

Maria Midjo Rønning

Kompetanse og tillit i menneske- chatbot-interaksjon

En casestudie av offentlig sektor

Masteroppgave i Digital samhandling

Veileder: Elena Parmiggiani

Juni 2021

Maria Midjo Rønning

Kompetanse og tillit i menneske- chatbot-interaksjon

En casestudie av offentlig sektor

Masteroppgave i Digital samhandling
Veileder: Elena Parmiggiani
Juni 2021

Norges teknisk-naturvitenskapelige universitet
Fakultet for informasjonsteknologi og elektroteknikk
Institutt for datateknologi og informatikk



Kunnskap for en bedre verden

Sammendrag

Bruk av kunstig intelligens i offentlig tjenesteleveranse er stadig økende, og målet om å legge til rette for et digitalt førstevalg skal gjøre digital samhandling til normen for kontakt mellom innbygger og forvaltning. Flere offentlige aktører har tatt i bruk chatboter for innbyggerkontakt for å oppnå en høyere grad av selvbetjening, ressurs sparing og økt tilgjengelighet. For å sikre utbredt akseptanse av chatboter er offentlig sektor avhengig av at brukere (innbyggere) har tillit til tjenesten. Tidligere forskning foreslår en rekke faktorer som påvirker utvikling av tillit i en menneske-chatbot interaksjon, deriblant oppfattet kompetanse. Mye av forskningen er derimot gjort på chatboter for kundeservice og i kommersiell sammenheng, og det er et gap i litteraturen om tillit til chatboter i offentlig sektor.

Denne studien har som mål å bidra med ny empirisk innsikt om hvordan kompetanse påvirker brukeres tillit til chatboter med følgende forskningsspørsmål: *Hvordan påvirker oppfattet kompetansenivå tillit mellom bruker og chatbot?* For å besvare forskningsspørsmålet er det gjennomført en casestudie på bruk av chatbot i offentlig sektor, eksemplifisert med NAVs chatbot, Frida. Datainnsamlingen består av intervjuer gjennomført med informanter hos NAV samt en spørreundersøkelse besvart av innbyggere som har benyttet seg av chatbot for kontakt med offentlig sektor i Norge.

Funnene viser at tillit i stor grad påvirkes av oppfattet kompetansenivå, og at brukeres forventninger før bruk spiller en viktig rolle i hvordan kompetanse oppfattes. Når brukere har høye forventninger til en chatbots kompetanse, innhold og evne til å forstå, og disse ikke innfris ved bruk, svekkes tilliten fort. Forventninger som i større grad stemmer overens med en chatbots reelle kompetansenivå fører til at tilliten forblir uendret ved bruk. Videre presenteres tiltak som kan legge til rette for tillit i en menneske-chatbot interaksjon.

Abstract

The use of artificial intelligence in public service delivery is constantly increasing, and the goal of promoting a "digitalt førstevalg", or a digital first choice, makes digital interaction the norm for interaction between citizens and public administration. Several public agencies are using chatbots for citizen contact to facilitate a higher degree of self-service, increase accessibility, and save resources. In order to achieve a widespread acceptance of chatbots, the public sector is dependent on users (citizens) trusting the service. Previous research suggests several factors that influence the development of trust in a human-chatbot interaction, including perceived competence. Although much research has been done on chatbots in recent years, most of it is done on chatbots for customer service and in a commercial context, and there is a gap in the literature on trust in chatbots in the public sector.

This study aims to contribute new empirical insight into how competence affects users' trust in chatbots with the following research question: *How does the perceived level of competence affect trust between user and chatbot?* To answer the research question, a case study has been carried out on the use of chatbots in the public sector, exemplified by NAV's chatbot, Frida. The data collection consists of interviews conducted with informants from NAV as well as a survey answered by citizens who have used chatbots for contact with the public sector in Norway.

The findings show that trust is largely influenced by perceived level of competence and that users' expectations before use play an important part in how competence is perceived. When users have high expectations of a chatbot's competence, content, and ability to understand, and these expectations are not met during the interaction, the users' trust quickly weakens. Expectations that correspond to a chatbot's real level of competence lead to users' expectations being met during the interaction, and the trust remains unchanged. Furthermore, measures that can facilitate trust in a human-chatbot interaction are presented.

Forord

Denne masteroppgaven er skrevet som del av min mastergrad i digital samhandling ved Institutt for datateknologi og informatikk ved Norges teknisk-naturvitenskapelige universitet (NTNU). Oppgaven ble gjennomført vårsemesteret 2021.

Først og fremst vil jeg takke min veileder, førsteamanuensis Elena Parmiggiani ved Institutt for datateknologi og informatikk ved NTNU, for god veiledning og grundige tilbakemeldinger under arbeidet med masteroppgaven. Dine bidrag og innspill har vært svært verdifulle. Videre ønsker jeg å takke informanter i NAV for deres tid og for den nyttige innsikten dere bidro med. Til slutt ønsker jeg å takke de andre studentene involvert i Frida@Work for samarbeidet, og en spesiell takk til medstudent Knut Gjelle for gode diskusjoner og et godt samarbeid rundt datainnsamlingen.

Trondheim, juni 2021

Maria Midjo Rønning

Innhold

Sammendrag	iii
Abstract	v
Forord	vii
Innhold	ix
Figurer	xi
Tabeller	xi
Akronymer	xiii
1 Introduksjon	1
1.1 Bakgrunn og motivasjon	1
1.2 Forskningsspørsmål	4
1.2.1 Avgrensninger	4
1.3 Oppgavens oppbygning	5
2 Teoretisk rammeverk	7
2.1 Kunstig intelligens og chatboter	7
2.1.1 Utvikling og design av chatboter	8
2.2 Tillit til chatboter	10
2.2.1 Kompetanse	12
3 Casebeskrivelse	15
3.1 Chatbot i offentlig sektor	15
3.2 Chatbot i NAV	16
4 Metode	19
4.1 Forskningsdesign	19
4.2 Datainnsamling	20
4.2.1 Intervju	20
4.2.2 Spørreundersøkelse	22
4.3 Dataanalyse	23
5 Empiriske funn	29
5.1 Tilgjengelighet	29

5.1.1	Avbrutte henvendelser	30
5.2	Brukeropplevelse i chatbot	31
5.2.1	Opplæring av brukere	31
5.3	Chatbotens kompetansenivå	33
5.3.1	Videreutvikling av chatbot og innhold	33
5.3.2	Brukerforventninger	35
5.3.3	Oppfatning etter bruk	37
5.4	Tillit til chatboten	38
5.4.1	Chatbotens kompetanse og tillit	39
5.4.2	Design og utseende	40
5.4.3	Tillit til den offentlige aktøren	41
5.4.4	Personvern og sikkerhet	42
6	Diskusjon: Tillit til chatbot?	43
6.1	Oppfattet kompetansenivå	43
6.2	Rollen brukerforventninger spiller	46
6.3	Implikasjoner for design	49
7	Konklusjon	51
7.1	Begrensninger og fremtidig arbeid	52
	Bibliografi	53
A	Tilleggsmateriale	59
A.1	Intervjuguide	60
A.2	Spørreundersøkelse	61

Figurer

1.1	Skjerm bilde av artikkel publisert på digi.no 5.10.2017 (Lekanger, 2017).	2
4.1	Oversikt over de ulike perspektivene som dekkes av prosjektet Frida@Work.	20
4.2	Skjerm bilde av kodeprosessen i Nvivo.	24

Tabeller

4.1	Oversikt over datainnsamlingsaktiviteter gjennom studien	21
4.2	Oversikt over tema, konseptuelle kategorier og eksempler og utdrag fra datamaterialet	25

Akronymer

IA intelligent agent.

KI kunstig intelligens.

ML maskinl ring.

NAV arbeids- og velferdsforvaltningen.

NLP natural language processing.

NPM new public management.

Kapittel 1

Introduksjon

1.1 Bakgrunn og motivasjon

Avansert teknologi er blitt en naturlig del av samfunnet vårt, og hverdagen digitaliseres stadig mer. Digitale løsninger påvirker hvordan vi samhandler og måten vi tilegner oss informasjon på, og disse prosessene overlates i økende omfang til automatiserte løsninger og kunstig intelligens. Teknologi basert på kunstig intelligens (KI) handler om å utvikle datasystemer som kan lære av egne erfaringer og løse komplekse problemer i ulike situasjoner og miljøer (Tennøe, 2017). Ved hjelp av kunstig intelligens kan datasystemer tilegne seg egenskaper som vi tidligere har tenkt er forbeholdt og unike for mennesker, og teknologien utvikler seg raskt.

Offentlig sektor er i en digitaliseringsprosess både når det gjelder tjenesteleveranse og brukerkontakt. Stortingsmeldingen ”Digital Agenda” (2016) var en omfattende gjennomgang av Norges digitale tilstand, og meldingen påpekte at det offentlige skal legge til rette for et digitalt førstevalg (Kommunal- og moderniseringsdepartementet, 2016). Det innebærer at digital samhandling skal være hovedregelen for kontakt mellom innbygger (bruker) og forvaltning. Kunstig intelligens kan være et nyttig verktøy for å senke innbyggeres terskel for å ta i bruk digitale løsninger i offentlig sektor, for eksempel med virtuelle agenter eller talegjenkjenning som gir mulighet for å snakke direkte med maskiner på mange ulike språk (Tennøe, 2017). Dette er også i tråd med elementer i New public management (NPM), som økt fokus på effektivitet, resultat, konkurranse, marked og brukere. For brukere betyr dette økt tilgjengelighet, redusert tidsbruk og økt selvbetjening. For forvaltningen betyr dette at ressursene som tidligere ble brukt til enkel brukerkontakt heller kan brukes andre plasser.



Figur 1.1: Skjerm bilde av artikkel publisert på digi.no 5.10.2017 (Lekanger, 2017).

Digitaliseringsprosessen er godt i gang på flere områder i offentlig sektor, blant annet i form av digitale plattformer som en ny måte å yte offentlige tjenester på (Vassilakopoulou mfl., 2017). Innen helse jobbes det med en felles, nasjonal helseplattform samtidig som det skal innføres elektroniske pasientjournalssystem med beslutningsstøtte for spesialisthelsetjenesten og kommuner (KS, 2018). I Skatteetaten har et modernisert og digitalt Folkeregister forenklet prosesser for innbyggere ved adresseendring, innflytting, navnevalg, fødsel og andre tjenester Folkeregisteret tilbyr (Hoem, 2020). En annen etat som arbeider med digitalisering av sine tjenester, er Arbeids- og velferdsforvaltningen (NAV). I NAVs omverdensanalyse fra 2019 peker de på at mange av prosessene i NAV er egnet for selvbetjening og automatisering, og målet med å digitalisere tjenester er å skape bedre brukermøter (NAV, 2019). Dette kan enten skje ved at de digitale tjenestene i seg selv skaper bedre brukermøter, eller ved at kostnader reduseres slik at ressursene som frigjøres kan brukes til å skape bedre brukermøter på andre områder.

En av teknologiene som kan øke graden av selvbetjening og automatisering, og dermed frigjøre menneskelige ressurser, er chatbot. En chatbot er en intelligent agent (IA), en autonom entitet som kan utføre tjenester basert på omgivelser, brukerinput og erfaringer. Chatboter er dataprogrammer som imiterer menneskelig kommunikasjon og som brukere kan interagere med ved hjelp av naturlig språk (Shawar og Atwell, 2007). I de siste årene har flere offentlige tjenester i Norge tatt i bruk chatboter for kontakt med innbyggere, blant annet Skatteetaten, NAV og omkring 90 av landets kommuner. Felles for disse er at chatbot er implementert som del av førstelinjekommunikasjonen med innbyggere, og at chatbotene kun kan svare på enkle, generelle spørsmål. Chatboter benyttes også i kommersiell sammenheng for kundeservice, kjøpshjelp, banktjenester og lignende.

I følge bransjerapporter forhindrer manglende tillit i ikke-menneskelig interaksjon utbredt akseptanse av chatboter (Müller mfl., 2019). En forutsetning for at offentlige tjenester skal lykkes med brukerkontakt gjennom chatboter, er at innbyggere har tillit til tjenesten som tilbys. Tillit til chatboter påvirkes av en rekke faktorer, og i hvilken grad chatboten tolker brukernes spørsmål riktig og gir hjelpsomme og informative svar er identifisert som en av nøkkelfaktorene i etablering av tillit, i tillegg til et personlig og menneskelig utseende (Følstad mfl., 2018). Tidligere forskning på tillit i online miljøer har foreslått kompetanse som en påvirkende faktor i etablering av tillit. Menneskers tillit til et online objekt, for eksempel en nettside, baseres delvis på deres oppfatning eller opplevelse av objektets kompetanse (Corritore mfl., 2003). Men hvordan brukere oppfatter kompetansenivå vil også variere basert på forventninger før bruk og opplevelse ved bruk (Khadpe mfl., 2020).

Det har vært en økning i forskning på tillit til teknologi, og spesielt kunstig intelligens de siste årene. Fra et teoretisk ståsted er det likevel fortsatt et gap i litteraturen når det gjelder intelligente agenter og chatboter. I sammenheng med kompetanse vet vi at forventninger til kompetanse og hvordan brukere oppfatter chatbotens kompetansenivå har innvirkning på brukerens evaluering av systemet (Khadpe mfl., 2020), og at tilliten påvirkes av hvordan chatboten presenterer seg selv (Følstad mfl., 2018). Litteraturen sier derimot for lite om kompetansenivå og oppfattet kompetanse er direkte knyttet til etablering og utvikling av tillit i menneske-chatbot-interaksjon i offentlig sektor. Mye av forskningen har fokusert

på tillit til chatboter i kommersiell sammenheng eller privat sektor, og er ikke nødvendigvis overførbart til chatboter brukt i offentlig tjenesteleveranse. Forskning gjort på innføring av automatiserte tjenester og bruk av chatboter i offentlig sektor i andre land er også utfordrende å overføre til offentlig sektor i Norge. Med utgangspunkt i et brukersentrert perspektiv på AI i praksis skal dette masterprosjektet derfor se på rollen oppfattet kompetansenivå spiller i etablering og vedlikehold av tillit mellom bruker og chatbot i offentlig sektor.

1.2 Forskningsspørsmål

Denne oppgaven vil undersøke hvilken rolle oppfattet kompetansenivå spiller i etableringen av tillit mellom menneske og chatbot. Målet med studien er å beskrive interaksjonen mellom innbygger og chatbot, med hensikt om å kartlegge hvordan brukere av systemet oppfatter en chatbots kompetanse samt hvilken påvirkning dette har på utvikling av tillit. Forskningsspørsmålet for oppgaven er følgende:

Hvordan påvirker oppfattet kompetansenivå tillit mellom bruker og chatbot?

Forskningsprosjektet har som mål å bidra med ny empirisk innsikt i bruk av chatbot i offentlig tjenesteleveranse. Det er et tema som blir stadig mer aktuelt med digitalisering av offentlig sektor, men som vil til nå vet for lite om. Fokuset for prosjektet er på brukere (innbyggere), og hensikten er å bidra til en bedre forståelse av hvordan brukere oppfatter en chatbots kompetansenivå samt hvordan dette påvirker tilliten brukeren har til systemet. Det vil også bidra med innsikt i hvordan kunstig intelligens kan benyttes for å skape et mer tillitsfullt forhold mellom innbygger og etat.

1.2.1 Avgrensninger

Casestudien ser på bruk av chatbot i offentlig sektor, og dette er derfor fokuset for datainnsamlingen. Funnene beskrevet i kapittel 5 er begrenset av tidsrammen til masteroppgaven.

1.3 Oppgavens oppbygning

Opgaven er strukturert som følger:

Kapittel 2 definerer hovedkonseptene for oppgaven og presenterer bakgrunnsliteratur samt tidligere forskning som er relevant for analysen og diskusjonen.

Kapittel 3 beskriver oppgavens case og presenterer NAVs chatbot Frida som en instans av chatbot i offentlig sektor.

Kapittel 4 beskriver forskningsstrategien og presenterer metoder for datainnsamling og dataanalyse.

Kapittel 5 presenterer empiriske funn i form av konseptuelle kategorier fra data-materialet samlet gjennom intervjuer og spørreundersøkelse.

Kapittel 6 diskuterer funnene i sammenheng med eksisterende teori og forskningsspørsmål.

Kapittel 7 inneholder en kort oppsummering av funnene og diskusjonen samt oppgavens begrensninger og forslag til videre eller fremtidig arbeid.

Kapittel 2

Teoretisk rammeverk

Dette kapittelet beskriver den teoretiske bakgrunnen og skisserer det overordnede teoretiske rammeverket for oppgaven. Først presenteres kunstig intelligens på et overordnet nivå, før intelligente agenter og chatboter beskrives mer i dybden. Videre presenteres konseptet tillit i sammenheng med bruk av chatbot. Denne delen tar utgangspunkt i forskning gjort på tillit til teknologi, online objekter og chatboter i ulike kontekster. Som underkonsepter av tillit er kompetanse, antropomorfisme og sikkerhet valgt ut som de mest relevante for denne studien. Som følge av at det er gjort lite forskning på tillit til chatboter i offentlig tjenesteleveranse, vil delen som omhandler tillit i dette kapittelet også inkludere litteratur om tillit til kommersielle chatboter. Det kan argumenteres med at chatbot i offentlig sektor er en type kundeservice med innbyggere som ”kunder”, men det må likevel presiseres at funnene som presenteres her er ikke nødvendigvis overførbare til konteksten denne studien er gjort i.

2.1 Kunstig intelligens og chatboter

Kunstig intelligens (KI) er et bredt begrep og kan defineres som et systems evne til å korrekt tolke ekstern data, lære fra slik data og bruke det systemet har lært til å løse oppgaver og oppnå spesifikke mål oppgaver gjennom fleksibel tilpasning (Kaplan og Haenlein, 2019).

I KI er en intelligent agent (IA) en autonom entitet som kan utføre tjenester basert på omgivelser, brukerinput og erfaringer. Chatboter er eksempler på intelligente agenter, og teknologien hatt stor vekst både i privat og offentlig sektor de siste årene. En chatbot er et dataprogram som imiterer menneskelig kommunikasjon og gjør det mulig for brukere å kommunisere med et system ved hjelp av naturlig språk (Jære, 2018). Chatboter ble først utviklet for rundt 60 år siden, og målet den gang var å se om chatbot-systemer kunne lure brukere til å tro de var faktiske mennesker i den såkalte Turingtesten (Przegalinska mfl., 2019). I dag be-

nytt chatboter for samhandling med brukere i forskjellige settinger, blant annet i prosesser relatert til kundeservice, innen utdanning eller i offentlig tjenesteleveranse. Chatboter tolker spørsmål fra brukere og henter svarene sine fra en slags database med svar. For å kunne kommunisere effektivt med brukere og forstå hvilken type spørsmål som skal generere et gitt svar, må chatboter være i stand til å gjenkjenne mønstre i et bestemt datasett (Ringnes, 2020). Dette oppnås ved hjelp av maskinlæring (ML). Maskinlæring kan defineres som "en spesialisering innen kunstig intelligens hvor man bruker statistiske metoder for å la datamaskiner finne mønstre i store datamengder" (Tidemann og Elster, 2019). Dette innebærer at maskinen "lærer" ved hjelp av eksempler fremfor å bli programmert til å løse oppgaver ved hjelp av et sett instruksjoner.

Det spesielle med chatboter er at brukere kan kommunisere skriftlig ved hjelp av naturlig språk, som om samtalen var med et menneske. Modellering av det menneskelige språket ved hjelp av datamaskiner har vært et fokusområde innen KI de siste årene, og en måte å gjøre dette på er ved hjelp av "natural language processing (NLP)", eller språkbehandling. NLP ser på bruk av beregningsteknikker for å lære, forstå og produsere menneskelig språkinnhold i tekst- eller taleform (Hirschberg og Manning, 2015). Det er gjort store fremskritt innen fagområdet, men dette har også vist seg å være en svært krevende oppgave. Kompleksiteten i forståelse og generering av språk gjør at hindringen for datamaskiner ikke bare er å forstå betydningen av ord, men også å forstå de uendelige variasjonene i hvordan disse ordene brukes sammen for å kommunisere mening (Hill mfl., 2015).

2.1.1 Utvikling og design av chatboter

I utvikling av intelligente agenter og chatboter er det mange faktorer som må tas i betraktning for å sikre best mulig brukeropplevelse og adopsjon av teknologien. Universell brukbarhet er et viktig konsept innen menneske-maskin-interaksjon, og for en chatbot innebærer dette at utviklere gjør den i stand til å møte brukere med ulike kulturelle, etniske, funksjonelle og språklige bakgrunner. Mennesker søker, bruker og absorberer informasjon fra omgivelsene på ulike måter, og det vil også være individuelle forskjeller i hvordan denne informasjonen oppfattes, organiseres og evalueres (Jackson mfl., 2003).

Amershi m.fl. (2019) presenterte 18 retningslinjer for utvikling av systemer med kunstig intelligens. Disse inkluderer også retningslinjer for bruk over tid, så for denne studiens kontekst er følgende retningslinjer utvalgt som de mest relevante:

G1: Tydeliggjør hva systemet kan gjøre. Gjør det klart for brukeren hva systemet er i stand til.

G2: Tydeliggjør hvor godt systemet kan gjøre det der gjør. Gjør det klart for brukeren hvor godt systemet er til å gjennomføre oppgaver. Tydeliggjør sannsynligheten for mulige feil.

G4: Vis kontekstuell relevant informasjon. Fremvis informasjon som kan være relevant for brukerens nåværende handling eller miljø.

G10: Benytt alternativer ved tvil. Ved usikkerhet rundt brukerens mål, tilby alternativer fremfor å anta hensikt.

G11: Tydeliggjør hvorfor systemet gjorde som det gjorde. Gi brukeren tilgang til en forklaring på hvorfor systemet gjorde som det gjorde hvis det ønskes.

Mentale modeller

Et viktig hensyn i utvikling av intelligente agenter og chatboter er brukerens mentale modeller av systemet. Luger og Sellen (2016) gjennomførte en studie på bruk av talebaserte samtaleagenter, og fant et stort skille mellom brukeres forventninger til hva systemet kan gjøre og hva systemet faktisk kan gjøre. Brukere skapte mentale modeller av systemet som systemet ikke klarte å leve opp til ved bruk. I kombinasjon med mangel på tilbakemeldinger fra systemet under interaksjonen førte dette til vansker med å forstå hvorfor systemet ikke fungerte som forventet og hva brukerne eventuelt måtte gjøre for å endre dette. En måte å løse brukerforventninger på er å designe for å hinte frem eller avsløre intelligens og kapabilitet, og dermed motivere realistiske forventninger (Luger og Sellen, 2016).

Konseptuelle metaforer

En brukers mentale modeller kan blant annet påvirkes ved å benytte konseptuelle metaforer, som er et viktig verktøy i design av intelligente agenter og chatboter. Disse metaforene er korte beskrivelser knyttet til et system som antyder systemets funksjonalitet og intensjon (Khadpe mfl., 2020). Beskrivelsene er med på å forme brukerens forventninger til systemet, og forventninger før bruk påvirker både inn-

ledende atferd og langsiktig atferd, selv om selve systemet forblir uendret mens brukerforventningene varieres (Padgham og Winikoff, 2005). Brukere av et system vurderer systemet til å være bedre eller dårligere i henhold til om tidligere informasjon om systemets kvalitet er formulert i en positiv eller negativ ramme (Hartmann mfl., 2008). Metaforer som påvirker brukere til å forvente en IA som er høyst kompetent vil føre til at de evaluerer systemet annerledes enn om brukerne forventer en IA som er mindre kompetent. Khadpe m.fl. (2020) fant i sin studie av konseptuelle metaforers virkning at brukere er mer tolerante for feil i systemer med lav kompetanse enn i systemer med høy kompetanse. Metaforer som indikerer lav kompetanse fører også til en økning i opplevd brukervennlighet og ønske om å samarbeide, mens metaforer for høy kompetanse øker sannsynligheten for at en bruker tar i bruk systemet. Fremvisning av kompetanse kan altså tiltrekke nye brukere, men med mindre forventningene korrigeres med lavere kompetanse-metaforer ved bruk kan brukerne ende opp med å avvise chatboten (Khadpe mfl., 2020).

2.2 Tillit til chatboter

Brukeres tillit til en teknologi er en avgjørende faktor for at de skal ta i bruk teknologien (Corritore mfl., 2003), og for chatboter fant Nordheim (2018) at brukere som har tillit til chatboter også har en intensjon om å benytte seg av disse i konteksten kundeservice.

Tillit kan defineres som en psykologisk tilstand som innebærer at man godtar å være sårbar, basert på positive forventninger om andres intensjoner og atferd (Rousseau mfl., 1998). Historisk sett er tillit sett på i en mellommenneskelig kontekst hvor troverdighet, tro på hverandres dømmekraft og forutsigbar oppførsel er nøkkelfaktorer for en tillitsfull relasjon (Cassell og Bickmore, 2000). I en organisatorisk kontekst er nøkkelfaktorer for utvikling av tillit oppfattet velvilje, integritet og ekspertise eller evne (Mayer mfl., 1995). I sammenheng med automatisering kan tillit defineres som i hvilken grad en bruker er trygg på og villig til å handle på grunnlag av anbefalingene, handlingene og beslutningene til et kunstig intelligent beslutningshjelpemiddel (Madsen og Gregor, 2000). Tillit til teknologi er et om diskutert tema, spesielt etter som teknologi har utviklet seg fra å være maskiner som utfører fysiske oppgaver til å også utføre intellektuelt arbeid. Det virker like-

vel å være en enighet i fagmiljøet om at tillit kan tilskrives menneskelige relasjoner til komplekse teknologier, som for eksempel teknologier basert på maskinlæring (Taddeo, 2010).

Følstad m.fl. (2018) identifiserer fire faktorer som påvirker brukeres tillit til chatboter brukt i kundeservice:

- **Tolkning og rådgivning:** Chatbotens evne til å tolke brukerens spørsmål riktig samt evnen til å bidra med hjelpsomme og informative svar.
- **Menneskelighet:** Om chatbotens utseende og fremtoning er menneskelig, personlig og høflig.
- **Selvpresentasjon:** Klar kommunikasjon av hva chatboten kan gjøre og hvordan den kan hjelpe samt tydelighet rundt begrensninger.
- **Profesjonelt utseende:** Om chatbotens fremtoning er gjennomtenkt utviklet, med riktig staving og grammatikk.

Videre fremheves en rekke faktorer som ikke direkte omhandler chatboten, men som likevel vil virke inn på en brukers tillit; (1) brukerens oppfatning av og tillit til tjenestetilbyderen eller organisasjonen bak chatboten, (2) hvordan sikkerhet og personvern ivaretas og hvordan dette kommuniseres til brukeren og (3) oppfattet risiko for bruker ved bruk av chatboten (Følstad mfl., 2018).

Nordheim (2018) fant i sin studie av tillit til chatboter i konteksten av kundeservice fire chatbot-relaterte faktorer som påvirker en brukers tillit; (1) ekspertise, (2) rask respons, (3) antropomorfisme og (4) fravær av markedsføring. Men en brukers tillit påvirkes ikke kun av chatboten og dens attributter. Faktorer relatert til miljø eller omgivelser vil også ha innvirkning, og her trekkes (1) lav risiko, (2) merkevare og (3) tilgang til en menneskelig operatør også frem som aspekter ved tillit. Til slutt påpekes det at det også er individuelle forskjeller blant brukere av chatboter og at ikke alle er like tilbøyelige til å stole på teknologi, uten at det nødvendigvis har noe med de overnevnte faktorene å gjøre (Nordheim, 2018). Dette støttes av Müller m.fl. (2019), som fant at tilstedeværelse av enkelte personlighetstrekk hos brukere påvirket evnen til å utvikle tillit til en samtaleagent.

Corritore m.fl. (2003) utviklet en modell for tillit til informative nettsider, og fant tre påvirkende faktorer. *Brukervennlighet* trekkes frem som den første faktoren, og innebærer hvor enkel brukeren oppfatter den interaktive teknologien til å være. *Risiko* ble også funnet å være en påvirkende faktor, og omhandler brukeres oppfatning av sannsynligheten for et uønsket resultat. Til slutt vil *troverdighet* ha innvirkning på en brukers tillit. Troverdighet deles opp i ærlighet, ekspertise, forutsigbarhet og rykte eller anseelse (Corritore mfl., 2003). Til tross for at studien omhandler nettsider, og ikke chatboter, ses flere paralleller til litteratur om tillit til chatboter og påvirkende faktorer.

Videre i dette delkapittelet vil tre påvirkende faktorer utdypes videre; kompetanse, antropomorfisme og sikkerhet og personvern. Disse er alle funnet å være viktige faktorer i utvikling av tillit i menneske-chatbot-interaksjon, og studiens forskningsspørsmål er fokusert på kompetanseaspektet av brukerens tillit. De to andre faktorene henger tett sammen med hvordan en bruker oppfatter en chatbot og dens kompetanse og er derfor inkludert som relevant teori.

2.2.1 Kompetanse

Ekspertise er, som nevnt i avsnitt 2.2, en påvirkende faktor for utvikling av tillit i en menneske-chatbot-interaksjon. I sammenheng med chatboter kan ekspertise defineres som oppfattet kunnskap eller kompetanse (Nordheim mfl., 2019). Hvordan brukere oppfatter ekspertise og kompetanse i en chatbot har mye å si for tilliten brukerne utviser. I mellommenneskelige interaksjoner benytter vi en rekke indikatorer for å kommunisere og forstå hverandres intelligens og evner (Luger og Sellen, 2016). I interaksjon med chatboter vurderes kompetanse gjennom fire kategorier; (1) riktighet og relevans i svar gitt av chatbot, (2) tolkning av spørsmål samt kommunikasjon av hva chatboten ikke forstår, (3) korte, konkrete og tydelige svar og (4) velformulerte svar som fremstår profesjonelle og pålitelige (Nordheim mfl., 2019).

Chatboters manglende evne til å løse mer komplekse forespørsler og spørsmål fra brukere kan svekke tilliten. Chatbotteknologien er fortsatt relativt ny og har mange begrensninger, samtidig som den er ukjent for veldig mange. Det betyr at det er essensielt å kommunisere tydelig ovenfor brukere hva en chatbot kan og ikke kan gjøre for å sikre realistiske forventninger (Følstad mfl., 2018). Manglende

forståelse av en chatbots evner og begrensninger fra brukernes ståsted kan føre til en oppfatning av at chatboten er kapabel til mer enn den er, noe som igjen kan føre til frustrasjon ved bruk og svekket tillit til tjenesten (Luger og Sellen, 2016). Konseptuelle metaforer som sier noe om chatbotens kompetansenivå kan her bidra til å forme brukernes forventninger, og dermed også påvirke tilliten mellom bruker og chatbot (Khadpe mfl., 2020).

På den andre siden vil feilaktig automatisering og uriktige svar i chatbot påvirke tillit negativt, og faktisk i større grad enn dersom et menneske gjør en lignende feil (Dietvorst mfl., 2015). Misforståelser og opplevelse av å ikke bli forstått i samtale med chatbot kan føre til at en bruker avviser chatboten og velger å benytte andre midler for å nå sine mål. Unntaket er dersom chatboten aktivt forsøker å rette opp i misforståelser og tilby avklaringer fremfor å anta kontekst og mening (Sheehan mfl., 2020). Brukere vil først og fremst benytte chatboten til å utføre oppgaver de har tillit til at chatboten klarer å utføre (Luger og Sellen, 2016).

Chatboter snakkes ofte om i sammenheng med kunstig intelligens, noe som kan føre til at brukere tror chatbotprogrammer har evnen til å utøve skjønn. Det kan også føre til misforståelser eller feiloppfatninger knyttet til i hvilken grad chatbotenes svar er programmert av mennesker og hvordan KI benyttes i tjenesten (Aoki, 2020). Luger og Sellen (2016) fant i sin studie av bruk av samtaleagenter at flere brukere, bortsett fra de som var svært teknisk kyndige, følte usikkerhet knyttet til om samtaleagenten hadde evnen til å lære. Gjennomskinnelighet når det kommer til chatboters oppbygning og hva som ligger bak kan derfor være positivt for å utvikle tillit (Kocielnik mfl., 2019; Luger og Sellen, 2016).

Antropomorfisme

En avgjørende del av å skape tillit til intelligente agenter er relatert til antropomorfisering. Antropomorfisme er å tilegge noe ikke-menneskelig menneskelige egenskaper eller trekk (Duffy, 2003), og er en naturlig tendens hos mennesker (De Angeli mfl., 2001). Det å tillegge en chatbot menneskelige trekk, som språkstil, sosial tilstedeværelse eller navn, kan påvirke brukerens oppfatning av både chatboten og organisasjonen eller aktøren bak (Araujo, 2018). Mennesker har allerede en tendens til å reagere sosialt på datamaskiner, også når de er klar over at de kommuniserer med en maskin (Nass og Moon, 2000).

Det er store variasjoner i hvordan utviklere velger å designe chatboter. Mens noen har menneskelignende navn og ikoner, presenteres andre som roboter. Et felles menneskelig trekk for chatboter er evnen til å kommunisere på naturlig språk, og teknologien har utviklet seg raskt på dette området de siste årene. Denne utviklingen og forbedringen i språk i chatboter kan bidra til at brukere tillegger chatboter enda flere menneskelige egenskaper (Mone, 2016), som igjen kan føre til vanskeligheter med å skille mellom en chatbot og en menneskelig samtalepartner (Candello mfl., 2017). Eksempelvis kan bruk av humor i chatbotens svarbase forsterke brukerens oppfatning av chatbotens menneskelige egenskaper, noe som videre kan påvirke brukerens forventninger til chatbotens evner (Luger og Sellen, 2016).

I utgangspunktet kan en tenke at dersom en intelligent agent, som en chatbot, har menneskelige trekk er det mer sannsynlig at det skapes en bærekraftig og tillitsfull relasjon mellom bruker og agent (Przegalinska mfl., 2019). Dette er kjent som menneske-menneske-perspektivet. Her hevdes det at siden mennesker skaper sosiale forventninger til maskiner, vil antropomorfisering øke brukernes tillit til agenten. På den andre siden hevdes det at mennesker plasserer mer tillit i datastyrte systemer enn i andre mennesker. Mennesker er feilbare, mens automatisering og datamaskiner er det motsatte, og det å menneskeliggjøre en datamaskin kan derfor ha negative konsekvenser for brukerens tillit (Seeger og Heinzl, 2018).

Sikkerhet og personvern

Nordheim (2018) trekker frem oppfattet risiko som en av nøkkelfaktorene for utvikling av tillit i en menneske-chatbot-interaksjon. Hvordan brukere oppfatter sikkerhetsrisiko samt kommunikasjon av hvordan sikkerheten håndteres er essensielt (Nordheim, 2018). Dette gjelder spesielt for chatboter som håndterer personlige og sensitive data. I EU og Norge sikres håndtering av slike data gjennom loven for behandling av personopplysninger (GDPR) (Regjeringen, 2019). Denne loven skal sikre blant annet gjennomsjennelighet knyttet til hvordan personopplysninger prosesseres, at innsamlet data kun benyttes til formålet brukeren har samtykket til, at samtykke er dokumentert, og at brukeren har rett til å se egne data på forespørsel (Sağlam og Nurse, 2020).

Kapittel 3

Casebeskrivelse

Dette kapitlet er en beskrivelse av studiens case. Studien fokuserer på bruk av chatbot i offentlig sektor og casen eksemplifiseres med NAVs chatbot, Frida. Innledningsvis beskrives bruk av chatbot i offentlig sektor i dag. Videre presenteres chatboten Frida og dens rolle i etaten.

3.1 Chatbot i offentlig sektor

Chat er blitt en prioritert kanal for kundeservice både i privat og offentlig sektor, og chatboter er et resultat av dette. De ulike offentlige etatene i Norge tilbyr innbyggere en rekke kommunikasjonskanaler, men noen kanaler er mer ressurskrevende enn andre. Henvendelser per e-post, telefon eller personlig oppmøte krever typisk samhandling med veiledere eller saksbehandlere, mens en chatbot kan betjene flere brukere samtidig, døgnet rundt. Implementasjon av chatbot i offentlig sektor motiveres av sparte ressurser samtidig som det er i tråd med det digitale førstevalget for samhandling mellom innbygger og etat. For innbyggerne sin del er det ment som en enkel måte å få svar på generelle spørsmål uten behov for menneskelig interaksjon.

Rundt 90 norske kommuner har de siste årene tatt i bruk chatboten Kommune-Kari for innbyggerkontakt. Chatboten er tilgjengelig på kommunenes nettsider, kan svare på spørsmål knyttet til kommunale tjenester, åpningstider og søknadsfrister, og er åpen døgnet rundt. Kommune-Kari besvarer rundt 80 000 henvendelser landet rundt månedlig (Prokom, 2020), og dersom chatboten ikke kan besvare spørsmålet blir brukeren satt over til en servicekonsulent i kommunen i åpningstiden. Hensikten med chatboten er å avlaste kommunenes servicetorg ved å gi innbyggere konkrete, raske og relevante svar på generelle spørsmål som ikke krever samtale med en menneskelig veileder (Mærøe, 2019). Der innbyggere tidligere måtte lete seg frem på kommunenes nettsider eller andre kilder, er Kommune-Kari ment å bidra med en kortere vei til de relevante svarene. I nærmeste fremtid er

chatboten tenkt å skulle besvare personrelaterede spørsmål og utføre enkel saksbehandling (Prokom, 2020).

Skatteetatens nettsider er blant landets mest besøkte, og etaten har lansert en chatbot som tar imot henvendelser fra innbyggere som har spørsmål rundt Skatteetatens tjenester. Chatboten kan ikke besvare spørsmål knyttet til pågående saker, og svarene er begrenset til generell informasjon knyttet til skatt og Folkeregistret (Skatteetaten, 2021). Chatbotløsningen er levert av Boost.ai, et programvareselskap som tilbyr løsninger som bygger på kunstig intelligens og maskinlæring. Boost.ai står blant annet bak chatbotene til Posten og NAV, samt Helsedirektoratets korona-bot.

Felles for de offentlige tjenestene som har tatt i bruk chatbot for innbyggerkontakt er ønsket om å spare eller omplassere ressurser. Samfunnet digitaliseres stadig, og i takt med utviklingen blir også innbyggerne mer digitale. Måten vi interagerer på har endret seg drastisk de siste tiårene, og offentlig sektors satsning på chatboter, automatisering og selvbetjeningsløsninger er et resultat av dette. Ved at chatbotene tar seg av henvendelser av generell karakter, som enkle og rutinebaserte spørsmål, kan de ulike etatene heller fokusere på mer komplekse oppgaver som krever menneskelig interaksjon.

3.2 Chatbot i NAV

For å eksemplifisere casen fokuserer dette prosjektet på NAVs chatbot, Frida, som ble lansert i 2018. I 2017 fikk NAV tre millioner henvendelser fra brukere som var usikre på om de hadde krav på en ytelse eller om de hadde søkt på riktig måte. Mye av denne informasjonen finnes på nav.no (Ringnes, 2018). I 2018 gjorde NAV undersøkelser for å kartlegge brukerbehov, og de fant at mange brukere opplever nettsidene som et hav av informasjon, at de ofte er usikre på hva de har rett til og at de opplever språket på nav.no som vanskelig (NAV, 2018). Som del av digitaliseringen og for å øke tilgjengeligheten for innbyggere og avlaste veilederne som sitter på chat og telefon, har NAV derfor utviklet chatboten Frida. Frida kan svare på en rekke enkle spørsmål, er tilgjengelig døgnet rundt og lærer kontinuerlig av samtaler med brukere (innbyggere). Dersom Frida ikke kan svare på brukernes spørsmål, henvises de videre til en veileder (saksbehandler) i NAV, hvor de videre kan logge inn for å få svar på spørsmål rundt spesifikke saksforhold.

Arbeidet med Frida startet som et studentprosjekt tilbake i 2017, da en gruppe studenter utviklet den første svarbasen begrenset til temaet foreldrepenger. Frida ble lansert på NAVs hjemmesider i 2018, og var på det tidspunktet i stand til å besvare generelle spørsmål knyttet til foreldrepenger, engangsstønad og svangerskapspenge. Etter hvert som Frida kom i drift ble samtaleanalyser et viktig verktøy for å identifisere brukernes språk og for å gjøre Frida i best mulig stand til å besvare spørsmålene, og NAV har en rekke chatbot-trenere som daglig arbeider med å videreutvikle chatboten. Disse jobber med den automatiske dialogen på de ulike fagområdene, og er ansvarlige for å rette opp feil og mangler i dialogen samt oppdatere chatboten ved lovendringer eller endringer på nav.no.

NAV får daglig flere tusen henvendelser, og trykket har vært spesielt høyt etter utbruddet av covid-19. Mange innbyggere har blitt permittert eller mistet jobben og tatt kontakt med NAV for å se på mulighetene videre, i tillegg til alle de som er i kontakt med NAV av andre grunner. I denne perioden har Frida fungert som førstelinje-support og hjulpet mange brukere med spørsmål av generell karakter. Dette har også ført til at veilederne som sitter på telefon og chat har kunnet fokusert mer på å hjelpe innbyggere med behov for menneskelig rådgivning og mer komplekse saker.

For å oppmuntre innbyggere til å prøve chatboten, sender NAV ut tekstmeldinger til de som står i lange telefonkøer for å komme i kontakt med etaten. Frida tar seg av en stor andel av henvendelsene NAV får og kan svare på de fleste spørsmål knyttet til informasjon om NAV sine tjenester eller navigering på nettsidene. Chatboten er under kontinuerlig utvikling, og mates fortløpende med ny informasjon som kan være relevant for innbyggere. Samtaler med Frida er anonyme og krever ikke innlogging på nav.no. Samtale rundt spesifikke eller personlige saker må derfor gjøres med en av veilederne ved kontaktsenteret, enten ved hjelp av chat med innlogging eller gjennom telefon.

Kapittel 4

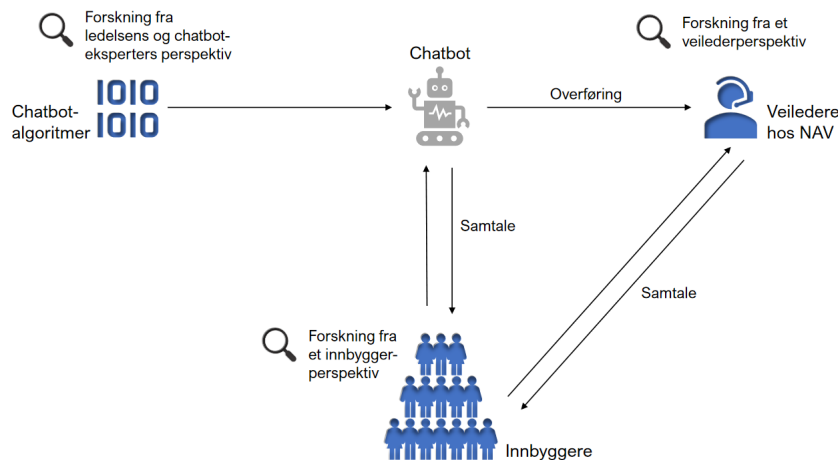
Metode

Dette kapitlet beskrives metodene brukt for å besvare oppgavens forskningsspørsmål. Her presenteres studiens forskningsdesign, endringer gjort i løpet av prosjektet samt aktiviteter knyttet til datainnsamling. Til slutt forklares prosessen med analyse av det innsamlede datamaterialet.

4.1 Forskningsdesign

Dette forskningsprosjektet omhandler mennesker og deres holdninger til bruk av chatboter i offentlig sektor, og for å besvare forskningsspørsmålet ble det valgt å gjennomføre en casestudie. Bakgrunnen for valget er ønsket om å fokusere på en instans av chatbot, med mål om å få en rik, detaljert innsikt i forholdet mellom bruker og chatbot, noe som gjør det egnet for en casestudie (Oates, 2005). Valget er delvis grunnet i praktiske årsaker, da NAV ønsket å få dypere innsikt i temaet og dermed var villige til å bidra med informanter til datainnsamlingen. Det at case velges mer pragmatisk på bakgrunn av tilgjengelighet er ikke nødvendigvis optimalt for størst mulig generaliserbarhet (Tjora, 2012), men NAV er i en unik posisjon med en stor, mangfoldig brukergruppe. Dette åpnet for en mulighet til å se på bruk av chatbot i en offentlig etat av betydelig størrelse og omfang.

Studien er en del av prosjektet Frida@Work som undersøker den kritiske rollen tillit spiller i bruk av chatboten Frida. Forskningsprosjektet er et samarbeid mellom NAV, Universitetet i Oslo, Universitetet i Agder og NTNU. Grunnet prosjektets natur er store deler av datainnsamlingen gjennomført i samarbeid med masterstudenter fra NTNU og Universitetet i Agder. Figur 4.1 viser de ulike perspektivene prosjektet Frida@Work tar for seg. Denne studien ser på innbyggerperspektivet, og fokuserer på relasjonen mellom chatbot og innbygger.



Figur 4.1: Oversikt over de ulike perspektivene som dekkes av prosjektet Frida@Work.

4.2 Datainnsamling

Dette delkapittelet beskriver metodene som ble benyttet for datainnsamling i forbindelse med forskningsprosjektet. Tabell 4.1 viser en oversikt over datainnsamlingsaktiviteter gjennomført i løpet av studien. Aktivitetene bestod av intervjuer med informanter fra NAV og en spørreundersøkelse for brukere av chatboter.

4.2.1 Intervju

For å få detaljert informasjon knyttet til forskningsspørsmålet, ble det valgt å gjennomføre intervjuer med informanter fra NAV. Intervjuer er godt egnet når man ønsker å få innsikt i komplekse spørsmål knyttet til fenomenet det forskes på, samt for å utforske erfaringer og følelser (Oates, 2005). Rekruttering til intervjuene foregikk gjennom en kontaktperson i NAV, som forsøkte å hente inn informanter i relevante roller knyttet til arbeidet med chatbot og brukerkontakt i etaten. Dette inkluderte både informanter på ledernivå med innsikt i overordnede kommunikasjonsstrategier og chatbot-trenere som jobber med den daglige driften og videreutviklingen av chatboten. Intervjuene bidro med dybdeinnsikt i informantens perspektiver på temaer som tillit, forventningsstyring og kompetanse hos chatboter, samtidig som de tilførte informasjon rundt chatbotens oppbygging.

Uke	Aktivitet	Lokasjon	Informant
2	Informasjonssamtale	Digitalt	Leder 1 (L1) og kontaktperson i NAV
7	Intervju	Digitalt	Chatbot-trener 1 og 2 (CT1 og CT2)
11-14	Spørreundersøkelse	Digitalt	120 anonyme respondenter
15	Intervju	Digitalt	Chatbotkoordinator 1 (CK1)
15	Intervju	Digitalt	Chatbot-trener 1 og 3 (CT1 og CT3)
15	Intervju	Digitalt	Chatbot-trener 2 og 4 (CT2 og CT4)
15	Intervju	Digitalt	Chatbot-trener 5 og 6 (CT5 og CT6)
16	Intervju	Digitalt	Leder 2 (L2)
16	Intervju	Digitalt	Chatbotkoordinator 2 (CK2)
17	Intervju	Digitalt	Leder 1 (L1)
17	Intervju	Digitalt	Leder 3 (L3)

Tabell 4.1: Oversikt over datainnsamlingsaktiviteter gjennom studien

Som følge av covid-19 ble alle intervjuene gjennomført digitalt på Teams. Alle intervjuene var semistrukturerte, og i flere tilfeller fikk informantene tilsendt intervjuguiden i forkant av intervjuet. Dette ble gjort for å best mulig kunne imøtekomme informantenes hektiske arbeidshverdag, men også for at kontaktpersonen vår i NAV kunne fordele de ulike temaene i intervjuguiden på informantene som

var best egnet til å besvare spørsmålene. Hvert intervju varte i rundt en time, og det var varierende hvor mange informanter som var til stede. Intervjuene ble tatt opp etter samtykke.

For å utnytte tiden vi hadde til rådighet med informantene best mulig, fant vi det hensiktsmessig å dele inn intervjuet etter prosjektene til de ulike studentgruppene. Alle masterprosjektene tilhørte Frida@Work, men vi fokuserte på ulike tema og så på forskjellige brukergrupper innad i etaten. Vi utviklet derfor hver vår intervjuguide og fordelte tiden tilgjengelig likt mellom de ulike gruppene. Jeg utarbeidet intervjuguidene i fellesskap med en medstudent, og spørsmålene ble forsøkt tilpasset informantens roller i etaten og tilknytning til arbeidet med chatboten, men de overordnede temaene var de samme i alle intervjuene. Vi forsøkte å stille enkle, fokuserte spørsmål som kunne gi oss en forståelse av hvordan de jobber og videreutvikler chatboten, samtidig som vi ønsket å få innsikt i hvilke rolle brukerne spiller i prosessen.

I starten av datainnsamlingen var spørsmålene i stor grad basert på litteraturen vi hadde lest innledende i prosjektet. Denne utviklet seg underveis, etter som vi fikk dypere innsikt i hvordan NAV videreutvikler chatboten samt problemstillingene de arbeider med daglig. Intervjuguiden ble revidert i flere runder, både etter hvert som vi hadde mer informasjon å jobbe med og etter hvert som nye spørsmål oppsto. Den endelige intervjuguiden finnes i vedlegg A.1.

4.2.2 Spørreundersøkelse

For å bedre belyse brukeres holdninger til bruk av chatbot i offentlig sektor ble det valgt å gjennomføre en spørreundersøkelse. Bakgrunnen for valget om å bruke en spørreundersøkelse i stedet for intervjuer handlet i stor grad om at tilgang til enkelte brukere ble utfordrende med hensyn til personvern og tillatelser. Videre så vi også nytten i å undersøke bredden og få så mange perspektiver som mulig, i motsetning til intervjuene med informantene i NAV som bidro med dybdeinformasjon. Undersøkelsen hadde som hensikt å samle informasjon fra mennesker som hadde vært i kontakt med en chatbot i offentlig sektor. Fokuset for undersøkelsen var på forventninger før bruk, opplevelse av kompetansenivå og etablering av tillit til systemet de interagerer med.

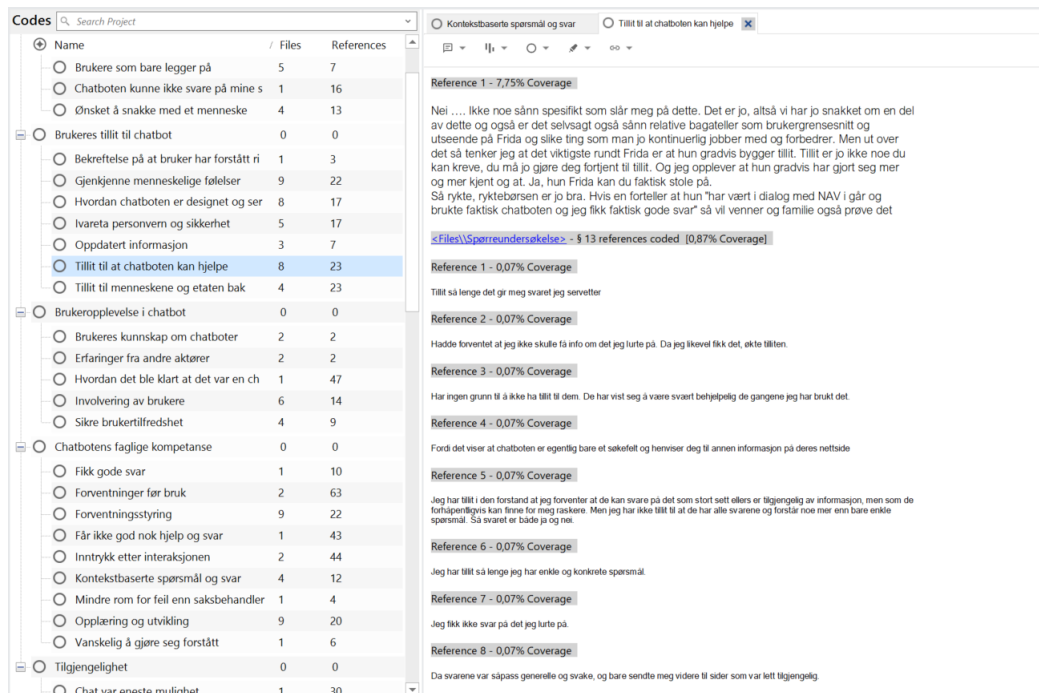
Undersøkelsen ble utarbeidet i samarbeid med en medstudent som også var del av prosjektet Frida@Work og inneholdt 18 spørsmål som hadde hensikten å fange informasjon som var relevant for begge prosjektene. Vi ønsket å få både kvalitative og kvantitative data, og valgte derfor en blanding av åpne og lukkede spørsmål. Dette for å få en rikere innsikt enn ved å bare benytte flervalgsspørsmål. Ulempen med åpne spørsmål er at undersøkelsen tar lengre tid å gjennomføre, og at det kan føre til at noen ikke tar seg tid til å besvare. Vi så likevel behovet for kvalitative data fra brukere, og fordelene ble vurdert store nok til at vi heller valgte å risikere færre svar. For å få flest mulig respondenter ble undersøkelsen delt på sosiale medier og på interne sider for NTNU, samtidig som vi fikk hjelp av veileder til å dele med ansatte ved instituttet og andre som kunne ha vært i kontakt med en chatbot. Spørreundersøkelsen finnes i vedlegg A.2.

Spørreundersøkelsen ble besvart av 120 respondenter, hvorav 58,3% var kvinner og 37,5% var menn. Aldersfordelingen var følgende: 15,8% mellom 18 og 25 år, 34,2% mellom 26 og 35 år, 20,8% mellom 36 og 45 år, 23,3% mellom 46 og 55 år, 5% mellom 56 og 65 år og 0,8% over 65 år. På spørsmål om høyeste fullførte utdanning svarte 0,8% grunnskole, 10% svarte videregående skole og 89,2% svarte høgskole eller universitet.

4.3 Dataanalyse

For å unngå premature konklusjoner og opprettholde en gjennomgående god systematikk, fulgte dataanalysen den stegvis-deduktive-induktive (SDI) modellen utviklet av Tjora (Tjora, 2012). SDI-modellen for analyse innebærer å jobbe med det empiriske materialet mot en teoretisk, konseptuell forståelse, samtidig som deduktive tilbakekoblinger sikrer at teorien kontrolleres mot det empiriske materialet. Målet med forskningen er konseptuell generalisering.

Den første fasen av analysen er induktiv, og starter med rådata fra intervjuer og spørreundersøkelse. I denne fasen måtte datamaterialet først og fremst være på det samme, tekstlige formatet. Dette innebar å transkribere alle intervjuene og hente ut de åpne svarene fra spørreundersøkelsen. Videre ble datamaterialet kodet. Her var målet å (1) ekstrahere essensen i det empiriske materialet, (2) å redusere materialets volum, og (3) å legge til rette for idégenerering på grunnlag av detaljer i empirien (Tjora, 2012). Målet var å få til en så åpen og empirinær



Figur 4.2: Skjerm bilde av kodeprosessen i Nvivo.

koding som mulig. Dette innebar å finlese hver linje i empirien for så å benytte koder som ligger tett på deltakerutsagn og dermed ivareta det spesifikke i materialet i stedet for å benytte koder fra eksisterende litteratur og forskning (Oates, 2005; Tjora, 2012). Videre grupperes kodene, med hensikt om å sitte igjen med et sett hovedkonsepter som analysen kan bygge videre på. Kodingen ble i dette prosjektet gjennomført i verktøyet NVivo. Bakgrunnen for dette var ønsket om en oversiktlig og strukturert kodeprosess med full oversikt over koblinger mellom koder og tilhørende empiri. Et skjermbilde av kodeprosessen i NVivo kan ses i figur 4.2.

Den andre fasen av analysen er deduktiv og består av en rekke tester som danner utgangspunkt for iterasjoner mellom stegene i modellen. Konseptene identifisert i første fase grupperes i nye konseptuelle kategorier, og denne prosessen gjentas for å utvikle tema og bevege seg mot teori (Tjora, 2012). For denne studien ble tre iterasjoner gjennomført før resultatet var et sett overordnede temaer og konseptuelle kategorier med tilhørende empiri. Tabell 4.2 viser en oversikt over temaer og utvalgte kategorier med eksempler på tilhørende rådata.

Tabell 4.2: Oversikt over tema, konseptuelle kategorier og eksempler og utdrag fra datamaterialet

Tema	Konseptuelle kategorier	Eksempler og utdrag
Tilgjengelighet	For å nå et menneske	”Er chatbot obligatorisk og jeg har et spesifikt spørsmål den ikke kan svare på, så handler det om å lure chatbot til å la meg snakke med et menneske.” (Anonym respondent, spørreundersøkelse).
	Brukere som bare legger på	”Men det betyr jo også at det er en del som sannsynligvis tar kontakt som kanskje ikke.. som da fort skjønner at det er en chatbot og ikke ønsker å ... og mener at dette er feil kanal.” (Leder, intervju).
Brukeropplevelse i chatbot	Brukeres kunnskap om chatboter	”Problemet er jo at det er jo ... folk ... veldig mange vet veldig lite om hvordan de skal kommunisere med en chatbot. Og derfor starter de ofte samtalen sånn som de starter med et menneske.” (Chatbot-trener, intervju).
Chatbotens faglige kompetanse	Forventningsstyring	”Og etter hvert har vi også bygd ut innholdet så mye i chatboten, at det å for eksempel gi en sånn liste på hva gjelder det, da måtte vi ha hatt det så overordnet at det hadde vært vanskelig å treffe uansett hva det gjelder.” (Leder, intervju).
	Kontekstbaserte spørsmål og svar	”Jeg føler ikke de skjønner problemet. Hvis det er enkle spørsmål som hvor finner man ditt og datt har jeg tiltro, men om jeg lurer på noe rundt skattemeldingen feks stoler jeg ikke på at de tar alle aspekter med i vurderingen” (Anonym respondent, spørreundersøkelse).

Utfordrende å gjøre seg forstått	”Den forstod ikke spørsmålet og sendte meg i ring på nettsider som jeg allerede hadde sjekket.” (Anonym respondent, spørreundersøkelse).
Fikk gode svar	”The chatbot actually resolved my inquiry. As I understood if the chatbot wasn’t able to resolve the inquiry it would automatically forward me to a human operator.” (Anonym respondent, spørreundersøkelse).
Brukeres tillit til chatbot	Tillit til at chatboten kan hjelpe ”Så, ja, jeg vil si at det er en del tillit til innholdet, men det er vanskelig å si hvor mye som går på tillit og hvor mye som går på et behov for å gjenta og bekrefte, som vi ikke helt klarer å løse i chatbot.” (Chatbot-trener, intervju).
Tillit til menneskene og etaten bak	”På samme måte som jeg har stor tillitt til det offentlige, ser jeg på chat-bot’en som en troverdig kilde for informasjon” (Anonym respondent, spørreundersøkelse).
Ivareta personvern og sikkerhet	”Usikker på om det jeg skriver ikke blir delt noen plasser som det ikke skal deles på” (Anonym respondent, spørreundersøkelse).
Hvordan chatboten er designet og ser ut	”... vi ser jo det at det er noen som tror at Frida er et menneske, og kanskje hvis at det hadde vært mer robotaktig, både utseende og navn da, at folk hadde kommunisert på en litt annen måte også.” (Chatbot-trener, intervju).

I den første delen av analysen samarbeidet jeg med en medstudent for å validere analyseprosessen. Vi var begge involvert i datainnsamlingen og hadde tilgang

til det samme datamaterialet, men vinklingene og forskningsspørsmålene var ulike. Diskusjoner rundt datamaterialet og kodingarbeidet opplevdes som svært nyttig, både med veileder og medstudent. Det ble likevel nødvendig å legge vekk samarbeidet videre i skriveprosessen for å unngå å ha for stor innflytelse på hverandres analyse og for å sikre individuelle vinklinger på oppgavene.

I henhold til SDI-modellen skal det underveis i analysearbeidet gjennomføres en rekke deduktive tester med mål om å ta ut potensialet i den empirien man har generert. Dette inkluderer blant annet kodetester som skal hjelpe til å unngå sorteringsbasert koding, og heller utvikle empirinære koder som beskriver hva som kommer frem i datamaterialet (Tjora, 2012). Hensikten er å sitte igjen med et sett koder man umulig kunne lagd før kodingen. Den første intervjuguiden tok utgangspunkt i litteraturen vi hadde lest forberedende, og fokuserte derfor på det vi så på som viktige konsepter og tema å undersøke. Denne intervjuguiden var begrenset til få spørsmål av en mer utforskende natur for å få innblikk i hvilke tema som burde undersøkes videre, og den ble gradvis mer detaljert etter hvert som vi fikk dypere innsikt i temaene gjennom intervjuer og spørreundersøkelse. De overordnede temaene funnet i litteraturen forble likevel de samme. Mange av kodegruppene jeg satt igjen med var derfor ikke overraskende, men heller forventet fra start med utgangspunkt i tidligere forskning på området. Den ferdige intervjuguiden er tilgjengelig i vedlegg A.1.

Kapittel 5

Empiriske funn

I dette kapitlet presenteres funnene fra datainnsamlingen. Her trekkes anonymiserte sitater fra intervjuer med informanter i NAV samt tekstlige besvarelser fra spørreundersøkelsen frem. Sitater som gjengis fra spørreundersøkelsen er kopiert fra svarskjemaet, og eventuelle skrivefeil vil derfor være slik de forekommer i svarene. Kapitlet er delt inn etter temaene funnet i analyseprosessen og uthevet tekst og underoverskrifter tilsvarende konseptuelle kategorier tilhørende de ulike temaene.

5.1 Tilgjengelighet

Raske svar og enkelt tilgjengelig. For å få innblikk i innbyggeres holdninger til bruk av chatboter i offentlig tjenesteleveranse, fikk respondentene innledende i spørreundersøkelsen spørsmål om hvorfor de valgte å benytte chatbot for kontakt med den aktuelle offentlige aktøren. Majoriteten svarte at de benyttet seg av chatbot fordi den var enkelt tilgjengelig, de ønsket å få raskt svar eller de ønsket ikke å stå i telefonkø. Mange så på chatbot som en lavterskel kommunikasjonskanal som muliggjorde kontakt døgnet rundt.

Chat var eneste mulighet. Andre opplevde at valget ble tatt på deres vegne ved at andre kommunikasjonskanaler var godt skjult eller vanskelige å finne, eller ved at chatbot var eneste alternativ fordi de ikke kom gjennom på telefon eller det var utenfor åpningstidene for andre kanaler. Brukere som benyttet seg av kommersielle chatboter oppga i større grad at de følte seg "tvunget" til å bruke chatbot da andre kommunikasjonskanaler enten ikke var tilgjengelige eller hadde begrenset kapasitet. I intervju med en av lederne i NAV ble det påpekt at de ikke ønsker å tvinge innbyggere til å bruke hverken den ene eller den andre kanalen, men heller legge til rette for flere kanaler slik at brukeren kan ta avgjørelsen selv basert på hvilken type henvendelse de har.

5.1.1 Avbrutte henvendelser

Brukere som bare legger på. I intervjuene kom det frem at NAV opplever at noen brukere avbryter interaksjonen med chatboten med en gang etter velkomstmeldingen, uten at de skriver noe. Flere av informantene påpeker at disse samtalene ofte ses på som uhell, at en bruker har sett seg rundt på nettsiden og har trykt på knappen av ren nysgjerrighet, uten at hensikten er å starte en samtale. En annen informant nevnte at det kan være mange årsaker til at brukere avslutter samtalen uten å skrive, og at det også kan være brukere som åpner chatboten, forstår at det ikke er et menneske, og så velger å ringe eller møte opp på kontoret i stedet. Denne gruppen er derimot vanskelig å kartlegge og få informasjon om, og det ble påpekt av flere informanter at etaten har for lite kunnskap om disse brukerne.

En annen gruppe brukere er de som ikke avslutter samtalen med chatboten på "korrekt" måte, altså de som avslutter uten å gi noe indikasjon på om svaret de fikk på sin henvendelse var nyttig eller riktig. Av respondentene i spørreundersøkelsen hadde flere avbrutt samtalen med chatbot. De fleste pekte på det at de ikke fikk svar på det de lurte på som grunnen til dette, eller at chatboten ikke forsto spørsmålet. Noen av respondentene opplevde å havne i en *"endeløs loop med unyttig informasjon"* (Anonym respondent, spørreundersøkelse), og endte med å ringe inn til NAV istedet.

For å nå et menneske. Videre kom det frem både i intervjuene og spørreundersøkelsen at mange brukere ønsker å snakke med en menneskelig veileder, og starter samtalen med chatboten med et ønske om å bli satt over til et menneske. De fleste offentlige aktørene som har tatt i bruk chatboter benytter disse som førstelinje brukerkontakt, i tillegg til telefon. Det innebærer at brukere som ønsker å snakke med et menneske via chat må overføres ved hjelp av chatboten. En av respondentene svarte følgende om kanalvalg; *"Er chatbot obligatorisk og jeg har et spesifikt spørsmål den ikke kan svare på, så handler det om å lure chatbot til å la meg snakke med et menneske."* (Anonym respondent, spørreundersøkelse). På spørsmål om bakgrunnen for ønsket om å snakke med et menneske, svarte en annen at *"Mennesker kan gi svar. Chatbots er bare til irritasjon. Målet var alltid mennesker. Det første jeg sier er alltid: "Jeg vil snakke med et menneske." I beste fall kan den sende meg til riktig person."* (Anonym respondent, spørreundersøkelse).

En av de andre lederne påpekte at NAV ser at brukere lærer seg en slags strategi for å komme raskest mulig til et menneske, gjerne uten å ha forsøkt å stille spørsmål i chatbot først. Videre nevnes det at det skal være lett for brukere å komme til et menneske når behovet er der, men at chatboten skal fungere som et filter for å stoppe de som ikke trenger å gå videre. Flere respondenter svarte at de allerede hadde søkt rundt på nettsidene til etaten, og at de var sikre på at chatboten ikke kunne tilføye noe ut over den generelle informasjonen de hadde funnet selv. Brukere som hadde spørsmål knyttet til spesifikke saksforhold så heller ikke chatboten som en hensiktsmessig kanal for henvendelsen. Det ble derimot påpekt gjennom intervjuene at mye av den informasjonen NAV oppfatter at brukere tror de trenger en menneskelig veileder og innlogging for å få, kan chatboten også svare på.

5.2 Brukeropplevelse i chatbot

NAV måler brukeropplevelse i chatbot gjennom at brukere får et kort tilbakemeldingsskjema etter samtale i Frida, i tillegg til at chatboten er del av den årlige brukerundersøkelsen. Tilbakemeldingsskjemaet er ment å gi indikasjoner på løsningsgrad, altså om brukeren fikk løst saken sin i chatbot, samt hva de ønsket informasjon om og hvordan brukeren opplevde interaksjonen med chatboten. Innbyggere som tar kontakt gjennom telefon og e-post får de samme spørsmålene, og hensikten er å kunne måle de ulike kanalene opp mot hverandre og videreutvikle innholdet.

5.2.1 Opplæring av brukere

Brukeres kunnskap om chatboter. I intervjuene kom det frem at mange brukere ikke nødvendigvis er vant til å interagere med chatboter, og at dette kan vanskeliggjøre samtale mellom bruker og chatbot. Chatbot-trenere opplever at noen brukere har en tendens til å skrive lange paragrafer med bakgrunnsinformasjon i møte med chatboten, fordi det er slik de er vant til å kommunisere. I et av intervjuene uttrykte en av informantene at *"...veldig mange vet veldig lite om hvordan de skal kommunisere med en chatbot. Og derfor starter de ofte samtalen sånn som de starter med et menneske. De har en historie å fortelle, ikke sant, det er en bakgrunn, før de kommer til poenget."* (Chatbot-trener, intervju). Dette gjør det svært vanskelig for chatboten å trekke ut nøkkelord og gi relevante svar i den gitte konteksten.

Dette kommer også frem på brukersiden. På spørsmål om hvordan de forventer at chatboten skal svare, skriver en av respondentene i spørreundersøkelsen følgende: *"På en måte som er relevant for den gitt situasjonen jeg presenterer for chatboten. Siden den ikke har evnen til å se det bredere skopet, er den heller ikke veldig skikket for å gi råd, som ultimat sett er som oftest det man trenger i slike tilfeller."* (Anonym respondent, spørreundersøkelse). Svaret tyder på at respondenten har gått inn i samtale med chatboten med håp om å få råd i en gitt situasjon, og at dette ikke skjedde. Det kan også tyde på at respondenten har presentert chatboten med en viss mengde informasjon og hatt en forventning om at den så skulle forstå hvilken kontekst brukeren stilte spørsmålet i, samt ta hensyn til annen relevant bakgrunnsinformasjon i sitt svar. En annen respondent skriver at *"jeg sitter med en følelse at jeg må skrive de riktige ordene, "game" systemet heller enn å beskrive det jeg lurer på. kanskje pga. lite erfaring med det. litt som å google "smart".* (Anonym respondent, spørreundersøkelse). Som respondenten påpeker vil lange beskrivelser av situasjonen fungere dårlig i samtale med en chatbot, da det gjør det utfordrende å hente ut nøkkelord og definere kontekst.

Det trekkes frem at NAV har sett en bedring på dette punktet, og det er tydelig at brukere blir mer vant til å kommunisere med chatboter og derfor også flinkere til å formulere spørsmål. En informant nevner at det også kan være generasjonsforskjeller som spiller inn: *"Men dette tror jeg også er et generasjonsskifte. At den unge befolkningen forstår dette, mens den eldre kanskje aldri vil forstå. At vi må leve med det i noen år og så ... Mens de som er i midten, de kan lære seg."* (Leder, intervju). NAV jobber også med å integrere en video og veiledningsmateriale i chatboten som sier noe om hvordan den fungerer og hvordan brukeren skal skrive for å få best mulig svar.

Erfaringer fra andre aktører. I flere av intervjuene ble innbyggernes erfaringer med chatboter på generelt basis trukket frem som et hinder for å ta i bruk teknologien. I Norge benytter flere av de offentlige aktørene den samme chatbotleverandøren, men det er stor variasjon i hvor langt de ulike aktørene er kommet i utvikling av sin chatbot, f.eks. utvikler NAV sin egen frontend i motsetning til mange andre aktører. De offentlige etatene i Norge som jobber med chatbot har et samarbeidsforum hvor de møtes for å utveksle erfaringer og kunnskap. Det at flere jobber med den samme teknologien ble trukket frem som gunstig i intervju-

ene, fordi det betyr at Norges innbyggere får trening i å forstå og vite hvordan chatboter virker og hvordan de skal stille spørsmål for å få best mulig hjelp. Det ble likevel påpekt at dette kan ha negative konsekvenser, da brukere kan bli preget av at andre aktører har kommet kortere og leverer dårligere løsninger enn andre, slik at tilliten til chatboter generelt blir lavere.

Da chatboter først ble lansert var det også store forskjeller i hvilken teknologi som lå bak og hvor god denne teknologien var. Videre benytter mange private aktører seg av chatboter for kundeservice, kjøpshjelp o.l., og erfaringer med disse chatbotene kan også påvirke det overordnede inntrykket innbyggere har av chatboter. En av lederne påpekte at *"... det gjør også at terskelen, tror jeg, for å ta i bruk dette, kan være stor hos mange, at mange synes dette er en dårlig løsning."* (Leder, intervju). Informanten uttrykker at dårlige erfaringer med chatboter fra tidligere, både i offentlig og privat sektor, kan føre til at brukeren ikke ønsker å benytte seg av chatbot for kommunikasjon med etaten, uavhengig av om tidligere erfaringer er med den aktuelle chatboten.

5.3 Chatbotens kompetansenivå

Intervjuene og spørreundersøkelsen ble brukt for å forsøke å få innsikt i hvordan NAV og innbyggere som har brukt chatbot i offentlig sektor ser på chatbotenes faglige kompetansenivå. Intervjuene avdekket informasjon knyttet til hvordan chatboten er bygd opp samt hvordan videreutvikling og opplæring foregår, samt hvilke brukerhensyn som tas i denne prosessen. Spørreundersøkelsen avdekket informasjon knyttet til brukernes forventninger til en chatbot sitt kompetansenivå, samt oppfatning etter faktisk bruk og erfaring.

5.3.1 Videreutvikling av chatbot og innhold

Opplæring og utvikling. For å videreutvikle chatboten og dens innhold, benytter NAV seg av chatbot-trenere. Disse jobber kontinuerlig med å forbedre kvaliteten på innholdet, sikre at svarbasen er oppdatert i henhold til gjeldende regelverk samt utvide chatbotens svarbase med nye fagområder. Chatbot-trenerne er også selv i direkte kontakt med innbyggere som veiledere på chat eller telefon. I intervju med en av lederne ble følgende sagt: *"Jeg tror det er ganske mange andre virksomheter, de har folk som sitter kun på fulltid og mater chatboten, og har knapt*

nok truffet en bruker, men vi har valgt en helt annen strategi på det.” (Leder, intervju). Det at chatbot-trenerne ikke utelukkende jobber med å utvikle innholdet i chatboten, men også møter brukere i ulike kanaler og dermed har innsikt i hva brukere søker informasjon om og hvordan de formulerer seg i spørsmål, trekkes frem som en fordel.

Når Frida skal utvides med et nytt fagområde, benytter NAV samtaleanalyse for å finne ut hva brukerne spør om og hva chatboten burde kunne svare på. Dersom chatboten ofte får spørsmål den ikke forstår eller svarer feil på, gir det en indikasjon om at det er noe som burde implementeres i svarbasen. Utvikling av et nytt fagområde i chatbot starter alltid med en manuell chat med en menneskelig veileder. Dette gir chatbot-trenerne innsikt i hvilke spørsmål som går igjen, hvordan brukere formulerer spørsmål og hva som er enkelt nok til at chatboten kan svare. Chatloggene fra den manuelle chatten brukes deretter til å utvikle svarbasen på det nye fagområdet.

Kontekstbaserte spørsmål og svar. NAV har den siste tiden jobbet med å få chatboten til å bedre forstå hvilken kontekst brukere tar kontakt i. Brukere befinner seg i ulike situasjoner, og dersom en bruker tar kontakt med chatboten for å stille spørsmål om sykepenger, vil svaret være et annet hvis brukeren er arbeidsgiver enn om brukeren spør på egen vegne. I den sammenheng har NAV implementert et filter i Frida hvor brukeren kan velge om hen tar kontakt som arbeidsgiver eller personbruker, slik at chatboten kan gi svar basert på brukerens faktiske situasjon.

En av respondentene i spørreundersøkelsen uttrykte at *”Problemet er ikke ”faglig” kompetanse, problemet er å greie å finne fram til informasjonen som er relevant for meg i den gitte konteksten”* (Anonym respondent, spørreundersøkelse). Flere svar i spørreundersøkelsen indikerer at brukere er usikre på om svarene de får gjelder i deres situasjon, at det er en usikkerhet knyttet til om chatboten tar alle hensyn i betraktning og at det oppleves som tryggere å snakke med et menneske som kan bekrefte dette.

5.3.2 Brukerforventninger

Forventningsstyring. I oppstartsfasen av denne studien, i januar 2021, ble velkomstmeldingen i NAVs chatbot endret. Tidligere sa den at brukere måtte stille korte, konsise spørsmål og gjerne nevne hvilken stønad eller ytelse det gjelder for at chatboten best skulle forstå henvendelsen. Denne ble endret til en beskjed om behandling av personopplysninger som brukere må trykke vekk før samtalen med chatboten starter. Beskjeden sier at dette er en anonym chat hvor du kan få svar på generelle spørsmål om NAV sine ytelser og tjenester, samtidig som den understreker at brukeren må unngå å skrive sensitive opplysninger av hensyn til egen sikkerhet. I intervjuene ble bakgrunnen for denne endringen forklart med et ønske om å holde introduksjonen kort og konsis. Flere av chatbot-trenerne opplevde at det å møte brukeren med en vegg med tekst idet interaksjonen starter virket mot sin hensikt, og at mange lot være å lese den. I et av intervjuene ble følgende sagt om den nye løsningen: *"Så det vi har landet på er jo å få en ryddig og oversiktlig velkomstmelding som brukerne faktisk kanskje leser da."* (Chatbot-trener, intervju).

Forventninger før bruk. Svarene i spørreundersøkelsen viser at brukere har ulike forventninger til chatbotens kompetansenivå før de starter en samtale. Noen har forventninger om å få svar på spørsmål som går ut over generell informasjon, for eksempel at chatboten kan gi informasjon om status i brukerens sak. Svarene i spørreundersøkelsen viser at omkring 20% av respondentene tok kontakt med chatbot for oppfølging av søknad eller for å sende inn eller opprette ny søknad. En av respondentene svarte på spørsmål om forventninger før bruk *"At den kunne svare på spørsmål jeg ikke fant frem til via et medium grundig Google-søk."* (Anonym respondent, spørreundersøkelse). Et annet svar på det samme spørsmålet var *"De erstatter mennesker, og jeg forventer da gode svar - les kompetente svar."* (Anonym respondent, spørreundersøkelse).

En annen gruppe brukere har forventninger om å få svar på spørsmål av generell karakter i chatbot. Disse brukerne forventer hjelp til å finne frem til et skjema, at chatboten kan vise til informasjon på nettsidene eller svare kort på helt enkle spørsmål. Omtrent halvparten av respondentene i spørreundersøkelsen tok kontakt med chatboten med mål om å få generell informasjon. Dette oppleves også som en ønskesituasjon av NAV, at brukere har forventninger som stemmer over-

ens med chatbotenes reelle kompetansenivå. I av intervjuene ble det uttrykt at *”Vi ønsker at de skal forstå at det kun er generelle spørsmål, om regelverk og skjema og sånne ting. De kan ikke spørre om sin sak.”* (Chatbotkoordinator, intervju).

Andre hadde lave forventninger til chatbotens kompetansenivå, og gikk inn i samtalen med lite håp om å få svar på det de lurte på. Enkelte påpeker at de har så lite tro på å få svar på sin henvendelse i chatbot at de ikke ønsker å bruke tid på en samtale med chatboten i det hele tatt, og heller forsøker å bli overført til en menneskelig veileder. På spørsmål om forventninger til chatbotens faglige kompetanse svarte en av respondentene at *”Null forventninger. Jeg ønsker å bli satt over til et menneske så fort som mulig.”* (Anonym respondent, spørreundersøkelse). For NAV er chatboten ment å fungere som et filter før brukere eventuelt blir satt over til veileder. Brukere som ber om å bli satt over til veileder med en gang får dette ønsket oppfylt, men da får ikke chatboten mulighet til å forsøke å svare på spørsmålene de har først.

Av de med lave forventninger var det også noen som trakk frem usikkerhet rundt hva chatboten kunne bidra med som bakgrunnen for dette. De gikk inn i interaksjonen med lave forventninger om å få svar på det de lurte på, men valgte likevel å initiere samtale og stille spørsmål. En respondent uttrykte at hen ikke forventet å få hjelp i chatbot, og at det var en positivt overraskelse når chatboten ga et hjelpsomt svar. Denne gruppen brukere skiller seg fra den forrige ved at de har lave forventninger, men likevel forsøker å få svar i chatbot før de eventuelt ber om å bli satt over til et menneske.

Flere av informantene påpeker i intervjuene at det er en utfordring å få brukere til å forstå hva de kan få svar på i chatbot. De ser at det oppstår misforståelser rundt hva generell informasjon innebærer, og at det ikke nødvendigvis er klart for alle at dette utelukker muligheten til å f.eks. kunne se status i egen sak gjennom chatbot. Det er likevel en oppfatning i NAV om at de fleste brukere forstår at en chatbot kun kan hjelpe på et overfladisk plan, og at brukere som ønsker informasjon knyttet til mer komplekse og personlige saker må overføres til et menneske for å oppnå dette. Fra spørreundersøkelsen kom det frem at mange brukere har høyere forventninger til chatboter i offentlig sektor enn de nødvendigvis klarer å innfri i interaksjoner. Det handler ikke nødvendigvis om at brukere ønsker svar på

spørsmål knyttet til egen sak eller med komplekse saksforhold, men at chatboten enkelt og greit ikke klarer å tolke forespørselen og trekke ut riktige nøkkelfaktorer for å gi et korrekt svar. Undersøkelsen viser også at mange blir satt over til et menneske eller avbryter samtalen fordi de ikke får svar på sine spørsmål i chatbot.

5.3.3 Oppfatning etter bruk

Brukernes oppfatning av chatbotens faglige kompetansenivå etter bruk var også varierende. Mange oppfattet kompetansenivået som grunnleggende, enkelt og begrenset. En respondent svarte på spørsmål om inntrykk av chatboten etter bruk at *"Dersom oppgavene er enkle og spesifikke nok så har jeg erfart at de klarer å gjøre jobben greit."* (Anonym respondent, spørreundersøkelse). Flere påpeker at de som regel fungerer godt til å besvare generelle spørsmål til brukere som ikke har lest seg opp på forhånd, eller som sliter med å finne riktig informasjon.

Utfordrende å gjøre seg forstått. Svarene i spørreundersøkelsen viser at flere ser at hjelpen de får avhenger av hvordan de selv formulerer henvendelsen, og at de får gode svar når de stiller enkle og konkrete spørsmål. En av respondentene skriver at *"Man må tilpasse sin skriftlig språk for å få ut informasjon man ønsker."* (Anonym respondent, spørreundersøkelse). Enkelte er bevisst egen språkbruk og formulering i møte med chatbot, og oppfatter at den beste måten å få svar på er å forenkle spørsmål og justere skrivemåten. På spørsmål om sin opplevelse av chatbotens kompetansenivå skrev en annen respondent at *"Den forstod ikke spørsmålet og sendte meg i ring på nettsider som jeg allerede hadde sjekket."* (Anonym respondent, spørreundersøkelse). Det kan tyde på at brukeren hadde behov for svar som gikk ut over den generelle informasjonen chatboten kunne tilby. Flere av de som oppfatter kompetansenivået som dårlig etter interaksjonen, skriver også at målet med samtalen var oppfølging av søknad eller status på egen sak, og at de endte med å avslutte samtalen uten svar eller ble overført til et menneske. En av respondentene som oppfattet kompetansenivået som dårlig skrev videre at *"De har all verdens informasjon tilgjengelig; problemet er å ikke greie å analysere situasjonen nok til å finne det som er relevant, eller i nevneverdig grad kunne gjøre anbefalinger for å hjelpe til å ta en avgjørelse."* (Anonym respondent, spørreundersøkelse).

Fikk gode svar. Svarene i spørreundersøkelsen indikerer også at mange brukere får tilfredsstillende svar i chatbot. Flere svarer at de opplevde kompetansenivået som godt nok, og at det stemte overens med forventningene de hadde før interaksjonen. En respondent skrev; *”Har stort sett fått svar på det jeg lurte på de gangene jeg har brukt det.”* (Anonym respondent, spørreundersøkelse). En annen påpeker at hen ble positivt overrasket over at chatboten kunne svare på spørsmålet, men at hen også hadde forstått at samtalen uansett ville blitt overført til en veileder som kunne hjelpe dersom det ikke var tilfelle.

5.4 Tillit til chatboten

Innledende i spørreundersøkelsen ble respondentene spurt om de har tillit til automatiserte tjenester i det offentlige, og chatbot nevnes som eksempel på dette. Fordeelingen mellom de som svarte ”ja”, det som svarte ”nei” og de som ikke visste var forholdsvis lik, med en viss overvekt på ”ja”. Av de som oppgir at de har tillit til automatiserte tjenester i det offentlige, trekker flere frem tillit til teknologien som bakgrunn for dette. En av respondentene skriver at *”Jeg tror at svarene genereres ut fra algoritmer som er pålitelige”* (Anonym respondent, spørreundersøkelse). Det at automatiserte tjenester som chatboter opererer etter regelbaserte algoritmer pekes på som positivt for utvikling av tillit, og det nevnes også at datamaskiner har mindre rom for personlige feil enn en menneskelig motpart.

Tillit til at chatboten kan hjelpe. Gjennom intervjuene kommer det frem at flere av informantene oppfatter at brukere flest har tillit til chatboten. Det påpekes også at chatboten ikke står alene som etatens kommunikasjonskanal, men at den er en del av en større kanalmiks som skal gi brukeren verdi. Om dette sa en av lederne at ønsket er å tilby brukerne en komplett kanal hvor de er sikre på at de får svar på det de lurte på. Ønskesituasjonen er at brukere skal kunne starte med Frida og ha en holdning om at; *”Okey, det er greit for meg å starte med Frida, jeg kan teste Frida først også ser vi hvor langt vi kommer.”* (Leder, intervju). Dersom chatboten ikke kan hjelpe, kan brukeren uansett overføres til en menneskelig veileder, i alle fall i åpningstiden. Chatboten er ikke på et sted i dag der den kan svare på alle typer henvendelser, men det oppfattes ikke nødvendigvis som et problem av NAV fordi de sakene som ikke løses i chatbot kan løses i andre deler av kanalmiksen.

Videre ser informantene i NAV også at mange brukere tilsynelatende slår seg til ro med den informasjonen de får i chatbot. Det påpekes at det ikke er noen garanti for at de ikke ringer eller tar kontakt gjennom en annen kanal senere, men mange brukere indikerer at de er fornøyd med svarene i slutten av samtalen. Her er det også gjort tiltak for å få chatboten til å oppfattes mer tillitsvekkende, blant annet bruk av avklarings spørsmål for å kunne gi brukere best mulig svar. I intervjuene kom det frem at NAV ser at brukere ofte har behov for å bekrefte at de har forstått riktig, både i chatbot og i samtale med menneskelige veiledere. Ved å tilby avklarings spørsmål er målet at chatboten skal bli mer treffsikker og sikre at brukere får relevante svar i den gitte konteksten. Dette trekkes frem som et tiltak for å trygge brukere på at den informasjon de får er gjeldende i den situasjonen de er i, og kan bidra til at brukere ikke føler behov for å ta kontakt i andre kanaler for å bekrefte informasjonen de fikk i chatbot.

5.4.1 Chatbotens kompetanse og tillit

Av de som svarte ”nei” på spørsmål om tillit til automatiserte tjenester i det offentlige, trekkes systemets funksjonalitet og kompetanse frem som påvirkende faktorer av flere respondenter. Opplevelse av å ikke få hjelp til det de etterspør og begrensede svar pekes på som grunner til manglende tillit. En respondent svarer at *”Det er altfor dårlig til å oppfatte det rette spørsmålet og gir dermed svar som ikke er relevant eller direkte feil.”* (Anonym respondent, spørreundersøkelse). En annen skriver at hen ikke har tillit *”Fordi de bare linker til FAQ og ikke svarer på spørsmålene man faktisk stiller.”* (Anonym respondent, spørreundersøkelse).

Respondentene i spørreundersøkelsen ble også spurt om i hvilken grad inntrykket av chatbotens kompetanse førte til at tilliten til systemet endret seg. Her svarte 67,9% at det var ingen endring, 30,2% svarte at tilliten ble svekket og 1,9% svarte at tilliten økte. På spørsmål om hvorfor tilliten endret seg, peker flere respondenter på utilstrekkelige svar eller mangel på svar som grunnen til dette. Av de få som svarte at tilliten økte, svarte en av respondentene følgende på spørsmål om bakgrunnen for endringen: *”Hadde ikke forventet at jeg skulle få info om det jeg lurte på. Da jeg likevel fikk det, økte tilliten.”* (Anonym respondent, spørreundersøkelse).

Flere informanter oppfatter at brukeres tillit til chatboten svekkes fort når de ikke får den informasjonen de ønsker. Svarene i spørreundersøkelsen viser det samme og flere svarer at de opplever å bli frustrerte og irriterte når de opplever å ikke få svar på det de lurer på eller chatboten misforstår spørsmålet. Dette fører til at de gir opp og ber om å bli satt over til et menneske eller velger å ringe i stedet.

5.4.2 Design og utseende

Av de offentlige aktørene som benytter chatbot i Norge, er det varierende hvordan de har valgt å designe og presentere sine chatboter. NAV sin chatbot Frida er gitt et menneskelig navn, mens Skatteetaten sin chatbot heter Skatteetatens chatrobot. Både Skatteetatens chatrobot, Frida og Kommune-Kari presenteres som chat-roboter i velkomstmeldingene. Fra intervjuene kommer det frem at valget bak navngivningen av Frida stammer fra da det startet som et studentprosjekt. I senere diskusjoner rundt navnet og om det burde endres for å presisere at det er en chatbot, ble det konkludert med at merkevaren Frida allerede var etablert og at det var mest hensiktsmessig å ikke gjøre endringer.

Videre er det forskjell i hvilke ikoner som er valgt for å representere chatbotene. Mens noen går for et "robot-aktig" utseende, har både Frida og Kommune-Kari fått ikoner som ligner på mennesker. Om bakgrunnen for valget om å gi Frida et menneskelig utseende nevnte en av lederne at *"NAV og offentlig og stat kan virke veldig kald og litt sånn kynisk ut mot bruker. Så vi ønsket på en måte å ha litt mer varme, og ikke være en sånn mekanisk robot som driver å svarer deg og henviser deg og så videre. At Frida fremstår som litt mer human da, uten at vi skal gi henne for mange menneskelige trekk."* (Leder, intervju).

For å sikre et tydelig skille mellom når brukere snakker med Frida og når de snakker med menneskelige veiledere, brukes farger og skifte av ikoner. I intervju med chatbot-trenere nevnes det at de likevel ser at enkelte brukere tror Frida er et menneske, og at veiledere opplever å få spørsmål fra brukere om de også er en robot. Svarene i spørreundersøkelsen viser at de fleste respondentene var klar over at de snakket med en chatbot før eller i starten av samtalen. Et fåtall opplevde at de var usikre på om de snakket med en robot eller en person og uttrykte at de på et tidspunkt kan ha snakket med en robot uten at de var bevisst på dette.

5.4.3 Tillit til den offentlige aktøren

Flere svar i spørreundersøkelsen indikerer at brukere har tillit til chatboten fordi de har tillit til menneskene eller aktøren bak. På spørsmål om tillit skrev en av respondentene at *"Tenker at folka som har skrudd sammen boten er dyktige og flinke folk."* (Anonym respondent, spørreundersøkelse). Videre skriver en annen at *"På samme måte som jeg har stor tillit til det offentlige, ser jeg på chat-bot'en som en troverdig kilde for informasjon."* (Anonym respondent, spørreundersøkelse). Flere viser forståelse for at chatboten ikke lever sitt eget liv, men styres av mennesker som lager innholdet og svarene, og at de har tillit til disse menneskenes kompetanse.

Respondentene ble i spørreundersøkelsen spurt om bruken av chatbot førte til en endring i tilliten til den offentlige tjenesten. 75,5% svarte ingen endring, 1,9% svarte at tilliten økte og 22,6% svarte at tilliten ble svekket. På spørsmål om hvorfor tilliten ble svekket, svarte en av respondentene at *"Hvis de ikke har kompetanse nok til å lage eller skaffe seg en god nok chatbot, har de da god nok kompetanse på andre områder?"* (Anonym respondent, spørreundersøkelse). En annen skriver at grunnen til at tilliten til den offentlige tjenesten ble svekket var *"Fordi tjenesten er så informasjonsmessig dårlig, og gjør kronglete tekst ut av noe som burde vært løst med bedre søkefunksjon."* (Anonym respondent, spørreundersøkelse). Andre nevner at tilliten ble svekket fordi de oppfattet det som at etaten eller tjenesten ikke ønsket å lønne og utdanne folk, at chatboter er en meningsløs sperre i veien for faktisk veiledning og at etaten undervurderer innbyggernes behov for kompetente svar da chatboten bare kunne besvare enkle spørsmål. En av respondentene skrev at den offentlige tjenesten hen ønsket å kontakte hadde fjernet telefoni som kommunikasjonskanal, og erstattet dette med en chatbot. Respondenten mente chatboten ikke var trent nok til å erstatte andre kanaler, og ble derfor irritert av interaksjonen og følte det var vanskelig å få hjelp.

En av informantene påpekte i intervju at en stor andel innbyggere tar kontakt med NAV fordi de er i en uheldig situasjon, om det er som arbeidsledig eller på grunn av sykdom, og at tilliten også kan påvirkes av dette. Videre uttrykker informanten at mange innbyggere har et dårlig forhold til etaten fra tidligere, og at disse kan være vanskelige å møte på riktig måte i chatbot. En av chatbotkoordi-

natorene påpeker at *”Og de som allerede har en dårlig opplevelse med NAV er jo litt mer sånn kinkig å få snudd med en chatbot, som ikke på en måte kan fange opp de her mellommenneskelige ... altså sånn ja, man får liksom ikke den der menneskelige touchen i en chatbot enda hvertfall.”* (Chatbotkoordinator, intervju).

5.4.4 Personvern og sikkerhet

NAV har gjort flere tiltak for å forsikre brukeren om at personvern og sikkerhet er ivaretatt i møte med chatboten. Som nevnt i avsnitt 5.3.2, ble velkomstmeldingen endret i starten av 2021. Denne presiserer at tjenesten er anonym og at brukeren ikke må skrive personlig informasjon, samtidig som den viser til en lenke med mer informasjon rundt håndtering av data og personvern i NAV. Også chatboten til Skatteetaten og Kommune-Kari presiserer i sine velkomstmeldinger at brukeren ikke må dele sensitiv data eller personlige opplysninger i samtalen. I intervju med en av lederne ble det nevnt at *”Det ene er at vi skal på en måte være ryddig og håndtere ting på en ordentlig måte, og så skal brukerne være trygge på at vi håndteres deres data på en god måte.”* (Leder, intervju). Her påpekes det at det er spesielt viktig for aktører i offentlig sektor å følge lovverket når det kommer til personvern og sikkerhet, og at det også er viktig å formidle dette til brukeren før samtalen starter.

I spørreundersøkelsen ble respondentene spurt om det har tillit til automatiserte tjenester i det offentlige. Der svarte flere at de har tillit, og at enkelte skrev at grunnen til dette er fordi de har tillit til at offentlige tjenester har høye krav til sikkerhet i sine tjenester. På spørsmål om hvorfor de har tillit til automatiserte tjenester i det offentlige, svarer en respondent *”Profesjonelle leverandører som har fokus på personvern og kvalitet”* (Anonym respondent, spørreundersøkelse).

Kapittel 6

Diskusjon: Tillit til chatbot?

I dette kapitlet diskuteres de mest interessante funnene presentert i kapittel 5 i kontekst av litteraturen og forskningsspørsmålet. Det første delkapitlet omhandler rollen oppfattet kompetansenivå spiller i etablering og vedlikehold av tillit i en menneske-chatbot-relasjon. Videre diskuteres skillet mellom forventninger og faktisk opplevelse i chatbot med hensyn til kompetanse, samt hvordan dette påvirker tillit til chatbot.

Mye av litteraturen som er presentert i kapittel 2 og som inngår i diskusjonen omhandler tillit til kommersielle chatboter for kundeservice, fordi mesteparten av forskningen på menneske-chatbot-interaksjon er gjort nettopp i den konteksten. Tillit er en nøkkelfaktor i menneske-chatbot-interaksjon uansett bakgrunn og mål, men for offentlig sektor er det også et underliggende samfunnsansvar som må opprettholdes i møte med innbyggere. Dette ansvaret gjør tillit essensielt for at etatene skal lykkes med innbyggerkontakt ved hjelp av chatboter. Brukere av disse chatbotene ønsker informasjon og bistand i ulike situasjoner og hendelser i livet, og for at chatbot skal oppleves som en trygg kanal å søke svar i må de ha tro på at de får god hjelp og korrekt informasjon. Det å sikre at brukere har tillit til tjenesten er derfor essensielt for at chatbot skal benyttes som kommunikasjonskanal mellom innbygger og etat i offentlig sektor.

6.1 Oppfattet kompetansenivå

Funnene i denne studien viser at opplevelsen av et lavt kompetansenivå kan medføre svekket tillit til en chatbot. Dette støtter eksisterende litteratur på tillit til chatboter som definerer oppfattet ekspertise og kompetanse som avgjørende i utvikling av tillit (Følstad mfl., 2018; Nordheim, 2018; Nordheim mfl., 2019). Brukeres oppfatning av kompetansenivå påvirkes av en rekke faktorer. Først og fremst trekkes korrekte svar frem som indikator på kompetanse, men en nødvendighet for å gi korrekte svar er at chatboten tolker brukernes spørsmål korrekt. Dette

fremkommer også i litteraturen, hvor tolkning og rådgivning samt korrekte svar er blant nøkkelfaktorene som påvirker brukeres tillit til chatboter (Corritore mfl., 2003; Følstad mfl., 2018; Nordheim mfl., 2019). Funnene i denne studien viser at brukere i stor grad vurderer chatbotens kompetansenivå som dens evne til å tolke spørsmål korrekt, og når chatboten ikke tolker spørsmål riktig eller misforstår hva brukerne mener oppfatter brukere dette som at chatboten er lite kompetent. Opplevelsen av å ikke bli forstått kan føre til at brukers tillit til chatboten svekkes, noe som videre kan bidra til at brukeren avviser chatboten og heller benytter andre midler eller kanaler for å oppnå sine mål (Dietvorst mfl., 2015).

Det er tydelig at det særlig oppstår irritasjon der brukere opplever å bli sendt rundt i ring på nettsidene av chatboten, eller får feil svar eller feil kontekst om og om igjen. Dersom en bruker tar kontakt i chatbot for å avklare eller forstå informasjon fra etatens nettside, og så bare blir sendt tilbake til den samme nettsiden av chatboten skaper dette frustrasjon. Brukere som opplever dette oppfatter chatboten som lite kompetent, og det påpekes av flere at meningen bak spørsmålet eller hensikten med forespørselen sannsynligvis hadde blitt fanget opp raskt dersom samtalen hadde vært med et menneske. For å unngå misforståelser og en opplevelse av å ikke bli forstått kan det være fordelaktig å inkludere brukeren i bestemmelse av situasjonen fremfor å anta kontekst og gi feilaktige svar som ikke er relevante for brukers henvendelse (Amershi mfl., 2019; Sheehan mfl., 2020). Dersom misforståelsene ikke kan løses med alternativer og kontekstbestemmelse vil neste steg være å tilby brukeren å snakke med en veileder fremfor å fortsette interaksjonen i chatbot. NAV har jobbet mye med å bestemme terskelen for når brukere skal settes over til menneskelige veiledere, og dette er et viktig faktor for å sørge for god brukeropplevelse og legge til rette for utvikling av tillit (Nordheim, 2018). Det å være tydelig på når chatboten ikke kan svare på spørsmålet og heller videreføre til veileder før brukere når et punkt og frustrasjon kan være med på å opprettholde et godt inntrykk og en eventuell tillit fra brukers side.

For noen brukere førte ikke opplevelsen av chatbotens kompetansenivå og brukeropplevelsen til at tilliten til chatboten ble svekket, men heller til at tilliten til etaten bak og ”merkevaren” ble lavere. Det at chatbotens kompetansenivå ble oppfattet som lavt ble pekt på som en direkte konsekvens av jobben den offentlig

aktøren gjorde. Det må påpekes at siden chatboten er del av etatens kanalmiks og ikke er frittstående i den forstand, er det sannsynlig at svekket tillit til etaten også fører til svekket tillit til chatboten. Dette kommer også frem i Nordheim m.fl. (2018), hvor merkevare eller organisasjonen bak chatboter trekkes frem som en påvirkende faktor i utvikling av tillit, og i Følstad m.fl. (2018) hvor brukeres oppfatning av og tillit til tjenestetilbyderen ble funnet å ha innvirkning på tilliten til chatboten.

Dette tyder også på at brukere oppfatter at chatboten styres av mennesker. Luger og Sellen (2016) fant at brukere kan være usikre på om hvorvidt teknologi som benytter kunstig intelligens, som for eksempel intelligente agenter og chatboter, har evne til å lære og i hvilken grad de er styrt av mennesker. Bruk av kunstig intelligens kan også føre til usikkerhet knyttet til utførelse og om chatboten kan utøve skjønn (Aoki, 2020). Funnene i denne studien tyder derimot på at brukere er bevisst på at chatboten ikke lever sitt eget liv, og at det er mennesker som står bak både algoritmene og innholdet selv om chatboten i seg selv er en robot. Brukere har også høy tillit til at selve innholdet i chatboten er korrekt og til at sikkerheten ivaretas, nettopp fordi det er en offentlig aktør som står bak. Dette kan ha sammenheng med at de offentlige aktørene har vært nøye med å tydeliggjøre ovenfor brukere at chatbotene er roboter, samtidig som det understrekes at innhold både utvikles og kvalitetssikres av mennesker.

Funnene viser tydelig at brukere flest har tillit til svarene chatboten gir i den forstand at de oppfatter at informasjonen er den samme som på nettsidene til etaten og at de som skriver svarene har god kunnskap på området. Det som likevel trekkes frem som vanskelig er når brukere er usikre på om informasjonen de får gjelder i situasjonen de befinner seg i. Ofte presenterer brukere spørsmål i en gitt kontekst når de henvender seg til en chatbot, og flere uttrykker usikkerhet rundt hvorvidt chatboten tar alle aspekter i betraktning i sine svar. Dette kan handle om at informasjonen på nettsidene er vanskelig å forstå og at språket som benyttes oppleves som byråkratisk og komplekst, men for å unngå at brukere forlater samtalen usikre eller tar kontakt i andre kanaler for avklaring, kan det derfor være hensiktsmessig å tilby en type bekreftelse i chatbot. Dette kan eksempelvis gjøres i form av eksempler på situasjoner knyttet til stønader og ytelser, eller ved å overføre informasjonen til en mer hverdagslig kontekst for å tydeliggjøre hva den

innebærer.

Det kan også vurderes i hvilken grad det at brukeren ikke får ønsket svar i chatbot er negativt, dersom en bruker uansett får mulighet til å overføres til et menneske. Chatbot i offentlig sektor er en del av en større kanalmiks i de ulike etatene og tjenestene, og det vil være viktig å se kanalmiksen i helhet og se koblingene mellom de ulike punktene for brukerkontakt. En av faktorene som påvirker tillit til chatboter er tilgang til en menneskelig operatør (Nordheim, 2018). Om svaret kommer ved hjelp av en menneskelig veileder i stedet for direkte i chatbot, er det sannsynlig at brukeren likevel sitter igjen med et godt inntrykk av tjenesten som helhet. Det er likevel også mulig at brukers opplevelse av å ikke få hjelp i chatbot fører til at de kanskje ikke ønsker å benytte seg av chatboten neste gang, og at de heller går direkte til en annen kanal (Dietvorst mfl., 2015).

6.2 Rollen brukerforventninger spiller

Majoriteten av brukerne kartlagt i denne studien oppgir å ha forventninger som langt på vei stemmer overens med det reelle kompetansenivået i chatboter i offentlig sektor. De aller fleste uttrykker at det er minimale forskjeller i hva de forventer av interaksjonen før bruk og hva som møter dem i samtalen med chatbot, og for de fleste brukerne var det heller ingen forskjell i tilliten de følte ovenfor chatboten før og etter bruk. De opplever at chatboten forstår målet med henvendelsen og tilbyr relevante svar. Svarene beskrives gjerne som generiske og enkle, men for brukere som ønsker svar på det som kan kalles generelle spørsmål oppleves dette som tilfredsstillende. Det var også tilfeller der respondenter rangerte forventningene til kompetansenivået før bruk noe høyere enn hvordan de oppfattet kompetansenivået etter bruk, men der det likevel ikke påvirket tilliten negativt. Dette tyder på at chatboten til en viss grad kan prestere dårligere enn forventet uten at dette påvirker tilliten. Det var også svært få som opplevde en økning i tillit etter bruk, til tross for at de rangerte kompetansenivået høyere etter bruk enn i forventningene. Dette kan selvsagt bety at tilliten var på et høyt nivå fra før, og at det ikke var noen innlysende årsak til at den skulle øke ytterligere etter interaksjonen.

Der det derimot er store sprik i forventningene før bruk og den faktiske opplevelsen, er det tydelig at tilliten til chatboten svekkes. Flere brukere har til dels urealistiske forventninger til kompetansenivå før bruk og opplever at forventnin-

gene ikke innfris i interaksjon med chatbot Dette kommer også frem i Luger og Sellen (2016) sin studie, hvor brukere hadde forventninger som langt overgikk agentens evner. Mens de fleste brukere kartlagt i denne studien forventet at chatboten skulle være i stand til å svar på enkle og generelle spørsmål, forventet andre et høyere kompetansenivå både med tanke på kompleksitet i ønsket informasjon og forståelse i chatbot. Forventninger om å få svar knyttet til mer komplekse spørsmål eller personlige saker vil naturligvis føre til at brukeren opplever å ikke få svar som ønsket i chatbot, og brukerne som hadde slike forventninger før bruk oppga at tilliten ble svekket som følge av dette. Det at tilliten påvirkes av opplevelsen av å ikke få hjelp i chatbot beskrives også i Følstad m.fl. (2018) og Nordheim m.fl. (2019). Denne gruppen brukere var likevel liten, noe som tyder på at de fleste oppfatter at det er en anonym tjeneste uten innlogging, og at det derfor ikke er mulig for chatboten å svare på brukerens sak uten å videreføre til en menneskelig veileder.

Brukere som har høye forventninger til chatbotens kompetansenivå i den forstand at de forventer et høyt nivå av forståelse i chatbot er en annen utfordring. Det handler altså ikke bare om hvilket nivå chatboten er på i sammenheng med informasjonen den tilbyr, men også i hvilken grad den er i stand til å tolke brukeres henvendelser. En chatbots evne til å tolke spørsmål riktig er en nøkkelfaktor for utvikling av tillit (Følstad mfl., 2018). Brukere som forventet et høyt nivå av forståelse og opplevde at chatboten ikke innfridde forventningene oppga at dette svekket tilliten til chatboten. Det at brukere har høye forventninger til kompetansenivå før bruk kan også ha med brukerens egne kompetanse rundt chatboter å gjøre. Chatbotteknologien er fortsatt ny for mange, og brukere som har lite erfaring med chatboter kan ha utfordringer med å vite hvordan de skal skrive til en chatbot. Dersom brukere henvender seg til chatboter som om de var et menneske, med lange situasjonsbeskrivelser og mye tekst, er det sannsynlig at chatboten ikke oppfatter hva henvendelsen egentlig omhandler. Denne måten å formulere seg på skaper mye støy, noe som videre vanskeliggjør kontekstbestemmelse og kan føre til unøyaktige svar. Mens noen brukere oppfatter dette og tilpasser seg chatboten ved å stille korte, konsise spørsmål med beskrivende nøkkelord, er det tydelig at andre brukere i større grad forventer at teknologien skal tilpasse seg dem.

Brukernes forventninger til chatboter kan påvirkes av en rekke faktorer. Først og fremst kan det tenkes at brukernes erfaringer med chatboter på generelt basis påvirker forventningene. Chatbot for kundeservice i kommersiell sammenheng blir mer og mer vanlig, og det er sannsynlig at også slike chatboter former brukernes forventninger til chatboter i offentlig sektor. Funnene viser at brukeres erfaring med kommersielle chatboter, for eksempel kundeservice eller kjøpshjelp, er at de i større grad fungerer som en erstatning for menneskelige rådgivere, og at chatbot oppleves som eneste kommunikasjonsalternativ. Her har offentlig sektor et annet og større samfunnsansvar som også må reflekteres i bruk av chatbot for innbyggerkontakt. Det at enkelte brukere uttrykte en følelse av å måtte "lure" systemet til å overføre de til et menneske kan være en konsekvens av erfaring med kommersielle chatboter, for i de offentlige trenger brukeren kun å uttrykke et ønske om å snakke med et menneske for å bli overført. Videre er det sannsynligvis store forskjeller i hvor langt de ulike aktørene er kommet i utvikling av chatbot. Dette kan føre til at forventningene som skapes ved bruk av en chatbot ikke nødvendigvis innfris i en annen chatbot, noe som kan føre til at tilliten svekkes.

Videre kan design og utseende spille en rolle i forventninger til en chatbot. Brukere av chatboter kan ha vanskeligheter med å skille mellom en chatbot og en menneskelig samtalepartner (Candello mfl., 2017). Funnene i denne studien viser at alle brukere oppfattet at de kommuniserte med en chatbot enten før eller i løpet av samtalen, og at de eventuelle menneskelignende aspektene av chatbotene ikke var villedende i den forstand. Men det at utviklingen har kommet lengre og språket i chatboter har blitt bedre fører også til at brukere kan tillegge chatboter enda flere menneskelige egenskaper (Mone, 2016). For eksempel kan det at en chatbot er utformet med et menneskelignende ikon og navn føre til at brukere forventer en annen fleksibilitet enn hos chatboter som er utformet som roboter, fordi de i større grad sammenlignes med mennesker. Dette kan bidra til urealistiske forventninger til kompetansenivå ved at brukere i større grad forventer en chatbot på samme nivå som en menneskelig veileder. Antropomorfisme er likevel et viktig verktøy i utvikling, og for en etat som ofte oppfattes som kynisk eller kald kan en viss grad av antropomorfisme gjøre at en chatbot oppfattes som varmere (Khadpe mfl., 2020). Funnene viser at det for etaten er ønskelig å fremstå som varm i møte med brukere, uten at det påvirker opprettholdelse av et profesjonelt utseende.

6.3 Implikasjoner for design

Denne studien har identifisert faktorer som påvirker hvordan brukere oppfatter kompetansenivå i chatboter, samt hvordan dette påvirker tilliten. Videre presenteres tiltak i utvikling og design av chatboter som kan legge til rette og skape gode forutsetninger for tillit i en menneske-chatbot-interaksjon.

Tilrettelegg for realistiske brukerforventninger. Gi brukere indikatorer og hint som skaper en ide av hva chatboten kan gjøre før interaksjonen (Khadpe mfl., 2020; Luger og Sellen, 2016). Tydeliggjør chatbotens begrensninger.

Tilby avklaringer og alternativer. I tilfeller hvor det er usikkert hva brukeren ønsker med sin henvendelse kan det være hensiktsmessig å tilby avklaring eller alternativer for å bestemme kontekst og dermed gi best mulig svar. Feilantagelser i situasjonsbestemmelse kan føre til frustrasjon for brukeren, og for å unngå misforståelser og at chatboten gir svar i feil kontekst kan det å presentere alternativer for hva chatboten tror brukeren ønsker være positivt for brukeropplevelse og tillit (Amershi mfl., 2019).

Gi eksempler. For å betrygge bruker som søker bekreftelse på at informasjonen de får i chatbot gjelder i deres spesifikke situasjon og at alle aspekter er tatt i betraktning, kan det være fordelaktig å tilby eksempler med informasjonen de gis. Dette kan bekrefte ovenfor brukeren om informasjonen er gjeldende for konteksten de tar kontakt i.

Kommuniser tydelig når chatboten ikke kan hjelpe. Når brukere ønsker hjelp eller svar på spørsmål som går utenfor chatbotens kompetanseområde, gi brukeren klar beskjed om dette og tilby overføring til menneskelig veileder for videre samtale og rådgivning.

Tilby veiledningsmateriale. Brukere som har lite eller ingen erfaring med chatboter kan ha problemer med å formulere seg på en måte som tydeliggjør for chatboten hva de ønsker svar på. For å unngå at brukeren skriver lange avhandlinger med mange ord som skaper støy og gir unøyaktige svar, kan det være hensiktsmessig å tilby veiledningsmateriale som sier noe om hvordan brukeren bør formulere seg for å få best mulig hjelp i chatbot.

Kapittel 7

Konklusjon

Denne studien har tatt utgangspunkt i bruk av chatbot i offentlig sektor og undersøkt brukeres tillit til chatboter og hvordan den påvirkes av oppfattet kompetansenivå. Data er samlet inn ved hjelp av intervjuer med informanter fra NAV og en spørreundersøkelse besvart av innbyggere som har benyttet seg av chatbot i offentlig sektor.

Funnene viser at oppfattet kompetansenivå påvirker en brukers tillit til chatbot, men at dette også avhenger av hvilke forventninger brukeren gikk inn i interaksjonen med. Selv om chatbotteknologien har utviklet seg mye de siste årene, kan den fortsatt ikke likestilles med menneskelige veiledere i offentlig sektor, og hvorvidt en bruker kan få svar på sin henvendelse i chatbot avhenger av henvendelsens natur. Brukere som går inn i interaksjonen med en forventning om å få svar på enkle spørsmål knyttet til stønader og tjenester opplever kompetansenivået som tilfredsstillende, og uttrykker ingen endring i tillit til chatboten. Brukere som har forventninger som overgår chatbotens faktiske kompetansenivå og evner, opplever chatbotens kompetansenivå som lavt, og uttrykker at tilliten svekkes som følge av dette. Disse brukerne har forventninger knyttet både til chatbotens innhold og evne til forståelse som ikke kan innfris i interaksjonen. Dette handler både om at offentlig sektor ikke tilbyr innlogget chatbot for innsyn i personlige saker og at chatbotteknologien ikke er kommet dit at de er i stand til å tolke språk på måten noen brukere forventer. Dette kan også handle om brukeres kompetanse om chatboter og hvordan de skal kommunisere. For brukere flest ser forventningene ut til å justeres ved bruk, og oppfatningen av en chatbots kompetansenivå endres med økt erfaring og innsikt i hvordan teknologien fungerer. På samme måte er en brukers tillit dynamisk, og vil være i stadig endring gjennom interaksjon med chatboter. Teknologien blir stadig bedre, og brukere får mer og mer erfaring med denne måten å kommunisere på ved at flere aktører tar i bruk chatbot for innbygger- og kundekontakt.

7.1 Begrensninger og fremtidig arbeid

Det må anerkjennes at studien har sine begrensninger. Først og fremst er den begrenset av den empiriske datainnsamlingen; intervjuer med informanter i NAV og en spørreundersøkelse for brukere av chatbot i offentlig sektor. Intervjuene var gjennomført med et begrenset antall mennesker i NAV, og datamaterialet gjenspeiler disse informantenes meninger. Det må også nevnes at NAV bare er en av flere offentlige etater som benytter chatbot for innbyggerkontakt, og at resultatene sannsynligvis hadde vært noe annerledes ved involvering av informanter fra andre etater. Videre må det påpekes at av de 120 respondentene i spørreundersøkelsen hadde nesten 90% høyere utdanning og antageligvis en viss erfaring med teknologi på forhånd. Dette kan problematisere generalisering av funnene, da undersøkelsen ikke dekker alle kategorier av reelle chatbotbrukere i offentlig sektor. En spørreundersøkelse vanskeliggjør også dybdeinnsikt i en brukers holdninger til chatboter. Til tross for en blanding av flervalg og åpne spørsmål, er det opp til hver enkelt respondent hvor langt de ønsker å skrive og hvor utdypende de ønsker å være i svarene sine.

Funnene i denne studien åpner også for ytterligere forskning på fagområdet. Med hensyn til videre arbeid anbefales mer empirisk forskning på bruk av chatbot i offentlig sektor, og spesielt på brukersiden. Det å få en dypere innsikt i brukeres holdninger til chatboter vil være nyttig for å forstå hvordan denne teknologien kan brukes videre for å skape verdi for innbyggere. Det vil også være hensiktsmessig å gjennomføre flere studier hvor ulike etater inkluderes, eksempelvis Skatteetaten og kommuner som benytter seg av chatbot for innbyggerkontakt. En annen tilnærming kan være å se på et teoretisk rammeverk for design og utvikling av chatboter i offentlig sektor, samt hvordan og om dette skiller seg fra chatboter benyttet i kommersiell sammenheng. En mer uniform bruk av chatbot i offentlig sektor på tvers av etater kan være nyttig for å sikre at brukeres forventninger i mindre grad påvirkes av store ulikheter.

Bibliografi

- Amershi, S., Weld, D., Vorvoreanu, M., Fourney, A., Nushi, B., Collisson, P., Suh, J., Iqbal, S., Bennett, P. N., Inkpen, K. mfl. (2019). Guidelines for human-AI interaction. *Proceedings of the 2019 chi conference on human factors in computing systems*, 1–13.
- Aoki, N. (2020). An experimental study of public trust in AI chatbots in the public sector. *Government Information Quarterly*, 37(4), 101490.
- Araujo, T. (2018). Living up to the chatbot hype: The influence of anthropomorphic design cues and communicative agency framing on conversational agent and company perceptions. *Computers in Human Behavior*, 85, 183–189.
- Candello, H., Pinhanez, C. & Figueiredo, F. (2017). Typefaces and the perception of humanness in natural language chatbots. *Proceedings of the 2017 chi conference on human factors in computing systems*, 3476–3487.
- Cassell, J. & Bickmore, T. (2000). External manifestations of trustworthiness in the interface. *Communications of the ACM*, 43(12), 50–56.
- Corritore, C. L., Kracher, B. & Wiedenbeck, S. (2003). On-line trust: concepts, evolving themes, a model. *International journal of human-computer studies*, 58(6), 737–758.
- De Angeli, A., Johnson, G. I. & Coventry, L. (2001). The unfriendly user: exploring social reactions to chatterbots. *Proceedings of The International Conference on Affective Human Factors Design, London*, 467–474.
- Dietvorst, B. J., Simmons, J. P. & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*, 144(1), 114.
- Duffy, B. R. (2003). Anthropomorphism and the social robot. *Robotics and autonomous systems*, 42(3-4), 177–190.
- Følstad, A., Nordheim, C. B. & Bjørkli, C. A. (2018). What makes users trust a chatbot for customer service? An exploratory interview study. *International conference on internet science*, 194–208.

- Hartmann, J., De Angeli, A. & Sutcliffe, A. (2008). Framing the user experience: information biases on website quality judgement. *Proceedings of the SIGCHI conference on human factors in computing systems*, 855–864.
- Hill, J., Ford, W. R. & Farreras, I. G. (2015). Real conversations with artificial intelligence: A comparison between human–human online conversations and human–chatbot conversations. *Computers in human behavior*, 49, 245–250.
- Hirschberg, J. & Manning, C. D. (2015). Advances in natural language processing. *Science*, 349(6245), 261–266.
- Hoem, T. (2020). Norges digitalisering har fått en ny grunnmur. <https://www.skatteetaten.no/presse/nyhetsrommet/norges-digitalisering-har-fatt-en-ny-grunnmur/>
- Jackson, L. A., Von Eye, A., Biocca, F. A., Barbatsis, G., Fitzgerald, H. E. & Zhao, Y. (2003). Personality, cognitive style, demographic characteristics and Internet use-Findings from the HomeNetToo project. *Swiss Journal of Psychology/Schweizerische Zeitschrift für Psychologie/Revue Suisse de Psychologie*, 62(2), 79.
- Jære, L. (2018). Chatbots blir fortrolige venner og psykologer. <https://www.sintef.no/siste-nytt/chatbots-blir-venner-og-psykologer/>
- Kaplan, A. & Haenlein, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, 62(1), 15–25.
- Khadpe, P., Krishna, R., Fei-Fei, L., Hancock, J. T. & Bernstein, M. S. (2020). Conceptual metaphors impact perceptions of human-ai collaboration. *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW2), 1–26.
- Kocielnik, R., Amershi, S. & Bennett, P. N. (2019). Will you accept an imperfect ai? exploring designs for adjusting end-user expectations of ai systems. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–14.
- Kommunal- og moderniseringsdepartementet. (2016). Digital agenda for Norge: IKT for en enklere hverdag og økt produktivitet (Meld. St. 27 (2015-2016)). <https://www.regjeringen.no>

- KS. (2018). Helseplattformen. <https://www.ks.no/fagomrader/digitalisering/utviklingsprosjekter/digitalisering-i-helse-og-omsorgsektoren-e-helse/nasjonale-prosjekter/helseplattformen/>
- Lekanger, K. (2017). Vil "robotisere" kommune-Norge. <https://www.digi.no/artikler/vil-robotisere-kommune-norge/408800>
- Luger, E. & Sellen, A. (2016). Like Having a Really Bad PAThe Gulf between User Expectation and Experience of Conversational Agents. *Proceedings of the 2016 CHI conference on human factors in computing systems*, 5286–5297.
- Madsen, M. & Gregor, S. (2000). Measuring human-computer trust. *11th australasian conference on information systems*, 53, 6–8.
- Mayer, R. C., Davis, J. H. & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of management review*, 20(3), 709–734.
- Mone, G. (2016). The edge of the uncanny. *Communications of the ACM*, 59(9), 17–19.
- Müller, L., Mattke, J., Maier, C., Weitzel, T. & Graser, H. (2019). Chatbot Acceptance: A Latent Profile Analysis on Individuals' Trust in Conversational Agents. *Proceedings of the 2019 on Computers and People Research Conference*, 35–42.
- Mærøe, A. (2019). Gevinster fra en kommunal chatrobot. <https://medium.com/digitalisere/gevinster-fra-en-kommunal-chatrobot-ba94ab0954a3>
- Nass, C. & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of social issues*, 56(1), 81–103.
- NAV. (2018). nav.no og Ditt NAV for personbrukere - en helhetlig brukeropplevelse. <https://memu.no/wp-content/uploads/2018/07/Brukeropplevelser-p%5C%C3%5C%A5-nav.no-kortversjon-2018.pdf>
- NAV. (2019). NAVs omverdensanalyse 2019. Utvikling, trender og konsekvenser fram mot 2030 (NAV-rapport 2019:1. <https://www.nav.no/no/nav-og-samfunn/kunnskap/analyser-fra-nav/analyser/navs-omverdensanalyse-2019>
- Nordheim, C. B. (2018). *Trust in chatbots for customer service—findings from a questionnaire study* (Masteroppgave).
- Nordheim, C. B., Følstad, A. & Bjørkli, C. A. (2019). An initial model of trust in chatbots for customer service—findings from a questionnaire study. *Interacting with Computers*, 31(3), 317–335.

- Oates, B. J. (2005). *Researching information systems and computing*. Sage.
- Padgham, L. & Winikoff, M. (2005). *Developing intelligent agent systems: A practical guide* (Bd. 13). John Wiley & Sons.
- Prokom. (2020). Kommune-Kari. <https://prokom.no/produkter-og-tjenester/chatbots/kari/>
- Przegalinska, A., Ciechanowski, L., Stroz, A., Gloor, P. & Mazurek, G. (2019). In bot we trust: A new methodology of chatbot performance measures. *Business Horizons*, 62(6), 785–797.
- Regjeringen. (2019). Ny personopplysningslov. <https://www.regjeringen.no/no/tema/statlig-forvaltning/personvern/ny-personopplysningslov/id2340094/>
- Ringnes, I. F. (2018). Brukerne strever med nav.no. <https://memu.no/artikler/brukerne-strever-med-nav-no/>
- Ringnes, I. F. (2020). Frida jobber døgnet rundt. <https://memu.no/artikler/frida-jobber-dognet-rundt/>
- Rousseau, D. M., Sitkin, S. B., Burt, R. S. & Camerer, C. (1998). Not so different after all: A cross-discipline view of trust. *Academy of management review*, 23(3), 393–404.
- Sağlam, R. B. & Nurse, J. R. (2020). Is your chatbot GDPR compliant? Open issues in agent design. *Proceedings of the 2nd Conference on Conversational User Interfaces*, 1–3.
- Seeger, A.-M. & Heinzl, A. (2018). Human versus machine: Contingency factors of anthropomorphism as a trust-inducing design strategy for conversational agents. *Information systems and neuroscience* (s. 129–139). Springer.
- Shawar, B. A. & Atwell, E. (2007). Chatbots: are they really useful? *Ldv forum*, 22(1), 29–49.
- Sheehan, B., Jin, H. S. & Gottlieb, U. (2020). Customer service chatbots: Anthropomorphism and adoption. *Journal of Business Research*, 115, 14–24.
- Skatteetaten. (2021). Chat med oss. <https://www.skatteetaten.no/kontakt/chat-med-oss/?21152=21152>
- Taddeo, M. (2010). Trust in technology: A distinctive and a problematic relation. *Knowledge, Technology & Policy*, 23(3-4), 283–286.
- Tennøe, T. (2017). Kunstig intelligens: smarte eller skremmende? <https://teknologiradet.no/kunstig-intelligens-smart-eller-skremmende/>

- Tidemann, A. & Elster, A. C. (2019). Maskinl ring. <https://snl.no/maskinl%C3%A6ring>
- Tjora, A. (2012). Kvalitative forskningsmetoder i praksis.
- Vassilakopoulou, P, Grisot, M., Blegind Jensen, T, Sellberg, N., Eltes, J., Thorseng, A. A. & Aanestad, M. (2017). Building national ehealth platforms: the challenge of inclusiveness. *Proceedings of the International Conference on Information Systems-Transforming Society with Digital Innovation, ICIS 2017*.

Vedlegg A

Tilleggsmateriale

A.1 Intervjuguide

Spørsmål til NAV/ansatte om Frida

Vi er to studenter som undersøker brukeres tillit til chatboter. Den ene oppgaven omhandler rollen kompetanse spiller i utvikling og etablering av tillit, mens den andre ser på sammenhengen mellom brukeropplevelse og tillit til chatboter.

Praktisk informasjon, opptak og samtykke

Om utvikling (ikke innhold) og brukerinvolvering

- Med tanke på Fridas utseende og hvordan hun presenterer seg selv. Hvilke valg er tatt? Hvilke vurderinger ligger bak avgjørelsene?
- Konkrete eksempler på tiltak som er gjort for å skape bedre tillit blant brukerne?
- Gjør dere noe med tanke på brukere som avbryter interaksjonen med en gang de har startet chat? Hva med de som ikke avslutter samtalen på "korrekt" måte?
- Foruten å ta utgangspunkt i tidligere chatter når dere lager nye svar til Frida. Er det andre måter dere kunne tenke dere å involvere brukere på i forbindelse med utvikling?

Om tillit

- I samtalen mellom Frida og brukere - oppfatter dere at brukere har tillit til Frida? På hvilken måte?
- Hvordan selger dere inn Frida til brukere?
- Hvilke tiltak har dere gjort for å få Frida til å oppfattes mer tillitsfull av brukere? Hvilke fordeler tror dere dette vil ha?
- Hvordan ser NAV på sammenhengen mellom tillit og brukeropplevelse?
- Tillit til AI kan uttrykkes konseptuelt langs akser for varme og kompetanse. Hvor befinner Frida seg langs disse aksene?

Om kompetanse

- Frida sier ikke noe om hvordan brukeren burde stille spørsmål eller hva hun kan svare på i velkomsthilsenen. Hvilke vurderinger er gjort rundt dette?
- Hvordan behandler/styrer dere brukernes forventninger til Frida med tanke på hennes kompetansenivå?
- Hvilke forventninger ønsker dere at brukerne skal ha til Frida?

Om empati

- Vi har blitt fortalt at enkelte tema, f.eks. død gjør at du blir satt over til mennesker med en gang. Har NAV planer om at Frida i fremtiden skal kunne svare på denne typen henvendelser? Hva må på plass før hun kan det? Hvilket nivå av empati burde en chatbot av denne typen ligge på?

A.2 Spørreundersøkelse

Bruk av chatbot i offentlig sektor

Side 1

I denne spørreundersøkelsen ønsker vi å samle erfaringer fra brukere av chatboter innen offentlig sektor i Norge. Vi ønsker å undersøke hvordan brukernes tillit til selve chatboten og til tjenestetilbyderen blir påvirket av interaksjonene med denne typen chatteprogram. Eksempler på chatboter innen offentlig sektor er Kommune-Kari og Frida hos NAV.

Spørreundersøkelsen er utviklet av to studenter fra masterstudiet Digital Samhandling ved NTNU og utgjør en del av arbeidet med våre mastergrader. Vi har bakgrunn fra bachelorgradene Digital Forretningsutvikling og Informatikk fra NTNU.

Spørreundersøkelsen er anonym og samler ikke inn personlig informasjon. Den tar ca 5-10 minutter å gjennomføre.

Generelle spørsmål

1. Alder *

- 18-25
- 26-35
- 36-45
- 46-55
- 56-65
- 66+

2. Kjønn *

- Kvinne
- Mann
- Annet
- Ønsker ikke å spesifisere

3. Høyeste fullførte utdanning? *

- Grunnskole
- Videregående
- Høgskole/universitet

Erfaringer med chatbot

4. Har du erfaringer med bruk av chattekanaler for kontakt med kundestøtte eller det offentlige? *

- Ja, for kundestøtte
- Ja, for kontakt med det offentlige
- Ja, begge
- Nei

4a. Hvorfor valgte du chat som kommunikasjonskanal for kundestøtte? *

- 1** Dette elementet vises kun dersom alternativet «Ja, for kundestøtte» er valgt i spørsmålet «4. Har du erfaringer med bruk av chattekanaler for kontakt med kundestøtte eller det offentlige?»

Husk å ikke oppgi personlig informasjon.

4b. Hvorfor valgte du chat som kommunikasjonskanal for kontakt med det offentlige? *

- 1** Dette elementet vises kun dersom alternativet «Ja, begge» eller «Ja, for kontakt med det offentlige» er valgt i spørsmålet «4. Har du erfaringer med bruk av chattekanaler for kontakt med kundestøtte eller det offentlige?»

Husk å ikke oppgi personlig informasjon.

5. Har du hatt kontakt med en chatbot i offentlig sektor i Norge de siste tre årene? *

i Dette elementet vises kun dersom alternativet «Ja, begge» eller «Ja, for kontakt med det offentlige» er valgt i spørsmålet «4. Har du erfaringer med bruk av chattekanaler for kontakt med kundestøtte eller det offentlige?»

En chatbot er en automatisert chat hvor man snakker med et program og ikke med et menneske. Eksempler på denne typen chatboter i offentlig sektor i Norge er Kommune-Kari og NAVs Frida.

Ja

Nei

6. Har du tillit til automatiserte tjenester i det offentlige? Som f.eks. chatboter. *

Ja

Nei

Vet ikke

6a. Hvorfor har du tillit til automatiserte tjenester i det offentlige? *

i Dette elementet vises kun dersom alternativet «Ja» er valgt i spørsmålet «6. Har du tillit til automatiserte tjenester i det offentlige? Som f.eks. chatboter.»

6b. Hvorfor har du ikke tillit til automatiserte tjenester i det offentlige? *

i Dette elementet vises kun dersom alternativet «Nei» er valgt i spørsmålet «6. Har du tillit til automatiserte tjenester i det offentlige? Som f.eks. chatboter.»

6c. Hva er du usikker på rundt automatiserte tjenester i det offentlige? *

- 1 Dette elementet vises kun dersom alternativet «Vet ikke» er valgt i spørsmålet «6. Har du tillit til automatiserte tjenester i det offentlige? Som f.eks. chatboter.»



7. Ta utgangspunkt i din kommunikasjon med en chatbot i offentlig sektor. På hvilket tidspunkt forsto du at du kommuniserte med en chatbot? *

- 1 Dette elementet vises kun dersom alternativet «Ja» er valgt i spørsmålet «5. Har du hatt kontakt med en chatbot i offentlig sektor i Norge de siste tre årene?»

Før samtalen

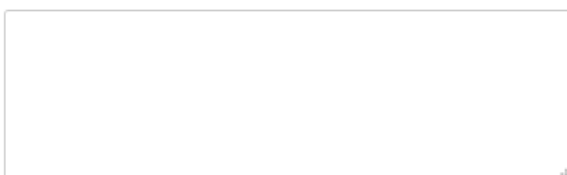
I begynnelsen av samtalen

Underveis i samtalen

Etter samtalen

7a. Hva gjorde at du forsto at du kommuniserte med en chatbot? *

- 1 Dette elementet vises kun dersom alternativet «Før samtalen», «Etter samtalen», «I begynnelsen av samtalen» eller «Underveis i samtalen» er valgt i spørsmålet «7. Ta utgangspunkt i din kommunikasjon med en chatbot i offentlig sektor. På hvilket tidspunkt forsto du at du kommuniserte med en chatbot?»



8. Hva ønsket du å få ut av samtalen med chatboten? *

i Dette elementet vises kun dersom alternativet «Ja» er valgt i spørsmålet «5. Har du hatt kontakt med en chatbot i offentlig sektor i Norge de siste tre årene?»

- Generell informasjon
- Sende inn eller opprette ny søknad
- Oppfølging av søknad/status for min sak
- Annet

9. Fikk du svar på det du lurte på fra chatboten? *

i Dette elementet vises kun dersom alternativet «Ja» er valgt i spørsmålet «5. Har du hatt kontakt med en chatbot i offentlig sektor i Norge de siste tre årene?»

- Ja
- Nei, jeg avbrøt samtalen
- Nei, jeg ble satt over til et menneske

9a. Hvorfor avbrøt du samtalen? *

i Dette elementet vises kun dersom alternativet «Nei, jeg avbrøt samtalen» er valgt i spørsmålet «9. Fikk du svar på det du lurte på fra chatboten?»

Husk å ikke oppgi personlig informasjon.

9b. Hvorfor ble du satt over til et menneske? *

i Dette elementet vises kun dersom alternativet «Nei, jeg ble satt over til et menneske» er valgt i spørsmålet «9. Fikk du svar på det du lurte på fra chatboten?»

- Jeg ble automatisk videresendt
- Jeg ba om å bli satt over til et menneske

9c. Hvorfor ønsket du å snakke med et menneske? *

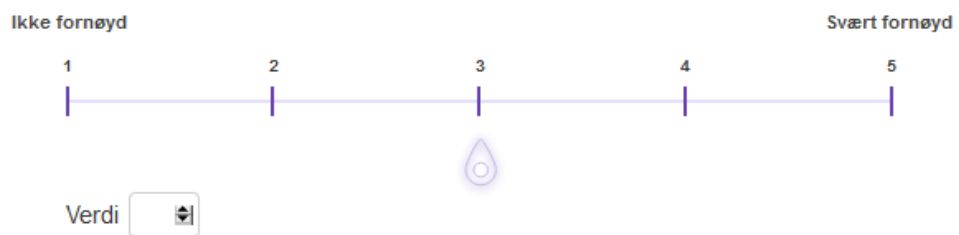
1 Dette elementet vises kun dersom alternativet «Jeg ba om å bli satt over til et menneske» er valgt i spørsmålet «9b. Hvorfor ble du satt over til et menneske?»

Husk å ikke oppgi personlig informasjon.

10. Hvor fornøyd var du med interaksjonen? *

1 Dette elementet vises kun dersom alternativet «Ja» er valgt i spørsmålet «5. Har du hatt kontakt med en chatbot i offentlig sektor i Norge de siste tre årene?»

(3 - nøytral)



i Dette elementet vises kun dersom alternativet «Ja» er valgt i spørsmålet «5. Har du hatt kontakt med en chatbot i offentlig sektor i Norge de siste tre årene?»

Kompetanse, brukeropplevelse og tillit

11. Hvilke forventninger hadde du til chatbotens faglige kompetanse i forkant av samtalen? *

i Dette elementet vises kun dersom alternativet «Ja» er valgt i spørsmålet «5. Har du hatt kontakt med en chatbot i offentlig sektor i Norge de siste tre årene?»

Med faglig kompetanse menes hva chatboten kan gjøre og hva den hjelpe med.

12. Hvilket inntrykk fikk du av chatbotens faglige kompetansenivå under samtalen? *

i Dette elementet vises kun dersom alternativet «Ja» er valgt i spørsmålet «5. Har du hatt kontakt med en chatbot i offentlig sektor i Norge de siste tre årene?»

(3 - nøytral)

Lite kompetent

Svært kompetent



Verdi

12a. Beskriv ditt inntrykk av chatbotens faglige kompetanse. *

i Dette elementet vises kun dersom alternativet «Ja» er valgt i spørsmålet «5. Har du hatt kontakt med en chatbot i offentlig sektor i Norge de siste tre årene?»

13. I hvilken grad førte ditt inntrykk av chatbotens faglige kompetanse til at tilliten til systemet endret seg? *

i Dette elementet vises kun dersom alternativet «Ja» er valgt i spørsmålet «5. Har du hatt kontakt med en chatbot i offentlig sektor i Norge de siste tre årene?»

Tilliten ble svekket

Ingen endring

Tilliten økte

13a. Forklar hvordan chatbotens faglige kompetanse førte til at din tillit til systemet endret seg.

i Dette elementet vises kun dersom alternativet «Tilliten økte» eller «Tilliten ble svekket» er valgt i spørsmålet «13. I hvilken grad førte ditt inntrykk av chatbotens faglige kompetanse til at tilliten til systemet endret seg?»

14. I hvilken grad førte bruken av chatboten til at din tillit til den offentlige tjenesten endret seg? *

i Dette elementet vises kun dersom alternativet «Ja» er valgt i spørsmålet «5. Har du hatt kontakt med en chatbot i offentlig sektor i Norge de siste tre årene?»

Tilliten ble svekket

Ingen endring

Tilliten økte

14a. Forklar hvorfor din tillit til den offentlige tjenesten endret seg.

- i** Dette elementet vises kun dersom alternativet «Tilliten ble svekket» eller «Tilliten økte» er valgt i spørsmålet «14. I hvilken grad førte bruken av chatboten til at din tillit til den offentlige tjenesten endret seg?»

15. Presenterte chatboten svarene sine på en hensiktsmessig måte? *

- i** Dette elementet vises kun dersom alternativet «Ja» er valgt i spørsmålet «5. Har du hatt kontakt med en chatbot i offentlig sektor i Norge de siste tre årene?»

Ja

Nei

15a. Hvordan forventer du at chatboten skal svare? *

- i** Dette elementet vises kun dersom alternativet «Nei» er valgt i spørsmålet «15. Presenterte chatboten svarene sine på en hensiktsmessig måte?»

