modellen ved å lære hvilke handlinger som leder mot målet (Datatilsynet, 2018, s. 9).

Dyplæring - Deep Learning

Dyplæring er et underfelt av maskinlæring, og er i dagens samfunn en viktig del av forskjellige teknologiske løsninger som blant annet talegjenkjenning, anbefalingssystemer og avviksdeteksjon (Regjeringen, 2020, s. 12). Sammenlignet med maskinlæring, er dyplæring mer presis, krever mer data og takler ustrukturerte data på en bedre måte (Yadav, 2018). Dyplæring er en form for læring som baserer seg på kunstige nevralt nettverk, og er i stand til å fange opp komplekse forhold basert på store datasett - både kontekstuell, tekstuell og visuell informasjon (Marr, 2018; Zhang et al., 2019, s. 2). På bakgrunn av dette kan man si at essensen med dyplæring er at datamaskinen skal lære seg noe nytt, noe den ikke vet noe om eller kan fra før (Marr, 2018). Dyplæringsalgoritmer trenes opp basert på ikke-veiledet læring; algoritmen må selv finne mønstre, og ta beslutninger basert på dette (Regjeringen, 2020, s. 11). En utfordring med enkelte dyplæringsalgoritmer er at de kan ses i sammenheng med Black Box-problematikken, hvor man ikke har innsyn i modellen, og kan dermed heller ikke forklare hvorfor en gitt dataverdi gir et gitt resultat (Regjeringen, 2020, s. 58).

2.3.5. Black Box-problematikken

Som nevnt er en av utfordringene med kunstig intelligens å forklare hvordan modellen har kommet fram til en spesifikk beslutning. Ofte vil modellen produsere et resultat som ikke kan forklares, og fordi man ikke vet hva som foregår inne i boksen/modellen blir dette omtalt som Black Box-problematikken, også kalt den sorte boksen (Datatilsynet, 2018, s. 12). Denne

problematikken blir særs relevant dersom en forsikringskunde blir merket med rødt flagg i henhold til risiko for svindel i prediksjonsmodellen og kunden ønsker en forklaring på dette. Forsikringsselskapet må i denne sammenheng vært best mulig skikket til å kunne gi en god forklaring på gitt resultat, som selskapet er pliktig til i henhold til kravet om åpenhet ovenfor kunden.

2.3.6. Bias

Bias, ofte kalt skjevhet eller metodefeil, kan i forskning føre til at resultater ikke samsvarer med virkeligheten (Staff, 2015). Begrepet refererer til hvilken som helst form for preferanse – rettfærdig eller urettferdig – det vil si, systematisk diskriminering av enkeltpersoner eller grupper basert på visse kjennetegn (Silberg & Manyika, 2019, s. 2). Den økende bruken av kunstig intelligens på sensitive områder – som blant annet ved behandling av personopplysninger, har vekket debatten om bias og rettfærdighet. Men menneskelig beslutningstaking kan også formes av individuelle eller samfunnsmessige skjevheter som ofte kan oppstå ubevisst. Disse menneskelige skjevhetene er det ikke mulig å komme utenom. Silberg og Manyika (2019, s. 1) stiller spørsmål om hvorvidt beslutninger som er basert på KI, er mindre partiske enn menneskelige beslutninger.

Bruken av kunstig intelligens kan i mange tilfeller identifisere og redusere effekten av menneskelig bias, da maskinlæringsalgoritmer kun lærer å vurdere variablene som kan forbedre deres prediktive nøyaktighet, basert på treningsdataene som blir brukt (Silberg & Manyika, 2019, s. 2). Man kan imidlertid stille spørsmål om hvorvidt algoritmene KI benytter seg av vil lære og dermed også bli påvirket av menneskelige og samfunnsmessige skjevheter (Silberg & Manyika, 2019, s. 3). Denne formen for skjevhet kalles algoritmisk bias, og går ut på at underliggende data er kilden til problemet. For å minimere sjansene for dette er det viktig at selskaper har fokus på å forbedre sine KI-systemer. Alt fra hvordan de utnytter data, hvordan data utvikles, distribueres og brukes, for å forhindre at KI kan adoptere de menneskelige skjevhetene, eller i noen tilfeller også danne sine egne skjevheter. For at dette skal kunne være mulig, er det nødvendig å ha tilgang på menneskelig dømmekraft og prosesser (Silberg & Manyika, 2019, s. 3).

2.4. Personvern

Som følge av digitaliseringen og økt bruk av digitale tjenester, har mengden persondata, også kalt personopplysninger, som genereres økt (Wasenden, 2020, s. 64). Samtidig som det genereres større mengder persondata til ulike formål, er personvern en av våre grunnleggende rettigheter som innbyggere i et demokratisk samfunn. Likevel regnes det som krevende å ha rådighet over egne personopplysninger som følge av digitaliseringen og det som i dag regnes som et informasjonssamfunn (Regjeringen, 2014). Det er tydelig at selskaper legger større vekt på å personalisere tilbud og tjenester, det blir derfor viktig å finne en balanse mellom å ivareta kunders personvern, samtidig som de kan drive med personalisert tilpasning (Wasenden, 2020, s. 64). Som følge av at det behandles større mengder personopplysninger både i offentlig og privat sektor, er det forankret i lov hvordan man er forpliktet til å behandle denne type data.

«Personvern handler om retten til et privatliv og retten til å bestemme over egne personopplysninger» (Datatilsynet, 2019). Alle som behandler personopplysninger, er dermed pliktet til å opptre i samsvar med de lovfestede personvernspriksippene (Datatilsynet, 2019). Personopplysninger er alle opplysninger som kan knyttes til en enkeltperson (Datatilsynet, 2018, s. 4). Vi ønsker dermed å presentere personvernspriksippene som er lagt fram i Personopplysningsloven Kapittel II, Artikkel 5, av Norsk Lov, og forskjellige personvernutfordringer i forbindelse med behandling av data;

Lovlig, rettferdig og gjennomiktig

«Personopplysninger skal behandles på en lovlig, rettferdig og åpen måte med hensyn til den registrerte» (Lovdata, 2018). Dette innebærer at enhver som behandler personopplysninger er pliktig til å behandle dataene lovlig med hensyn til at det skal eksistere et rettslig forankret grunnlag for at man skal behandle persondata, rettferdig med hensyn til personens interesser og forventninger, og til slutt gjennomiktig (åpen) i den forstand at persondataene skal være oversiktlige og forutsigbare med hensyn til personen, noe som skaper tillit mellom databehandler og enkeltperson (Datatilsynet, 2019). Mennesker kan oppleve å ikke vite hvordan data om den enkelte blir benyttet og sammenstilt, noe som kan føre til at man ikke forstår beslutninger som blir tatt. Dette kan føre til vegring for å bevege seg på internett, da man kan oppleve usikkerhet rundt hvilke aktører som innhenter informasjon om deg, og faktisk hva denne informasjonen blir brukt til (Bergsjø & Bergsjø, 2019, s. 83). Dette kan

være en utfordring for dette prinsippet, fordi behandling av data skal være oversiktlig og forutsigbar for den det gjelder (Datatilsynet, 2019).

Formålsbegrensning

«Personopplysninger skal samles inn for spesifikke, uttrykkelig angitte og berettigede formål og ikke videre behandles på en måte som er uforenlig med disse formålene» (Lovdata, 2018). Formålsbegrensning innebærer at persondataene som behandles skal spesifiseres og beskrives i den grad at det ikke foreligger noe tvil om hva dataene skal benyttes til, samt formålet med behandlingen. Denne begrensningen legger også vekt på at persondataene som benyttes til et konkret formål, ikke skal gjenbrukes til et annet formål som ikke er forenlig med det formålet som det opprinnelig ble benyttet til (Datatilsynet, 2019). På bakgrunn av at stordata i stor grad omhandler gjenbruk av data, kan dette prinsippet føre til utfordringer knyttet til kommersiell stordataanalyse (Bergsjø & Bergsjø, 2019, s. 82).

Dataminimering

«Personopplysninger skal være adekvate, relevante og begrenset til det som er nødvendig for formålene de behandles for» (Lovdata, 2018). Prinsippet om dataminimering handler om at databehandler skal begrense mengden persondata som samles inn i forbindelse med et spesifikt formål. Data som ikke bidrar til formålet regnes ikke som nødvendig, og skal derfor heller ikke samles inn (Datatilsynet, 2019). Ofte kan verdien av dataene ligge i deres fremtidige bruksmuligheter, noe som igjen kan gjøre at selskaper ikke har ønske om å slette data på grunn av potensiell fremtidig verdi, ny innsikt og potensielt mer fortjeneste (Bergsjø & Bergsjø, 2019, s. 83). Dette strider mot prinsippet om dataminimering, og kalles ofte datamaksimalisering.

Lagringsbegrensning

«Personopplysninger skal lagres slik at det ikke er mulig å identifisere de registrerte i lengre perioder enn det som er nødvendig for formålene som personopplysningene behandles for» (Lovdata, 2018). Begrenset lagring innebærer at data som er innhentet til et spesifikt formål enten skal slettes eller anonymiseres ved formålets utløp, eller på et tidspunkt hvor de anses som unødvendige for formålet (Datatilsynet, 2019).

Riktighet

«Personopplysninger skal være korrekte og nødvendig oppdaterte» (Lovdata, 2018). Dette innebærer at behandler plikter til å enten slette, eller endre, data som viser seg å være uriktige med hensyn til formålet de brukes til (Datatilsynet, 2019). Dette kan også ses i sammenheng med at beslutningstaking som har konsekvenser for den enkelte skal være basert på fakta, og dette er en utfordring med hensyn til stordataanalyse ettersom analysen ofte ikke tar hensyn til kontekst (Bergsjø & Bergsjø, 2010, s. 83).

Integritet og konfidensialitet

«Personopplysninger skal behandles på en måte som sikrer tilstrekkelig sikkerhet for personopplysningene, herunder vern mot uautorisert eller ulovlig behandling og mot utilsiktet tap, ødeleggelse eller skade, ved bruk av egnede tekniske eller organisatoriske tiltak» (Lovdata, 2018). Prinsippet innebærer at databehandler er pliktet til å sørge for at tap, ødeleggelse eller skade av innsamlet data ikke finner sted, og i den forbindelse fokusere på beskyttelse og konfidensialitet (Datatilsynet, 2019).

Ansvarlighet

«Den behandlingsansvarlige er ansvarlig for og skal kunne påvise at disse prinsippene overholdes» (Lovdata, 2018). Det siste prinsippet vektlegger viktigheten av å overholde prinsippene som er presentert over. Databehandler er pliktet til å vise ansvar i form av kunne dokumentere etterlevelse av prinsippene, samtidig som databehandler også har ansvar for å sikre at regelverket til enhver til følges gjennom ulike tiltak, både organisatorisk og teknisk (Datatilsynet, 2019).

Vi ser at det finnes utfordringer knyttet til ivaretagelse av personvernet som kan kobles direkte opp mot personvernprinsippene. Likevel finnes det også andre utfordringer som man må knytte til personvernprinsippene i sin helhet. Vi finner blant annet utfordringer i sammenheng med risiko for reidentifisering av enkeltpersoner og datadeterminisme som relevante med hensyn til stordataanalyse (Bergsjø & Bergsjø, 2019, s. 83-84). Risiko for reidentifisering innebærer at man ved stordatanalyse potensielt kan kunne identifisere enkeltpersoner ved å sammenstille datasett. Her kan man stille seg spørsmål om hvorvidt anonymisering av data i forbindelse med lagringsbegrensning har noe hensikt (Bergsjø & Bergsjø, 2019, s. 83). Datadeterminisme innebærer at stadig flere beslutninger er automatisert og basert på algoritmer og prediksjonsanalyse. Denne type beslutningstaking kan bidra til å

forsterke eksisterende fordommer, sosial ekskludering og lagdeling (Bergsjø & Bergsjø, 2019, s. 84), som igjen kan føre til diskriminering av enkeltpersoner.

2.4.1. GDPR

Som følge av utvikling av ny teknologi og økt mengde data i omløp, har fokus på håndtering av persondata blitt stadig viktigere. Allerede i 2010 ble det avdekket at det eksisterende lovverket knyttet til personvern ikke var godt nok tilpasset til vår nye digitale hverdag. Basert på dette ble det i EU utviklet en ny personvernforordning som trådte i kraft den 25.mai 2018 i EU, og i Norge fra den 1. juli 2018. Den nye personvernforordningen fikk navnet GDPR; *General Data Protection Regulation*, og hadde som formål å gjøre det mulig for enkeltpersoner å ha kontroll over hvilke data offentlige og private selskaper har registrert om den enkelte (Sørebø et al., 2020, s. 241). GDPR-forordningen legger til rette for at enkeltpersoner må gi samtykke til at selskaper kan innhente og behandle personopplysninger, slik at enkeltpersonen har råderett til å verne egne opplysninger (Sørebø et al., 2020, s. 242). De største konsekvensene som følge av forordningen fra 2018 er regler med hensyn til hvordan personopplysninger skal behandles i en digital hverdag, samt at konsekvenser for overtredelse er blitt ilagt større vekt i form av bøter. Endringen med hensyn til større fokus på overtredelse av loven har bidratt til at både offentlige og private selskaper som behandler data må ta direktivet mer på alvor og gi det betydelig mer innsats (Kingston, 2017, s. 429).

Ettersom automatiserte stordatanalyser baserer seg på algoritmer, er det viktig å innføre tiltak som verner enkeltindividets rettigheter, friheter og interesser, i tråd med GDPR. Med hensyn til automatisert beslutningstaking kan individet kreve at det er et menneske som tar den endelige beslutningen, og det er også mulig å klage på beslutningen som tas (Datatilsynet, 2019, s. 19). Dette kan være utfordrende for selskaper dersom beslutningssystemet i aller største grad baserer seg på stordataanalyse og algoritmer; høyere grad av autonomi i beslutningsprosessen fører til at spørsmål rundt hvem som er ansvarlig for beslutningen blir mer kompleks (Butterworth, 2018, s. 258), noe som igjen påvirker evnen til å forklare hvorfor en beslutning ble tatt.

2.5. Teoretisk relevans

Studien tar for seg hvordan forsikringsbransjen i Norge kan skape en felles stordataplattform for å bekjempe forsikringssvindel, og vi anså det derfor som relevant å starte med en

introduksjon til forsikring og begrepet svindel. Med hensyn til at vi studerer dette konseptet i lys av et spesifikt innovasjonsprosjekt, vurderte vi det derfor relevant å fremheve begrepet innovasjon, og ulike dimensjoner innenfor innovasjon. Innovasjonsprosjektet består av et samarbeid mellom konkurrerende selskaper i samme bransje, og vi anså det derfor hensiktsmessig å inkludere teori knyttet til samkonkurranse og næringsklynger.

Innovasjonsprosjektet bygger på en idé om at det skal være mulig å dele data på tvers av konkurrerende selskaper, og ta i bruk kunstig intelligens for å bedre kunne predikere og detektere svindel. Siden det er snakk om KI og Big Data anså vi det derfor nødvendig å presentere teori i sammenheng med disse begrepene. Vi antar at forsikringsselskapene i et slik prosjekt deler data som direkte kan knyttes til deres respektive kunder, og avslutningsvis har vi derfor valgt å inkludere teori knyttet til personvern og GDPR.

3. METODE

Vi ønsker å bruke metodekapittelet i denne oppgaven til å presentere og begrunne de metodiske valgene som er tatt for å best mulig kunne besvare oppgavens forskningsspørsmål. Vi legger fram valg av forskningsdesign, forskningstilnærming og forskningsmetode, samt beskrivelse av studien vi har gjennomført. Deretter vil vi gå nærmere inn på hvordan vi har samlet inn data og hvordan disse er bearbeidet. Avslutningsvis presenterer vi kritiske vurderinger med hensyn til studiens kvalitet.

3.1. Forskningsdesign og tilnærming

Hensikten med denne studien har vært å se på hvordan forsikringsselskaper tar i bruk kunstig intelligens og Big Data for å bekjempe forsikringssvindel, samt hvilke utfordringer som følger ved arbeidet med dette innovasjonsprosjektet. På bakgrunn av at dette er et innovasjonsprosjekt, som per dags dato befinner seg i en MVP-fase og er lite forsket på fra tidligere, har vi valgt å ta i bruk et eksplorativt design. Eksplorativt design innebærer at forskeren vet lite om forskingsområdet fra før, og ønsker å utforske dette temaet nærmere (Gripsrud et al., 2016, s. 47), noe som sammenfaller med våre forskningsspørsmål.

Det skilles mellom to typer forskningstilnærminger; induktiv tilnærming og deduktiv tilnærming. I den deduktive tilnærmingen handler det om at den beste framgangsmåten er

å først lage seg noen forventninger om hvordan virkeligheten ser ut, og videre samle inn data for å se om forventningene stemmer overens med virkeligheten. Denne tilnærmingen blir sett på som «lukket», da det er forskeren som definerer hva som er interessant å studere videre (Jacobsen, 2015, s. 25-26). Den induktive tilnærmingen tar for seg at det er lite hensiktsmessig å lage teorier om hvordan virkeligheten ser ut, og hvordan den henger sammen. Denne tilnærmingen tar derfor heller for seg viktigheten av å forstå det dynamiske og det unike. Forskeren samler her inn data med et helt åpent sinn, og målet med tilnærmingen er at den ikke skal begrense hvilken informasjon forskeren samler inn (Jacobsen, 2015, s. 28-29).

Vi har sett det hensiktsmessig å kombinere disse to tilnærmingene i vår studie da vi ønsket å ha en åpen tilnærming som legger til rette for å hente inn ny informasjon, og ikke kun informasjon som var basert på tidligere antakelser. Kombinasjonen av de to tilnærmingene blir kalt abduksjon, og all vitenskapelig tenking starter med observasjon. Man oppdager noe man blir overrasket over, som gjør at man stiller spørsmål. Dette spørsmålet ønsker man gjerne å finne svar på, som ender i spekulasjon. Disse spekulasjonene leder videre til et sett med antakelser eller hypoteser. Videre vil det være naturlig å teste om antakelsene stemmer, som gjør at spørsmålet må settes opp mot teori (Jacobsen, 2015, s. 34-35).

3.2. Forskningsmetode

For å best kunne være i stand til å svare på oppgavens problemstilling har vi benyttet en kvalitativ tilnærming for å få en bedre forståelse om temaet. Kvalitative metoder innebærer å samle inn data i en ikke-numerisk form, altså tekst, til forskjell fra kvantitativ metode (Yin, 2018, s. 55). Kvalitative metoder kjennetegnes også gjennom den interaktive og fortolkende prosessen som dataene utarbeides fra, og har en tendens til å være av eksplorativ natur, samtidig som metoden fokuserer på åpne spørsmål (Easterby-Smith et al., 2018, s. 173). Vi har valgt å gjennomføre dybdeintervjuer for å bygge opp under den abduktive og eksplorative tilnærmingen vi har valgt. Kvalitative intervjuer defineres som; «direkte samtaler rundt spesifikke spørsmål eller et spesifikt tema» (Egen oversettelse fra Easterby-Smith et al., 2018, s. 179). Disse intervjuene vil hjelpe oss til bedre forståelse, refleksjon og forklaringen rundt oppgavens tema og forskningsspørsmål. På bakgrunn av at noen av dybdeintervjuene vi gjennomførte var med flere enn et intervjuobjekt, anser vi disse som et gruppeintervju. Et gruppeintervju er når gruppen som blir intervjuet er naturlig satt sammen, i en spesifikk setting (Easterby-Smith et al., 2018, s. 181). I våre tilfeller bestod gruppeintervjuene våre av

respondenter fra samme selskap, men med forskjellige bakgrunner. Ett av gruppeintervjuene foregikk på engelsk, og funnene fra dette intervjuet er dermed gjengitt på engelsk.

3.3. Casestudieforskning

Forskningsspørsmålene i denne oppgaven er forklarende forskningsspørsmål, noe som gjør casestudieforskning relevant. Når casestudien er forklarende stilles det spørsmål om «hvordan» og «hvorfor», for eksempel; «Hvordan og/eller hvorfor en organisasjon fungerer slik den gjør, og hvordan og/eller hvorfor ulike prosesser/metoder har ledet til visse resultater» (Yin, 2018, s. 251). Definisjonen til Yin sammenfaller med oppgavens forskningsspørsmål som handler om hvordan forsikringsselskaper kan ta i bruk kunstig intelligens og Big Data for å bekjempe forsikringsvindel, samt videre undersøkelse av de tekniske, forretningsmessige og juridiske utfordringene som konsekvens. Casestudieforskning egner seg når studiets forskningsspørsmål undersøker nåværende omstendigheter, og når vi som forskere har begrensede forkunnskaper og innsikt om fenomenet som skal studeres. Det legges også vekt på at det som studeres er et moderne fenomen, og ikke historisk (Yin, 2018, s. 43). Innovasjonsprosjektet som studeres er et pågående prosjekt, og resultatene fra arbeidet er langt i fra konkludert og vi anser derfor prosjektet som nåværende og moderne.

Vi har gjort vurderinger knyttet til hvorvidt dette kan regnes som en single case eller et multiple case. Vurderingene er gjort på bakgrunn av at innovasjonsprosjektet som studeres består av flere individuelle selskaper, men ettersom multiple case i all hovedsak omhandler sammenligning av flere caser (Yin, 2018, s. 102) vil det være mer hensiktsmessig for studiens formål å undersøke innovasjonsprosjektet som en helhet i en single case. Selskaper som inngår i innovasjonsprosjektet er NCE Finance Innovation, Tryg Forsikring, Frende Forsikring, Fremtind Forsikring, Webstep og PwC. Innovasjonsprosjektet ledes av NCE Finance Innovation, hvor de tre forsikringsselskapene er ansvarlig for å dele data; Webstep som utvikler av teknologisk løsning; og PwC som juridisk rådgiver.

3.4. Valg av informanter

Ved valg av informanter til vår studie har vi fokusert på informanter som tett kan knyttes til innovasjonsprosjektet ettersom disse informantene bidrar til et bredt og detaljert datagrunnlag. På bakgrunn av at vi har kontakter i bransjen, valgte vi å starte med et ustrukturert intervju med to informanter fra Frende Forsikring som videre ga oss tips til hvilke andre informanter

og andre selskaper det var relevant å intervju, og vi fikk kontaktinformasjon til informant fra NCE Finance Innovation. Etter møtet med NCE Finance Innovation ble vi satt i kontakt med informanter fra Webstep, Tryg Forsikring, Fremtind Forsikring og PwC. Denne formen for innhenting av informanter kalles for snøballmetoden. Snøballmetoden starter med noen som oppfyller kriteriene i en studie, og som deretter nevner andre som kan være kvalifiserte å ta kontakt med (Easterby-Smith et al., 2018, s. 109). På eget initiativ tok vi kontakt med en skadesjef og en etterforsker i Tryg Forsikring. Bakgrunnen for dette var å innhente informasjon om hvordan skadebehandlingsprosessen og sviketterforskningen foregår i dag.

Utvalget av informanter kan også kategoriseres som kriteriebasert, hvor informantene er valgt ut ifra kriterier om de har kunnskap og kjennskap til det spesifikke innovasjonsprosjektet. Ved å ha disse kriteriene sikrer vi at informantene skal kunne gi detaljert informasjon på de områdene vi undersøker.

Selskap	Stillingstittel
Frende Forsikring	<ul style="list-style-type: none"> - Head of Business Intelligence Center (BICC)/Business Analyst - Finansdirektør
Tryg Forsikring	<ul style="list-style-type: none"> - Data scientist - Skadesjef - Etterforsker
Fremtind Forsikring	<ul style="list-style-type: none"> - Head of Machine Learning and AI
NCE Finance Innovation	<ul style="list-style-type: none"> - Prosjektleder
Webstep	<ul style="list-style-type: none"> - Avdelingsdirektør - Rådgiver
PwC	<ul style="list-style-type: none"> - Advokat m/ekspertise på personvern

Tabell 2: Intervjuobjekter.

3.5. Datagrunnlag

Det er datainnsamlingen som utgjør selve fundamentet for å kunne svare på oppgavens forskningsspørsmål. Det vil derfor være helt essensielt å ta stilling til hvordan man vil innhente data, hvordan man skal analysere data og hvordan man vil bruke dataen, før man starter med selve innsamlingen (Yin, 2014, s. 69).

I denne oppgaven har vi derfor tatt i bruk kvalitativ primær datainnsamling, supplert med kvalitativ sekundærdata og annen litteratur. På bakgrunn av lite forskning rundt temaet, vil primærdata gjennom semi-strukturerte intervjuer av informanter, ha en sentral rolle i denne oppgaven. Primærdata har blitt supplert og sett i sammenheng med litteratur, samt noe sekundærdata gitt av noen deltakende selskaper i prosjektet. Innsamling av litteratur har i all hovedsak foregått gjennom følgende søkemotorer; Google Scholar og Oria, hvor vi har benyttet oss av tidligere publiserte artikler og andre relaterte fagtekster. Det er også benyttet fagbøker hvor vi fant det hensiktsmessig.

3.5.1. Kvalitative sekundærdata og litteratur

Sekundærdata er skriftlige kilder til informasjon som produsert til et annet formål enn forskning, men som har en viss relevans for et gitt forskningsprosjekt (Easterby-Smith et al., 2018, s. 173-174). Sekundærdata kan inkludere rapporter fra selskaper og myndigheter, websider, arkivdata, avisartikler, bøker, m.m. Fordelen med å ta i bruk sekundærdata er at det er tid- og ressursbesparende sammenlignet med innsamling av primærdata. Selv om det er enklere å samle inn sekundærdata, er det viktig å ha et kritisk blikk på dataene som velges ut. Sekundærdataene kan ikke nødvendigvis la seg direkte overføre til vår studie, i tillegg til at den kan være biased, eller preget av personen som har innhentet eller produsert dataen tidligere (Easterby-Smith et al., 2018, s. 174). Sekundærdata som er brukt i denne studien er rapporter fra noen av de deltakende selskapene som primært inneholder en prosjektbeskrivelse, samt status for prosjektet på nåværende tidspunkt med hensyn til muligheter og utfordringer. Innsamlingen av litteratur har i all hovedsak bestått av rapporter, tidligere forskning, artikler og fagbøker. Ved å hente inn og ta i bruk data fra ulike kilder gjør at vi får et diversifisert grunnlag for vår studie, i tillegg til at vi gjerne vil få med både generell og mer spesifikk data som er relevant.

3.5.2. Kvalitative primærdata

I kontrast til forskning som tar sikte på innsamling av sekundære data, retter primærdata seg mot opprettelsen av egen data (Easterby-Smith et al., 2018, s. 175). Med andre ord vil dette si at vi som forskere selv må utvikle disse dataene. En slik prosess kan være tidkrevende, og man må vie mye tid på forberedelse, gjennomføring og transkribering av intervjuene (Easterby-Smith et al., 2018, s. 174). Innsamling av primærdata kan foregå ved forskjellige metoder; dagbokmetoden, skriftlig korrespondanse, eller ved kvalitative intervjuer, enten

individuelle intervju, gruppeintervju eller fokusgrupper. I vår studie har vi ansett det som mest hensiktsmessig å ta i bruk individuelle intervjuer av informanter fra selskapene som er delaktig i innovasjonsprosjektet som studeres, samt ressurser innenfor forsikringsbransjen som til daglig jobber med skadebehandling, svindeldeteksjon og etterforskning av svindel.

Intervjuene er hovedsakelig semi-strukturert; som er en struktur som er åpen, men samtidig styrt med utgangspunkt i en intervjuguide med tema og spørsmål som er ønsket å avdekke i løpet av intervjuene. Strukturen legger likevel til rette for at man skal kunne være både fleksibel og dynamisk i løpet av intervjuet (Easterby-Smith et al., 2018, s. 184).

I utgangspunktet ønsket vi å gjennomføre intervjuene med ansikt-til-ansikt metoden. Denne metoden bidrar til å bygge tillit, i tillegg til at man kan observere informantene i deres omgivelser, og også deres kroppsspråk som kan gi oss nyttig tilleggsinformasjon. Grunnet Covid-19 pandemien ble det utfordrende å gjennomføre intervjuene ansikt-til-ansikt, og vi måtte derfor se på alternative løsninger for å kunne gjennomføre intervjuene. Intervjuene ble derfor gjennomført digitalt over ulike sosiale kommunikasjonsplattformer, blant annet Zoom, Teams og Google Meet. Å gjennomføre intervjuer digitalt er en fleksibel metode som gjør at informantene kan bidra på ulike tidspunkt, i tillegg til at lokasjon for gjennomføring er mer fleksibel (Easterby-Smith et al., 2018, s. 180).

Før intervjuene har vi lagt vekt på å tilegne oss så mye kunnskap som mulig om selskapet vi skal intervjuer, slik at vi skulle stille forberedt. I tillegg til dette hadde vi utarbeidet en intervjuguide med noen spesifikke spørsmål vi ønsket å avdekke. Heriblant de ulike selskapenes og informantenes rolle og bidrag i innovasjonsprosjektet, samt deres perspektiver på prosjektet spesielt med hensyn til muligheter og utfordringer som følge av arbeidet med KI og Big Data i samarbeid med konkurrerende selskaper. Vi utarbeidet forskjellige intervjuguider til hvert intervju basert på bransje ettersom informantene kommer fra ulike bakgrunn og har ulike roller i prosjektet. Intervjuguidene ble basert på beskrivende, fortolkende og teoretiske spørsmål. Beskrivende spørsmål blir knyttet til konkrete hendelser eller handlinger, fortolkende spørsmål blir rettet mot informantens oppfattelse av hendelser, mens teoretiske spørsmål ble tatt i bruk for å avdekke årsaker til ulike handlinger (Easterby-Smith et al., 2018, s. 186). Ved å ta i bruk semistrukturerte intervjuer legger man til rette for flyt i samtalen utover de forhåndsdefinerte spørsmålene som er forberedt, og metoden gir også mulighet til å innhente informasjon som kanskje ellers ikke ville kommet frem.

Dato	Intervju og tittel
18.02.2021	Ustrukturert gruppeintervju – Frende Forsikring m/ Finansdirektør og Head of BICC
25.02.2021	Semistrukturert intervju – NCE Finance Innovation m/ Prosjektleder
25.02.2021	Semistrukturert intervju – Tryg Forsikring m/ Data Scientist
12.03.2021	Semistrukturert gruppeintervju – Webstep m/ Avdelingsdirektør og Rådgiver
17.03.2021	Semistrukturert intervju – PwC m/ Advokat
07.04.2021	Semistrukturert intervju – Fremtind Forsikring m/ Head of Machine Intelligence and AI
23.04.2021	Semistrukturert gruppeintervju – Tryg Forsikring m/ Skadesjef og Etterforsker

Tabell 3: Fremstilling av kvalitative primærdata

3.6. Analyse av data

Yin (2014, s. 164) anbefaler at analysen skal bestå av å undersøke, kategorisere, organisere og deretter kombinere data for å produsere empiriske funn. Dette kan være krevende da det ikke finnes en fremgangsmåte eller klar prosedyre for hvordan analysen skal gjennomføres.

Vi har i forkant av prosessen med å samle inn data spurt om samtykke fra informanter om å ta lydopptak av intervjuene, se vedlegg 9.1. Årsaken til at vi valgte å benytte lydopptak var at vi hadde et behov for tilgang på all informasjon som kommer frem i løpet av intervjuene når analysen skulle gjennomføres. For å få en mer systematisk orden på innsamlet data, har vi også foretatt transkribering av lydopptakene fra intervjuene. På bakgrunn av at gjennomsnittlig lengde på de gjennomførte intervjuene ligger på ca. én time, har dette resultert i store mengder transkriberte intervju. Det har dermed gått mye tid til å kategorisere, organisere og kombinere data.

Vi startet med å kategorisere funnene i henhold til teorikapittelet ved hjelp av fargekoder. Videre samlet vi fargekategorisert tekst inn i egne dokumenter for videre analyse. Her kunne vi hente fram viktig informasjon og sitater som er gjengitt i kapittel fem. På denne måten fikk vi en helhetlig oversikt over alle informantenes perspektiver til ulike temaer og deretter se disse i sammenheng med hverandre. Temaer som ble belyst under intervjuene var hovedsakelig innovasjon og digitalisering, teknologi, samkonkurranse og personvern.

Funnene er presentert som en historie på bakgrunn av at prosjektet befinner seg i en utforskningsfase. Til slutt ble disse funnene diskutert opp mot det teoretiske rammeverket.

3.7. Datamaterialets kvalitet

Ved innhenting av data i kvalitativ forskning, er det viktig å vurdere både primær- og sekundærdataens styrker og svakheter, hvor det samme gjelder for valg av litteratur. Ved innsamling av sekundærdata og litteratur er det viktig å være bevisst på hva man leter etter, og hvordan man bruker disse dataene. Det vil derfor være viktig å vurdere kildenes relevans opp mot egen studie, i tillegg til å være kildekritisk. Når det kommer til innsamling av primærdata, kan studiens tidsbegrensning by på utfordringer. Det kan også oppstå utfordringer rundt intervjuobjekters tilbakeholdenhet, eller at spørsmålene kan oppfattes som ledende. Dette delkapitlet vil vurdere dataens validitet og reliabilitet, samt vurdere datamaterialet opp mot etiske retningslinjer og personvern.

3.7.1. Validitet

Validitet måler gyldigheten og påliteligheten til resultatene av en studie (Gripsrud et al., 2016, s. 132). Ifølge Dahlum i Store Norske Leksikon kan validitet defineres som; «i hvilken grad man ut fra resultatene av et forsøk eller en studie kan trekke gyldige slutninger om det man har satt seg som formål å undersøke» (Dahlum, 2020, u.s.). Eksperimentelle metoder er spesielt opptatt av å sikre at resultatene gir nøyaktige refleksjoner av virkeligheten. Det skilles mellom intern og ekstern validitet.

Intern validitet

Intern validitet går ut på om resultatene kan oppfattes som korrekte. Om noe er korrekt, kan avhenge av flere forhold, men i dette tilfellet skal man konsentrere seg om informasjonen er sann, og om sammenhenger er reelle. Man stiller derfor spørsmål rundt om det er samsvar mellom virkeligheten og forskerens beskrivelse av virkeligheten (Jacobsen, 2015, s. 228). I vår studie har vi intervjuet informanter med relevante stillinger i sine representative selskaper. Disse informantene er derfor egnet til å kunne gi et godt bilde av tematikken vi har studert. Intern validitet i vår studie vil derfor gå ut på om vi kan betrakte informasjonen gitt av informantene som gjeldende.

Ekstern validitet

Ekstern validitet handler om i hvilken grad funnene fra en undersøkelse kan generaliseres til andre formål enn det man faktisk har undersøkt (Jacobsen, 2015, s. 237). En av utfordringene med å gjennomføre en casestudie er at resultatene kan være krevende å generalisere da de kun tar for seg den enkelte casen (Easterby-Smith et al., 2018, s. 116). For at noe skal kunne generaliseres, kreves det et stort utvalg. På bakgrunn av at det vi studerer er en enkeltstående case med få deltakere, vil det ikke per i dag kunne generaliseres. Studien vil likevel kunne bidra til å skape oppmerksomhet og større innsikt innenfor et tema som blir stadig mer relevant i dagens digitale samfunn.

3.7.2. Reliabilitet

Reliabilitet handler om at prosedyren i forbindelse med en undersøkelse, for eksempel dens datainnsamling, kan gjentas og oppnå samme resultat (Yin, 2018, s. 42). Det innebærer at hvis en annen forsker på et senere tidspunkt, gjennomfører samme prosedyre som beskrevet av en tidligere forsker og utfører den samme studien på nytt, skal den andre forskeren kunne opparbeide de samme funnene og ta de samme konklusjonene. Målet med reliabilitet er å minimere feil og bias. I en kvalitativ studie, og spesielt i en casestudie, kan det ofte være utfordrende å oppnå høy reliabilitet siden studien ofte er vanskelig å etterprøve eller gjenskape (Yin, 2018, s. 46). En populær karakteristikk av reliabilitet er å si at den viser hvor godt vi måler det vi måler, og hvor «godt» i denne sammenhengen kan oversettes til «konsistent» (Gripsrud et al., 2016, s. 135).

Utfordringen som Yin presenterer vedrørende at casestudier kan være vanskelige å etterprøve eller gjenskape, finner vi også i sammenheng med vår casestudie. Prosjektet som skal studeres i denne oppgaven er under en rivende utvikling, og det kan derfor være vanskelig å få lik respons og like konklusjoner på et senere tidspunkt. Ved at vi ønsker å gjennomføre dybdeintervjuer ved hjelp av semistrukturerte intervjuguider, kan det også bidra til å skape utfordringer med hensyn til reliabilitet. Semistrukturerte intervjuer legger til rette for at spørsmål og samtalen endres underveis etter hvert som respondenten svarer, og selv om vi har utarbeidet spesifikke spørsmål vi i alle fall ønsker svar på, kan responsen i sin helhet avvike ved en eventuell gjenskaping.

3.7.3. Etiske betraktninger og personvern

Forskningens troverdighet er avhengig av at forskerne følger de etiske prinsippene for forskning og at vi har et normsystem for forskningen (Regjeringen, 2020). Vi har ivaretatt informantenes interesser ved å sende ut samtykkeskjema i forkant av intervjuene. Skjemaet presiserer studiens formål, samt hva informantene har krav på når det kommer til innsyn, retting og sletting av informasjon de har delt. På denne måten sikres det at respondenten ikke tar skade av informasjonen som deles i forbindelse med studien. Skjemaet som har blitt sendt ut i forkant beskriver og sikrer taushetsplikten vi som forskere besitter, tillegg til hvordan navn på informanter og selskaper behandles. Vår studie har gjort at vi også er nødt til å ta hensyn til personopplysningsloven. Loven gjelder om en skal registrere, behandle eller lagre personopplysninger, og lovens formål er å verne om respondentenes personopplysninger. Vi har dermed en lovpålagt meldeplikt til personvernombudet for forskning; Norsk Senter for Forskningsdata (NSD). Plikten er overholdt ved at det er sendt inn meldeskjema for prosjektet i god tid før datainnsamlingen starter. Se vedlegg 9.2.

3.8. Studiens begrensninger

Gjennom arbeidet med vår masteroppgave har vi møtt på utfordringer knyttet til Covid-19 pandemien. Dette har ført til at intervjuene våre har foregått over digitale plattformer, noe som gjør at vi ikke har hatt mulighet til å gjennomføre ansikt til ansikt intervjuer. Dette kan ha påvirket intervjuenes kvalitet da det er vanskeligere å tolke og observere kroppsspråk over video. I tillegg til dette går man glipp av å observere intervjuobjektene på deres naturlige arbeidsplass. En annen begrensning ved studiet er at det er relativt få deltakere involvert, og enda færre ressurser fra hvert enkelt selskap som jobber dedikert med dette prosjektet. Likevel har vi fått et tilfredsstillende datagrunnlag for å kunne belyse studiens tematikk.

4. EMPIRISK KONTEKST

På bakgrunn av at vi studerer et spesifikt innovasjonsprosjekt ønsker vi å presentere dette prosjektet for å gi en dypere forståelse for hva prosjektet innebærer. Videre ønsker vi å presentere de ulike aktørene som inngår i prosjektet. Vi starter med å presentere NCE Finance Innovation som er leder. Videre presenteres de tre forsikringsselskapene som tar del i prosjektet, samt selskapet som utvikler den tekniske løsningen og selskapet bak juridisk rådgivning. Alle aktørene i innovasjonsprosjektet er medlemmer av finansklubben NCE Finance Innovation.

4.1. Presentasjon av innovasjonsprosjektet

Innovasjonsprosjektet, Fraud Detection Platform i regi av NCE Finance Innovation, legger fokus på å validere et konsept for trygg deling av data mellom konkurrerende selskaper for å bekjempe et felles samfunnsproblem, nemlig forsikringssvindel. Ved å ta i bruk maskinlæring og Big Data ønsker selskapene som er involvert å finne uregelmessigheter i skadeoppgjør. De ønsker også å identifisere fellesnevnerer som peker i retning av forsøk svindel, og på denne måten skal sammen prøve å avdekke flere svindeltilfeller. Målet med prosjektet er at plattformen for svindeldeteksjon skal bli en nasjonal og obligatorisk bransjestandard. Det er likevel viktig å presisere at prosjektet ikke har et ferdig produkt enda og at løsningen foreløpig er i en testfase.

4.2. NCE Finance Innovation

NCE Finance Innovation er en norsk non-profit fintech klynge som arbeider mot å legge til rette for teknologisk innovasjon og samarbeid innen teknologi og økonomi (Finance Innovation, u.å.). Målet med klyngen er å utvikle nye finansielle tjenester og innovative løsninger hos både nye og eksisterende selskaper. Klyngen består nå av 75 medlemsselskaper (Finance Innovation, u.å.), blant annet innen bank, forsikring, teknologi og forskning, samt startups og investorer (Innovasjon Norge, u.å.).

4.3. Tryg Forsikring

Tryg Forsikring operer i både Norge, Sverige og Danmark, og er Nordens nest største skadeforsikringselskap. I Norge har Tryg en markedsandel på 13,2% (Finans Norge, 2020), som gjør det til Norges fjerde største forsikringselskap. Tryg Forsikring Norge har hovedkontor i Bergen, og selskapet har om lag 1000 medarbeidere som alle jobber for en tryggere og enklere hverdag for sine kunder (Tryg, u.å.). Tryg er et selskap som satser på nytenking og innovative løsninger, og de har sett seg nødt til å tenke nytt for å overleve i et konkurransepreget marked (Tryg, u.å.). Svindeldeteksjon er et område hvor Tryg har oppdaget et stort potensial for forbedring. I 2020 ble det derfor implementert en mer avansert metode for svindeldeteksjon ved hjelp av algoritmer, som har som formål å hjelpe skadebehandlerne i å plukke ut saker som har større sannsynlighet for å inneholde svindel (Tryg Årsrapport, 2020, s. 14).

4.4. Frende Forsikring

Frende Forsikring er et resultat av at flere banker har fusjonert for å tilby kundene sine forsikring, i tillegg til respektive banktjenester. Bankene som i 2007 gikk sammen og dannet Frende Forsikring, er Sparebanken Vest, Sparebanken Øst, Fana Sparebank og Helgeland Sparebank (Frende, u.å.). På bakgrunn av at forsikringsselskapet er forholdsvis nytt har de likevel fått en markedsandel på 3,5% (Finans Norge, 2020), og har over tid vært Norges raskeste voksende forsikringsselskap (Frende, u.å.). Frende Forsikring har hovedkontor i Bergen, og selskapet har om lag 300 ansatte (Frende, u.å.). Frende Forsikring bidrar med å skape et tryggere samfunn ved innføre kundetiltak for å sikre at de ikke blir brukt som et ledd innen finansiell kriminalitet. Frende har blant annet satt i gang prosesser for å forbedre og videreutvikle svindeldetektering ved hjelp av maskinlæring (Frende Årsberetning, 2020, s. 6).

4.5. Fremtind Forsikring

Fremtind Forsikring er et forsikringsselskap som i 2019 vokste fram ved at DNB og Sparebank 1 fusjonerte, og er Norges tredje største forsikringsselskap med en markedsandel på 14,1% (Finans Norge, 2020). Fremtind Forsikring har hovedkontor i Oslo, og har om lag 900 ansatte (Fremtind, u.å.). Målet med fusjonen var å bygge et forsikringsselskap som skal kjennetegnes av digital innovasjon og fornyelse. En stor del av selskapets samfunnsansvar bygger på å utvikle gode løsninger innen svindel og skadeforebygging (Fremtind Årsrapport, 2019, s. 31).

4.6. Webstep

Webstep ble etablert i år 2000, og er et teknologi- og rådgivningsselskap som ønsker å skape verdi i samspill mellom mennesker og teknologi (Webstep, u.å.). Webstep er hele tiden ute etter å følge utviklingen i markedet, og tilpasser derfor sine tjenester deretter. Selskapet setter kunnskapsutvikling og kunnskapsdeling høyt, og er godt posisjonert for å utvikle nye digitale løsninger (Webstep, u.å.). Webstep tilbyr en rekke tjenester innenfor utvikling av stabile, fremtidsrettede og brukervennlige digitale løsninger. Blant annet tilbyr de rådgivning, systemutvikling, AI og maskinlæring. Innenfor tjenesten AI og maskinlæring har Webstep 30 ansatte som er spesialister på feltet (Webstep, u.å.).

4.7. PwC

PwC er et konsultentselskap som kan bistå forskjellige selskaper med rådgivning, revisjon, økonomi og juridiske tjenester (PwC, u.å.). Innen juridiske tjenester har de spesialister på forskjellige områder, blant annet innenfor GDPR, personvern og informasjonssikkerhet. Dette er et område som har fått mye oppmerksomhet i forhold til moderne forretningsdrift- og utvikling (PwC, u.å.).

5. ANALYSE

I følgende kapittel skal vi presentere empiriske funn. Funnene presenteres som en historie hvor vi ser på begrepet svindel og hvorfor det er regnes som et samfunnsproblem, hvordan digitaliseringen har påvirket forsikringsbransjen med hensyn til svindel, og dernest hvorfor innovasjonsprosjektet ble initiert. Videre ser vi nærmere på hvordan innovasjonsprosjektet konkretiseres underveis, og hva som foreløpig er presentert som *minimum viable product* (MVP). Nåværende MVP har tre underliggende mål knyttet til det forretningsmessige, det teknologiske og det juridiske. Disse målene legger grunnlaget for teknologien og strategien i prosjektet, og vi ser nærmere på tilhørende utfordringer som har påvirket arbeidet så langt.

5.1. Fra problem til MVP

Det kommer fram i samtlige intervjuer at forsikringssvindel regnes som en kriminell handling, og selskaper i forsikringsbransjen er dermed pliktig til å gjøre sitt ytterste for å avdekke denne type aktiviteter på lik linje med annen finansiell kriminell aktivitet; som for eksempel hvitvasking. Head of Machine Learning and AI fortalte at: «*Selskapsvisjonen og samfunnsoppdraget vårt handler om å begrense svindel og å prøve å finne dem som ønsker å utnytte systemet*».

Prosjektleder i NCE Finance Innovation forklarte bakgrunnen for dette på følgende måte:

(...) Det er ganske mye forsikringssvindel i Norge. Det seneste tallet jeg har, som er fra 2019, er at det ble avdekket skadeforsikringssvindel for 500 millioner kroner (...) og i tillegg så tror man jo at det er ganske store mørketall som man ikke har klart å oppdage (Prosjektleder, NCE Finance Innovation).

Avdekket skadeforsikringssvindel kan være mange forskjellige scenarioer, og våre informanter beskrev flere kreative metoder kundene benytter for å tilegne seg midler de egentlig ikke har krav på. Head of BICC/Business Analyst i Frende uttrykte at: *«De typiske sakene er eneulykker; altså der biler har blitt kjørt på vannet, men meldt stjålet og så finner man de på 20 meters dyp et eller annet sted»*. Finansdirektør i Frende la til:

Og så er det fungerte kollisjoner hvor man avtaler en kollisjon fordi man kanskje har lyst på en ny forskjerm, og man synes det er en bedre måte å løse det på enn å betale verkstedet for å reparere skaden (...) Et annet tilfelle er for eksempel dersom en skade oppstår kort tid etter man kjøpte forsikringen, det er en indikator på at noe er mistenkelig (Finansdirektør, Frende Forsikring).

Spesielt en av våre informanter legger vekt på at digitaliseringen har vært en avgjørende drivkraft bak endringer generelt i bransjen, men også når det gjelder kundeforventninger. På bakgrunn av dette uttalte etterforsker i Tryg at: *«Det stilles høyere krav til tempo i saksbehandlingen og raske oppgjør (...) og det er hele tiden denne avveiningen; jo bedre vi er på, kall det kundebehandling, jo større risiko eksponerer vi oss for»*. Etterforsker la imidlertid til at *«Dette er noe markedet legger på oss, men også noe vi delvis legger på oss selv fordi vi ikke ønsker å falle bakpå»*. Mye av risikoen ligger i at selskapene stadig prøver å tilpasse seg et marked i endring, og ikke vil tape markedsandeler sammenlignet med sine konkurrenter. Automatiserte prosesser som legger vekt på rask behandlingstid er ofte der man finner svindeltilfeller, og informantene uttrykte at denne risikoen vurderes i sammenheng med hvilke konsekvenser det derimot vil ha å ikke følge markedets forventninger.

Forsikringssvindel blir ansett som et samfunnsproblem, og flere aktører i finansbransjen ser et behov for å bekjempe svindlerne. Herfra vokste det fram et spørsmål som handlet om hvorvidt den norske finansbransjen har mulighet til å skape felles innovative stordataløsninger som både gir konkurransefortrinn, men som også er i stand til å møte internasjonal konkurranse. I regi av NCE Finance Innovation og forsikringsselskapene i klyngen, ble det derfor iverksatt et innovasjonsprosjekt med mål om å bekjempe forsikringssvindel gjennom samarbeid. Data scientist i Tryg forsikring fortalte: *«Du kan jo tenke deg at vi kan samarbeide om mye forskjellig, men det vi så var at svik er noe som angår alle forsikringsselskapene»*.

Ideen med innovasjonsprosjektet bygger på at det skal være mulig å lage et felles datasett hvor deltakerne deler historiske data knyttet til sine skadesaker, og ved å kombinere alle disse dataene skal selskapenes respektive prediksjonsmodeller for svindeldeteksjon ha et større datasett å trene på. Behovet for en slik løsning kommer av at Norge er et relativt lite land med en relativt liten befolkning, som også gjør at antall skader og svindeltilfeller er færre sammenlignet med andre land. Derfor finnes det også en risiko for at større internasjonale selskaper kan ta over markedsandeler.

The idea of meeting this unfair data advantage that the huge international giants have; like Google, Facebook and Amazon, and trying to use your superior knowledge of your local market, customers and culture, combined by sharing data with others who also have this extra insight to the market, is an interesting idea. It's an interesting idea to do this pooling of data who allows more companies to meet the competition on bigger data. And the concept of a sharing platform for data; it has this self-reinforcing value, like every addition creates value for everyone (Avdelingsdirektør, Webstep).

Som et resultat av digital innovasjon og ny teknologi har flere forsikringsselskaper begynt å ta i bruk KI og Big Data i interne prediksjonsmodeller for svindeldeteksjon. Likevel la flere informanter vekt på at datagrunnlaget for hvert selskap er for lite til at disse modellene fungerer optimalt alene.

(...) Det er ingen som har et veldig stort datagrunnlag for å trene de modellene som de bruker til å oppdage forsikringssvindel. Det er jo dataalgoritmer som trener på et datasett (...) og tester hvor mye de klarer å oppdage. Løsningen på det er jo å få samlet en haug med data som gir disse modellene et større datagrunnlag å trene på. Hvis de har det, er hypotesen at de kan bli mye mer effektive til å finne svindeltilfeller eller de røde flaggene som må undersøkes nærmere (Prosjektleder, NCE Finance Innovation).

Avdelingsdirektør i Webstep underbygget beskrivelsen til prosjektleder fra NCE Finance Innovation ved å beskrive at verdien i plattformen ligger i det faktum at du kan trene prediksjonsmodeller på større datasett noe som i teorien skal bidra til å lage bedre modeller med større prediksjonskraft.

Basert på behovet for større datasett, innebærer det også at konkurrerende selskaper må samarbeide med hverandre. Ved å dele data i en felles dataplattform håper forsikringsselskapene å få tilgang til mer fjerne observasjoner som det enkelte forsikringsselskap ikke besitter alene, men også gjerne mer heterogenitet i hvordan selve deteksjonen foregår. På denne måten kan hvert enkelt forsikringsselskap forbedre sine treningsdata ved å kombinere data gjennom samarbeid.

I starten av prosjektet favnet det så bredt at det inkluderte medlemmer fra både bank, liv- og skadeforsikring, ulike akademia, rådgivnings- og konsultentselskaper og et teknologiselskap. Mange medlemmer førte også til store prosjektmøter med mange involverte, noe som bidro til at det ble vanskelig å bestemme en felles retning og et felles fokus for prosjektet. Likevel førte det også til at prosjektet fikk tilgang på mange ulike ideer og kunnskap rundt hva som ville være mulig å få til.

Another challenge we had was that we had both bank and insurance companies, and both life and non-life insurance, and all the companies had their own agendas on what they wanted to solve. (...) Everybody was going in their own direction and wanted to solve their own problem, and that's not a very simple way to meet collaboration (...) In this phase we definitely lacked focus and a clearly defined goal, and then we also lacked commitment for everyone involved. It was very much people who had an interest within the area of data, but kind of came along for a meeting every couple of weeks and then went back to their own jobs, and nothing happened. So, it was not particularly productive in terms of finding a solution or going in a very structured direction (Avdelingsdirektør, Webstep).

Som et resultat av for mange «kokker på kjøkkenet» ble det besluttet at dersom prosjektet i det hele tatt skulle komme videre var det behov for å snevre inn fokusområdet betydelig. Herifra ble det bestemt at prosjektet skulle fokusere på bekjempelse av svindel innenfor bilforsikring. Derfor forsvant aktører som hovedsakelig jobbet med livsforsikring, men også akademia som ønsket å forske på dataene med ulike vinklinger da det foreløpig ikke er nok data å forske på. På dette tidspunktet sitter innovasjonsprosjektet igjen med NCE Finance Innovation som prosjektleder, Webstep som teknologiselskap, PwC som juridisk rådgiver, og forsikringsselskapene Tryg og Frende.

(...) Thankfully we got some speed and direction on this, and I think it has a lot to do with getting Finance Innovation as a project manager (...) as well as getting support from Innovasjon Norge which was really important (Avdelingsdirektør, Webstep).

Selv om det nå ble bestemt et fokusområde for prosjektet, viste det seg at det ikke var så enkelt å få fart på prosjektet likevel. Flere informanter beskrev forsikringsbransjen som en risikoavers bransje, noe som gjør at flere forsikringsselskaper har valgt å vente med å delta i prosjektet før de ser at en potensiell deltakelse vil påvirke dem positivt. En av våre informanter la også vekt på at en risikoavers holdning kan bidra til å hemme utviklingen i samfunnet. Det ble også avdekket at større forsikringsselskaper viser tendenser til å være mer risikoavers sammenlignet med mindre selskaper. Dette har bidratt til ulike inngangsbarrierer for forskjellige forsikringsselskaper i prosjektet. Store forsikringsselskaper har større dilemmaer knyttet til problemstillinger om kost-nytte, mens mindre forsikringsselskaper har mye å hente på et slikt samarbeid. Informant fra Fremtind uttalte at: «*De som sitter på minst data, er de som får mest og gir minst.*», og informant fra Tryg la til at: «*(...) Selskaper med mindre ressurser vil også ha mye mer igjen for å få tilgang til et stort datasett enn de store selskapene.*»

Herifra møter prosjektet utfordringer knyttet til å få nok momentum og data inn i prosjektet. På dette tidspunktet var det kun to skadeforsikringsselskaper involvert, og disse to selskapene hadde ikke store nok datasett for at prosjektet kunne skape verdi. På bakgrunn av dette ønsket ikke selskapene å legge inn for mye ressurser i arbeidet med prosjektet før det viste seg at dette kunne bidra til å gi en reell verdi for hvert av selskapene. Arbeidet med å finne et tredje forsikringsselskap gjorde derfor at prosjektet har gått over en mye lengre periode enn først antatt. Fremtind Forsikring var det tredje forsikringsselskapet som valgte å bli med på innovasjonsprosjektet, og derfra ble en MVP konkretisert.

One of the things that helped us pick up speed is that we got a third insurance company on board, and we lost the other guys on the way. That wasn't necessarily a good thing, but it happened, and I think it also helped us regain our focus (Avdelingsdirektør, Webstep).

I begynnelsen var vi veldig opptatt av at vi måtte få flere selskaper inn i samarbeidet for at det skulle ha en verdi i forhold til at datasettet ble stort nok til at det kunne gi

merverdi. Så et nøkkelpunkt var at vi fikk med et tredje selskap inn i samarbeidet (Head of BICC/Business Analyst, Frende Forsikring).

Selskapene som nå inngår i innovasjonsprosjektet, har forpliktet seg til en feasibility studie. Dette innebærer at alle relevante faktorer for prosjektet analyseres; dette kan være vurderinger knyttet til det teknologiske, forretningsmessige og juridiske. Med bakgrunn i feasibility studien er det også utarbeidet tre mål for MVP-en. Det første målet er et forretningsmål som innebærer å verifisere hypotesen om at deling av data vil resultere i bedre prediksjonsmodeller og bedre datakvalitet. Det andre målet er et teknologisk mål; her ønsker man å verifisere at det er mulig å utvikle en løsning hvor man kan dele data på tvers av konkurrerende selskaper, uten å gi innsyn i disse dataene til de konkurrerende selskapene. Det tredje målet er et juridisk mål hvor det handler om å verifisere at det er mulig å lage en slik løsning uten å bryte med GDPR, personverns- og konkurranselovgivning. Det er også innenfor disse målene at prosjektet har sett sine største muligheter, men også begrensninger.

5.2. Teknologisk løsning og utfordringer knyttet til denne

I løpet av intervjuene er det avdekket at kunstig intelligens og Big Data er i ferd med å bli viktige verktøy i kampen mot forsikringssvindel. Dette kan være utfordrende for mindre forsikringsselskaper ettersom kundegrunnet er for lite, og skadekravene og svindeltilfellene for små, til at man kan bruke denne type teknologi effektivt. På lang sikt kan det være en utfordring for norske forsikringsselskaper som kan tape markedsandeler til større internasjonale selskaper som sitter på større mengder data. Dette legger grunnlaget for innovasjonsprosjektets tekniske løsning som bygger på at norske forsikringsselskaper skal kunne utvide sitt datagrunnlag for svindel deteksjon ved å dele data. Prototypen som utvikles i prosjektet er begrenset til et delt datasett på bilskade.

(...) It's giving everybody involved possibilities of training algorithms on Big Data sets, which is going a good way towards levigating the difference between big international companies and their huge data advantage and smaller Norwegian companies and their data disadvantage. So, it's a very high-level concept that helps Norwegian small, or any smaller company basically, to combat bigger competitors by pooling data together (Avdelingsdirektør, Webstep).

Delingsplattformen skal legge til rette for at hvert forsikringsselskap skal kunne trene sine egne prediksjonsmodeller for svindel deteksjon på et felles dataunivers, uten å få innsyn i de konkurrerende selskaperes data. Løsningen bygges i en felles sky levert av Amazon Cloud Services, hvor hvert selskap deler sine historiske skadedata.

(...) Løsningen er jo at det foregår på Amazon Cloud Services hvor hvert selskap har deres egen «bucket» (...) hvor de kan laste opp data. Så er det en felles konto som Webstep foreløpig styrer i MVP-en (...) den kontoen kan lese de dataene som er tilgjengeliggjort i hvert selskap sin bucket, men selskapene kan ikke se hverandre sine data. Når selskapene har lastet opp data og felleskontoen klarer å lese hva som er i de boksene, kan selskapene sende opp deres prediksjonsmodeller til felleskontoen som da gjennom en applikasjon i clouden kan kjøre modellene på dette felles datasettet. Uansett hvor stor forskjell det er på prediksjonsmodellene, alle er jo treningsmodeller, så kan den kjøres og sendes tilbake til selskapet (...) tenk på det som et treningssenter – selskapene har mulighet for å gå på treningssenteret for å trene seg bedre enn de kunne gjort for seg selv (Prosjektleder, NCE Finance Innovation).

Som følge av at selskapene har ulike prediksjonsmodeller, er det også naturlig at de har ulikt begrepsapparat og ulike definisjoner på ting. Løsningen i Amazon Cloud bygger dermed også på det som kalles en felles Data Dictionary, hvor dataene fra de ulike selskapene oversettes inn i dataplattformen. Hvert forsikringsselskap setter sammen sitt datasett i henhold til kravene i Data Dictionary, verifiserer disse dataene gjennom den felles plattformen og gir tilgang til at andre prediksjonsmodeller i plattformen kan lese av de verifiserte dataene. Plattformen returnerer en ferdig trent prediksjonsmodell og resultatene av modellkjøringen. De ulike forsikringsselskapene kan ende opp med forskjellige resultater basert på hvordan deres modell både bygges og tolkes på bakgrunn av kompetanse og ressurser hos det enkelte selskapet.

Man må kunne å snakke et felles språk for å i det hele tatt klare å dele dataene. Så man har definert en felles Data Dictionary hvor vi tar våre data fra Tryg, Fremde sine data, Fremtind sine data og oversetter disse inn i denne strukturen, og deler de samme dataene (Data scientist, Tryg Forsikring).

(...) We are sure that everything is being shared on the Data Dictionary format that is accepted by the platform, is in compliance with GDPR and competitive laws (...) Every data set that you would try to introduce here, that is not in compliance with the Data Dictionary is automatically rejected, so it's not going to be exposed to the other participants. At the same time, you create a framework and a methodology for safely sharing data, and of course the idea is not only to save money but also to combat the social problem of fraud (Avdelingsdirektør, Webstep).

Ved hjelp av en felles Data Dictionary kan prediksjonsmodellene fra hvert selskap fange opp relevante sammenhenger som har vist seg å være knyttet til for eksempel svindel. Data Dictionary har blitt forutsetningen for at selskapene skal kunne danne et godt datagrunnlag, og informantene legger vekt på at et godt datagrunnlag er svært viktig for å kunne lykkes med maskinlæring. Informant fra Frende Forsikring hevdet at: «*Garbage in, garbage out har nok ikke vært mer sant enn nå*».

Likevel er svindel et komplekst begrep ettersom man bare kan bygge modeller på de tilfellene man allerede kjenner til, altså historiske data. Verdien av plattformen, og muligheten til å avdekke flere tilfeller av svindel, er antatt å øke med økt datamengde. Informant fra Tryg forsikring uttrykte at: «*Her har vi et problem hvor vi ikke kjenner til alle sviktilfellene, så det er jo noen svakheter og utfordringer for maskinlæringsalgoritmer som vi prøver å kompensere med å dele data*». Videre la informant fra Frende forsikring til at: «*(...) Man prøver å predikere noe som skjer en av tusen, eller en av ti tusen ganger, og da må man ha veldig mye data for å klare å lage gode modeller*».

Når man sitter på større mengder data og jobber med maskinlæring, vokser det også fram utfordringer knyttet til dataenes forklaringskraft. Sammenlignet med systemer for automatisert skadebehandling som bygger på regresjonsmodeller, vil ikke maskinlæring være like transparent, noe som generelt er en vanskelig problemstilling. Det vil kreve mye av en maskinlæringsmodell bestående av mange variabler å forklare hvorfor den har kommet fram til sin beslutning. Dette er omtalt som Black Box-problematikken innen KI. En kunde kan også kreve en forklaring på hvorfor deres skade ble gitt et rødt flagg i prediksjonsmodellen.

Det vil kreve mye av modellen å forklare hvorfor den har kommet til sin beslutning, og det vil kreve veldig mye av kundene for å kunne forstå det. Hva mener man egentlig

med en forklaring? Trenger kundene en mastergrad i statistikk? Eller handler det om å lage en historie rundt det? Det kan man jo alltid gjøre. Eller er det en forklaring som betyr at det skal påvise en kausalitet mellom de dataene du har gitt? (...) Så hva gjør man i praksis? Det er ingen svar på akkurat det temaet. Men det man gjør stort sett er at man synliggjør hvilke variabler som gjør utslag til en rådgiver, og så blir det rådgivers oppgave å fortelle historien til kunden (Head of Machine Learning and AI, Fremtind Forsikring).

Dataene fra andre selskap som våre modeller trener på er fullstendig administrert, og vi vil umulig kunne si hvilket selskap eller hvilken kunde som har disse dataene. Det er jo ikke det som er interessant heller, for det som er interessant er å lage mønstre som gjør at vi klarer å identifisere det (Etterforsker, Tryg Forsikring).

Til syvende og sist skal kunstig intelligens og stordataplatfformen fungere som verktøy for å bedre kunne predikere svindel på tvers av flere forsikringsselskaper. Ambisjonen er å øke etterforskningsdekningen til 100% ved å trene på data, teste dataene, og over tid ha dannet algoritmer som blir bedre og bedre på å predikere det man leter etter. Avdelingsdirektør i Webstep uttalte at: «(...) *You can imagine if the algorithms get sufficiently good, you can theoretically cover every fraud*».

5.3. Samkonkurransen

Det var først når arbeidet med MVP-en startet og alle deltakerne i prosjektet var bestemt, at man oppdaget flere utfordringer knyttet til samarbeid, og forskjeller mellom de tre involverte forsikringsselskapene. Når selskaper skal gå sammen i et slikt innovasjonsprosjekt er det viktig å ha et klart skille på hvor man samarbeider og hvor man konkurrerer.

(...) Vi er jo veldig bevisst på at vi representerer ulike selskaper og har til dels motstridende interesser. Men samtidig så er det jo en felles interesse her som bringer oss sammen. Jeg synes ikke at det har preget samarbeidet negativt, men vi har jo vært konservative og forsiktige i forhold til hva vi sier til hverandre (...) Vi har generelt vært forsiktig (Data scientist, Tryg Forsikring).

(...) Å begrense svik og på en måte prøve å ta folk som utnytter systemet. Det er jo noe som er viktig for oss, men det er også viktig for hele bransjen og noe vi ikke har noe

problem med å gå inn å hjelpe andre selskaper med å få til (Head of Machine Learning and AI, Fremtind Forsikring).

I think the reason that this is a fairly easy area for cooperation is that all the companies have a common interest in identifying the fraudsters. So, this is a problem for the entire industry, and none of the companies has the business idea to make money on having less fraud than other companies. They all want to kill fraud. So, I think there is a common interest in making it really hard to get away with insurance fraud. So, I think fraud is an area that is easy to create a competition on (Avdelingsdirektør, Webstep).

Selv om innovasjonsprosjektet handler om å bekjempe et felles problem, kan de forskjellige forsikringselskapene gjerne ha forskjellige perspektiver på hva det er som gir verdi, og hva som vil være forretnings- og konkurransesensitivt å samarbeide om. Her så vi at selskapene pekte i litt forskjellige retninger, noe som igjen har ført til at innovasjonsprosjektet har gått over en mye lengre periode enn ønsket. Ett av selskapene hadde for eksempel en annen oppfatning om hva som faktisk er mest sensitivt å dele mellom selskapene, og ønsket i utgangspunktet en annen tilnærming til prosjektet;

Tryg var veldig opptatt av at man skal dele data, og så holde modelleringen skjult. Mens vi ser på det litt motsatt, at det er større risiko knyttet til å eksponere data både på en forretningsmessig, men også med hensyn til personvern. Men det å eksponere modeller som vi har utviklet er mindre forretnings sensitivt. Så vi har hatt et litt annet perspektiv på det enn de to andre selskapene, som jeg tror har balansert ting litt (Head of Machine Learning and AI, Fremtind Forsikring).

Perspektivet til Fremtind om å samarbeide om modellering er noe som ikke er utelukket å samarbeide om fra de andre selskapene. Men på bakgrunn av at innovasjonsprosjektet manglet fokus og momentum i starten, er prosjektet foreløpig avgrenset til å se på mulighetene til å samarbeide om å dele data. Dersom man får gode resultater av dette innovasjonsprosjektet understreker informantene at man ikke skal se bort i fra at selskapene vil se på andre måter å samarbeide på i fremtiden.

(...) Så er det veldig mange andre ting som vi også må finne ut av. Hvordan skal prosessene være når det kommer endringer i ulike selskapers datasett? Hvordan skal prosessene være i forhold til å øke det felles datasettet? Hvordan skal ting integreres inn i våre operasjonelle systemer? Skal vi kanskje samarbeide om å kjøpe eksterne data som alle forsikringsselskapene trenger? Kan dette være en arena for å samarbeide om å kjøpe data spesifikk knyttet til svikprediksjon? (Head of Machine Learning and AI, Fremtind Forsikring).

I tillegg til at deltakerne i innovasjonsprosjektet har hatt ulikt perspektiv på hva som potensielt kan gi mest verdi, blir det også lagt vekt på at det eksisterer forskjeller knyttet til agenda, språk og kultur. For at den tekniske løsningen i det hele tatt skal få verdi har selskapene som nevnt arbeidet mye med å oppnå et felles språk – kalt Data Dictionary. Informant fra Fremtind forsikring uttalte at: *«(...) En annen utfordring er jo den åpenbare med at det er vanskelig å samarbeide tre selskaper som har litt forskjellig agenda, snakker litt forskjellig språk og har litt forskjellig kultur»*. Dette ble underbygget av Data scientist i Tryg Forsikring:

(...) Det har også vært komplisert å diskutere det praktiske i forhold til hvordan man skal oppnå dette felles språket, en felles Data Dictionary, hvordan man skal oversette dataene og være trygg på at man faktisk gjør det på rett måte (Data scientist, Tryg Forsikring).

På bakgrunn av at det er konkurrerende selskaper som skal samarbeide om å dele data har selskapene stått ovenfor komplekse sammenhenger som må løses i fellesskap. Slike arbeidsoppgaver kan ofte ta mye tid basert på ulike perspektiver og interesser. Til tross for flere ulikheter har det ikke oppstått store konflikter i samarbeidet mellom de konkurrerende selskapene. Informant fra Fremtind forsikring beskrev at: *«Det har ikke vært noen store konflikter i prosjektet. Utfordringene har lagt mer rundt hva man skal gjøre, og hvordan man skal løse dette problemet»*.

At NCE Finance Innovation fungerer som prosjektleder mener flere har vært til fordel for arbeidet med prosjektet på bakgrunn av at næringsklyngen ikke har en finansiell interesse, men kun en genuin interesse for å løse et samfunnsproblem. Klyngen har derfor heller ikke noe økonomisk eierandel i resultatet, og vil kunne fungere som en megler mellom de

involverte selskapene. Avdelingsdirektør i Webstep beskrev at: «(...) *It works as a mediator that is a trusted partner, everybody can be sure that the focus is on the part of the innovation and creating things rather than purely financial or profitable*».

Selv om selve løsningen av prosjektet er en forholdsvis enkel løsning, var det ikke like enkelt å komme seg dit. En av informantene fortalte oss at prosessen med innovasjon ikke alltid kan ses på som en rett linje. Det kan lages en plan om hvordan innovasjonsprosjektet skal foregå, og avstanden fra A til B i prosjektet ser gjerne ikke så lang og krevende ut. Men det er først når du faktisk skal gå fra A til B, at veien endrer seg dramatisk og at ting gjerne tar lengre tid enn man først hadde antatt. De forskjellige forsikringsselskapene har gjerne forskjellige tråkkfrekvenser i sine arbeidsprosesser, som gjør det vanskelig å få alle i samme retning på samme tid.

There are quite big differences between the companies and how they work, so actually getting them all in the same direction at the same pace is a fairly difficult action. Everything would be easier with one company, but then it wouldn't be a data collaboration initiative and we wouldn't have this coopetition case at all
(Avdelingsdirektør, Webstep).

Informantene vektla at forutsetningen for et godt samarbeid mellom konkurrerende selskaper er gjensidig tillit. En fordel for prosjektet er at det gjennomføres i Norge hvor samfunnet generelt er preget av høy grad av tillit. Samtidig er det også forankret strenge lover og regler knyttet til samarbeid og konkurranse, noe som har bidratt til et stort fokus på etisk forretningsdrift.

I think the level of trust in companies, between companies and to regulators is higher than in every other country. So, we started at a better point then you might have done in other countries. (...) So, this trust is something that is certainly key and absolutely necessary for something like this to be created (Avdelingsdirektør, Webstep).

På bakgrunn av at det er et prosjekt som har utfoldet seg over en ganske lang periode, har det blitt opparbeidet tillit mellom de involverte partene gjennom disse årene. Dette har gjort at selskapene i prosjektet føler seg sikker på at sine samarbeidende konkurrenter ikke utnytter prosjektet til egen vinning. Samfunnsproblemet de samarbeider om å bekjempe er viktig for

selskapene alene, men også bransjen som en helhet. Derfor la våre informanter vekt på at de ikke ser på det som konkurransesensitivt å hjelpe andre selskaper med å lykkes. Likevel utdypet de at det er viktig å huske på at dataene som deles kan brukes til andre forretningssensitive områder i de ulike selskapene, og det er svært viktig at dataene ikke brukes til andre formål enn det prosjektet har bestemt.

Formålet med dette prosjektet er jo å bli bedre på svikdeteksjon. Og så finnes det et regelverk som regulerer hva du ikke har lov til. Og det er jo klart det, man kan kanskje tenke seg at det finnes et potensiale for å misbruke dataene til andre formål. Men det er jo da ikke tillatt (Data scientist, Tryg Forsikring).

Det er jo en utfordring at de samme dataene som kan brukes til dette her, brukes også til andre typer ting i selskaper som er mer forretningssensitivt og som har en veldig stor verdi som man ønsker å verne om (Head of Machine Learning and AI, Fremtind Forsikring).

Intervjuobjektene fra forsikringsselskapene la imidlertid til at de ikke er bekymret for at noen av de andre deltakerne i prosjektet kommer til å utnytte denne formen for deling av data. Det viktigste med samarbeidet, og som alle deltakerne har interesse av, er å bekjempe et felles problem.

Gjennom arbeidet med å sikre forsikringsselskap som deltakere i prosjektet, har det også vært diskutert om flere skadeforsikringsselskap skal inn i prosjektet. Per nå er ikke dette aktuelt for innovasjonsprosjektet, da de involverte deltakerne anser sitt felles datasett som stort nok for å verifisere hypotesen om at deling av data vil resultere i bedre prediksjonsmodeller og bedre datakvalitet. Informantene legger også vekt på at dersom de skulle inkludert flere i prosjektet per i dag, ville nok ting tatt enda lengre tid og de gjerne ville havnet i samme situasjon som i startfasen av prosjektet. Informant fra Fremtind Forsikring la til at: «(...) *Da tror jeg det blir enda tyngre å dra prosessene. Det blir vanskelig å skape konsensus. Det blir flere møter, flere mailer og mer administrasjon*».

Fokuset i innovasjonsprosjektet ligger derfor på at de selskapene som er med til nå skal klare å lande et godt konsept og bærekraftig modell som gjør det enklere for andre selskaper å kunne bli med på et senere tidspunkt. Forsikringsselskapene som er involvert i prosjektet

jobber i dag med testdata for å verifisere at løsningen faktisk bidrar til å avdekke flere svindeltilfeller sammenlignet med eksisterende automatiserte prosesser. Viser det seg at løsningen gir gode resultater, blir det mer aktuelt å involvere andre selskaper på sikt.

Med hensyn til at innovasjonsprosjektet foregår i et miljø preget av samkonkurransen og det arbeides med store mengder data, henholdsvis også persondata, er det et naturlig fokus på personvernrettslige og konkurranserettslige regler. At dataene deles i en felles dataplattform gjør at prosjektets deltakere har vært nødt til å legge til rette for at samarbeidet ikke strider mot konkurranseloven, og at datagrunnlaget ikke strider mot personvernet. Det er på bakgrunn av dette at PwC ble knyttet til prosjektet som juridisk rådgiver.

Vår rolle har vært å sørge for at det som blir gjort er lovlig i henhold til både personvernrettslige, konkurranserettslige, finansregulatoriske regler, og å utarbeide avtaleverk mellom partene. Vi har også vært med i diskusjonene rundt hvordan hele strukturen til innovasjonsprosjektet kan settes opp (Advokat, PwC).

Informanten fra PwC fremhevet at det er viktig å gjøre vurderinger knyttet til størrelsen på selskapene som er involvert i et slikt samarbeid. Dersom samarbeidet primært hadde bestått av store forsikringsselskaper, kunne det blitt ansett som i strid med konkurranselovgivningen. Dette er fordi de store selskapene sitter på større mengder data, og de ville dermed fått en naturlig fordel sammenlignet med mindre forsikringsselskaper som ikke hadde vært inkludert i samarbeidet. Likevel la informanten vekt på at man bør kunne se forbi dette dersom modellene de samarbeider om, ikke har direkte utspring i hvordan forsikringsproduktene prises ut mot forsikringskunder. Dette kan ses i sammenheng med at de ulike forsikringsselskapene som er med i prosjektet har egne prediksjonsmodeller for svindel, og resultatet til hvert enkelt selskap vil variere. Resultatene av prediksjonsmodellene vil igjen påvirke det enkelte selskapets prismodeller.

For det første så vi jo at det var store forsikringsselskaper involvert, og at dette automatisk strider mot konkurransereglene fordi de aller største selskapene vil få en fordel foran alle andre forsikringsselskaper i Norge (...) Men så lenge modellene kommer opp med et resultat som ikke gir god nok informasjon til at det kan være avgjørende for hvordan forsikringstjenestene skal prises ut mot forbrukerne, så vil det trolig være innenfor konkurranselovgivningen (...) Da er ikke fokuset på hvordan ting

skjer men hva resultatet blir. Hvis sluttresultatet blir at de sitter på et mye bedre beslutningsgrunnlag og priser forsikring ut ifra det, så blir det konkurransebrytende (Advokat, PwC).

5.4. Personvern

Prosjektet bygger hovedsakelig på ideen om datadeling. Ettersom forsikring består av store mengder data som potensielt kan knyttes til enkeltpersoner, har ivaretagelse av personvernet også preget rådgivningen fra PwC i stor grad.

(...) Det er veldig mange krav i personvernregelverket som skal tas hensyn til, og det som har vært den største utfordringen har vært å sette det opp på en måte som etterlever de jevne kravene, men samtidig ikke gir en redusert funksjonalitet eller redusert output (Advokat, PwC).

Fram til lansering av MVP har det vært et tidkrevende arbeid å avgrense dataene som deles i dataplattformen slik at det ikke strider med enten personvern eller GDPR. Dataene som deles i MVP-en er derfor av en så lav terskel som overhodet mulig, bare for å i det hele tatt kunne avklare om løsningen er mulig å gjennomføre innenfor det juridiske regelverket.

Dataene som deles nå er av den typen slik at man absolutt ikke kan identifisere noen med den, med mindre du peker deg inn på en spesifikk linje og det er eksempelvis en som eier en Rolls Royce i Vik i Sogn. Det er utrolig liten sjanse for at man klarer å få noe ut av det. Dersom man ikke kjenner datastrukturen, vil det være vanskelig å skjønne hva dataene forteller dersom du ikke vet hva det passer inn i (Prosjektleder, NCE Finance Innovation).

På grunn av compliance, GDPR og konkurranserettslige problemstillinger så har vi vært veldig forsiktige i forhold til hva vi har valgt å dele nå i første omgang. Det er et veldig lite datasett vi deler nå. (...) Det vi gjør med MVP-en nå er jo ekstremt konservativt, (...) og inneholder få felter. Så dette er jo noe vi har lyst til å utvide (Data scientist, Tryg Forsikring).

Nå har vi vært veldig restriktive på hvilke data vi utveksler. Bakgrunnen for dette er at hvis vi klarer å finne datasett som mer eller mindre kan sendes på e-post, så vil det gå

mye raskere å finne ut og målt effekten av dette på (Head of Machine Learning and AI, Fremtind Forsikring).

It's a very low risk with sharing data in this phase. We kept the person data to an absolute minimum in the MVP. So, at the moment you could say that we have no direct person identifying data, but by combining different valuables in the dataset you would be able to identify events and persons directly (Avdelingsdirektør, Webstep).

Et av kravene i personvernlovgivningen er at det skal eksistere et behandlingsgrunnlag når man oppbevarer persondata. En av våre informanter fortalte at grunnlaget for å oppbevare disse dataene absolutt er til stede ettersom hensikten er å bekjempe forsikringssvindel som en kriminell handling og et samfunnsproblem.

Det er mye som taler for at man kan bruke persondata til dette her. (...) Bortsett fra de som gjør noe kriminelt, så har de fleste en egeninteresse av at man tar de som prøver å svindle forsikring. (...) Vi har forholdsvis godt behandlingsgrunnlag hvor det vil være lett å gjøre gode interesseavveininger (Head of Machine Learning and AI, Fremtind Forsikring).

Forsikringssvindel er egentlig en kriminell handling. Det er det samme som hvitvasking og annen kriminell aktivitet som finansbransjen har en plikt til å gjøre sitt ytterste for å avdekke. I den settingen så er det sannsynligvis greit å bruke disse dataene til det, men et prissamarbeid er for eksempel ikke lov (BICC/Business Analyst, Frende Forsikring).

Hvordan dataene skal behandles har blitt ansett som en av de største utfordringene så langt i prosjektet. Derfor har mer eller mindre alle midlene prosjektet ble tildelt av Innovasjon Norge, gått til å dekke utgifter knyttet til juridisk rådgivning og veiledning fra PwC.

We had a whole lot of problems or discussions and clarifications related to GDPR, privacy laws and competition laws. There is a whole lot of work in the legal direction. (...) The one thing that might really be a competitive advantage for this project, is that we have done the legal work and the legal work is so heavy on a solution like this that

any startup couldn't lift it. You really need the financial muscle and the legal partner to do that (Avdelingsdirektør, Webstep).

Alle involverte parter i prosjektet informerte om at arbeidet med personvern og GDPR har vært veldig komplisert, og mye av diskusjonene har bestått av å finne ut om nye små detaljer som legges til vil stride mot regelverket. Det gjelder hovedsakelig temaer knyttet til rettferdighet, Black Box, åpenhet og dataminimering. Rettferdighet handler i stor grad om at algoritmene i prediksjonsmodellene ikke skal diskriminere eller forskjellsbehandle enkeltpersoner basert på diverse forutsetninger. Likevel vil algoritmene trene på de dataene som er flagget som potensiell risiko for svindel, og hvis disse dataene fungerer som kjennetegn på en viss kundegruppe kan det føre til diskriminering og/eller forskjellsbehandling selv om ikke alle i denne gruppen forsøker å svindle.

A very small portion of cases are fraud, and also a very small portion of policies have claims. So, you have a very skewed selection on datapoints, and that will influence what is a good or poor choice of algorithm (...) You could create an algorithm that went and picked every 18-year-old driver out there, and in return there could be fraud (...) You can create any bias algorithms you want, and it's up to the insurance companies to be ethical and correct about it (Avdelingsdirektør, Webstep).

Innen bilskade kan man jo i enkelte tilfeller se at for eksempel unge menn med en viss type bil eller en viss type skadehendelse, er mer suspekt enn andre. Så de tingene har jo prediktiv kraft, og vi vil jo benytte oss av de innenfor regelverket selvfølgelig (Data scientist, Tryg Forsikring).

Utfordringene knyttet til Black Box handler om at beslutningsprosessen i prediksjonsmodellene er vanskelige å forklare. Forsikringskundene har i henhold til personvernlovgivningen krav på å vite hvorfor akkurat deres krav ble flagget som risiko for svindel, og forsikringsselskapene kan måtte gi ut informasjon om hele beslutningsprosessen. Dette kan skape utfordringer for forsikringsselskapene som må være åpne og transparente i dialog med kundene.

Personvernregelverket oppstiller blant annet krav til at man skal behandle persondata transparent, og gi informasjon om alt man gjør. Det innebærer jo da at når man kan

begynne med maskinlæring så blir det rett og slett vanskelig å fortelle hva som skjer. Det er ikke alltid de vet det selv en gang (Advokat, PwC).

Prinsippet om dataminimering har vært en nøkkelfaktor i arbeidet med prosjektet til nå, og kan ses i sammenheng med at MVP-en inneholder et minimum av persondata. Det er fordi at personvernlovgivningen legger vekt på at man skal benytte seg av så få persondata som mulig, samtidig som at man ikke skal beholde dataene lengre enn nødvendig. Det er likevel opp til hvert enkelt selskap hvor lenge de oppbevarer sine data.

Dataene som er lastet opp har veldig lave konsekvenser med hensyn til personvern, eller risiko knyttet til det. De er mer eller mindre så godt som anonymiserte. Jeg liker jo ikke det begrepet «anonyme data», for det finnes i utgangspunktet ikke (...) Man trenger ikke være så streng på sletting, man har på en måte, man kjøper seg mer slack ved å minimere personverkonsekvensene (Head of Machine Learning and AI, Fremtind Forsikring).

Informanten fra PwC understrekte at personvernlovgivningen er uklar, og at det derfor er vanskelig å vite akkurat hvor grensene går. Det er Datatilsynet som vurderer hva som er lov og ikke lov, men dette kan være utfordrende for innovative prosjekter som dette, å få avklart i forkant. Informant fra Fremtind forsikring la til at: «Jeg tror det er viktig i alle slike typer prosjekter at man starter i utgangspunktet litt forsiktig». Advokat fra PwC underbygget dette på følgende måte:

Det gjør at vi har lagt oss på den trygge linjen, og det kan derfor hende at vi har lagt oss langt under linjen og får egentlig mye mindre verdi enn vi potensielt kunne ha fått. Men det er rett og slett fordi vi ikke vet helt hvor grensen går (Advokat, PwC).

I forbindelse med at finansteknologi er blitt stadig mer relevant, og flere jobber med problemstillinger knyttet til finans og teknologi i samspill, har Datatilsynet i 2020 opprettet det de kaller for en regulatorisk sandkasse. Informantene beskrev denne sandkassen som en plattform hvor ulike selskaper kan utvikle og teste nye teknologier i samsvar med personvernlovgivningen, og med veiledning fra Datatilsynet. Innovasjonsprosjektet søkte opptak til denne sandkassen i desember 2020, men havnet dessverre ikke gjennom nåløyet denne gang. Advokat fra PwC beskrev den regulatoriske sandkassen som en arena hvor:

«Datatilsynet gir mulighet for at selskapene forklarer hva de har tenkt og hvordan, for å så få input fra Datatilsynet om dette vil være lovlig eller ikke». Viktigheten og relevansen for en slik arena ble underbygget av Head of BICC/Business Analyst i Frende Forsikring:

Hvis du setter et produkt i produksjon, og så finner Datatilsynet ut at du har gjort noe ulovlig så kan du få en ganske solid bot. Men i en sånn sandkasse kan det hende at det du gjør, og for så vidt det du tester ut på ordentlig, blir konkludert med at er ulovlig, men det vil i så tilfelle ikke være føre med seg noe straff. Du vil ikke kunne fortsette, og finne andre løsninger. Men det er ingen trussel om straff, hvis det viser seg at noe ikke er innenfor regelverket (...) De er opprettet for å fremme innovasjon for at streng regulering og lovverk ikke skal hindre innovasjon altfor mye (Head of BICC/Business Analyst, Frende Forsikring).

For å være på den trygge siden, og kunne teste løsningen mest mulig, er det som nevnt ikke mange persondata inkludert i dataene som behandles. Likevel fortalte samtlige informanter at de ser for seg at løsningen kan bli en nasjonal standard dersom testingen av MVP-en blir en suksess.

(...) The long-term goal, or kind of the big hairy goal of the solution, is to become a national platform for fraud detection in insurance. A fully covering platform that every insurance company is a part of. We've had different talks about this, and Finanstilsynet are supporting this idea and are saying if such platform can be created and it works, we may look into making it mandatory for every insurance company (Avdelingsdirektør, Webstep).

Avdelingsdirektør i Webstep uttalte videre at dersom innovasjonsprosjektet skal bli en nasjonal plattform for å detektere forsikringssvindel i Norge, så må den styres av en gjensidig tredjepart. To forslag blir fremmet; Finans Norge og Finanstilsynet. Dette ble støttet av informant fra Tryg og Fremtind som begge ønsker at prosjektet skal ende opp som en obligatorisk og nasjonal løsning, og at det blir sett på som en naturlig utvikling av prosjektet. Det fremmes likevel at det vil være en del utfordringer som må løses før dette kan bli aktuelt.

There is to viable options; Finans Norge which is the collaboration of insurance companies and financial institutions in Norway, and they already have a lot of these

common data solutions for insurance companies. Or Finanstilsynet who will have it as an owning to insure continue of the fraud detection for every insurance company in the market and equal accessibility to data for training algorithms to every competitor in the market (Avdelingsdirektør, Webstep).

Jeg har i hvert fall et ønske om at dette skal ende opp som en obligatorisk nasjonal løsning som håndteres av Finans Norge. Vi har jo vært og diskutert dette med dem, og vi har jo presentert det for styret i Finans Norge. Jeg tror de vil være positive til det, men det er jo veldig langt frem. Det er en mengde med problemstillinger som man må gå nærmere inn på; for eksempel om man vil at alle skal være med, i tillegg til nyetablerte selskaper og små selskaper uten mye relevant data. Jeg tror først og fremst at vi må gjennomføre MVP-en, og få en beslutning om vi faktisk ønsker å gå videre vi tre som er med (Data scientist, Tryg Forsikring).

Det kan være veldig interessant og veldig spennende som en nasjonal standard. Forsikringsbransjen, eller finansbransjen generelt, har jo på en måte en god del sånne type ordninger fra før hvor Finans Norge eier og drifter systemer som medlemmene bruker og betaler for vedlikehold. Jeg ser på det som en naturlig utvikling gitt bransjens historie, vi pleier å løse felles utfordringer (...) Det er avhengig av at vi finner en god og bærekraftig modell, og at vi klarer å vise at her er det gevinster og effekter å få ut av dette (Head of Machine Learning and AI, Fremtind Forsikring).

De la også vekt at det kan være aktuelt å skalere datasettet til å inneholde andre forsikringsprodukter i fremtiden, men at de da vil stå ovenfor mye større utfordringer med hensyn til personvern- og konkurranselovgivning. Informant fra Fremtind la fram at det er flere retninger man kan skalere innovasjonsprosjektet i; datasettene kan gjøres bredere, eller man kan inkludere andre forsikringsprodukter.

Det er jo et interessant spørsmål om vi tre selskapene skal prøve å skalere. Det er flere retninger man kan skalere dette her i. Det ene er at man kan gjøre datasettene bredere, for nå har vi vært veldig restriktive på hvilke data vi utveksler (...) det er en måte å skalere på, ta inn flere og flere sensitive variabler, både i personvernmessig forstand, men også forretningssensitiv forstand, så ser vi hvordan det går og om det øker prediksjonskraften. Den andre måten å skalere på er å inkludere andre

oppgjørshendelser eller andre produkter (Head of Machine Learning and AI, Fremtind Forsikring).

Advokat i PwC informerte om at: «Så lenge man klarer å finne en god teknisk løsning skal det være mulig å skalere, men hvis man ikke klarer det kan man få problemer». Informanten anbefalte at i den grad det er mulig å gjennomføre analyser på anonymiserte datasett, er det det som bør gjøres. Dersom man gjennomfører analyser på anonymiserte datasett, vil opplysningene ikke lenger klassifiseres som personopplysninger, og man kan gjøre hva man vil med dem.

(...) Derfor prøver jeg alltid å si at man bør anonymisere først, men da er ofte tilbakemeldingene at det ikke kan anonymiseres fordi det vil redusere kvaliteten på datasettene. Okay greit sier jeg da, hva med å pseudonymisere? Det vil si at man bytter for eksempel «bil» med «362» og «mann» med «821» (...) Dette har også blitt diskutert, men er vanskelig å implementere (...) Dersom man ikke klarer å anonymisere eller pseudonymisere sitter man i en situasjon hvor man egentlig har persondata i klartekst, og da er det en del begrensninger for hva man kan gjøre. Dette blir plutselig veldig high risk (Advokat, PwC).

5.5. Oppsummering av funn

Oppsummert så har denne ideen har utviklet seg til et innovasjonsprosjekt kalt *Fraud Detection Platform*, som har mål om å validere et konsept for trygg datadeling mellom konkurrerende selskaper. Ambisjonene for prosjektet er å validere at det enkelte forsikringsselskap sine egne prediksjonsmodeller vil være mer presise og har mulighet til å avdekke flere svindeltilfeller dersom de trenes på større datasett. Større datasett krever som nevnt deling av data mellom konkurrerende selskaper, og løsningen må derfor tilfredsstillende finansregulatoriske krav i lys av konkurranse- og personvernlovgivningen. Analysen har avdekket at prosjektet har støtt på sine største utfordringer når det kommer til datasettets forklaringskraft, samarbeid og konkurranse i samspill, i tillegg til ivaretagelse av personvernet.

Fraud Detection Platform	
Fra problem til MVP	<ul style="list-style-type: none"> - Svindel anses som et samfunnsproblem. - Digitalisering som drivkraft for økt konkurranse, økte kundeforventninger og økning i svindeltilfeller. - Norske forsikringsselskaper besitter for små datamengder til å kunne predikere og detektere svindel godt nok alene.
Teknologi	<ul style="list-style-type: none"> - KI og Big Data anses som nyttige verktøy i kampen mot forsikringssvindel. - Felles stordataplattform på tvers av konkurrerende selskaper for å danne større datagrunnlag. - Felles Data Dictionary for å skape kvalitet i dataene. - Black Box utfordrer forklaringskraften.
Samkonkurranse	<ul style="list-style-type: none"> - Skille mellom konkurranse og samarbeid. - Ulik agenda, språk, ambisjoner og kultur blant deltakerne. - Næringsklynge som viktig uavhengig tredjepart. - Gjensidig tillit som en forutsetning for godt samarbeid.
Personvern	<ul style="list-style-type: none"> - Streng lovgivning begrenser datagrunnlaget. - Grunnlag for å behandle persondata fordi forsikringssvindel anses som kriminell aktivitet. - Bias hos både maskin og mennesker. - KI og Big Data utfordrer flere av de lovfestede personvernprinsippene uten at de nødvendigvis strider imot.

Tabell 4: Oppsummering av funn

6. DISKUSJON

I dette kapittelet skal vi koble funnene i analysen med litteratur som er lagt til grunn i det teoretiske rammeverket. Diskusjonen er fremstilt på et overordnet nivå hvor vi først diskuterer behovet for innovasjon i bransjen for å kunne bekjempe forsikringssvindel. Deretter hvorfor nye digitale innovasjoner, teknologi og samkonkurranse kan være gode bidragsytere i forbindelse med å bekjempe forsikringssvindel, samt tilknyttede utfordringer.

6.1. Innovasjon

Digitaliseringen kan beskrives som en av årsakene at forsikringsbransjen opplever større økonomiske tap i forbindelse med svindel (Viaene & Dedene, 2004, s. 313). Det kommer fram av analysen at forsikringsselskaper må møte økende kundeforventninger og konkurranse fra internasjonale aktører, samtidig som de gjør viktige vurderinger knyttet til risiko for svindel. Digitaliserte og automatiserte prosesser i bransjen legger til rette for at det kan være enklere å svindle – det blir dermed viktig å utvikle seg, men også minimere risiko på samme tid. Felles for alle aktører i forsikringsbransjen er at det eksisterer et ansvar og en plikt til å gjennomføre tiltak som skal avdekke finansiell kriminalitet. Finansiell kriminalitet har store konsekvenser for bransjen, forbrukerne og samfunnet, og det eksisterer et tydelig behov for nye løsninger som kan få bukt med finansiell kriminalitet.

Som følge av nye digitale innovasjoner og teknologier har internasjonale teknologiselskaper som for eksempel Google og Facebook bygget enorme datagrunnlag. Datagrunnlaget baserer seg i all hovedsak på bevegelsene vi foretar oss i digitale rom, og det er ikke lenger en fremmed tanke at denne type selskaper også kan begynne med finansielle tjenester. Med hensyn til at innovasjon legger grunnlaget for utvikling og økonomisk vekst (Njøs & Sjøtun, 2016, s. 4), ser forsikringsbransjen seg nødt til å utforske nye løsninger for å møte kundeforventninger, internasjonal konkurranse og redusere økonomiske tap i forbindelse med svindel. Det eksisterer dermed et behov for å innovere, også innen prediksjonsmodeller for svindel i forsikringsbransjen, som er relevant for denne oppgaven. Nye digitale innovasjoner som kunstig intelligens og Big Data fremheves i våre funn som teknologier forsikringsbransjen finner interessante når det kommer til å forbedre prediksjonsmodeller for svindel, og samtidig levere på andre forretningsmål. Likevel kommer det fram i analysen at norske forsikringsselskaper ikke sitter på tilfredsstillende mengder data til å kunne optimalisere prediksjonsmodeller for svindel på egen hånd, og at de dermed kan komme til kort i møte med økte forventninger og økt konkurranse. Dette bygger grunnlaget for ideen om

at norske forsikringsselskaper skal kunne dele data og trene sine prediksjonsmodeller på et felles datasett.

6.2. Teknologi – KI og Big Data

For at konkurrerende selskaper skal kunne samarbeide om å dele data, må det også etableres en felles teknisk løsning som tilrettelegger for korrekt deling. Basert på vår analyse er analytisk intelligens vurdert til den formen av kunstig intelligens som lønner seg med hensyn til å predikere og detektere svindel. Analytisk intelligens innebærer at teknologien lærer og tilpasser seg systematisk basert på den dataen den har tilgjengelig gjennom et gitt datasett. Analytisk intelligens vil behandle informasjon og trene på et gitt datasett for å foreta logiske resonnement (Huang & Rust, 2018, s. 158). Forslaget vi har studert i denne casen bygger på en løsning hvor hvert selskap deler relevante data, og at algoritmene dermed får et betydelig større grunnlag å trene sine prediksjonsmodeller på. For at den tekniske løsningen ikke skal være konkurransehemmende, er det viktig at hvert forsikringsselskap har utarbeidet egne prediksjonsmodeller for svindel basert på egne forutsetninger. På denne måten vil også de ulike selskapene ende opp med ulike resultater selv om de trener på det samme datasettet. Dersom forsikringsselskaper klarer å predikere større mengder svindelforsøk vil det naturligvis føre til lavere kostnader knyttet til forsikringsutbetalinger, noe som igjen kan påvirke prismodellene til hvert enkelt selskap. At prediksjonsmodellene til hvert selskap ikke deles og bygger på ulike forutsetninger er spesielt viktig med hensyn til konkurranselovgivning ettersom selskapene da ikke står i fare for ufrivillig prissamarbeid.

Konseptet om en stordataplattform for å bedre kunne predikere og detektere svindel legger også til rette for bruk av maskinlæring. Maskinlæring skal bidra til at modellen, her de forskjellige selskapenes prediksjonsmodeller for svindel, skal finne mønstre i datasettet som kan predikere fremtiden bedre. På denne måten kan prediksjonsmodellene optimalisere seg selv basert på tidligere erfaring og kontinuerlig læring. Selv om hypotesen for denne casen er at intelligente systemer og større datasett bedre skal kunne predikere svindelforsøk, er det ikke gitt at det stemmer. Svindel er et komplekst begrep, og svindlerne blir stadig mer kreative og finner andre svindelmetoder som forsikringsselskapene gjerne ikke har avdekket enda. En begrensning for løsningen er at modellene kun kan trene på data knyttet til svindeltilfeller de allerede kjenner til. Ved at konkurrerende selskaper velger å dele data med

hverandre, kan det kompensere for noen av disse ukjente tilfellene ettersom hvert enkelt selskap kan ha avdekket svindeltilfeller som et annet selskap ikke har gjort, og omvendt.

Maskinlæringsalgoritmer, som er aktuelt for denne casen, kan lære av et datasett på tre forskjellige måter jf. kapittel 2.3.4. I et datasett hvor en algoritme skal predikere et utfall og selv finne mønstre hvor det ikke eksisterer en fasit slik som ved svindel, blir det aktuelt å beskrive læringen i dette tilfellet som ikke-veiledet læring (Datatilsynet, 2018, s. 8). Vi kan likevel diskutere om løsningen ikke bare bygger på ikke-veiledet læring, men også semi-strukturert læring. Dette synspunktet bygger på at vi kan anse data som tilhører ett forsikringsselskap som kategoriserte og «kjente» for det aktuelle selskapet, men som ikke-kategorisert og «ukjent» for et annet selskap. Det blir dermed opp til hvert selskap å bygge prediksjonsmodeller på både kjente, men også ukjente data de tilegner seg gjennom den felles stordataplattdformen.

En av de største utfordringene knyttet til bruk av kunstig intelligente systemer er å forklare hvordan modellen har kommet fram til en spesifikk beslutning. Denne utfordringen blir heller ikke noe mindre ved at datamengden øker. Denne forklaringsproblematikken blir ofte omtalt som Black Box, ettersom man ikke ser hva som foregår inne i selve modellen (Datatilsynet, 2018, s. 12). Når man jobber med større mengder data, i tillegg til data som ikke er basert på egne forutsetninger i selskapet, øker utfordringene knyttet til dataenes forklaringskraft. Dersom et tilfelle blir avdekket som risiko for svindel, har eieren av tilfellet krav på å vite hvorfor. Det vil kreve mye av modellen å kunne forklare hvorfor den har tatt en gitt beslutning, og det vil også kreve mye av utenforstående å forstå bakgrunnen for beslutningen. Ett av grepene som er tatt for å øke forklaringskraften til dataene i plattformen, er å implementere en felles Data Dictionary for hvordan dataene i plattformen skal kategoriseres. En felles Data Dictionary kan bidra til at et risikotilfelle kan bli enklere å forklare ved at dataene er basert på et felles språk og til dels felles kategorisering mellom selskapene. Likevel vil problematikken knyttet til ukjente og kjente svindeltilfeller og dens kategorisering fortsatt være problematisk knyttet til forklaringskraft.

6.3. Samkonkurransen

Som nevnt foreligger det et stort behov i forsikringsbransjen for å bekjempe forsikringssvindel og konsekvensene dette medfølger. Samtidig blir det også nevnt at norske forsikringsselskaper ikke har mulighet til å få bukt med dette alene basert på de enkelte

selskapenes individuelle størrelse. Forsikringsselskaper tilhører en bransje som i lang tid har blitt ansett som både tradisjonell og byråkratisk, men i likhet med andre bransjer opplever også forsikringsbransjen endringer og må tilpasse seg deretter. Det har derfor blitt mer vanlig å se utenfor egne organisatoriske grenser for å sikre utvikling og fleksibilitet i selskapenes forretningsmodeller. Analysen fremhever at tradisjonelle arbeidsmetoder og lukket innovasjon ikke lenger er bærekraftig, noe som underbygger teorien til Chesbrough (2003) om åpen innovasjon jf. kapittel 2.2.1. Det er avdekket at det blir stadig viktigere å samarbeide med andre for å legge et godt grunnlag for innovasjon på tvers av bransjen. At selskaper i bransjen har skiftet fokus og ser etter nye arbeidsmetoder- og strategier for å optimalisere egen forretning, kan ses i sammenheng med åpen innovasjon og samkonkurranse.

Som følge av et større fokus på at samarbeid og deling av kompetanse og ressurser har det vært en fremvekst av næringsklynger på tvers av flere bransjer. Deltakelse i en næringsklynge kan være en form for åpen innovasjonsaktivitet ved at selskaper kan dra nytte av kompetente ressurser utenfor egne organisatoriske grenser, samt gjennomføre inngående og utgående aktiviteter som til sist gagnar egen forretningsmodell. Samtidig som deltakelse i en næringsklynge gjør at man kan dele på kompetanse og ressurser, tilrettelegger det også for prosjektbasert samkonkurranse. Det stilles strenge finansregulatoriske krav til samarbeid mellom konkurrerende selskaper, i tillegg til at det er viktig at prosjektet har klare retningslinjer for hva som skal samarbeides om, men også på hvilke områder det fortsatt skal konkurreres. Samkonkurranse som bidrar til å hemme fri konkurranse er ulovlig, og desto viktigere er det at den tekniske løsningen et slikt konsept bygger på ikke strider med disse reglene. Et konsept for datadeling på tvers av forsikringsselskaper kan fort havne i strid med lovverket dersom det ikke er godt nok regulert. En måte å luke unna denne utfordringen er at deltakende forsikringsselskaper ikke deler interne prediksjonsmodeller, forutsetninger eller har innsyn i hverandre sine data, slik som det er lagt til grunn for caset i denne oppgaven. På denne måten kan selskapene samarbeide om å dele data uten at resultatet vil ha en konkurransehemmende effekt og havne i strid med lovverket.

Samtidig som et konsept for datadeling på tvers av ulike selskaper kan ha utfordringer knyttet til konkurranselovgivning, kan det også ha utfordringer knyttet til gjennomføring av selve samarbeidet. Ved samkonkurranse foreligger det et samarbeid mellom i utgangspunktet konkurrerende selskaper, og en ser ofte at de ulike selskapene har med seg ulik kompetanse, ressurser, språk og intensjoner i samarbeidet. Denne type ulikheter kan bidra til skjevhet i

samarbeidet, og i verste fall føre til mistillit og konflikter mellom partene. Gjennom arbeidet med denne studien har vi erfart viktigheten av at et slik samarbeid har et felles mål og felles ambisjoner. Dersom det ikke eksisterer et felles utgangspunkt for hvorfor et gitt prosjekt gjennomføres, er det lett for å miste momentum og engasjement, noe som kan true samarbeidsevnen mellom selskapene jf. kapittel 2.2.2. For et innovasjonsprosjekt som baserer seg på samkonkurranse blir det derfor viktig å avklare et felles mål og en felles strategi for prosjektet som skal gjennomføres. Det er viktig at deltakerne i samarbeidet ser en verdi av å delta, og bygger samarbeidet på felles interesser, tillit og god kommunikasjon.

6.4. Personvern og GDPR

I et innovasjonsprosjekt hvor målet er å kombinere data fra konkurrerende selskaper i forsikringsbransjen for å danne et større datagrunnlag, vil det naturligvis føre til utfordringer knyttet til personvern. Vårt digitale samfunn slik vi kjenner det i dag bygger på store mengder data, også persondata. Personvernet er en av våre grunnleggende rettigheter i et demokrati, men nye teknologier og digitale innovasjoner har gjerne gjort det vanskeligere for oss å ha oversikt over alle dataene våre. Det er likevel viktig at selskaper som behandler persondata, opptrer i samsvar med de lovfestede personvernprinsippene jf. kapittel 2.4. Det er avdekket at Datatilsynet ikke fungerer som et støtteapparat for selskaper som driver med digitale innovasjoner og nye teknologier som bygger på persondata. De fører heller tilsyn i etterkant, og selskapene vil dermed enten få et klarsignal om at de kan bygge videre på den aktuelle løsningen eller ikke. Tilsynet kan i verste fall føre til bøter av signifikante størrelser dersom det ikke tilfredsstillende kravene til personvernprinsippene på en god nok måte. For forsikringsselskaper som skal dele data i en felles stordataplattform blir det derfor viktig å utarbeide en løsning som etterlever de juridiske kravene, men som samtidig ikke gir en redusert funksjonalitet. Flere er derfor enige om at dersom man skal behandle persondata i et innovasjonsprosjekt, er det viktig å verifisere at løsningen fungerer basert på et datagrunnlag som er så lav-terskel som overhodet mulig.

Ved bruk av kunstig intelligens og Big Data i løsninger som behandler persondata, er det spesielt ett av personvernprinsippene som blir spesielt viktig; prinsippet om at persondata skal behandles på en lovlig, rettferdig og åpen måte (Lovdata, 2018). Dette innebærer blant annet at det må eksistere et behandlingsgrunnlag for at et selskap skal kunne oppbevare persondata. En teknisk løsning basert på deling av data mellom forsikringsselskaper for å bekjempe et samfunnsproblem kan identifiseres som et gyldig grunnlag for behandling av persondata.

Vurderingen av gyldigheten er gjort på bakgrunn av at det eksisterer et problem som regnes som en kriminell handling, i tillegg til at det foreligger en egeninteresse for å bekjempe svindlerne fra samfunnet for øvrig.

Personvernprinsippet legger også vekt på at persondata skal behandles rettferdig med hensyn til individets interesser og forventninger (Lovdata, 2018). Ved bruk av KI og Big Data i beslutningsmodeller, som for eksempel en prediksjonsmodell for svindel, kan det oppstå utfordringer knyttet til diskriminering og/eller forskjellsbehandling. Algoritmene vil automatisk trene på data som er merket som risiko for svindel, og dersom flere kjennetegn knyttet til en spesifikk kundegruppe er risikomerket kan det føre til diskriminering og/eller forskjellsbehandling av nettopp denne gruppen. Et relevant eksempel er den unge, mannlige bilføreren på 18 år med en god, gammel «rånebil»; denne gruppen er gjerne mer utsatt for å havne i ulykker og/eller ha skade på bil, og kan gjerne dermed ha større risiko sammenlignet med andre kunder. Denne problematikken kan knyttes opp mot bias som enten rettferdig eller urettferdig, driver med systematisk diskriminering av individer eller grupper basert på gitte kjennetegn jf. kapittel 2.3.6. Selv om bias er et omdiskutert tema i sammenheng med kunstig intelligens og trening av algoritmer, er det også slik at mennesker kan diskriminere systematisk. Ved bruk av kunstig intelligens og Big Data i kampen mot forsikringssvindel er det derfor viktig at hvert enkelt forsikringsselskap utarbeider rutiner og forbedrer systemer for å handle så etisk og korrekt som mulig med hensyn til biased algoritmer.

Det siste punktet i personvernprinsippet som omhandler gjennomsiktighet vil kunne være vanskelig å opprettholde på bakgrunn av Black Box-problematikken. Prediksjonsmodeller som baserer seg på kunstig intelligente systemer og Big Data utfordrer systemets forklaringskraft, og beslutningene modellen tar kan derfor være vanskelig å begrunne. Som nevnt vil man gjerne trenge dyp innsikt i hvordan selve prediksjonsmodellen fungerer for å i det hele tatt kunne forklare hvorfor en beslutning er tatt. Dette er kompetanse de aller fleste enkeltindivider, her relevant forsikringskunder, ikke har. Det kan derfor diskuteres om dette vil være i strid med personvernprinsippet om gjennomsiktighet ettersom behandlingen av data skal være oversiktlig og forutsigbar for den det gjelder (Datatilsynet, 2019). Black Box-problematikken kan dermed bidra til at forsikringsselskaper kan ha utfordringer med å være åpne og transparente i dialog med sine kunder.

Et annet viktig personvernprinsipp er at mengden data skal holdes til et absolutt minimum, og det skal heller ikke oppbevares data som ikke er relevante (Datatilsynet, 2019). Det er avdekket at flere oppfatter personvernlovgivningen som uklar, og det er derfor vanskelig for forsikringsselskapene å vite nøyaktig hvor grensen går ved oppbevaring av persondata. Dette fører til at forsikringsselskaper gjerne er ekstra forsiktig ved utvikling av nye løsninger, spesielt med hensyn til hvilke data som inkluderes og deles i en slik stordataplattform. Forsikringsbransjen anses samtidig som en risikoavers bransje hvor selskaper vanligvis ikke foretar handlinger preget av mye risiko. I sammenheng med at Datatilsynet ikke fungerer som en sparringspartner, men heller som et overvåkende organ i etterkant av utarbeidet løsningsforslag, gjør det at bransjens utvikling potensielt kan hemmes. Gjennom denne studien er det erfart at selskapene heller velger å være ekstra forsiktige, enn å eksperimentere med noe som kan gi negative konsekvenser senere.

6.5. utfordringer knyttet til skalering

Innovasjonsprosjekter som bygger på kunstig intelligens, Big Data og samkonkurranse for å bekjempe forsikringssvindel er foreløpig ikke implementert i forsikringsbransjen. Det er imidlertid avdekket i denne studien at prosjekter som bygger på denne type teknologi og strategi potensielt kan skaleres til en nasjonal plattform for svindeldeteksjon i fremtiden. Dersom et slikt prosjekt skal kunne skaleres er det naturlig at flere deltakere involveres og at forsikringsproduktene de ønsker å avdekke svindel for, utvides.

Hvis et prosjekt velger å enten inkludere flere deltakere, eller utvide målet med prosjektet, kan det både gi store muligheter, men også nye utfordringer. Ved flere deltakere blir det desto viktigere å administrere samarbeidet med hensyn til kommunikasjon, engasjement og retning, men også ved ivaretagelse av motivasjonen og verdien av samarbeidet. Studien har også avdekket behovet for at en bestemt aktør leder samarbeidet for å sikre momentum og effektivitet – en prosjektleder blir desto viktigere dersom skalering i fremtiden blir aktuelt. Innen forsikringsbransjen kan relevante aktører for eksempel være store finansregulatoriske selskaper som Finans Norge eller Finanstilsynet, fungere som prosjektledende, gjensidige tredjeparter. Finans Norge eier og drifter allerede lignende plattformer, mens det er forsikringsselskapene som bruker og betaler for vedlikehold.

En kan også utvide et slikt innovasjonsprosjekt ved å for eksempel gjøre datasettet bredere; enten ved å hente data fra flere aktører eller type produkter. Kunstig intelligens og Big Data

som benytter maskinlæring for å predikere mønstre kan bli enda bedre dersom systemet får ytterligere større datamengder å trene på. Likevel kan det bli utfordrende å involvere flere aktører ettersom det eksisterer et klart behov for en felles Data Dictionary; altså språk, for å få en slik løsning til å fungere optimalt. Flere deltakere (aktører) bidrar til økt administrasjonsarbeid for å sikre kvalitet på dataene som deles i plattformen, og det kan potensielt bli vanskeligere å bli enige om et felles språk som passer for alle.

Å utvide datasettet kan også bidra til utfordringer i forbindelse med Norges strenge personvernlovgivning som allerede er en problematikk i arbeidet med KI og Big Data. Ved å utvide til flere forsikringsprodukter for å kunne detektere svindel på flere områder, gjør dette også at selskapene må dele flere, og gjerne mer sensitive data med hverandre.

Produktområder innenfor skade, som for eksempel bil, hus og innbo, kan regnes som mindre personsensitive sammenlignet med for eksempel personforsikringer. En kan derfor diskutere om det i det hele tatt er reelt å for eksempel utvide produktområdet til personforsikringer hvor det er helt nødvendig å dele personsensitive data.

Vi opplever det som avgjørende for å sikre innovasjon i forsikringsbransjen at det eksisterer en strategi og klare retningslinjer for bruk av KI og Big Data i forbindelse med håndtering av persondata. Den regulatoriske sandkassen er et initiativ av Datatilsynet som legger til rette for at innovasjonsprosjekter skal kunne teste løsninger uten å være bekymret for negative konsekvenser. Likevel er det begrenset tilgang til sandkassen, og alle prosjekter som søker om å bli med får nødvendigvis ikke tilgang når de trenger det. Dette kan potensielt bidra til å hemme innovasjon i denne bransjen som allerede beskriver seg selv som risikoavers.

7. KONKLUSJON

I denne masteroppgaven har vi valgt å studere bruken av KI og Big Data for å bekjempe svindel i forsikringsbransjen. Vi har i den forbindelse ønsket å avdekke hvordan forsikringsselskaper kan ta i bruk disse teknologiene og samarbeide for å bekjempe forsikringssvindel. Samtidig ønsket vi å avdekke om det eksisterte noen utfordringer knyttet å utvikle en slik felles løsning. På bakgrunn av dette utformet vi følgende forskningsspørsmål:

Hvordan kan forsikringsbransjen i Norge skape en felles stordataplattform for å bekjempe forsikringssvindel?

Hva er de største utfordringene i utviklingen av en felles stordataplattform som potensielt kan forbedre arbeidet med å detektere og predikere forsikringssvindel?

Oppgavens første forskningsspørsmål ønsker å redegjøre for «*hvordan forsikringsbransjen i Norge kan skape en felles stordataplattform for å bekjempe forsikringssvindel*». I denne forbindelse tok vi utgangspunkt i et innovasjonsprosjekt i regi av NCE Finance Innovation som hadde ambisjoner om å validere et konsept for trygg datadeling mellom konkurrerende selskaper for å bekjempe et felles problem. Innovasjonsprosjektet bygger på en idé om at norske forsikringsselskaper vil være nødt til å dele data på tvers av organisatoriske grenser, for å forbedre egne prediksjonsmodeller for svindel. Ved å dele data og på denne måten etablere et større datagrunnlag, legger det også til rette for at norske forsikringsselskaper i større grad blir konkurransedyktige i møte med internasjonale aktører. Hypotesen er at man ved hjelp av kunstig intelligens og Big Data skal kunne trene prediksjonsmodeller på en felles stordataplattform for i større grad kunne avdekke risiko for svindel. Forsikringsbransjen i Norge må derfor benytte seg av samkonkurransen og kunstig intelligente systemer for å bekjempe dette samfunnsproblemet.

Vårt andre forskningsspørsmål bygger videre på forskningsspørsmål én, hvor det er avklart et behov for samkonkurransen og kunstig intelligente systemet for å predikere og detektere svindel. Som følge av KI, Big Data og samkonkurransen dukker det opp utfordringer spesielt knyttet til selve teknologien, dimensjoner ved samkonkurransen og juridiske aspekter. Vår studie avdekker vesentlige utfordringer relatert til;

- 1) *Teknologi* → Kunstig intelligens og maskinlæring kan bidra til å gjøre arbeidet med å detektere og predikere forsikringssvindel enklere, og på den måten også løse et omfattende samfunnsproblem. Maskinlæringsalgoritmer som benyttes i eksempelet denne oppgaven baserer seg på, lærer og optimaliserer seg på tidligere erfaring. Den største utfordringen til kunstig intelligente systemer og maskinlæring, er at det kan være vanskelig å gi en begrunnelse for en beslutning som er tatt av maskinen. Siden teknologien og modellene behandler persondata kan løsningen havne i strid med personvernprinsippet om åpenhet. Dette blir sett i sammenheng med Black Box som beskriver forklaringskraften til kunstig intelligente systemer som et problem. Forsikringsselskaper har krav om å opptre i henhold til personvernlovgivning ved behandling av denne type data, og spesielt dersom en kunde krever en forklaring på hvorfor den er blitt flagget som risiko for svindel. Black Box-problematikken kan potensielt gjøre det vanskelig for skadebehandlere eller etterforskere å gi en forklaring for beslutningen. For å forbedre forklaringskraften til modellene har den felles Data Dictionary vært avgjørende for å gjøre det enklere å forklare en prosess i modellen selv om den fortsatt ikke gir full innsikt. En annen begrensning ved løsningen er at prediksjonsmodellene kun trener på et datasett som består av svindeltilfeller selskapene allerede kjenner til. Ved at konkurrerende selskaper velger å dele data med hverandre, kan det kompensere for noen av disse ukjente tilfellene ettersom hvert enkelt selskap kan ha avdekket svindeltilfeller som et annet selskap ikke har gjort, og omvendt.

- 2) *Samkonkurranse* → Ved samkonkurranse i en næringsklynge, slik som ved dette prosjektet, er det viktig å skille mellom samarbeid og konkurranse. Selskapene i et slik samarbeid får ikke samarbeide om løsninger som potensielt kan gi konkurransehemmende effekt. For forsikringsselskaper kan det for eksempel være prissamarbeid. Utfordringen for et slikt prosjekt blir derfor å optimalisere en felles stordataplatteform, uten at resultatet ved å ta i bruk en slik plattform har direkte effekt på prismodellene til hvert enkelt selskap. Viktig for dette prosjektet blir at hvert selskap har respektive prediksjonsmodeller for svindel som baseres på egne forutsetninger fra hvert enkelt selskap. På denne måten forhindrer man at selskapene kan ende opp med like resultater. I tillegg til konkurranselovgivning, vil samkonkurranse også oppleve utfordringer knyttet til holdninger, kultur og språk. Forskjellige selskaper har ofte med seg ulike kompetanse, ressurser, språk og

intensjoner i et samarbeid, og det blir dermed viktig å avklare en felles retning for samarbeidet. Samtidig er det også viktig å administrere arbeidet slik at det blir mest mulig effektivt. Ved slike samarbeid som baserer seg på samkonkurranse viser det seg spesielt viktig å ha en uavhengig tredjepart som prosjektleder. Felles retning, ambisjoner, språk og ledelse blir desto viktigere dersom prosjektet skaleres til å inkludere flere selskaper.

- 3) *Personvern* → En stordataplattform som deler persondata på tvers av forskjellige forsikringsselskaper vil naturligvis oppleve en del problematikk knyttet til håndteringen av disse dataene. De aller største utfordringene knyttet til kunstig intelligente systemer og personvern/GDPR er potensiale for diskriminering og/eller forskjellsbehandling, gjennomsiktighet og dataminimering. Ettersom kunstig intelligente systemer og maskinlæring lærer og predikerer basert på tidligere erfaring, er det fare for systematisk diskriminering, også kalt bias. Bias oppstår som følge av at enkeltpersoner eller grupper diskrimineres eller forskjellsbehandles basert på enkelte kjennetegn; for eksempel en spesielt kostbar bil. Gjennomsiktighet ses i sammenheng med åpenhet, og Black Box problematikken som også er et problem knyttet til selve den tekniske løsningen. Forsikringsselskapene kan derfor ha utfordringer med å være åpne og transparente i dialog med sine kunder på bakgrunn av den systemets lave forklaringskraft. Dataminimering handler om at selskaper skal begrense mengden data de besitter, og dette kan skape utfordringer med hensyn til å optimalisere en potensiell løsning på et gitt problem som for eksempel forsikringssvindel. Dette er spesielt relevant hvis prosjektet ønsker å skalere i form av å utvide produktgrunnlaget de detekterer svindel for.

7.1. Faglige implikasjoner

Denne studien belyser en stadig mer relevant tematikk i norsk forsikringsbransje, nemlig utvikling og anvendelse av nye digitale teknologier på tvers av konkurrerende selskaper. Caset som studeres utforsker bruk av kunstig intelligens, Big Data og samkonkurranse for å bekjempe forsikringssvindel – samspillet av disse begrepene er noe vi anser som nytt i bransjen, og som per i dag ikke har en opplevd effekt. Oppgavens formål er dermed å gi leseren økt kunnskap innen et forholdsvis nytt felt, fremme teknologier som blir viktige for fremtiden og hvilke utfordringer dette kan medføre. Samkonkurranse og ny teknologi står sentralt i studien, og det er avdekket flere utfordringer som selskaper gjerne ikke er vant til å

håndtere og lovverk som ikke er tilpasset denne type innovasjon. Vi anser derfor studiens som relevant for å belyse viktige endringer i bransjen med hensyn til behovet for innovasjon.

7.2. Praktiske implikasjoner

Denne masteroppgaven fremhever at streng personvern- og konkurranselovregulering kan hemme forsikringsbransjens muligheter og evne til å innovere. Dette kan ses i sammenheng med at det per i dag er begrenset bruk og effekt av kunstig intelligens, Big Data og samkonkurranse for å bekjempe forsikringssvindel. Vi anser det som en stor fordel for forsikringsbransjen i Norge dersom det implementeres et mer tilgjengelig regulatorisk rammeverk som tilrettelegger for økt innovasjon. Dagens retningslinjer og lovverk kan skape utfordringer med hensyn til selskapers muligheter og evne til å innovere, noe som også har preget arbeidet mot en felles løsning for å bekjempe forsikringssvindel. Retningslinjer og lovverk som tilrettelegger for innovasjon kan bidra til at selskaper enklere kan etterleve kundenes forventninger i lys av digitalisering, økt internasjonal konkurranse og redusere økonomiske tap i forbindelse med svindel.

7.3. Metodiske begrensninger og forslag til videre forskning

Arbeidet med studien startet med at vi leste om innovasjonsprosjektet i avisen, og det vekket en umiddelbar interesse om å potensielt studere prosjektet nærmere. Etter hvert initierte vi en samtale med en av de deltakende aktørene i prosjektet for å få et bedre innblikk, samt drøfte potensielt tema og forskningsspørsmål for vår studie. Samtalen gjorde oss oppmerksom på prosjektets kompleksitet, samtidig som vi ble opplyst om at det ikke finnes lignende prosjekter i Norge til nå. For å få en bedre forståelse rundt konseptet brukte vi derfor mye tid på å tilegne oss kunnskap rundt de teknologiske løsningene som blir tatt i bruk i innovasjonsprosjektet. I tillegg til dette leste vi oss opp på lignende konsept utført i utlandet, som viste seg å ha gode resultater. På bakgrunn av dette har vi selv formet innholdet i oppgaven og koblet inn teori som vi mener er relevant for å kunne svare på oppgavens tema og forskningsspørsmål.

En kjent begrensning når det kommer til studier som dette er at studien kun går over ett semester, noe som kan utfordre studiens omfang. Innovasjonsprosjektet som studeres i denne studien er foreløpig kun i en oppstartsfase, og det kunne vært givende å følge prosjektet videre over lengre tid for å se den virkelige effekten av løsningen. På bakgrunn av at vi har

studert et pågående innovasjonsprosjekt, har vi etter beste evne forsøkt å spre intervjuene våre så mye som mulig for å innhente informasjon på ulike stadier i prosjektet. Dette for å se om prosjektet ville få resultater tidligere enn antatt, og på denne måten prøve å minske effekten av oppgavens tidsbegrensing.

Gjennom arbeidet med studien har vi lært mye nytt, og oppdaget mange interessante og spennende funn innenfor et felt som var relativt ukjent for oss fra før. Til videre forskning finner vi det interessant å studere innovasjonsprosjektet over lengre tid og på denne måten kunne følge utviklingen av prosjektet gjennom flere faser, og effekten av potensielle resultat. Gitt at prosjektet blir en suksess og blir implementert som nasjonal plattform, kan det på sikt være enklere å belyse alle utfordringene som har oppstått underveis og hvordan disse best kan håndteres i fremtiden for lignende prosjekter.

8. REFERANSELISTE

Ames, H., Holte, H.H., Said, M.A., Heintz, M., & Berg, R. (2018). Tiltak mot trygde- og forsikringssvindel. *Folkehelseinstituttet*, 1-48

https://fhi.brage.unit.no/fhi-xmlui/bitstream/handle/11250/2507497/Ames_2018_Til.pdf?sequence=4

Amit, R. & Zott, C. (2012) Creating Value Through Business Model Innovation. *MIT Sloan Management Review*. Spring,. 1-9.

Andersen, E. & Sannes, R., (2017). Hva er digitalisering? *Magma*, 18-24.

Andersen, E. & Sannes, R. (2018). Er du klar for digitalisering? *Praktisk økonomi og finans* (3), 196-213.

Bergsjø, L. & Bergsjø, H. (2019). Digital etikk: big data, algoritmer og kunstig intelligens. Universitetsforlaget.

Bengtsson, M. & Kock, S. (2000). «Coopetition» in Business Networks – to Cooperate and Compete Simultaneously. *Industrial Marketing Management*, Vol. 29, Issue 5, 411-426.

Bouncken, R., Gast, J., Kraus, S. & Bogers, M. (2015). Coopetition: a systematic review, synthesis, and future research directions. *Springer-Verlag*, 577-601.

Butterworth, M. (2018). The ICO and artificial intelligence: The role of fairness in the GDPR framework. *Computer Law & Security Review: The International Journal of Technology Law and Practice*, 34(2), 257-268.

Cambridge. (2021). Data. *Cambridge Dictionary*.

<https://dictionary.cambridge.org/dictionary/english/data>

Chesbrough, H. (2003). *Open Innovation: The New Imperative for Creating and Profiting from Technology*. Harvard Business School Press.

Chesbrough, H. (2005). *Open Innovation: A New Paradigm for Understanding Industrial Innovation*. Oxford University Press.

Chesbrough, H. (2012). Open Innovation: Where We've Been and Where We're Going. *Research Technology Management*, 55: 4, 20-27.

Christensen, C., Bartman, T., & Bever, D. (2016). The Hard Truth About Business Model Innovation. *MIT Sloan Management Review*. Fall, vol. 58, 31-40.

Cortis, D., Debattista, J., Debono, J. & Farrell, M. (2019). InsurTech. Lynn, T., Mooney, J., Rosati, P. & Cummins, M. *Disrupting Finance: FinTech and Strategy in the 21st Century*.

Dahlum, S. (2020). Validitet. I *Store Norske Leksikon*.

<https://snl.no/validitet>

Datatilsynet. (2013). Big Data – personvernprinsipper under press. *Datatilsynet*, 1-59.

<https://www.datatilsynet.no/regelverk-og-verktoy/rapporter-og-utredninger/big-data/>

Datatilsynet. (2018). Kunstig intelligens og personvern. *Datatilsynet*, 1-29.

<https://www.datatilsynet.no/regelverk-og-verktoy/rapporter-og-utredninger/kunstig-intelligens/>

Datatilsynet. (2019). Personvernprinsippene. *Datatilsynet*.

<https://www.datatilsynet.no/rettigheter-og-plikter/personvernprinsippene/>

Døving, I. (2018). Forsikringshistorie. I *Store Norske Leksikon*.

<https://snl.no/forsikringshistorie>

Døving, I. & Loen, R. (2018) Forsikring. I *Store Norske Leksikon*.

<https://snl.no/forsikring>

Goodwin, M. (2020). *AI: Myten om maskinene*. Humanist Forlag.

Gripsrud, G., Olsson, U. H., & Silkoset, R. (2016). *Metode og Dataanalyse*. 3 utg. Oslo: Cappelen Damm

Easterby-Smith, M., Thorpe, R., Jackson, P. R., & Jaspersen, L. J. (2018). *Management & Business Research*. 6 utg. London: SAGE Publications Ltd.

Enkel, E., Gassmann, O. & Chesbrough, H. (2009). Open R&D and Open Innovation: exploring the phenomenon. *Blackwell Publishing, vol 39, no. 4*, 311-316.

Fagerberg, J. (2003). Schumpeter and the revival of evolutionary economics: an appraisal of the literature. *Springer-Verlag, vol.13*, 125-159.

Finance Innovation. (Uten år). About.

<https://financeinnovation.no/about>

Finance Innovation. (Uten år). Fintech Hub.

<https://financeinnovation.no/fintech-hub>

Finans Norge. (2020). Markedsandeler.

<https://www.finansnorge.no/statistikk/skadeforsikring/markedsandeler/>

Fichman, R. G., Dos Santos, B. L., & Zheng, Z. (2014). Digital Innovation as a Fundamental and Powerful Concept in the Information Systems Curriculum. *MIS Quarterly*, 38(2), 329–353.

Fremtind Forsikring. (Uten år). Om oss.

<https://www.fremtind.no/om-oss/>

Fremtind Årsrapport. (2019). Årsrapport Fremtind.

<https://www.fremtind.no/siteassets/dokumenter/fremtind-forsikring-as-aarsrapport-2019.pdf>

Frende Forsikring. (Uten år). Om oss.

<https://www.frende.no/om-oss/>

Frende Årsberetning. (2020). Styrets årsberetning for Frende Skadeforsikring AS.

<https://www.frende.no/om-oss/finansiell-informasjon/frende-skadeforsikring/>

Hamel, G., Doz, Y. & Prahalad, C. K. (1989). Collaborate with Your Competitors – and Win. *Harvard Business Review*, January-February, 133-140.

Holtan, J. (2018). Kunstig intelligens i forsikring.

<https://www.forsikringsforeningen.no/wp-content/uploads/2018/10/Jon-Holtan.pdf>

Huang, M. & Rust, R. (2018). Artificial Intelligence in Service. *SAGE Publications*, vol 21(2) 155-172.

IESE. (2018, 24.april). *10 Ways Artificial Intelligence Is Transforming Management*.

<https://www.iese.edu/stories/10-ways-artificial-intelligence-is-transforming-management/>

Innovasjon Norge. (Uten år). Norwegian Innovation Clusters.

https://www.innovasjon norge.no/no/subsites/forside/Om_NIC/

Innovasjon Norge (2019). Om klyngeprogrammet.

https://www.innovasjon norge.no/no/subsites/forside/om_klyngeprogrammet/

Innovasjon Norge. (Uten år). NCE Finance Innovation.

https://www.innovasjon norge.no/no/subsites/forside/om_klyngeprogrammet/kart/nce/nce-finance-innovation/

Iriondo, R. (2021). Machine Learning (ML) vs. Artificial Intelligence (AI) – Crucial Differences. *Towards AI*.

Jacobsen, D. I., (2015). *Hvordan gjennomføre undersøkelser?* 3 utg. Oslo: Cappelen Damm

Kim, K. (2018). Coopetition: Complexity of cooperation and competition in dyadic and triadic relationships. *Organizational Dynamics*, 100683 (49), 1-6.

Kingston, J. (2017). Using artificial intelligence to support compliance with the general data protection regulation. *Springer Science + Business Media*, vol 25, 429-443.

Konkurranseloven. (2004). Lov om konkurranse mellom foretak og kontroll med

foretakssammenslutninger (LOV-2019-06-14-23).

<https://lovdata.no/dokument/NL/lov/2004-03-05-12>

Luo, Y. (2007). A coopetition perspective of global competition. *Journal of World Business* 42, 129-144.

Marr, B. (2018). The Key Definitions of Artificial Intelligence (AI) That Explain Its Importance. *Forbes Magazine*.

<https://www.forbes.com/sites/bernardmarr/2018/02/14/the-key-definitions-of-artificial-intelligence-ai-that-explain-its-importance/?sh=63de62964f5d>

McCarthy, C., Carleton, P., Krumpholz, E. & Chow, M. (2018). Accelerating Innovation Through Coopetition. *Wolters Kluwer Health Inc. Vol. 42, no. 1*, 26-34.

Mending, J., Decker, G., Hull, R., Reijers, H., & Weber, I. (2018). How do Machine Learning, Robotic Process Automation, and Blockchains affect the Human Factor in Business Process Management? *Communications of the AIS, vol. 43*, 1-23.

Minelli, M., Chambers, M., & Dhiraj, A. (2013). *Big Data, Big Analytics: Emerging business intelligence and analytic trends for today's business*. John Wiley & Sons Inc.

Naqa, I. & Murphy, M. (2015). *What Is Machine Learning? Springer International Publishing Switzerland*.

Nesse, S. (2018). Hvordan sikre innovasjon ved å samarbeide med en konkurrent? *MAGMA, mai 2018*, 61-70.

Njøs, R. & Sjøtun, S. (2016). *Innovasjon: Ei kortfatta innføring i sentrale begrep og tenkemåter*. Høgskolen i Bergen, Senter for nyskaping.

Osmundsen, K., Iden, J. & Bygstad, B. (2018). Hva er digitalisering, digital innovasjon og digital transformasjon. *NOKOBIT, 26(1)*, 1-15.

Osmundsen, K., Iden, J. & Bygstad, B. (2018). Digital Transformation: drivers, success factors, and implications. *MCIS 2018 Proceedings*, 1-16.

Personopplysningsloven. (2018). Lov om behandling av personopplysninger (LOV-2018-12-20-116). Lovdata.

<https://lovdata.no/dokument/NL/lov/2018-06-15-38?q=personopplysningsloven>

Porter, M. (1998). Clusters and the New Economics of Competition. *Harvard Business Review*. Vol. November-December.

PwC. (Uten år). GDPR – personvern og informasjonssikkerhet.

<https://www.pwc.no/no/tjenester/forretningsjus/personvern.html>

Regjeringen. (2008-2009). Stortingsmelding, nr.7. Et nyskapende og bærekraftig Norge.

<https://www.regjeringen.no/no/dokumenter/stmeld-nr-7-2008-2009-/id538010/sec2>

Regjeringen. (2014). Fremtidsrettet digitalt personvern.

<https://www.regjeringen.no/no/aktuelt/fremtidsrettet-digitalt-personvern/id750005/>

Regjeringen. (2020). Nasjonal strategi for kunstig intelligens. *Kommunal- og moderniseringsdepartementet*, 1-67.

<https://www.regjeringen.no/no/dokumenter/nasjonal-strategi-for-kunstig-intelligens/id2685594/>

Sabanovic, S., Milojevic, S., & Kaur, J. (2012). John McCarthy history. *IEEE Robotics & Automation Magazine*, vol 19, no. 4, 99-106.

Schumpeter, J. (1934). *The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest, and the Business Cycle*. Harvard University Press.

Silberg, J. & Manyika, J. (2019). Notes from the AI frontier: Tackling bias in AI (and humans). *McKinsey Global Institute*.

Staff, A. (2015). Bias. De nasjonale forskningsetiske komiteene.
<https://www.forskningsetikk.no/ressurser/fbib/uavhengighet/bias/>

Sørebø, Ø., Fredriksen, J., Simnica, F. & Mollestad, H. (2020). EUs personvernforordning (GDPR). *Universitetsforlaget AS*, 240-256.

Theodoridis, S. (2015). *Machine Learning : A Bayesian and Optimization Perspective*, Elsevier Science & Technology. *ProQuest Ebook Central*,

Tidd, J. & Bessant, J. (2018). *Managing Innovation: Integrating technological, market and organizational change* (6.) Great Britain: TJ International Ltd.

Tidemann, A. (2020). Kunstig intelligens. I *Store Norske Leksikon*.
https://snl.no/kunstig_intelligens

Tryg Forsikring. (Uten år). Fakta om Tryg.
<https://www.tryg.no/om-tryg/fakta-om-tryg/index.html>

Tryg Forsikring Årsrapport. (2020). Annual Report.
<https://tryg.com/en/downloads-2020>

Usman, M. & Vanhaverbeke, W. (2017). How start-ups successfully organize and manage open innovation with large companies. *Esmerald Publishing*, vol. 20, no.1, 171-186

Viaene, S. & Dedene, G. (2004). Insurance Fraud: Issues and Challenges. *Blackwell Publishing Ltd*, Vol. 29, no. 2, 313-333.

Wasenden, O. (2020). Digitalt personvern – kunnskap, bekymring og atferd. *MAGMA* (2), 64-73.

Webstep. (Uten år). Om oss.
<https://www.webstep.no/om-oss/#fakta>

West, J. & Bhattacharya, M. (2016). Intelligent financial fraud detection: A comprehensive review. *Computers & Security*, vol. 57, 47-66.

Yadav, P. (2018). Machine Learning vs. Deep Learning. *Medium*.

<https://medium.com/@mail2princeyadav/machine-learning-vs-deep-learning-b5c5a4fc5c>

Yin, R. K. (2014) *Case Study Research and Applications*. 5 utg. London: SAGE Publications

Yin, R. K. (2018). *Case Study Research and Applications*. 6 utg. London: SAGE Publications

Yoo, Y., Lyytinen, K., J., Boland, R. & Berente, N. (2010). The Next Wave of Digital Innovation: Opportunities and Challenges. *A Report on the Research Workshop 'Digital Challenges in Innovation Research'*. 1-37.

Zaslavsky, A., Perera, C., Georgakopoulos, D., (2013). Sensing as a Service and Big Data. *Cornell University*. 1-8.

Zhang, S., Yao, L., Sun, A., & Tay, I. (2019). Deep Learning Based Recommender System: A Survey and New Perspectives. *ACM Comput*, vol. 5, 1-38.

Zhao, W. & Kapoor, G. (2011). Detecting evolutionary financial statement fraud. *Elsevier B.V. Vol 5, issue 3, 570-575*

9. VEDLEGG

9.1. Vedlegg 1 – Samtykkeskjema NSD

Vil du delta i forskningsprosjektet

«Kunstig intelligens og Big Data i forsikringsbransjen»?

Dette er et spørsmål til deg om å delta i et forskningsprosjekt hvor formålet er å avdekke hvordan forsikringsselskaper bruker maskinlæring og datalæring i arbeidet mot å sammen redusere forsikringsvindel. I dette skrevet gir vi deg informasjon om målene for prosjektet og hva deltakelse vil innebære for deg.

Formål

Forskningsprosjektet gjennomføres i forbindelse med masteroppgave ved Høgskulen på Vestlandet. Masteroppgaven løper over ett semester – januar til juni 2021.

Vi ønsker å undersøke ett spesifikt samarbeid mellom ulike selskaper i forbindelse med ett spesifikt innovasjonsprosjekt; hvor samarbeidet har som formål å redusere forsikringsvindel som et samfunnsproblem.

Hvem er ansvarlig for forskningsprosjektet?

Høgskulen på Vestlandet (HVL) er ansvarlig for prosjektet.

Hvorfor får du spørsmål om å delta?

Du får spørsmål om å delta fordi vi anser deg som en relevant person innenfor dette feltet som kan gi gode bidrag til vår masteroppgave. Utvalget av intervjuobjekter i oppgaven er basert på personer som aktivt jobber med innovasjonsprosjektet og jobber i selskapene som inngår i samarbeidet.

Kontaktopplysninger er opparbeidet gjennom eget nettverk og andre personer som inngår i samarbeidet/innovasjonsprosjektet vi undersøker.

Hva innebærer det for deg å delta?

Hvis du velger å inngå i prosjektet, innebærer det at du tar del i et personlig intervju. Intervjuet har ingen bestemt varighet, og avhenger av hvordan samtalen utvikler seg underveis. Vi antar at du må sette av ca. 1-1,5 time. Det blir tatt lydopptak av intervjuet. Grunnet Covid-19 vil intervjuet gjennomføres digitalt.

Det er frivillig å delta

Det er frivillig å delta i prosjektet. Hvis du velger å delta, kan du når som helst trekke samtykket tilbake uten å oppgi noen grunn. Alle dine personopplysninger vil da bli slettet. Det vil ikke ha noen negative konsekvenser for deg hvis du ikke vil delta eller senere velger å trekke deg.

Ditt personvern – hvordan vi oppbevarer og bruker dine opplysninger

Vi vil bare bruke opplysningene om deg til formålene vi har fortalt om i dette skrevet. Vi behandler opplysningene konfidensielt og i samsvar med personvernregelverket. Ansvarlige studenter som skriver denne masteroppgaven, vil ha tilgang til svar/lydopptak som er utarbeidet i forbindelse med intervjuprosessen. Denne informasjonen beholdes av

studentene fram til sensurfrist. Veileder av masteroppgaven og sensor vil også ha tilgang til oppgaven.

På bakgrunn av Covid-19 situasjonen har vi ikke tilgang på opptaksenheter/lagringsenheter levert av utdanningsinstitusjon. Lydopptak vil oppbevares på studentenes private låste enheter.

På bakgrunn av at innovasjonsprosjektet som studeres er offentlig er det også sannsynlig at data som opparbeides gjennom intervjuer kan kobles til individer/selskaper. Vi nevner ikke navn i oppgaven, men stillingstittel og arbeidssted kan fremkomme.

Hva skjer med opplysningene dine når vi avslutter forskningsprosjektet?

Opplysningene anonymiseres når prosjektet avsluttes/oppgaven er godkjent, noe som etter planen er 11.06.2021. Personopplysninger og lydopptak slettes ved prosjektslutt.

Dine rettigheter

Så lenge du kan identifiseres i datamaterialet, har du rett til:
innsyn i hvilke personopplysninger som er registrert om deg, og å få utlevert en kopi av opplysningene,
å få rettet personopplysninger om deg,
å få slettet personopplysninger om deg, og
å sende klage til Datatilsynet om behandlingen av dine personopplysninger.

Hva gir oss rett til å behandle personopplysninger om deg?

Vi behandler opplysninger om deg basert på ditt samtykke.

På oppdrag fra Høgskulen på Vestlandet har NSD – Norsk senter for forskningsdata AS vurdert at behandlingen av personopplysninger i dette prosjektet er i samsvar med personvernregelverket.

Hvor kan jeg finne ut mer?

Hvis du har spørsmål til studien, eller ønsker å benytte deg av dine rettigheter, ta kontakt med: Høgskulen på Vestlandet ved Kjersti Berg Danilova, Førsteamanuensis II. Tlf. +4755959215.

E-post: kjersti.danilova@nhh.no

Høgskulen på Vestlandet ved Frida Nord Arnesen, student. Tlf. +4799396561.

E-post: Frida.nord.arnesen@gmail.com

Høgskulen på Vestlandet ved Maiken Johannesen, student. Tlf. +4792688147.

E-post: johannesen.maiken@gmail.com

Vårt personvernombud: Anne-Mette Somby. Fagansvarlig for forskningsetikk og personvern ved Høgskulen på Vestlandet. Tlf. +4755587748. E-post: anne-mette.somby@hvl.no

Hvis du har spørsmål knyttet til NSD sin vurdering av prosjektet, kan du ta kontakt med: NSD – Norsk senter for forskningsdata AS på epost (personverntjenester@nsd.no) eller på telefon: 55 58 21 17.

Med vennlig hilsen

Frida Nord Arnesen & Maiken Johannesen
Masterstudenter ved Høgskulen på Vestlandet

Samtykkeerklæring

Jeg har mottatt og forstått informasjon om prosjektet «*Bruk av kunstig intelligens og Big Data i forsikringsbransjen*», og har fått anledning til å stille spørsmål. Jeg samtykker til:

å delta i personlig intervju

å delta i gruppeintervju

at opplysninger om meg publiseres slik at jeg potensielt kan gjenkjennes

at mine personopplysninger lagres fram til prosjektslutt; 11.06.2021

Jeg samtykker til at mine opplysninger behandles frem til prosjektet er avsluttet

(Signert av prosjektdeltaker, dato)

9.2. Vedlegg 2 – Godkjenning fra NSD

13.5.2021

Meldeskjema for behandling av personopplysninger



NSD sin vurdering

Prosjekttittel

Masteroppgave innovasjon og ledelse

Referansenummer

204856

Registrert

16.02.2021 av Frida Nord Arnesen - 587821@stud.hvl.no

Behandlingsansvarlig institusjon

Høgskulen på Vestlandet / Fakultet for økonomi og samfunnsvitenskap / Institutt for økonomi og administrasjon

Prosjektansvarlig (vitenskapelig ansatt/veileder eller stipendiat)

Kjersti Berg Danilova, Kjersti.Danilova@nhh.no, tlf: +4755959215

Type prosjekt

Studentprosjekt, masterstudium

Kontaktinformasjon, student

Frida Nord Arnesen, frida.nord.arnesen@gmail.com, tlf: +4799396561

Prosjektperiode

11.01.2021 - 21.05.2021

Status

17.02.2021 - Vurdert

Vurdering (1)

17.02.2021 - Vurdert

Det er vår vurdering at behandlingen av personopplysninger i prosjektet vil være i samsvar med personvernlovgivningen så fremt den gjennomføres i tråd med det som er dokumentert i meldeskjema med vedlegg 17.2.2021. Behandlingen kan starte.

MELD VESENTLIGE ENDRINGER

Dersom det skjer vesentlige endringer i behandlingen av personopplysninger, kan det være nødvendig å melde dette til NSD ved å oppdatere meldeskjemaet. Før du melder inn en endring, oppfordrer vi deg til å lese om hvilke type endringer det er nødvendig å melde:

https://nsd.no/personvernombud/meld_prosjekt/meld_endringer.html

Du må vente på svar fra NSD før endringen gjennomføres.

TYPE OPPLYSNINGER OG VARIGHET

Prosjektet vil behandle alminnelige kategorier av personopplysninger frem til 11.6.2021.

LOVLIG GRUNNLAG

Prosjektet vil innhente samtykke fra de registrerte til behandlingen av personopplysninger. Vår vurdering er at prosjektet legger opp til et samtykke i samsvar med kravene i art. 4 og 7, ved at det er en frivillig, spesifikk, informert og utvetydig bekreftelse som kan dokumenteres og som kan trekkes tilbake.

Lovlig grunnlag for behandlingen vil dermed være de registrertes samtykke, jf. personvernforordningen art. 6 nr. 1 bokstav a.

PERSONVERNPRINSIPPER

NSD vurderer at den planlagte behandlingen av personopplysninger vil følge prinsippene i personvernforordningen om:

- lovlighet, rettferdighet og åpenhet (art. 5.1 a), ved at de registrerte får tilfredsstillende informasjon om og samtykker til behandlingen
- formålsbegrensning (art. 5.1 b), ved at personopplysninger samles inn for spesifikke, uttrykkelig angitte og berettigede formål, og ikke viderebehandles til nye uforenlige formål
- dataminimering (art. 5.1 c), ved at det kun behandles opplysninger som er adekvate, relevante og nødvendige for formålet med prosjektet
- lagringsbegrensning (art. 5.1 e), ved at personopplysningene ikke lagres lengre enn nødvendig for å oppfylle formålet

DE REGISTRERTES RETTIGHETER

NSD vurderer at informasjonen om behandlingen som de registrerte vil motta oppfyller lovens krav til form og innhold, jf. art. 12.1 og art. 13.

Så lenge de registrerte kan identifiseres i datamaterialet vil de ha følgende rettigheter: innsyn (art. 15), retting (art. 16), sletting (art. 17), begrensning (art. 18) og dataportabilitet (art. 20).

Vi minner om at hvis en registrert tar kontakt om sine rettigheter, har behandlingsansvarlig institusjon plikt til å svare innen en måned.

FØLG DIN INSTITUSJONS RETNINGSLINJER

NSD legger til grunn at behandlingen oppfyller kravene i personvernforordningen om riktighet (art. 5.1 d), integritet og konfidensialitet (art. 5.1. f) og sikkerhet (art. 32).

For å forsikre dere om at kravene oppfylles, må dere følge interne retningslinjer og eventuelt rådføre dere med behandlingsansvarlig institusjon.

OPPFØLGING AV PROSJEKTET

NSD vil følge opp ved planlagt avslutning for å avklare om behandlingen av personopplysningene er avsluttet.

Lykke til med prosjektet!

Kontaktperson hos NSD: Lasse Raas
Tlf. personverntjenester: 55 58 21 17 (tast 1)

9.3. Vedlegg 3 - Intervjuguide

Intervjuguide 1 – Forsikringselskap

Hvordan starte intervjuet?

- Ønske respondenten velkommen og skape en god atmosfære – spørre om det er greit å ta opptak av samtalen slik at arbeidet videre med oppgaven blir enklere. Informere om samtykkeerklæringen som er sendt ut.
- Introduksjon av prosjektansvarlige
- La respondenten introdusere seg selv og sitt ansvarsområde

Kan du fortelle oss litt om ditt ansvarsområde i *selskap* hva du jobber med?

Fortell oss gjerne mer om innovasjonsprosjektet og *selskap* sitt bidrag:

- Hvordan kom dere i gang med innovasjonsprosjektet og hvorfor er dette prosjektet viktig for dere? Var det noen spesiell grunn for at dere kom senere inn i prosjektet enn de to andre selskapene?
- Hva anser du/dere som de største fordelene med denne formen for teknologi?
- Er det noen ulemper/utfordringer med prosjektet? I så tilfelle hva er disse og hvorfor er det utfordrende? Hva kan eventuelt gjøres for å gjøre det lettere?
- Hvilken verdi kan optimaliseringen av dette prosjektet gi fremtidige forsikringskunder?
- Hvilken betydning tror du det vil ha om flere forsikringselskaper involveres i prosjektet? Kan det ha effekt på datagrunnlaget og nytteverdien av disse dataene? For eksempel om de ulike selskapene koder sine saker på ulikt grunnlag og dermed skaper skjevhet. Henvis til felles Data Dictionary.
- Kan du fortelle oss litt mer om den felles Data Dictionary? Hvilken betydning har denne for å optimalisere løsningen?

Hvilke typer data er det som deles i denne plattformen/skyen?

Hvilken type kunstig intelligens brukes i prosjektet?

Hvordan fungerer MVP-en til nå? Tror du dette er et prosjekt dere kommer til å ville fortsette med?

Hvordan jobber Fremtind med å få datagrunnlaget best mulig?

Kan du gi oss et eksempel på hvordan svik oppdages gjennom denne typen teknologiske løsninger?

Ser du noen utfordringer rent teknisk knyttet til personvern, GDPR og datalekkasje?

Hvordan behandler dette verktøyet personopplysninger sammenlignet med slik dere håndterer det i dag? Hva er forskjellene?

- Dataene om kundene blir jo ikke sett av de andre forsikringselskapene – hvorfor er det da en utfordring å dele personopplysninger i plattformen?
- Hvor lenge oppbevares kundedata i skyen sammenlignet med hvor lenge dere oppbevarer data internt?

- Hvilken effekt ser du for deg at dette verktøyet/plattformen vil ha på skadebehandlingsprosessen slik dere kjenner den i dag?

Hvordan har denne nye arbeidsmåten – samarbeid, men samtidig konkurrere vært for Fremtind? Hvilken verdi, eventuelt utfordringer ser du i denne type arbeidsmetode?

- Hvor viktig har klyngesamarbeidet vært i denne sammenheng? Kunne et forsikringsselskap eller et teknologiselskap klart dette alene?

Hvordan har innovasjonsprosjektet/samarbeidet påvirket ditt arbeidssted og deres arbeidsmetoder?

Hvilke juridiske og tekniske utfordringer har dere møtt/opplevd i forbindelse med dette innovasjonsprosjektet/samarbeidet?

Har du en formening om dette prosjektet kan optimaliseres og videreutvikles så lenge man ikke kan dele personopplysninger?

- Har du noen tanker om hvorvidt dette kan bli en nasjonal og obligatorisk løsning for alle forsikringsselskap?

Intervjuguide 2 – Teknologiselskap

Hvordan starte intervjuet?

- Ønske respondenten velkommen og skape en god atmosfære – spørre om det er greit å ta opptak av samtalen slik at arbeidet videre med oppgaven blir enklere. Informere om samtykkeerklæringen som er sendt ut.
- Introduksjon av prosjektansvarlige
- La respondenten introdusere seg selv og sitt ansvarsområde

Kan du/dere fortelle litt om *selskap* og hva dere jobber med.

Fortell gjerne litt mer om innovasjonsprosjektet og *selskap* sitt bidrag:

- Hvordan er IT-løsningen utformet?
- Hva anser du/dere som de største fordelene med denne formen for teknologi og samarbeid?
- Har dere møtt på noen utfordringer underveis med utviklingen av den teknologiske løsningen?
- Hvilken betydning tror du det vil ha om flere forsikringsselskaper involveres i prosjektet? Kan det ha effekt på datagrunnlaget og nytteverdien av disse dataene? For eksempel om de ulike selskapene koder sine saker på ulikt grunnlag og dermed skaper skjevhet. Er IT-løsningen lagt opp til skalering?
- Har dere noen resultater på hvordan MVP-en har fungert til nå?
-

Hvordan sikrer deres tekniske løsning at ikke personvernopplysninger kan komme på avveie?

Hvordan ser dere at prosjektet fungerer til nå? Positivt/negativt?

Er samarbeidspartnerne fornøyde med løsningen dere har kommet opp med?

Er det noe form for utgruppering av data dersom det deles data som ikke burde deles?

Intervjuguide 3 – Næringsklynge

Hvordan starte intervjuet?

- Ønske respondenten velkommen og skape en god atmosfære – spørre om det er greit å ta opptak av samtalen slik at arbeidet videre med oppgaven blir enklere. Informere om samtykkeerklæringen som er sendt ut.
- Introduksjon av prosjektansvarlige
- La respondenten introdusere seg selv og sitt ansvarsområde

Kan du fortelle oss litt om *selskap* og hva dere jobber med?

Fortell oss gjerne mer om innovasjonsprosjektet og *selskap* sitt bidrag:

- Hvordan kom dere i gang med innovasjonsprosjektet og hvorfor er dette prosjektet viktig?
- Hva anser du/dere som de største fordelene med denne formen for teknologi?
- Hvilken verdi kan optimaliseringen av dette prosjektet gi fremtidige forsikringskunder?
- Hvilken betydning tror du det vil ha om flere forsikringsselskaper involveres i prosjektet? Kan det ha effekt på datagrunnlaget og nytteverdien av disse dataene? For eksempel om de ulike selskapene koder sine saker på ulikt grunnlag og dermed skaper skjevhet. Henvis til felles Data Dictionary.

Fortell oss om selve svindeldeteksjonsplattformen. Hva gjør denne rent praktisk?

- Hvilke typer data er det deles i denne plattformen/skyen?
- Hvilken type kunstig intelligens brukes i prosjektet?
- Hvordan fungerer MVP-en til nå?

Har du kjennskap til Black Box-problematikken og hvordan dette potensielt kan påvirke prosjektet?

- Kan man for eksempel forklare beslutninger som tas ved hjelp av svindeldeteksjonsmodellen?
- Hvordan tar man hensyn til personvern dersom man skal være mest mulig åpen om beslutningsprosessen?

Fortell oss gjerne litt mer om ulike utfordringer dere har møtt med hensyn til personvernlovgivningen og konkurranseloven.

- Dersom du har en kontaktperson i PwC ønsker vi veldig gjerne en prat med vedkommende.

Intervjuguide 4 – Juridisk rådgiver

Hvordan starte intervjuet?

- Ønske respondenten velkommen og skape en god atmosfære – spørre om det er greit å ta opptak av samtalen slik at arbeidet videre med oppgaven blir enklere. Informere om samtykkeerklæringen som er sendt ut.
- Introduksjon av prosjektansvarlige
- La respondenten introdusere seg selv og sitt ansvarsområde

Hvilken rolle har *selskap* hatt i forbindelse med dette prosjektet?

Har du noe oppfatning om hvilke data som kunne vært inkludert for å optimalisere løsningen, men som ikke inkluderes med hensyn til personvern?

Vi vet at modellen bruker ca. 30 treningsdata per i dag. Tror du at dette er nok, eller kan løsningen optimaliseres på så få data?

Hvordan tar man hensyn til personvern dersom man skal være mest mulig åpen om beslutningsprosessen?

Har du kjennskap til Black Box-problematikken og hvordan dette potensielt kan påvirke prosjektet?

Vil kunder i de ulike forsikringsselskapene ha mulighet til å reservere seg fra at sine data deles i denne type felles datasjø?

- Hvordan tror du i så tilfelle dette kan påvirke beslutningsmodellenes kvalitet?

For at prediksjonsmodellene/beslutningsmodellene skal få trene på mest mulig data over tid, så strider det gjerne mot at man ikke skal lagre i lengre perioder. Er dette en utfordring for prosjektet?

Hvilke data benyttes for at prediksjonsmodellen ikke skal stride med prinsippet om rettferdighet? For eksempel vil gjerne en spesifikk type bil eller område være flagget som høy-risiko – hvordan kan kunder i dette segmentet unngå å bli diskriminert eller utsatt for forskjellsbehandling?

Ambisjonen er å etablere en nasjonal plattform for svikdeteksjon i forsikring med anvendelse på flere produkter, både innenfor skade, men også personforsikringer. Hvordan kan man ta hensyn til personvern dersom man skal kunne detektere svik på personforsikring hvor det finnes mange sensitive opplysninger?

Har du noen tanker om hvilke eksterne data, utenom data fra forsikringsselskaper, som kunne bidratt til å optimalisere svikdeteksjonsmodellen?

Vi har hørt litt forskjellig om den regulatoriske sandkassen. Noen sier at prosjektet er med, mens andre sier det ikke er det.

- Har prosjektet mulighet til å teste og trene data i sandkassen?
- Hvilke fordeler gir en slik løsning?
- Kan prosjektet optimaliseres dersom det ikke kan teste og trene data i sandkassen?

Intervjuguide 5 – Skadebehandler og etterforsker

Hvordan starte intervjuet?

- Ønske respondenten velkommen og skape en god atmosfære – spørre om det er greit å ta opptak av samtalen slik at arbeidet videre med oppgaven blir enklere. Informere om samtykkeerklæringen som er sendt ut.
- Introduksjon av prosjektansvarlige
- La respondenten introdusere seg selv og sitt ansvarsområde

Kan du fortelle oss litt om hva du jobber med, og hva din rolle er i *selskap*?

Har du noe kjennskap til innovasjonsprosjektet i regi av Finance Innovation som handler om å bruke kunstig intelligens og datadeling for å bekjempe forsikringssvindel?

- Kan du i så fall si noe om hvordan dette prosjektet har påvirket ditt arbeid så langt?
- Erfaringer?
- Hva anser du som de største utfordringene?

Hvordan vil du definere begrepet svindel innenfor forsikring?

Kan du fortelle oss litt om hvordan dere predikerer og etterforsker svindel per i dag?

- Dersom dere jobber med prediksjonsmodeller. Kan du fortelle oss om hvordan disse fungerer per dags dato?

Har du gjort deg opp noen tanker om hvordan kunstig intelligens og datadeling vil påvirke dagens arbeid med svindelsaker/skadebehandling/etterforskning?

- Hvilken effekt ser du tror dette prosjektet, hvis det realiseres, vil ha på skadebehandlingsprosessen slik dere kjenner den i dag?

Ser du for deg at denne type teknologi vil ha større inngripen på personvernet sammenlignet med hvordan dere behandler persondata i dag?

- Vet du noe om hvor lenge kundedata oppbevares hos dere per i dag, og om dette vil være til stor forskjell for hvor lenge det oppbevares i en potensielt delt datasky?
- Hvordan jobber dere i dag med behandlingsformål, forklaringsplikt og dataminimering?

Har du noen tanker om hvorvidt svikdeteksjon ved hjelp av kunstig intelligens og datadeling vil minimere kostnader knyttet til for eksempel tilfeller hvor det blir avdekket at det ikke er svindel? For eksempel med hensyn til arbeidskraft, effektivitet, tid, m.m.

Har dere utfordringer knyttet til GDPR og personvern i dagens løsning for å oppdage svindel?

Har du noe formening om hvordan et større datasett vil påvirke kvaliteten på prediksjonsmodellene for svindel i deres selskap sammenlignet med dagens datasett?

Har samarbeidet med andre forsikringsselskaper påvirket dere i noe som helst grad så langt?