

Knut Yngve Barstad Gjelle

Chatboter, tillit og brukeropplevelse i det offentlige

En casestudie av forholdet mellom tillit til chatboter og brukeropplevelsen

Masteroppgave i Digital Samhandling

Veileder: Elena Parmiggiani

Juli 2021

Knut Yngve Barstad Gjelle

Chatboter, tillit og brukeropplevelse i det offentlige

En casestudie av forholdet mellom tillit til chatboter og brukeropplevelsen

Masteroppgave i Digital Samhandling
Veileder: Elena Parmiggiani
Juli 2021

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



Kunnskap for en bedre verden

Sammendrag

Bruk av AI-kontrollerte chatboter blir stadig vanligere både i privat og offentlig sektor. I Norge er adaptasjonen innenfor offentlig sektor en del av en overordnet digitaliseringsstrategi som skal gi mer effektive og bedre tjenester til landets innbyggere. For at innbyggerne skal ta i bruk de nye digitale tjenestene som dukker opp som følge av digitaliseringen, er det viktig at tjenestene skaper og ivaretar tilliten hos befolkningen. Innbyggernes brukeropplevelse i møte med chatboter er viktig for å ivareta og styrke denne tilliten.

Chatboter er i dag tatt i bruk av flere statlige og kommunale organisasjoner og fungerer som en del av deres førstelinje-kundestøtte. De er alltid tilgjengelige og har stor kapasitet til å svare på henvendelser fra innbyggerne. Dette frigjør arbeidstid fra mennesker som ellers måtte besvare de samme henvendelsene. Chatboter som brukes på denne måten er oppgave-orienterte og det er avgjørende for brukeropplevelse og tillit at de klarer å svare godt på brukernes henvendelser.

Tidligere forskning har avslørt at enkelte grupper blant innbyggerne får problemer i møte med denne teknologien. Enkelte forstår ikke at de snakker med en AI-kontrollert chatbot, andre er ikke villige til å gi chatboten en sjanse og ber om å få snakke med menneskelige veiledere med gang. En tredje gruppe interagerer ikke med chatboten i det hele tatt etter å ha tatt kontakt.

Studien tar utgangspunkt i eksisterende teori om tillit til AI generelt og AI-kontrollerte chatboter, både i en privat og offentlig setting, samt brukeropplevelser i møte med chatboter.

Denne forskningen har som mål å bidra med empirisk innsikt i hvordan en statlig etat jobber med videreutvikling av chatbot-teknologien, med hensyn til tillit og brukeropplevelser og bygge videre på eksisterende teori innen feltet menneske-maskin interaksjon. Forskningsspørsmålet som besvares er: *Hvordan påvirker tillit til et system brukeropplevelsen*. Systemet det er snakk om er i denne sammenhengen er chatboten Frida, som representerer NAV.

Forskningen er gjennomført som en deskriptiv casestudie, og har et kvalitativt forskningsdesign for å undersøke forholdet mellom tillit og brukeropplevelsen av chatboter innen offentlig sektor. Datainnsamling er gjort ved hjelp av semistrukturerte intervjuer med ansatte hos NAV som jobber med Frida. I tillegg er det gjennomført en spørreundersøkelse i mars-april for å samle erfaringer fra brukere av chatboter innen offentlig sektor. Dataanalysen er gjennomført i henhold til en stedvis-deduktiv metode.

Funnene fra undersøkelsen beskriver hvordan NAV jobber med videreutvikling av chatboten sin, hvilke tanker og vurderinger de gjør med tanke på å legge til rette for en god brukeropplevelse og skape tillit blant brukerne. Svar fra spørreundersøkelsen viser respondentens holdninger og erfaringer med digitalisering og automatisering av offentlige tjenester. Det avslører et bredt spekter av forventninger, og hvordan disse påvirker brukeropplevelsen i de individuelle tilfellene. Dette danner et grunnlag for om tilliten til systemet styrkes eller svekkes i etterkant av møte med chatboter.

Abstract

The use of AI-controlled chatbots is expanding both in the private and public sectors. In Norway, the adaption of these types of systems is part of an overarching strategy for digitalization of the public sector. One of the goals of this ongoing process is to provide better and more efficient services to the public at large. It is important to create and maintain public trust to ensure that new services that appear because of the ongoing digitalization is adopted. The public's user experiences from meetings with chatbots is important to maintain and strengthen this trust.

Chatbots are adopted by several public organizations and work as part of their first line of customer support. They are always available and have unlimited capacity to answer inquiries from the public. This frees up available workhours from human employees that would otherwise be needed to answer public inquiries. Chatbots that are used in this manner are task-oriented and it is crucial for the user experience they provide that they respond well to these inquiries.

Prior research has revealed that parts of the public struggle when they encounter AI-controlled chatbots. Some do not understand that they are talking to an AI. Others are not willing to interact with the chatbot and request a human councilor immediately. A third group contact the chatbot without engaging with it at all.

This study is based on prior research into general trust in AI and trust in AI-controlled chatbots, both in a private and public setting, as well as research into user experiences regarding chatbots.

This thesis aims to contribute with empiric insight into how a government agency work with continual development of chatbot technology regarding trust and user experiences, and to further develop existing theory in the field of human-machine interaction. To do this, the following research question is explored and answered: How does trust in a system affect user experience? The system in this case refers to Frida, a chatbot employed by NAV, the government agency running Norway's welfare programs.

The research is done as a descriptive case study and employs a qualitative research strategy to examine the relationship between trust and user experience of chatbots in the public sector. Data was gathered through interviews with employees that work with Frida at NAV. Additionally, an online questionnaire was run in March-April to gather experiences from users of public sector chatbots. The data was analyzed using a stepwise deductive-inductive method.

Findings from the interviews describe how NAV work with continual development of their chatbot, their thoughts and judgements regarding creation of a good user experience and creating and maintaining trust among the public. Findings from the questionnaire show attitudes and user experiences with digitization and automation of public services. It reveals a wide range of user expectations and show how these affect user experience in individual cases. This forms a basis for whether trust in the system is strengthened or weakened after a user's interaction with a chatbot.

Forord

Denne masteroppgaven markerer slutten på et to-års masterstudium i Digital samhandling ved Institutt for datateknologi og informatikk. Oppgaven ble skrevet i løpet av vår-semesteret 2021 i tilknytning til prosjektet Frida@work. Prosjektet er et samarbeid mellom NAV, NTNU, UiA og UiO.

Først vil jeg takke veileder for oppgaven, førsteamanuensis Elena Parmiggiani, for konstruktive tilbakemeldinger, gode diskusjoner om tema og motivasjon underveis.

Takk til medstudent Maria Midjo Rønning for samarbeid om datainnsamling, samt verdifulle diskusjoner om funn.

Takk til alle informanter og andre involverte fra NAV for tiden dere satte av, og innspillene dere kom med!

En siste takk går til Ban, Jon og Svein for hjelp med korrekturlesning og formatering.

Trondheim, juli 2021

A handwritten signature in black ink, reading "Knut Gjelle", is displayed on a light gray rectangular background. The signature is written in a cursive, flowing style.

Knut Yngve Barstad Gjelle

Innhold

Sammendrag	i
Abstract	ii
Forord	iii
1. Innledning.....	1
1.1 Forskningsspørsmål.....	3
1.2 Oppgavens struktur.....	3
2. Teori	5
2.1 Teoretisk bakgrunn	5
2.2 AI-styrte chatboter	5
2.3 Tillit til AI	6
2.4 Brukeropplevelse	9
3. Casebeskrivelse	11
4. Metode	13
4.1 Forskningsstrategi	13
4.2 Forskningsparadigme	14
4.3 Datainnsamling	14
4.3.1 Intervju	14
4.3.2 Spørreundersøkelse	15
4.4 Dataanalyse.....	16
5. Empiriske Funn	19
5.1 Behov	19
5.2 Brukeropplevelser.....	21
5.3 Forventninger	23
5.4 Kompetanse.....	27
5.5 utfordringer	30
5.6 Tillit	34
6. Diskusjon	39
6.1 Tillit	39
6.2 Brukeropplevelser.....	41
6.3 Forholdet mellom tillit og brukeropplevelser	42
6.4 Praktiske implikasjoner	43
7. Konklusjon	44
7.1 Begrensninger og fremtidig arbeid	45

8.	Referanser	46
9.	Vedlegg	49
	Vedlegg A: Intervjuguide innledende intervju	49
	Vedlegg B: Intervjuguide ansatt.....	50
	Vedlegg C: Intervjuguide leder	51
	Vedlegg D: Spørreundersøkelse	52

Figurer

Figur 1	Illustrasjon av interaksjonene mellom chatbot og omgivelsene rundt.	11
Figur 2	Til venstre: informasjon om personopplysninger før man får tilgang til chaten på Nav.no. Til høyre: Velkomstmeldingen fra Frida etter at man har lest informasjonen om personopplysningene.	12
Figur 3	Oversikt over forskningsdesign.....	13
Figur 4	Tidlig koding av svar fra spørreundersøkelsen.	17
Figur 5	Tillit til automatiserte tjenester i det offentlige.	35
Figur 6	Illustrasjon av forhold mellom forventninger, tillit og brukeropplevelser.	43

Tabeller

Tabell 1	Oversikt over intervjuobjekt i case-organisasjonen	15
Tabell 2	Oversikt over hovedtema med underkategorier.....	18

1. Innledning

I 2012 la regjeringen i Norge frem en plan for digitalisering av offentlige tjenester. Et av målene for digitaliseringen er at offentlig forvaltning skal være tilgjengelig via nettbaserte løsninger (Standard Norge, 2012). Dette er siden blitt videreført gjennom sittende regjerings digitaliseringsstrategi for offentlig sektor 2019-2025 (*En digital offentlig sektor*, 2019). Digitaliseringen av Norge har to overordnede mål. For det første er det besparende for det offentlige ved at mange digitale løsninger er mer effektive enn det ville vært å levere den samme tjenesten analogt. Som eksempel på en god digital løsning kan man se på skattemeldingene som tidligere hadde et stort apparat for å skrive ut forklaringer i tillegg til selve skattemeldingen og distribuere disse til innbyggerne. For det andre kan digitale løsninger også være mer brukervennlige og effektive å benytte for innbyggerne. Ulike former for informasjonssøk nevnes spesielt da man før dette var digitalt, måtte benytte telefon eller personlige møter for å finne frem til informasjonen og få denne overbrakt.

Det overordnede målet med denne digitaliseringen er en enklere hverdag både for innbyggerne og for bedrifter, samt effektivisering i det offentlige og for å legge til rette for produktivitetsøkning i samfunnet for øvrig (Meld. St. 27, (2015–2016)).

Digitaliseringen av det offentlige innebærer at måten innbyggerne kommuniserer med det offentlige på, endrer seg. Samtidig er det viktig for demokratiet at disse kommunikasjonslinjene fungerer godt, ettersom dette også inkluderer det sosiale sikkerhetsnett vi er vant til (Arbeids- og velferdsetaten, U.Å.). En nøkkel for å lykkes med målene som er satt for digitaliseringen er å ivareta innbyggernes tillit til staten, og de digitale løsningene man velger å bruke (Alzahrani, Al-Karaghoulouli & Weerakkody, 2017). Et aspekt av dette, som gjelder spesielt i forbindelse med AI-teknologi, handler om hvorvidt man kan stole på teknologien i seg selv. Dette henger sammen med de store mulighetene og risikoen som ligger i teknologien (Floridi et al., 2018; Thiebes, Lins & Sunyaev, 2020).

Overgangen til digitale tjenester stiller krav til ferdigheter hos brukerne av tjenestene og det er viktig at samfunnet også legger til rette for, og ivaretar, brukere som mangler tilstrekkelig digital kompetanse. Den digitale kompetansen til hver enkelt bruker kan være avgjørende for om man lykkes med det man ønsker å oppnå. Brukere som har kunnskap om en søkemotors algoritmer, har et mer balansert syn på informasjonen de finner (Cotter & Reisdorf, 2020). Uten denne kunnskapen er det fare for at brukeren kan tro informasjon de kommer over på internett er objektiv sann, med de implikasjonene det kan ha.

For at man skal dra nytte av digitale tjenester i det offentlige er det nødvendig at innbyggeren forstår hvordan man bruker tjenesten, og samtidig har en forståelse av hvilke konsekvenser de valgene man tar innebærer (Bratteteig, Saplacan, Soma & Svanes Oskarsen, 2020).

En ny teknologi som har gjort sitt inntog både i det private næringsliv og etter hvert også i det offentlige, er bruken av AI-kontrollerte chatboter som en del av førstelinje-kundestøtte. Innenfor det offentlige i Norge har vi chatboter som NAVs Frida, Kommune-Kari og Skatteetatens chatbot (NAV, U.Å.; Sem & Stenersen Prokom, U.Å.; Skatteetaten, U.Å.). Kommune-Kari er i dag tilgjengelig i 90 kommuner, og betjener med det omtrent 1,7

millioner av Norges innbyggere (Sem & Stenersen Prokom, U.Å.). Alle disse tre chatbotene er oppgave-orienterte chatboter.

Chatboter, eller virtuelle agenter som teknologien også kalles, er en form for kunstig intelligens (AI) og benytter seg av maskinlæring og natural language processing (NLP) for å lære seg å forstå språk og brukernes intensjoner. Deretter finner de frem til den riktige responsen i en database og gir denne til brukeren.

Oppgave-orienterte chatboter er godt egnet for å gjøre informasjonssøk mer effektive for brukerne ved at man gjennom samtale med chatboten kan følge en logisk vei til informasjonen man er ute etter. For brukerne er det dermed nødvendig at man viser tillit til chatboten og gir den en sjanse til å gjøre jobben sin.

For at brukerne skal få til å utnytte mulighetene som ligger i teknologien, er det essensielt at man har tillit til mer enn informasjonen AI kan gi brukeren. Det vil si at man forsøker å maksimere gevinstene og minimere risiko ved bruk av teknologien. For å sikre dette har EU utarbeidet et sett med syv prinsipper for etisk og ansvarlig bruk av kunstig intelligens (Ala-Pietilä et al., 2019). Prinsippene er utarbeidet av en uavhengig gruppe eksperter innen AI (ibid.). Regjeringen i Norge legger de samme prinsippene til grunn for utvikling og bruk også i Norge (*Nasjonal strategi for kunstig intelligens*, 2020). Pålitelig kunstig intelligens er et begrep innen forskningen og kommer fra ideen om at tillit er fundamentet vi bygger sivilisasjonen vår rundt og at tillit derfor er essensielt for å nå potensialet som ligger i kunstig intelligens (Thiebes et al., 2020).

Forskning viser at graden av tillit chatboter får er avhengig av mange faktorer. Hvor lik et menneske chatboten er i språk, utseende (avatar) og selve interaksjonen mellom mennesker og chatboten er viktige. Forskning på bruk av chatboter for kundestøtte kommersielt viser at tillit til organisasjonen bak chatboten også påvirker potensielle brukeres tillit til chatboten (Følstad, Nordheim & Bjørkli, 2018).

Tidligere forskning har vist at enkelte brukere har problemer med å forstå at de møter en slik teknologi, de tror de snakker med et menneske. Andre gir opp øyeblikkelig, uten at man vet hvilke årsaker som ligger bak. NAV hadde i overkant av 855 000 henvendelser som ble svart av chatbot i løpet av 2020 (T. Simonsen, 2021). I tillegg kommer rundt 30% (ca. 255 000) som aldri gjør mer enn å åpne chatvinduet (Intervju Subjekt 1, 2021).

For å forstå hva som gjør at noen avslutter samtalen med en gang, samt andre utfordringer brukerne kan møte i samtalene med slike chatboter må man studere interaksjonene mellom mennesker og maskin, i dette tilfellet AI (Ågerfalk, 2020). Selve brukeropplevelsen er en sentral bit i denne interaksjonen (Følstad & Brandtzaeg, 2020; Law, Roto, Hassenzahl, Vermeeren & Kort, 2009).

Norge er spesielt godt egnet for denne typen forskning ettersom landet er et av de mest digitaliserte i verden, også når det kommer til det offentlige (Mikalef, Fjørtoft & Torvatn, 2019).

Denne studien bidrar med dypere innsikt om bruk av virtuelle agenter innen offentlig sektor og bygger videre på eksisterende teori innen menneske-maskin interaksjon med hensyn til tillit og brukeropplevelse ved bruk av denne teknologien.

Innenfor teori om systemutvikling og brukerinvolvering viser casestudien hvordan brukerinvolvering kan gjøres indirekte og bidrar med dette med ny teori (Walsham, 1995). De praktiske implikasjonene studien avdekker, kan være nyttige i kontekst av bruk og utvikling av lignende system innen andre områder av offentlig forvaltning (ibid.).

Forskningsdesignet for denne oppgaven er tredelt. Teori som omhandler tillit til virtuelle agenter, bruk av disse innen offentlig sektor og om brukeropplevelser og brukerinvolvering i systemutvikling, legges til grunn for forskningene.

Data er samlet inn fra casebedriften gjennom semistrukturerte intervju med tolv informanter som jobber direkte med NAVs chatbot eller har en lederrolle med påvirkning for chatbotens arbeidsområde. Dette er med på å danne et grunnlag for å si noe om hvordan organisasjonens bruk av en virtuell agent støttes av teori om brukeropplevelser og tillit til virtuelle agenter.

Det er også gjennomført en spørreundersøkelse (se vedlegg D) som samler erfaringer fra brukere av chatboter innen offentlig sektor. Funn fra spørreundersøkelsen sammenlignes med og diskuteres i lys av funn fra intervjuene og gjeldende teori.

1.1 Forskningsspørsmål

I de senere årene er det forsket mye både på tillit til chatboter og også en del på brukeropplevelser i forbindelse med chatboter. Det mangler likevel forskning som drar paralleller mellom disse områdene og ser hvordan de påvirker hverandre. I denne oppgaven ser vi nærmere på hvordan tillit og brukeropplevelser påvirker hverandre og svarer på spørsmålet:

Hvordan påvirker tilliten til et system brukeropplevelsen?

Dette må sees i kontekst av virtuelle agenter som opererer innenfor offentlig sektor som denne studien omhandler.

1.2 Oppgavens struktur

Oppgaven er delt i 7 kapitler og følger denne strukturen:

Kapittel 1 består av en innledning med bakgrunnsinformasjon om studien sammen med motivasjon for forskningsspørsmålet.

Kapittel 2 presenterer teoretisk bakgrunn for studien og relevant teori som brukes til å besvare forskningsspørsmålet.

Kapittel 3 definerer casen som er grunnlaget for oppgavens studier.

Kapittel 4 forklarer oppgavens forskningsdesign, metoder for datainnsamling og dataanalyse.

Kapittel 5 presenterer funn fra intervju og spørreundersøkelse i henhold til hovedtema fra datanalysen.

Kapittel 6 diskuterer funnene i henhold til teorien og presenterer teoretiske og praktiske implikasjoner av det som er oppdaget.

Kapittel 7 avslutter oppgaven med en konklusjon og noen avsluttende tanker om forskningsbegrensninger og mulig fremtidig arbeid.

2. Teori

I dette kapitlet presenteres teorien resten av oppgaven bygger på. Leseren introduseres til konseptet kunstig intelligens, dets plass i samfunnet og en forklaring av tre hovedgrupper kunstig intelligens. Deretter blir grunnleggende informasjon om hvordan chatboter fungerer presentert, før vi går dypere inn i tillit tilknyttet AI og chatboter samt brukeropplevelse tilknyttet denne teknologien.

2.1 Teoretisk bakgrunn

Begrepet kunstig intelligens handler om datasystemer som har egenskaper som tidligere har vært unike for oss mennesker (Valevatn, 2017). Det er flere grunner til at denne typen system for alvor har blitt tatt i bruk det siste tiåret. For det første har bruken av internett gjort at store mengder data er tilgjengelig, uansett hvor man er i verden. Den teknologiske utviklingen av dataprosessering gjør at man kan utføre mer komplekse databeregninger enn tidligere i tillegg til at måten datamaskiner er i stand til å lære på (maskinlæren) er blitt forbedret (ibid.). Samtidig som utviklingen av kunstig intelligens har gått fremover, har den til gode å bli like god eller bedre enn mennesker for alle typer oppgaver.

Rollen til kunstig intelligens i den moderne økonomien er ikke å erstatte mennesker, men frigjøre de menneskelige ressursene som tidligere har vært bundet til gjentakende oppgaver (Mikalef et al., 2019). Innen offentlig forvaltning kan dette gjelde oppgaver som å svare på generelle spørsmål fra brukerne, eller sørge for at brukeren får riktig søknadsskjema for deres situasjon. Dette øker organisasjoners evne til å bruke disse ressursene til oppgaver som problemløsning og kreativ tenkning (ibid.). For NAV betyr det at mer av tiden veiledere er på jobb, kan brukes på mer kompleks saksbehandling. Løst på en god måte drar man nytte av komplementære styrker fra både mennesker og kunstig intelligente systemer (ibid.). Både skatteetaten og NAV benytter kunstig intelligens i sine chatboter for å gi menneskelige veiledere mer tid til å håndtere kompliserte arbeidsoppgaver.

2.2 AI-styrte chatboter

Glikson og Woolley (2020) skiller mellom tre ulike typer AI; robot, virtuell og innebygd AI. AI-Roboter skiller seg fra de andre ved at den har en fysisk komponent mennesker kan forholde seg til. Virtuell og innebygd AI i motsetning er kunstig intelligens som ikke har fysisk tilstedeværelse. Virtuelle agenter har like fullt en tydelig identitet ved at den har et navn, en avatar eller lignende. Chatboter, som denne oppgaven ser nærmere på, befinner seg i dette segmentet. Innebygd AI er usynlig for brukerne, og finnes blant annet i systemene bak søkemotorer på internett og GPS-er.

En chatbot eller chatrobot er et dataprogram som finner informasjon fra fritekst og svarer på en tilfredsstillende måte (Goyal, Pandey & Jain, 2018). Denne formen for AI er mye smalere enn hva man vanligvis tenker på i forbindelse med AI, som gjerne defineres som programmer som kan utføre oppgaver på tilsvarende nivå som mennesker (Aoki, 2020).

Chatboter kan også separeres ut ifra hva slags arbeidsoppgaver de løser. På den ene siden har vi oppgaveorienterte chatboter som løser konkrete brukerbehov. Chatbotene som nevnes i innledningen, Frida, Kommune-Kari og Skatteetatens chatbot tilhører denne kategorien. På den andre siden har vi chatboter som er designet som sosiale agenter eller

samtalepartnere. Microsofts Xiaoice er et eksempel på en svært vellykket sosial agent (R. Liao, 2020).

Måten mennesker kommuniserer med denne formen for AI har fellestrekk med hvordan mennesker kommuniserer seg imellom ved hjelp av tekstmeldinger eller instant message program. Denne kommunikasjonsformen kalles computer mediated communication (CMC) og har trekk fra både skriftlig og muntlig kommunikasjon (Hill, Ford & Ferreras, 2015). Forskningen til Hill et al. (2015) fant ut at de menneskelige brukerne av en chatbot sendte flere meldinger til chatboten enn de gjorde i menneske-menneske samtaler. Et kjennetegn ved samtale med chatboten ligger i dens begrensninger for å ha lengre diskusjoner om et tema, og muligheten for å bruke erfaring fra tidligere samtaler. Samtalene er preget av å være enklere enn det som er mulig mellom mennesker, fordi teknologien enda ikke kan simulere rekkevidden og kompleksiteten i menneskelig kommunikasjon (ibid.). Disse begrensningene kan gjøre det vanskelig for brukere som ikke er kjent med teknologien å få svar på sine henvendelser. En AIs evne til å utføre arbeidsoppgaven den er satt til, er viktig for innbyggernes tillit til systemet (Aoki, 2020).

2.3 Tillit til AI

Tradisjonelt defineres tillit som at man er villig til å gjøre seg sårbar overfor en annen part (Mayer, Davis & Schoorman, 1995). Når man snakker om tillit innenfor digital samhandling, må begrepet sees i kontekst av omgivelsene og aktørene som er involvert. Olson og Olson (2000) sier at tillitt er viktig for samarbeid på tvers av fysiske lokasjoner. Denne tilliten bygges over tid med utgangspunkt i det forfatterne kaller "felles bakgrunn" (eng: common ground), kunnskap individene har felles og som alle parter vet at de har felles. Et eksempel på dette kan være prat om været, personer som oppholder seg på samme plass vil ha mye lettere for å forstå hva som legges i et utsagn, enn noen som deltar i møte digitalt og oppholder seg på andre siden av Atlanterhavet. Chatboter deler sånn sett noe av imaget sitt med organisasjonen den representerer (Følstad et al., 2018). Tillit til offentlige etater er, sammen med tekniske faktorer, risikofaktorer og aspekter rundt den individuelle innbyggeren, avgjørende for om innbyggeren vil ta i bruk en offentlig digital tjeneste (Alzahrani et al., 2017).

Konseptet tillit er et komplekst fenomen og hva som inngår i pålitelig kunstig intelligens diskuteres fortsatt (Thiebes et al., 2020). For at en AI skal kunne defineres som pålitelig, må den ifølge forfatterne følge fem etiske prinsipper; velvilje, ikke-ondskapsfullhet, autonomi, rettferdighet og forklarbarhet (Floridi et al., 2018; Thiebes et al., 2020). Velvilje i lys av at AI-en er utviklet, distribuert og brukes på en måte som er bærekraftig og fremmer miljø og menneskerettigheter. Hensyn til miljø og bærekraft kan knyttes til utbygging og drift av datasenter, noe også det offentlige har behov for og ett spesielt ansvar for. For chatboten i studien ser vi at menneskerettighetsperspektivet er ivaretatt gjennom informasjonen i velkomstmeldingen. Denne utdyper hvordan personvernet opprettholdes og data fra samtalen behandles i ettertid.

Ikke-ondskapsfull er nært beslektet det første prinsippet, men er mer spisset mot personvern og fysisk sikkerhet. Autonomi handler om både kunstig intelligens evne til å ta avgjørelser selv, og menneskers mulighet til å overstyre og ta egne avgjørelser. Rettferdighet går på bruk av AI for å bekjempe ulikhet, skapelse av og fordeling av goder

og hindring av nye urettferdigheter som følge av AI. Forklarbarhet handler om at teknologien må være forståelig (gjennom bruk av AI-modeller) samtidig som den beholder sin effektivitet og treffsikkerhet. En del av dette innebærer også at hvem som er ansvarlig for avgjørelser AI-en tar er klargjort.

Knowles og Richards (2021) argumenterer for at enkeltpersoner ikke skal være ansvarlige for å avgjøre om en AI er til å stole på, men at det er et ansvar som faller på personer med tilstrekkelig kunnskap til å forstå AI-en. For at allmennheten skal kunne ha tillit til ulike AI-er som brukes i samfunnet, argumenterer Knowles og Richards for at det må et regelverk på plass, samt ressursene og myndighetsgjøring av individer til å kunne håndheve dette.

Aoki (2020) definerer offentlig tillit som innbyggernes opplevde sannsynlighet for at en maskin utfører arbeidsoppgaven sin på en gunstig måte. Det vil si at man har tro på at maskinen eller programmet kan gjøre jobben det er laget for, og at den lykkes i dette ved positiv oppførsel. Tilliten innbyggerne har til et system vil også endre seg ut fra hva man har hørt om systemet, til man har prøvd det selv, og også etter at interaksjonen er ferdig.

I forbindelse med automatiserte tjenester defineres tillit som viljen til å bruke tjenesten sammen med forventninger om at man som bruker kan nå behovet man prøver å dekke. I dette tilfellet bruken av virtuelle agenter til kundebehandling i det offentlige.

Den grunnleggende tilliten til chatboter brukt av det offentlige kan sees i lys av de tre faktorene hensikt, prosess og utførelse (Aoki, 2020). I forbindelse med chatboter som brukes av offentlig forvaltning, er det hensikten myndighetene kommuniserer til befolkningen som blir viktig. Når myndighetene introduserer intensjonene bak bruken på en velvillig og åpen måte, skaper dette tillit til teknologien uten at innbyggerne stiller store spørsmål. I dette tilfellet handler prosess om ansvarliggjørelse for anskaffelser overfor egen befolkning. Dette er ekstra viktig i et demokrati hvor de som styrer gjør dette på vegne av folket (Levitan, 1946). I Norge sikres det at anskaffelsesprosessen i det offentlige løses på en god måte gjennom loven om offentlig anskaffelser (Anskaffelsesloven, 2016). Åpenhet rundt personvern og databehandling der AI-støttede systemer er involvert i beslutningstaking kan også være en kilde til tillit for det offentlige. Samtidig vil kunnskapen om algoritmene og teknologien bak være utenfor interessen til store deler av befolkningen (Aoki, 2020). Dette har sammenheng med at teknologien bak kan være svært avansert, og kan sammenlignes med å eie en bærbar pc eller nettbrett. Så lenge det virker trenger man ikke vite hvordan de er skrudd sammen teknisk. Sluttes de å virke trenger man hjelp av personer med ekspertkunnskaper. I denne casen tar heller ikke chatboten noen avgjørelser og prosessen rundt anskaffelsen av chatboten blir derfor heller ikke avgjørende for tilliten i befolkningen. Dersom den hadde tatt avgjørelser, ville dette medført større konsekvenser for den enkelte innbyggeren som kommer i kontakt med systemet. Hvis prosessen rundt anskaffelsen ikke er ryddig, kan det i så fall få brukeren til å stille spørsmål ved systemets integritet.

En chatbots utførelse spiller på dens evne til å svare tilfredsstillende på brukernes spørsmål (Aoki, 2020). Med mangel på informasjon om bruken av chatboten kan dette telle negativt for brukernes tillit til systemet. Andre kan ha høye forventninger som ikke blir innfridd. Empati og evnen til å svare med det riktige nivået av empati er viktig i de tilfellene hvor chatboten svarer på sensitive tema.

Menneskers tillit til kunstig intelligens kan også skilles mellom kognitiv og emosjonell tillit (Glikson & Woolley, 2020). Kognitiv tillit er basert på rasjonelle valg og vurderinger gjort av brukeren. Denne tilliten utvikler seg typisk fra høy til lav og fra lav til høy når brukerne blir kjent med system og opplever at det under- eller overpresterer sammenlignet med forventninger. Håndgripelighet eller forståelse rundt den virtuelle agentens algoritmer med forklaringer på hvorfor beslutninger tas, skaper også tillit ved å balansere forventninger. Pålitelighet i svarene den kunstige intelligensen gir, er også viktig. Opplevelsen av et upålitelig svar svekker tilliten, mens transparens som gir brukeren kunnskap om at AI-en kan gi upålitelige svar også kan øke tilliten. Personalisering og teknikk for overtalelse spiller også inn på brukernes tillit i de tilfellene dette er en del av den kunstige intelligensens arbeidsoppgaver. Når det gjelder evnen til overtalelse er disse mer effektive når man får eksempler av AI-en, istedenfor argumentasjon for at noe er riktig eller eksempler sammen med argumentasjon (Kraus, 2012). Både tilfellet om personalisering og overtalelse henger sammen med kognitive beslutninger fra brukerne i møte med den virtuelle AI-en.

Emosjonell tillit kan være mer irrasjonell og variere ut fra brukerens følelser og humør (Glikson & Woolley, 2020). Håndgripeligheten til den kunstige intelligensen har mer å si for den emosjonelle enn den kognitive tilliten. For eldre brukere skaper spesielt tilstedeværelsen av en personifisering av systemet tillit ved en reduksjon av nervøsitet (Chattaraman, Kwon, Gilbert & Li, 2014). Dersom personifiseringen gjøres ved å gi den virtuelle AI-en en menneskelignende avatar med muligheter for tale, risikerer man å havne i uncanny valley, noe som vil redusere tilliten (Ciechanowski, Przegalinska, Magnuski & Gloor, 2019). Uncanny valley er et begrep for menneskelignende teknologi som vekker følelse av uro eller avsky hos mennesker fordi de oppfattes som nesten-men-ikke-helt menneskelignende (Mori, MacDorman & Kageki, 2012).

Bruken av antropomorfe trekk hos virtuelle AI-er kan gi brukerne urealistiske forventninger ved å øke forventningene ut over det systemet er designet for samtidig som personaliserte AI-er som er tilpasset kulturen de befinner seg i, kan øke deres appell hos brukerne (Glikson & Woolley, 2020). Et antropomorf trekk kan i dette tilfellet dreie seg om en avatar som ser menneskelig ut.

For chatboter som jobber med kundestøtte er det fem områder som påvirker brukerens tillit til systemet (Følstad et al., 2018). Effektiv tjenesteløsning, åpenhet om begrensninger, menneskelignende samtaler, tillit til merkevaren og sikkerhet (ibid.). Effektiv tjenesteløsning handler om å løse brukerens behov på en måte som er fordelaktig sammenlignet med ander alternativ. Evnen til å løse arbeidsoppgaver går igjen både med tanke på tillit og gode brukeropplevelser, som tas opp i delkapittel 2.4.

Konseptuelle metaforer er korte beskrivelser av et systems funksjonalitet og bruksområder som kan påvirke brukernes forventninger og evalueringer i ettertid (Khadpe, Krishna, Fei-Fei, Hancock & Bernstein, 2020). Brukere vil oppføre seg ulikt og ha ulike forventninger i møte med en ung og nyansatt service-arbeider, sammenlignet med en eldre arbeider som gir inntrykk av å ha hatt den samme jobben i flere tiår. Konseptuelle metaforer som beskriver kunstige intelligenser kan sees i lys av en psykologisk modell som plasserer metaforene i to dimensjoner, en for varme og en for kompetanse (Cuddy, Fiske & Glick, 2008; Fiske, Cuddy, Glick & Xu, 2002). Varme handler om medmenneskelighet og evnen til

å vise empati. Kompetansen kommer til uttrykk gjennom chatbotens evne til å vise intelligens, ansvar og dyktighet (Khadpe et al., 2020).

Funn fra Aoki (2020) viser at det kan skape mer tillit å kommunisere fordeler som tilgjengelighet utenfor arbeidstid enn å ikke gjøre det. Samtidig viste studien at det ikke gjorde stor forskjell på om man omtaler chatboten som et effektivt hjelpemiddel for innbyggerne, kontra hjelpemiddel for de ansatte hos organisasjonen.

Samtidig kan saker som gjelder innbyggernes liv, ofte være så kompliserte og sammensatte at det for individet er vanskelig å vite hvor man skal starte. I tilfeller med denne typen innviklede saker er det viktig for brukerens autonomi at man kan få hjelp til å dekonstruere problemet slik at man ser de ulike mulige løsninger og kan veie disse mot hverandre (Bratteteig et al., 2020). Innbyggerne er avhengig av denne muligheten for å kunne være autonome. Sett i sammenheng med måten man kommuniserer med chatboter på, kan dette skape utfordringer for brukerne (Hill et al., 2015).

Det krever en viss digital kompetanse for at innbyggerne skal klare å nyttiggjøre seg de mulighetene digitale tjenester i offentlig sektor åpner for. Dette gjelder både for å forstå hva tjeneste går ut på, og hva valgene man tar innebærer for fremtiden (Bratteteig & Verne, 2012). Oppretter man digital postkasse i Norge betyr det at man må sjekke denne for å motta informasjon som ellers ville kommet i den analoge postkassa. For å lykkes med automatisering av offentlige tjenester gjennom bruk av virtuelle agenter, vil det være viktig at disse tjenestene gjøres så enkle og transparente som mulig. På den måten vil innbyggerne ha bedre forutsetninger for å finne hjelpen de trenger autonomt (Bratteteig & Verne, 2012). Samtidig som man da får evnene til å hjelpe seg selv, må man ha tillit til at de digitale tjenestene man benytter kan løse brukerens behov. Kartlegging av brukerbehov er tett knyttet til kunnskap om gode brukeropplevelser (Følstad & Brandtzaeg, 2020; Quiñones, Rusu & Rusu, 2018).

2.4 Brukeropplevelse

Brukeropplevelse defineres som en brukers oppfatning og svar som følge av bruk og/eller forventet bruk av et system (International Organization for Standardization, 2019). Brukeropplevelsen er individuell, selv om den kan bli påvirket av andre gjennom konteksten den skjer (Law et al., 2009). Det vil si at om en gruppe mennesker sammen benytter et system, vil de kunne påvirke brukeropplevelsen til hverandre, men det er ikke gitt at alle vil beskrive sin brukeropplevelse på samme måte.

Brukeropplevelser kan sees i lys av et rammeverk utviklet av Hassenzahl (2018). Dette fokuserer på motivasjon og hvorfor mennesker handler som de gjør i en menneske-maskin-interaksjon. Modellen beskriver dette ved å svare på hvordan, hva og hvorfor vi handler som vi gjør. Som eksempel kan vi se på en tenkt interaksjon med en chatbot som bestiller ferieturer. Hvordan handler i dette tilfellet om de fysiske handlingene menneske må gjøre for å gjennomføre interaksjonen med chatboten, man skriver på tastaturet og bruker musen for å navigere på nettsider. Hva i dette eksemplet er bestillingen av ferieturen. Svar på hvorfor er gode følelser og forventninger om ferieturen.

Følstad og Brandtzaeg (2020) har forsket på brukeropplevelser med chatboter ved hjelp av rammeverket fra Hassenzahl (2018) hvor brukeropplevelsene blir sett i lys av pragmatiske

og hedonistiske egenskaper. Pragmatiske egenskaper hos en chatbot er tett koblet til det å utføre oppgaver for brukerne og hjelpe dem med konkrete behov (Følstad & Brandtzaeg, 2020). Positive opplevelser inkluderer å få den hjelpen man håpet på, mens negative kan knyttes til språkforståelse. For oppgaveorienterte chatboter er deres evne til å utføre oppgavene de er satt til, ved å hjelpe brukerne til å nå deres mål, avgjørende for å skape gode brukeropplevelser.

Disse egenskapene er spesielt viktig i forbindelse med oppgave-orienterte chatboter, som denne casen omhandler. Hedonistiske egenskaper kobles tettere til chatboter som jobber som sosiale agenter (ibid.). Positive opplevelser inkluderer det å bli underholdt, mens negative kan handle om hendelser der chatboten sier noe upassende eller støtende. Selv om pragmatiske og hedonistiske egenskaper er mer fremtredende i visse typer virtuelle agenter, vil en høy tilstedeværelse av begge gi den beste brukeropplevelsen (ibid.). Følstad og Brandtzaeg (2020) forskning ble gjennomført ved hjelp av en spørreundersøkelse hvor man samlet både positive og negative opplevelser fra respondentene.

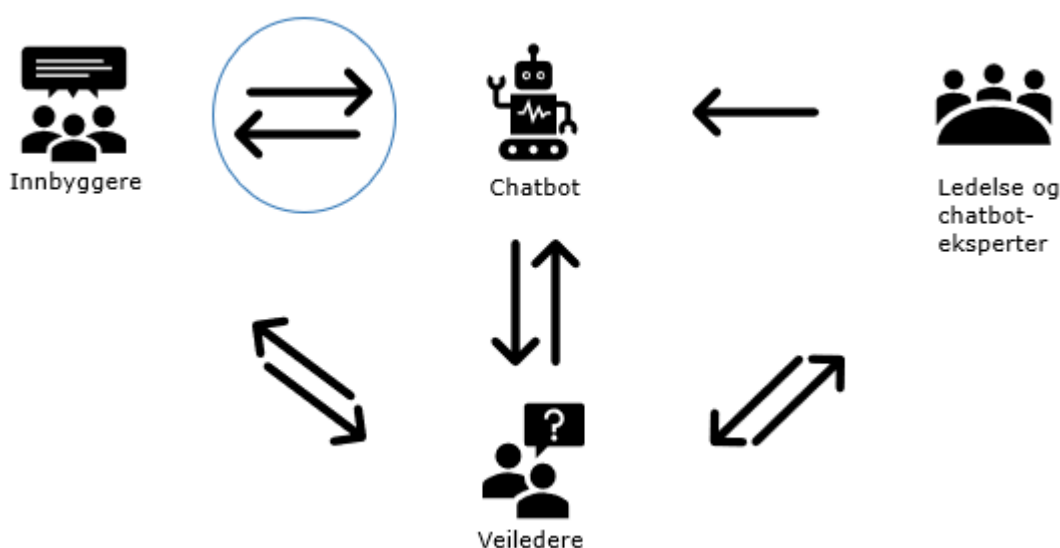
Alder kan også påvirke brukererfaringene, der eldre brukere er mer opptatte av effektivitet og treffsikkerhet mens yngre brukere reagerer mer på lekne og emosjonelt engasjerende chatbots (Følstad & Brandtzaeg, 2020).

Brukernes forståelse av teknologien bak et AI-system kan påvirke brukeropplevelsen. Dette henger sammen med hvordan teknologien er forklart og om brukeren ser nytteverdien i å benytte seg av systemet (Q. V. Liao, Gruen & Miller, 2020). Q. V. Liao et al. (2020) har samlet informasjon fra utøvere innen feltet Explainable-AI, om hva potensielle brukere av AI-system spør om, og hvordan man kan forklare AI-en på en folkelig måte. Dette henger sammen med at algoritmene i maskinlære-AI ofte kan være for kompliserte for lekmenn å forstå. En av informantene fra Q. V. Liao et al. (2020) arbeidet med brukeropplevelser i forbindelse med chatboter og ga uttrykk for at kunnskap om teknologien bak kunne gjøre at brukerne lettere stilte chatboten gode spørsmål. Forklaringer av teknologien bak var ifølge den samme informanten fraværende for brukerne av denne typen AI.

Involvering av sluttbrukere i design og utviklingsarbeidet av informasjonssystem har bred aksept hos praktikanter og forskere innen fagfeltet for menneske-maskin interaksjon (Svanæs & Gulliksen, 2008) og øker brukertilfredsheten til det ferdige systemet (Alzahrani et al., 2017). Dette sikrer brukervennlighet i det ferdige systemet (Svanæs & Gulliksen, 2008), og danner en del av grunnlaget for brukeropplevelsen (Quiñones et al., 2018).

3. Casebeskrivelse

Dette forskningsprosjektet har pågått i løpet av vårsemestret 2021 og er en deskriptiv case-studie (Oates, 2006). NAVs chatbot Frida er utgangspunktet for studien. Valget om å studere Frida i forbindelse med bruken av virtuelle agenter i offentlig sektor er pragmatisk av karakter. Tilgang til informasjon om Frida, samt informanter for intervju er fasilitet gjennom et prosjekt mellom NTNU, UiA, UiO og NAV. Prosjektet dreier seg om Fridas plass i organisasjonen og de ulike universitetene står for ulike tema for prosjektet. NTNU representerer innbyggerne i ulik grad. Denne oppgaven ser spesifikt på forholdet mellom tillit og brukeropplevelse ved bruk av intelligente agenter i offentlig sektor. Se figur 1 for en illustrasjon av aktørene og interaksjonene mellom dem. Den blå sirkelen markerer området for oppgaven.



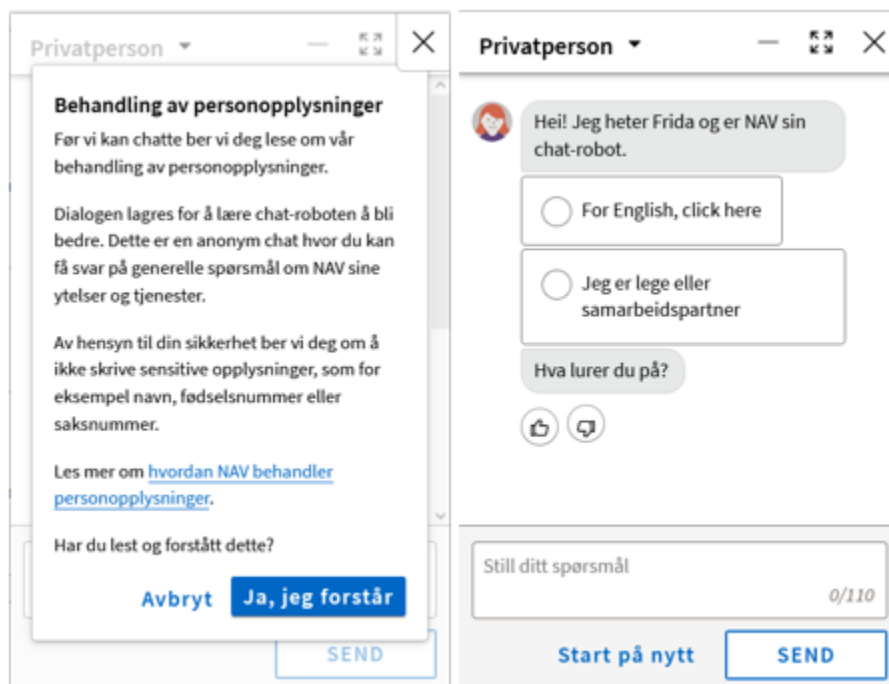
Figur 1 Illustrasjon av interaksjonene mellom chatbot og omgivelsene rundt. Sirkelen representerer oppgavens tema. Pilene viser veiene interaksjonen går. Illustrasjonen er utviklet sammen med medstudent Maria Midjo Rønning.

Som vi ser, er chatboten i sentrum av interaksjonene rundt den. Innbyggernes interaksjoner med både chatbot og veileder er toveis-kommunikasjon gjennom selve chaten. Vi ser også at det er kommunikasjon mellom veilederne og chatboten. Dette skjer i forbindelse med at innbyggernes chat blir overført fra Frida til en veileder. Veilederne bruker også Frida som et verktøy for å raskere slå opp informasjon de trenger å dobbeltsjekke. Veilederne kommuniserer også med ledelsen og chatbot-eksperter. Denne gruppen tar avgjørelser om chatboten basert på tilbakemeldinger som blant annet blir innhentet via veilederne.

Chatboten Frida er en del av førstelinje-kundestøtte hos NAV og ett av flere kontaktpunkt for innbyggere som søker kontakt med etaten. Hun er på jobb døgnet rundt og tilbyr derfor innbyggere hjelp også utenfor vanlig kontortid. Chatboten ble lansert i 2018 som et verktøy for dem som hadde spørsmål om foreldrepenger (Ringnes, 2018). Siden den gang har Frida fått stadig nye arbeidsområder og tar i dag unna generelle spørsmål for det meste man

kontakter NAV for (Ringnes, 2020). Den siste utvidelsen av arbeidsområder for Frida har skjedd som følge av Covid-19-pandemien, og et økt behov fra NAV sin side for å besvare henvendelser.

Under pandemien svarte Frida på det meste på henvendelser tilsvarende en arbeidsmengde for 220 menneskelige veiledere. Dette utgjorde på det meste, i snitt over 5400 henvendelser hver dag (T. Simonsen, 2021). Av de som snakker med chatboten blir omtrent 25% satt over til menneskelige veiledere (ibid.). Dette gjelder typisk spørsmål som ikke er av generell karakter, eller handler om oppfølging av egen sak.



Figur 2 Til venstre: informasjon om personopplysninger før man får tilgang til chatten på Nav.no. Til høyre: Velkomstmeldingen fra Frida etter at man har lest informasjonen om personopplysningene.

Interaksjon med chatboten fungerer ved at innbyggeren går inn på Nav.no og velger chat med oss. Tidligere fikk man da en kort velkomstmelding fra chatboten som ga brukeren en introduksjon til hvordan man skulle bruke systemet. Instruksen inkluderte oppfordring om å skrive korte spørsmål eller stikkord. I januar 2021 ble denne velkomstmeldingen byttet ut med en plakat om behandling av personopplysninger. Når man aksepterer denne, kommer man videre til chatten med Frida (se figur 2). Dersom chatboten ikke kan hjelpe blir man tilbudt å bli satt over til chat med en menneskelig veileder, så lenge henvendelsen skjer innenfor vanlig kontortid.

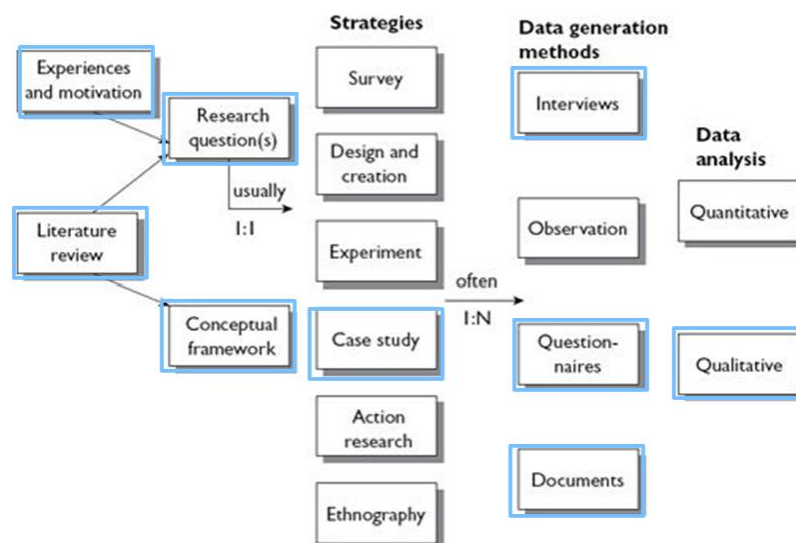
4. Metode

I dette kapitlet presenteres forskningsstrategien for oppgaven og forklart hvorfor denne er egnet for å finne svar på forskningsspørsmålet. Det blir videre forklart hvordan datainnsamlingen ble gjort ved hjelp av semistrukturerte intervjuer og en online spørreundersøkelse. Kapitlet avsluttes med en gjennomgang av dataanalysen som er benyttet.

4.1 Forskningsstrategi

Casestudier brukes for å undersøke et tilfelle av en ting eller fenomen man vil se nærmere på (Oates, 2006). Oppgaven er gjennomført som en deskriptiv casestudie, som medfører en rik og detaljert analyse av et fenomen og dets omgivelser (ibid.). I dette tilfellet undersøker oppgaven tillit til chatboter brukt av det offentlige i Norge. Fenomenet som undersøkes er sammenhengen mellom folks tillit til chatboten og deres brukeropplevelse. Gjennom forskningsstrategien som legges til grunn i dette kapitlet, vises det hvordan masteroppgaven er løst for å finne svar på forskningsspørsmålet. En viktig ramme for den valgte strategien var at forskningen skulle gjennomføres i løpet av et semester. Se figur 3 for en grafisk fremstilling av oppgavens forskningsstrategi.

Bruken av casestudie som forskningsstrategi ble valgt fordi den er godt egnet til å gi innsikt i sammenhenger og en forståelse av hvordan caset henger sammen med verden rundt (Oates, 2006) og er med på å avgrense forskningens omfang (Tjora, 2018). I dette tilfellet er caset også en typisk representant for fenomenet og et pragmatisk valg som følge av samarbeidsprosjekt mellom NAV og universitetene (Oates, 2006). Dessuten er casestudier godt egnet for å finne underliggende sammenhenger på hvordan noe skjer (Baxter & Jack, 2008), som forskningsspørsmålet forsøker å avdekke. En medvirkende årsak for valget av metoden er det pågående prosjektet Frida@work mellom NAV, NTNU, UiA og UiO. Prosjektet gir tilgang til caset på en måte som ikke ville vært gjennomførbart uten det. Dets overordnede mål er å undersøke betydningen av tillit i interaksjonen mellom menneske og chatbot (Arbeids- og velferdsetaten, 2021).



Figur 3 Oversikt over forskningsdesign. Modifisert figur fra Oates (2006)

Erfaring og motivasjonen som ledet til forskningsspørsmålet er dekket i innledningen av oppgaven. Dette ledet til en gjennomgang av mulig teori. Den relevante teorien ble plukket up og danner utgangspunktet for teorikapitlet.

4.2 Forskningsparadigme

Studien følger et fortolkende paradigme, det vil si at man er opptatt av å forstå den sosiale konteksten av et informasjonssystem (Oates, 2006), i dette tilfellet chatboter. For å vurdere fortolkende forskning må man vurdere forskningenes troverdighet, bekreftbarhet, pålitelighet, kredibilitet og overførbarhet (ibid.). Forskningens **troverdighet** er ivaretatt gjennom kapittel 3, 4 og 5 som viser hvordan forskningen er gjennomført, og at den er til å stole på. **Bekreftbarheten** kommer til syne gjennom presentasjonen av funn og diskusjonen av disse. Med tilgang til både intervjuguider og spørreundersøkelse kan forskningen som er gjennomført etterprøves for å teste dens **pålitelighet**. Forskningens **kredibilitet** er ivaretatt gjennom innsamling av data fra flere kanaler og bruk av flere metoder. Sammen gir dette overbevisende bilde av helheten i det observerte fenomenet. Casebeskrivelsen gir et grundig innblikk i caset og legger til rette for at man kan vurdere i hvilken grad konklusjonene som dras er **overførbare** til deres kontekst.

4.3 Datainnsamling

Datainnsamlingen i denne oppgaven er todelt. Det ble brukt intervju for å skaffe bakgrunnsinformasjon fra personer som har praktisk erfaring i feltet, og det ble benyttet en spørreundersøkelse for å samle brukeropplevelser fra brukere av chatboter. Innledningsvis i utformingen av oppgaven var det planlagt intervju med brukere av Frida, men det lot seg ikke gjøre å skaffe informanter til dette gjennom prosjektet. Spørreundersøkelsen ble brukt for å skaffe innsikt i innbyggernes forhold til chatboter generelt. Datainnsamlingen er dekket av NSD-søknad for Frida@work-prosjektet. Spørreundersøkelsen er anonym og lagrer ikke data om brukerne og er derfor unntatt søknaden.

Bakgrunns-teorien og det teoretiske rammeverket som er presentert i kapittel 2, er samlet inn og gått gjennom for å gi en dypere innsikt i chatboter. Den gjennomgatte teorien bidrar til å gi et bilde av hva slags forskning som er gjort innenfor bruk av virtuelle agenter, både for kommersielle og private aktører, men også av det offentlige. Søkene etter teori tok utgangspunkt i teori brukt i en masteroppgave som så på hva som ble vanskelig for brukerne når man snakket med en chatbot (L. Simonsen, 2019). Dette la grunnlag for videre søk ved hjelp av Google Scholar og Oria (universitetsbiblioteket), i tillegg til at kilder henvist til i artiklene fra de første søkene også ble funnet og fulgt opp.

I forbindelse med intervjuene og spørreundersøkelsen ble det samarbeidet om datainnsamling med en annen student som skriver masteroppgave innenfor samme overordnede tema som denne oppgaven. Konkret handlet samarbeidet om utvikling av intervjuguide (vedlegg A, B og C), gjennomføring, transkripsjon og diskusjoner rundt den tidlige kodingen av innsamlet data. Intervjuene ble utført i forbindelse med prosjektet mellom case-bedriften og NTNU. Informanter ble plukket ut fra dette prosjektet

4.3.1 Intervju

Hensikten ved å gjennomføre intervju var å samle detaljert bakgrunnsinformasjon om hvordan NAV jobber med virtuelle agenter. De åpne spørsmålene og den detaljerte

informasjonen vi var ute etter, gjorde at intervju ble ansett som godt egnet (Oates, 2006). Det var spesielt interessant å finne ut hvilke avgjørelser som ble tatt med tanke på brukeropplevelse og avveininger rundt tillit brukerne må ha til chatboten. Intervjuene ble gjennomført som semistrukturert over Microsoft Teams, hvor informantene hadde fått tilgang til intervjuguiden på forhånd. Intervjuguiden for hovedrunden med intervju ble utarbeidet etter det innledende møtet med representanter for case-bedriften. Se vedlegg A for spørsmål fra dette møtet. Spørsmålene fra møtet ble lagt til grunn for utvelgelse av informanter blant ansatte innenfor roller som chatbot-trenere, chatbot-koordinatører og lederstillinger med ansvarsområder som omhandler chatboten. Hele intervjuprosessen var tilknyttet Frida@work-prosjektet, både med tanke på utvalg og tid til gjennomføring. På grunn av tidsbegrensninger ble to andre studentgrupper også involvert i gjennomføringen av intervjuene. De varte en time, og hver studentgruppe hadde cirka 20 minutter til å stille spørsmål. Dette medførte at studentene fikk muligheten til å reflektere over andre perspektiver i den større konteksten rundt egen studie. I utgangspunktet ble det stilt tilnærmet de samme spørsmålene til alle informantene, men noen tilpasninger ble gjort i forhold til intervjuobjektets rolle (se intervjuguidene i vedlegg B, ansatt og C, leder). For eksempel fikk de som jobbet direkte med chatboten, spørsmål om hva som er gjort, hvilke valg og vurderinger som var gjort, mens leder-rollene fikk spørsmål om overordnet strategi. Intervjuene ble gjennomført med introduksjon, oppvarmingsspørsmål, hoveddel og avslutning (Oates, 2006; Tjora, 2018). Dette kommer ikke til syne i intervjuguidene, annet enn en kort introduksjon til tema for oppgaven.

Fordi intervjuobjektene brukte ulik tid på sine svar, ble det tatt høyde for at man ikke kom til å rekke gjennom alle spørsmålene i guiden i hvert intervju. Dette ble gjort slik at man fikk svar på alle spørsmålene fra guiden. Se tabell 1 for en oversikt over hvilke roller som fikk hvilke spørsmål.

Tabell 1 Oversikt over intervjuobjekt i case-organisasjonen

Informant-ID	Rolle	Spørsmålsoversikt
IS1	Mellomleder	A - G, i, ii, iii, iv
IS2	Leder	A - G, 4, 8, 9
IS3	Chatbot-trener	1, 3, 4, 6, 7, 8, 9, 10, 11, 12
IS4	Chatbot-trener	2, 3, 4, 6, 8, 9, 10, 12
IS5	Chatbot-koordinator	1, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12
IS6	Chatbot-trener	1, 3, 6, 7, 8, 9, 10, 11, 12
IS7	Chatbot-trener	2, 3, 4, 6, 8, 9, 10, 12
IS8	Chatbot-trener	3, 4, 6, 7, 8, 9, 10, 12
IS9	Chatbot-trener	3, 4, 6, 7, 8, 9, 10, 12
IS10	Mellomleder	1, 3, 4, 6, 7, 8, 9
IS11	Chatbot-koordinator	1, 2, 3, 4, 5, 6, 8, 9, 10, 11, 12, iii
IS12	Leder	ii, iii, v, vi

4.3.2 Spørreundersøkelse

Da det ikke lot seg gjøre å intervju brukere av chatboten direkte, ble en online spørreundersøkelse benyttet for å samle erfaringer fra brukere. Spørreundersøkelser er godt

egnet for å nå mange respondenter, og spørsmålene var utformet slik at man ikke trengte forkunnskaper om tema for å svare (Oates, 2006).

Undersøkelsen ble utviklet med utgangspunkt i gjeldende teori, og var en blanding av åpne og lukkede spørsmål. Vi valgte en struktur med et lukket spørsmål som ble fulgt av et åpent spørsmål som ba brukeren om å forklare svaret han eller hun hadde gitt i det lukkede spørsmålet. Formålet med dette var å spore brukerne inn på ulike tema og så samle kvalitative data gjennom de åpne tekstsvarene. Disse tiltakene var ment å sikre spørreundersøkelsens innhold og struktur for å sikre god forskning (Oates, 2006). Påliteligheten til svarene kan være utfordrende fordi undersøkelse distribueres over nettet og er anonym. Oates (2006) påpeker at en utfordring ved å gjennomføre dette over nettet er at respondentene kan miste fokus og velge å ikke fullføre hvis det blir for mange spørsmål som tar lang tid. Derfor ble det en konstant avveining mellom hvilke spørsmål som skulle tas med, da vi ikke ville ha en undersøkelse som tok veldig lang tid å gjennomføre.

Spørreundersøkelsen var mer generell enn hva dybdeintervju ville ha vært. Den dekket tema knyttet til brukeropplevelse, forventninger, kompetanse og tillit. Undersøkelsen ble publisert online i uke 12 med publiseringer via Facebook, Twitter og oppslagstavle på innsida (intern nettsiden for studenter og ansatte ved NTNU). Den var aktiv til og med uke 15 og fikk totalt 120 svar.

Utvalget i undersøkelsen avslører raskt at det ikke er representativt for Norges befolkning, da over 90% av de som svarte oppga at deres høyeste grad av utdanning var universitet/høgskole. Dette kan forklares ved at mange av de som har sett oppslagene på Facebook, Twitter og Innsida er tilknyttet utdanningsmiljøet rundt NTNU gjennom deres bekjentskap av oss som publiserte undersøkelsen.

4.4 Dataanalyse

Dataanalysen benyttet en Stegvis-deduktiv induktiv tilnærming (Tjora, 2018). Det vil si at man har en tilnærming der man jobber stegvis induktivt fra innsamlede data (empiri) mot teorien, og deduktivt fra teorien mot empirien for å finne forklaringer på det man finner. De innsamlede dataene blir sortert og testet mot gitte kriterier for å sikre validiteten i forskningen (Tjora, 2018), og man kan bruke teorien for å forklare det man oppdager eller lage et grunnlag for nye hypoteser. Koding og sortering i henhold til tema ble gjort i NVivo som er et verktøy for kvalitativ dataanalyse. Figur 4 viser et eksempel på tidlig koding, dette er hentet fra svarene på spørsmål om hvorfor brukeren avbrøt samtalen med en chatbot i det offentlige. Koden er valgt for å vise essensen i det informanten svarer, og lignende svar blir tilordnet den samme koden.



Figur 4 Tidlig koding av svar fra spørreundersøkelsen hentet fra NVivo.

Etter at alt av tekst fra både intervjuene og spørreundersøkelsen var ferdig kodet, ble disse sortert i ulike kategorier ut fra den enkelte kodens tema. Denne sorteringen la grunnlaget for valget av hovedtema og underkategorier som kapittel 5 er organisert etter (se tabell 2). Det er tette bånd mellom enkelte av hovedtemaene, og flere av datapunktene kunne vært plassert under mer enn en kategori. Dette er en konsekvens av at temaene for denne oppgaven henger tett sammen. Grupperingen er slik at de ulike gruppene ivaretar en indre konsistens, og skiller seg fra de andre gruppene i henhold til SDI-metoden (Tjora, 2018).

Der det i senere kapittel refereres til funn fra forskningen, gjelder henvisninger til informanter informasjon fra intervjuene, mens henvisninger til respondenter kommer fra spørreundersøkelsen.

Tabell 2 Oversikt over hovedtema med underkategorier og beskrivelser fra den gjennomførte forskningen.

Hovedtema	Underkategorier	Beskrivelse
Behov	<ul style="list-style-type: none"> • Kartlegging av behov • Brukere søker menneskelig kontakt • Chatbotens arbeidsoppgaver 	Virtuelle agenter har som mål å dekke behov for brukerne og for organisasjonen som benytter dem. Kategorien dekker disse perspektivene samt arbeid som gjøres i forbindelse med utvikling og bruk av agenten.
Bruker-opplevelser	<ul style="list-style-type: none"> • Ansatte • Positive opplevelser • Misforståelser mellom bruker og chatbot 	Omhandler opplevelser knyttet til utvikling (fra organisasjonens side) og bruk av virtuelle agenter som ikke bare dreier seg om man får riktig eller galt svar fra chatboten.
Forventninger	<ul style="list-style-type: none"> • Forventningene som ikke stemmer • Forventninger som stemmer med det teknologien kan levere • Brukernes forventninger innfris • Forventningsstyring 	Beskrivelser av ulike forventninger respondenter har til bruk av chatboter sammen med arbeid NAV gjør for å møte eller guide disse. Forventninger som stemmer med det teknologien og brukernes forventninger innfris kan sees på som realistiske forventninger, og de man helst ønsker at en bruker til Frida skal ha.
Kompetanse	<ul style="list-style-type: none"> • Demonstrerer kompetanse • Opplevd kompetanse • Optimalt kompetansenivå 	Svar på tanker og opplevelser intervjuobjekt og respondenter har om kompetansen til chatboter generelt og chatboten i caset spesielt.
Utfordringer	<ul style="list-style-type: none"> • Utenfor casebedriftens ansvarsområde • Frustrasjon blant brukerne • Høye forventninger • Forbedringspotensialet 	Ulike utfordringer ved bruk av chatboter som casebedriften må vurdere. Mange av utfordringene som diskuteres knyttes til brukeropplevelser som ender med ulik grad av frustrerte brukere.
Tillit	<ul style="list-style-type: none"> • Tillit til en chatbot • Tillit svekkes • Tiltak for å skape tillit 	Tar for seg organisasjonens syn på og arbeid med å styrke tillit tilknyttet deres chatbot. Ser også på ulike situasjoner som har ført til at respondenter har opplevd svekket tillit til en virtuell agent.

5. Empiriske Funn

I dette kapitlet presenteres funnene fra intervjuene med ansatte i caseorganisasjonen og funn fra den gjennomførte spørreundersøkelsen. Kapitlet er delt inn etter hovedtemaene som kom frem i dataanalysen. Hvert tema dekker i tillegg to-fire underkategorier markert i tabell 2. For hvert av temaene og underkategoriene vil det gå klart frem hvilke data som har opphav fra intervjuene og hva som kommer fra spørreundersøkelsen. Det er viktig å merke seg at svarene fra spørreundersøkelsen ikke kobles til Frida i seg selv, men er av mer generell karakter om folks holdninger og opplevelser med virtuelle agenter.

5.1 Behov

Dette temaet handler om brukernes behov for et system og dekker undertemaene kartlegging, etterlysninger fra brukerne og chatboters arbeidsoppgaver.

Kartleggingen av behov er gjerne et av de første skrittene når man skal anskaffe eller utvikle ny teknologi som en virtuell agent. Det betyr at man må identifisere mulige brukere og finne ut hvordan de vil dra nytte av teknologien som blir presentert. Fra intervjuene finner vi at opphavet til chatboten organisasjonen bruker var et sommerprosjekt utført av studenter i sommerjobb.

"Nei, i utgangspunktet da, så startet dette her opp som et prosjekt i Oslo. Vi har jo en sånn prosjektavdeling, [...] i Oslo, og det var på en måte sånn ... sånt sommerjobbprosjekt da, for en studentgruppe." (IS3, chatbot-trener).

Kunnskapen om mulige brukeres behov står sentralt i dette. I dette caset kommer mye av den kunnskapen fra tidligere interaksjon med brukerne og veilederne som tidligere har svart på henvendelsene den virtuelle agenten nå har tatt over.

"At den skal svare på et område den ikke ... Frida aldri har svart på før. Så starter vi alltid med å etablere en manuell chat hvor det sitter veiledere og svarer på de spørsmålene brukerne har, også bruker vi den databasen for å lære opp Frida fortløpende. Også har vi kanskje bare den manuelle chatten i tre måneder, også kobler vi på Frida i forkant når hun kan nok om området." (IS4, chatbot-trener).

Disse veilederne jobber på tvers av de ulike kontakt-kanalene organisasjonen har, og har dermed kunnskap om brukere fra ulike bakgrunner. Organisasjonen drar i tillegg nytte av en årlig oversikt over hvilke temaer som er mest aktuelle og en årlig brukerundersøkelse. Brukerundersøkelsen samler tilbakemeldinger på tvers av organisasjonens kontaktpunkt. Til sammen danner disse kildene brukerinvolveringen i måten systemet blir utviklet og kontinuerlig forbedret på.

En brukergruppe det kan være utfordrende å forstå behovene til, er de som åpner chatten, men så går ifra denne. Denne gruppen med brukere kommer vi tilbake til senere.

"På toppen av de 861 000 henvendelsene vi har, så har vi cirka 30% til. Det varierer da litt mellom 25-30 prosent til som starter Frida, men aldri skriver noe. Disse har vi trukket fra i de tallene som vi på en måte opererer med og går ut med sånn at vi ikke blåser opp tallene. 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. Så det er jo ett.. en utfordring som vi absolutt må jobbe med." (IS1, mellomleder).

Som vi ser av uttalelsene over, involverer casebedriften sine brukere indirekte i arbeidet med utvikling og videreutvikling av chatboten. Brukerne er direkte involvert i den forstand at de har kontakt med ansatte som videre representerer deres behov i den videre utviklingen.

"Og det er det jo sånn vi i NAV Kontaktsenter gjør noe, og så gjør NAV noe i helhet. Så det NAV gjør som helhet det er jo at chatbot er jo med på den årlige brukerundersøkelsen. Som på en måte tar for seg brukerservice i NAV, stort da, hvor det blir en av kanalene." (IS10, mellomleder)

Organisasjonen gjennomfører også en årlig brukerundersøkelse som kan gi svar på hvor godt de treffer brukermassens behov. Denne dekker alle kanalene for kontakt, som chat er en stadig større del av. En brukergruppe det er uvisst om får dekket sine behov, er de som tar kontakt med chatboten uten å skrive noe.

Spørreundersøkelsen viser at flere **brukere søker menneskelig kontakt** når de tar kontakt med en chatbot. Dette er brukere som kan være i situasjoner som søker ekstra empati, eller de vet de må snakke med en saksbehandler fordi de skal ha informasjon om egen saksbehandling. Oppfølging av enkeltsaker er arbeidsoppgaver chatboten i casebedriften ikke er egnet til å håndtere, samtidig som det er uvisst om brukerne her henviser til denne eller andre chatboter innen offentlig sektor. En av respondentene poengterer at denne typen tjenester er mer direkte og mer uformell, sammenlignet med epost-utvekslinger, mens en bruker valgte chat som kanal for å ha skriftlige bevis på samtalen i etterkant.

"Jeg tenker at det i særlige sårbare situasjoner som når man trenger helsehjelp eller kontakt med NAV er spesielt viktig å få snakket med et ekte menneske. Automatisering fungerer i min opplevelse veldig dårlig. Og det skaper frustrasjon" (spørreundersøkelse)

Sitatet over viser et sentralt synspunkt fra potensielle brukere ved at det sier noe om de brukergruppene virtuelle agenter ikke skal hjelpe. Samtidig kan man se kontrastene mot sitatet fra lederen under, som sier at en av chatbotens arbeidsoppgaver er å filtrere ut brukerne som ikke trenger å snakke med et menneske.

Enkelte brukere av casebedriftens chatbot ber om å bli satt over til menneskelige veiledere etter at de har fått svar fra chatboten, og noen av disse stiller også de samme spørsmålene for andre gang. Dette kan ha sammenheng med et brukerbehov om å få bekreftet at man har forstått informasjonen fra chatboten korrekt.

Virtuelle agenter og oppgavene disse løser er ikke statiske i den forstand at det er en teknologisk utvikling innenfor språkforståelse og mulighetene AI-en har til å forstå følelser, og kontekst. Ettersom teknologien bak systemet gjør nye arbeidsoppgaver mulige å gjennomføre, har organisasjonen ambisjoner om å utvide **chatbotens arbeidsoppgaver**.

"Vi har som mål om å gjøre Frida mer. Både mer intelligent, og mer i stand til å forstå setting, forstå følelser, forstå situasjonen og avhengig av hvordan den utviklingen går så vil Frida få flere oppgaver." (IS12, leder).

I intervju-rundene kom det også frem en del arbeidsoppgaver for den virtuelle agenten i casebedriften som ikke er beskrevet i teorien. Chatboten blir beskrevet som et filter som skal fange opp og besvare de henvendelse som ikke trenger hjelp fra en menneskelig veileder. Dette vil i hovedsak handle om henvendelser av generell karakter.

"Det skal være lett å komme til et menneske når behovet er der, men samtidig skal Frida være et godt filter til å stoppe [...] menneskene som ikke trenger å gå videre." (IS1, mellomleder).

Som nevnt innledningsvis i oppgaven er chatboten i caset oppgaveorientert. Den hjelper brukerne ved å svare på generelle spørsmål og å hjelpe de frem til riktig informasjon. Samtidig kommer det også frem at den fungerer som et filter som hjelper brukere som ikke trenger bistand fra et menneske for å løse sin henvendelse.

5.2 Brukeropplevelser

Her beskrives brukeropplevelser av ulik karakter, både fra ansatte i caseorganisasjonen og funn fra spørreundersøkelsen. Felles for brukeropplevelsene er at de representerer bruk av virtuelle agenter sett fra både organisasjonen perspektiv og brukere som ikke er tilknyttet organisasjonen utover det behovet de forsøker å dekke.

Intervjuene ga innsikt i hvordan casebedriften jobber med både utvikling og videreutviklingen av den virtuelle agenten. Den avdekket også momenter når det kommer til antropomorfisme og prosessene rundt både utvikling og videreutvikling av chatboten.

Ansatte opplever at det er deler av chatbotens design de ikke kan gjøre noe med, uten å spesifisere hva de eventuelt ville ha endret på.

"Og vi kan ikke gjøre noe med boksene, for vi har jo krav til universell utforming [...]. Både størrelse på tekst og knapper og sånne ting [...] er vi låst på. Så derfor ble det som det ble." (IS3, chatbot-trener).

Når chatboten skulle få nye arbeidsområder innenfor organisasjonen, ble det først slått på manuell chat, slik at man kunne få innsikt i hva brukerne ville ha hjelp med. Før det nye arbeidsområde ble åpnet for allmenne brukere, gjennomførte de ansatte brukertester for å avdekke eventuelle feil eller mangler.

"Men vi brukte de spørsmålsformuleringene, for å se hva det er mye ... hva. Hvilke spørsmål som går igjen og hvor mengden ligger." (IS8, chatbot-trener).

"Og så tester vi selv, selvsagt. Alle i teamet går inn og spør Frida om det og det området. Prøver så godt vi kan, å finne feil og mangler." (IS11, chatbot-koordinator).

Brukeropplevelsene fra de ansatte som jobber med chatboten i casen, viser at det finnes begrensninger for hva de kan gjøre med tanke på design av løsningen de bruker for chatbot. De gir også innsikt i utviklingsprosessen når chatboten skal få et nytt arbeidsområde. Som nevnt tidligere er det verd å merke seg at brukere utenfor organisasjonen ikke er direkte involvert i denne prosessen.

Mange av respondentene i spørreundersøkelsen har hatt **positive opplevelser** med bruk av chatboter. Av disse er det flere som melder at den var tidsbesparende og lett å komme i kontakt med. Mange melder at de visste de snakket med en chatbot fordi den enten fortalte det selv eller at det var tydelig merket.

"Fordi det er enkelt og raskt. Synes det er bedre å skrive enn å snakke i telefon. Sender man e-post tar det lengre tid." (spørreundersøkelse).

Andre positive opplevelser inkluderer det å bli positivt overrasket over chatbotens oppførsel, at den klarer å hjelpe eller fordi man foretrekker chat som kanal fremfor alternativ som telefon.

"Chatboten jeg bruker som eksempel her henvendte seg til ungdom, og hadde en vennlig og humoristisk sjargong, og jeg ble overrasket over at man kunne bli såpass "sjarmert" av en chatbot." (spørreundersøkelse).

"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" (spørreundersøkelse).

"Fordi det er enkelt og raskt. Synes det er bedre å skrive enn å snakke i telefon. Sender man e-post tar det lengre tid." (spørreundersøkelse).

Et overraskende funn var at chatboter også kan senke terskelen for å stille dumme spørsmål eller spørsmål som av ulike grunner oppleves som ubehagelige.

"Enkelt å spørre om ting som kan føles litt dumme eller ubehagelig." (spørreundersøkelse).

Fra caseorganisasjonen sin side virker det som de fleste som får hjelp av chatboten er fornøyde med hjelpen de har fått, samtidig som det kan være umulig å vite hva brukeren gjorde i ettertid. Det å legge på i sitatet under er å avslutte samtalen med chatboten ved å krysse seg ut av chatvinduet.

"Det virker jo som de ... flertallet der de får riktig svar så virker det jo som de slår seg til ro med de svarene. At det er litt vanskelig å vite hva som skjer etterpå for da er det ... sant når en avslutter så er det vanskelig å vite: fulgte de lenken som de fikk [...] Bare la de på uten å gi noe indikator også tok opp telefonen og ringte? Det er vanskelig å vite, men det er jo mange som sier takk og legger på når de har fått riktig svar." (IS4, chatbot-trener)

Mange av de positive brukeropplevelsene beskrevet her, omfatter tidsbesparing, effektivitet og komfort. Tidsbesparelser og effektivitet er sentrale salgspunkt for bruk av virtuelle agenter. At komfort ble trukket frem, i den sammenheng at chatbot var mer behagelig enn telefon eller gir en mer behagelig arena for å stille "dumme eller ubehagelige" spørsmål, var mer overraskende. At de fleste får vellykket hjelp fra casebedriftens chatbot var ventet, og det er bra å få bekreftet det.

Flere av svarene fra spørreundersøkelsen kan tyde på at det har oppstått **misforståelser mellom bruker og chatbot**. Andre respondenter opplever frustrasjon fordi de ikke har fått svar på sitt konkrete spørsmål. Disse brukerne blir sendt til informasjon de allerede hadde

funnet selv, og sammenligner chatboten med en søkemotor. Noen ganger oppfatter ikke chatboten hva brukeren egentlig er ute etter og gir svar som brukere opplever som feil, upassende eller støtende.

*"Spurte om utenlands reise, chatbot insisterte at jeg ønsket innenlands."
(spørreundersøkelse).*

Litt overraskende benyttet enkelte brukere seg av chatboten som en snarvei for å komme i kontakt med menneskelige veiledere, mens andre så på den som et hinder man måtte forsere for å få hjelp til problemet sitt. I tilfeller hvor man skal ha hjelp utover generell informasjon vil sistnevnte være riktig, men reflekterer muligens holdninger til systemet som ikke er helt heldige.

"Kommer an på kompleksiteten av problemet jeg trenger hjelp til. 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." (spørreundersøkelse).

Enkelte opplever også chatboten som den eneste muligheten for å ta kontakt, og er kritiske dersom dette har erstattet andre kanaler.

"Fordi de hadde erstattet telefon med chatbot, uten å i mine øyne trene den godt nok til at det var et alternativ, da blir man irritert i interaksjonen og føler det er vanskelig å få hjelp." (spørreundersøkelse).

Både misforståelsene og de negative opplevelsene respondentene her forteller om kan ha sammenheng med at man ikke forstår kapabilitetene til chatboten man snakker med godt nok. Samtidig som det er lett å bortforklare mange av hendelsene, er opplevelsene virkelig for brukerne og skaper frustrasjon.

Vi ser også at enkelte brukere vet at deres henvendelse ikke kan løses av chatboten, og har for så vidt riktig i at den blir et hinder på veien til menneskelig kontakt. Samtidig kan man spørre hvor lang køen for å komme i kontakt med et menneske på chat hadde vært dersom chatboten ikke var et obligatorisk stopp på veien. Det er uvisst hvilken chatbot respondenten i spørreundersøkelsen mener har erstattet mennesker. Caseorganisasjonens chatbot fungerer som en brikke i deres kanalstrategi, sammen med blant annet telefon og chat med veileder.

5.3 Forventninger

En av puslespillbitene for å forklare brukeropplevelsene beskrevet i delkapittel 5.2 handler om forventninger. Dette er brukernes antakelser om den virtuelle agentens kapabiliteter og hva brukeren kan få ut av møtet. Samtidig kan forventningene påvirkes utenfra, både av organisasjonen bak chatboten, brukerens omgivelser og samfunnet for øvrig.

En del av de uheldige brukeropplevelsene kan forklares ved at brukeren har hatt forventninger til systemet som ikke matcher kapabilitetene i det aktuelle systemet. Disse **forventningene som ikke stemmer** kan henge sammen med at teknologien er ny og brukerne ikke har funnet frem til de beste måtene å benytte seg av den. I dagligdags samtale bruker mennesker ofte å starte med bakgrunnsinformasjon før man kommer til poenget.

"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. De har en historie å fortelle, ikke sant, det er en bakgrunn, før de kommer til poenget. Og en chatbot klarer jo ikke å forholde seg til den ... en liten historie på samme måte som et menneske gjør. Og der kommer vi tilbake til det med korte spørsmål, ikke sant [...]." (IS3, chatbot-trener)

En mulig forklaring for brukere som tar kontakt med casebedriften og som tenker at de trenger å fylle på med bakgrunnsinformasjon, kan være valget av å kalle knappen man trykker på for å starte chaten "chat med oss" fremfor å kalle knappen noe som gir assosiasjoner til chatboten. Samtidig vil dette være en nødvendig avveining ettersom man må gå den samme veien om man vil chatte med en veileder.

"Og grunnen til at vi har valgt «chat med oss» er jo for at det skal virke litt mer sånn, lavere terskel for at du skal snakke med et menneske bak dette her. Selv om du da kan også bli skuffet, fordi at du ikke møter et menneske som du kanskje trodde når det står «chat med oss», for det virker jo mye hyggeligere og mye ... ja, det høres ikke ut som du møter en chatbot rett og slett. Så vi har nok et potensial der for å jobbe mer med, hvordan på en måte design og layout og hvilken informasjon du får i forkant, at du liksom møter da de forventningene." (IS1, mellomleder).

Blant tilbakemeldingene fra brukerne er det mange som forventer at chatboten skal gjøre det den blir bedt om, noe som kan tyde på at de har forventninger ut over det teknologien kan levere. Et viktig poeng en bruker trekker frem er at dersom de erstatter mennesker, så må de kunne svare like godt. En annen bruker ser ut til å kunne ha veldig høye forventninger til å bli guidet frem til riktig svar.

"Først gjengi "sin tolkning" av hva jeg har spurt om, og gi flere alternative svar hvis det er flertydig. Den burde svare på en slik måte at man penses aktivt inn mot hensiktsmessige oppfølgingsspørsmål, ikke bare komme raskest mulig til en avslutning." (spørreundersøkelse).

En bruker trekker spesielt frem tjenester i det offentlige, og forventer mindre feil enn hva en saksbehandler kunne gjort. Akkurat hva brukeren legger i personlige feil i en offentlig tjeneste er uvisst, men det kan tyde på høye forventninger om kvaliteten på avgjørelser. For virtuelle agenter som bare formidler informasjon og hjelper brukerne med å finne frem, virker det mindre relevant, samtidig viser det at brukeren stiller høye krav til det offentlige.

"Tjenester i det offentlige ser jeg på som tjenester som har mindre rom for personlige feil enn det feks en saksbehandler ville hatt" (spørreundersøkelse).

Fra intervjuene får vi inntrykk av at enkelte innbyggere som bruker Frida, ikke er vant til å snakke med chatboter, og har samme tilnærming til samtalen som de ville hatt med et menneske. Det betyr at man fyller på med bakgrunnsinformasjon for å gi et bilde av konteksten rundt spørsmålet. Et annet sentralt funn her, fra spørreundersøkelsen, er at noen respondenter gir uttrykk for høye forventninger til chatbotens kapabiliteter, både innenfor måten den kommuniserer på og også at enkelte tar avgjørelser.

Blant respondentene i spørreundersøkelsen var det også mange som oppgir at de har **forventninger som stemmer med det teknologien kan levere**. Flere påpeker at de får raskt svar, mens de fleste oppgir at de har lave eller ingen forventninger i møte med teknologien. En respondent påpeker at måten man bruker chatboter på utvikler seg over tid ettersom vedkommende ble mer kjent med teknologien.

"Til slutt lærer du hvordan du skal skrive til den for å forstå, men ikke alltid jeg får ønsket/nødvendig svar." (spørreundersøkelse).

Enkelte trekker også frem sikkerhet når det er snakk om forventninger til chatboter, samtidig som det er flere som er usikre på om de vil få riktig svar på sine spørsmål. God tilgjengelighet blir også trukket frem blant forventninger til denne typen system, noe som stemmer godt ved at chatboten alltid er tilgjengelig.

Uten å være sikker på at dette gjelder casebedriften var det også enkelte som påpekte at dersom chatboten ikke var i stand til å hjelpe, ville den sette brukeren i kontakt med et menneske.

"Om boten ikke kan svare, forventer jeg at den setter meg videre til et menneske, etter at jeg har sagt "ja, det vil jeg"". (spørreundersøkelse).

Brukere med lave eller realistiske forventninger når det gjelder hvordan man kommuniserer effektivt med en chatbot, og at man får hjelp av mennesker dersom man trenger det, viser at det finnes potensielle brukere som har et godt utgangspunkt for å ha god nytte av interaksjonene med chatboter.

Gjennom det kontinuerlige arbeidet med å forbedre chatboten, har casebedriften økt mulighetene for at **brukernes forventninger innfris**. Samtidig gir informantene uttrykk for at det er noen forventninger de ønsker at brukerne var enige om. Flere av svarene fra spørreundersøkelsen handler om å få svar på generelle spørsmål. Hos casebedriften har de i løpet av våren fått på plass et system hvor chatboten vet hvilket område av nettstedet man tok kontakt ifra. Denne konteksten er med på å gjøre det enklere å svare riktig på spørsmål som kan ha flere svar, selv om alle svarene gjelder generell informasjon.

"At vi både jobber med det kontekstbaserte, at chatboten skal forstå hvor du er hen på NAV.no og sånn sett svare riktig i forhold til hvem du er som bruker." (IS1, mellomleder).

Begrepet generell informasjon kan også være en utfordring. Der en av informanten henviser til brukere som har vært ute etter informasjon om egen saksbehandling, noe chatboten ikke vil kunne svare på fordi det ikke er generell informasjon. Generell informasjon i denne forstanden vil være informasjon som er lik, uavhengig av hvem det gjelder.

"Nei, det hadde jo vært fint hvis vi kunne fått de til å forstå det her med at det er generelle spørsmål Frida kan besvare. Men vi vet jo at det er en utfordring for folk å forstå hva som ligger i begrepet generelle spørsmål." (IS6, chatbot-koordinator).

Blant informantene er det flere som forteller at bedriften jobber med designet rundt chatboten og språket den benytter. Dette henger sammen med et ønske om at brukerne

oppfatter at de snakker med en chatbot. Samtidig er det viktig for organisasjonen at språket den bruker ikke blir for lite folkelig.

"Vi har jobbet en del med at Frida skal [...] både oppleves som menneskelig, men også skal være en robot. Så den er vanskelig, den der. [...] Men vi skal ha en tone, språk og stil som er folkelig, også i chatbot. Samtidig som vi må være tydelig på at det er en chatbot og ikke et menneske som hjelper deg." (IS1, mellomleder).

Her får vi et innblikk i hvor komplekst arbeidet med chatbot er. Begrep som generell informasjon tolkes ulikt av brukere. Samtidig jobber casebedriften med å møte forventninger om at brukerne skal bli møtt på en folkelig måte når de tar kontakt.

For å passe på at brukerens forventninger til systemet stemmer overens med mulighetene systemet faktisk gir, er det viktig at det jobbes med **forventningsstyring**. En av de mest synlige plassene dette skjer er i velkomstbeskjeden en bruker får når man åpner en chat. Hos case-organisasjonen er denne endret flere ganger i forsøk på å tydeliggjøre at det er en chatbot samt å gi noen indikasjoner om hvordan man får den beste brukeropplevelsen ut av interaksjonene med den.

IS5 "For at ikke det skal bli så ... det er noe med å si nok, men samtidig ikke så mye at når du åpner den chatboba så får du en hel sånn vegg med tekst. Også blir det til at, man scroller jo bare ned og klikker sånn ja, ok, greit ... eller hvis man må ... Det blir til at man ikke leser det, man må liksom ha frem essensen, også må det være kort og presist." (IS5, chatbot-koordinator).

Samtidig blir brukerne mer vant til virtuelle agenter, ettersom flere organisasjoner innenfor både i det offentlige og det private tar slike i bruk. Dette er med på å gjøre teknologien allment kjent. Løsningene de ulike bedriftene går for vil påvirke forventningene til lignende løsninger hos andre bedrifter.

"Det ene vi ser er at vi i NAV, vi lever jo ikke på en isolert øy. Det betyr at de forventningene som skapes når [...] andre har gode løsninger i brukermøte, er det også med å forme forventningene mot NAV. Og brukerne lærer seg nå veldig fort teknologi og dette med chatbot." (IS12, leder).

I januar 2021 førte regler om informasjon om personvern til at man måtte endre velkomstmeldingen, og man endte opp med et vindu med tekst hvor man må trykke "ok" for å komme videre til chaten.

"Med den løsningen vi har nå så er det jo faktisk ikke så mye styring, nå er det litt for å se, om de endrer måten de spør på [...] Nå baseres det jo vil jeg tenke ... egentlig på at det at det er en chatrobot i seg selv skal styre forventningene litt. Vi har jo igjen det med at hvis de får dette som vi kaller et nonsvar da. At Frida ikke klarer å få et svar da sier vi mer om at det er best med korte spørsmål og si gjerne stønader. Prøv med stikkord." (IS4, chatbot-trener).

Brukernes forventninger til systemet er også dynamiske, og endres av at man har gode eller dårlige opplevelser. I den forbindelse blir tilbakemeldingene chatboten gir når den ikke forstår en bruker også viktige. Sett sammen med hvor høye eller lave forventninger brukeren hadde innledningsvis vil dette kunne påvirke tilliten til systemet.

"Det er viktig å ikke legge forventningene for høyt, fordi vi ser jo ofte at det ikke går så bra, at de bommer. Så det er viktig å ha en sånn realistisk forventning til hva chatboten kan hjelpe med, og en god hjelp er jo nettopp at brukeren skriver kort og konkret. Og bakgrunnen for den omformuleringen, den som vi kaller for «unknown» da, hvor ikke Frida finner et svar, kan du skrive det på en annen måte, så har vi jo sett at ... det er en effekt i å stille det spørsmålet da, at mange da tenker seg om, også skriver de kanskje mer presist og mer stikkordsform, som gjør det lettere for Frida å finne rett svar hvis vi har et rett svar." (IS3, chatbot-trener).

Chatbotens utseende er også med på å forme brukernes forventninger. Et mer menneskelig utseende gjør at brukernes forventninger blir høyere. Om man kaller den virtuelle agenten for chatbot eller chatrobot kan også skape ulike assosiasjoner blant brukerne. Utsende og væremåte er tema som går igjen blant informantene i casebedriften.

"Og gitt da på en måte Frida det menneskelige trekket, både med utseende og navn. Som da veldig mange mener er veldig uheldig da, når det kommer til chatboter. Men vi har jo ikke hatt noe innflytelse på det, på den utformingen. Vi har jo holdt på lenge bare for å få lov til å kalle den en chatrobot i stedet for en chatbot." (IS3, chatbot-trener).

Blant respondentene i spørreundersøkelsen er det flere som gir uttrykk for at de forsto at de snakket med en virtuell agent enten via agentens navn eller ved at den selv identifiserte seg som chatbot. Flere oppgir ulike forventninger avhengig av om man har et innlogget system eller ikke. En av respondentene viser også at man kan ha ulike forventninger basert på om man er logget inn eller ikke. En innlogging kan i utgangspunktet gjøre det mulig for en virtuell agent å hente informasjon om søknadsprosesser og kundebehandling for den aktuelle kunden, uten at det er snakk om saksbehandling.

"Forventer at den kan svare på generelt grunnlag hvis jeg treffer på den utenfor innlogging." (spørreundersøkelse).

Arbeidet som gjøres med å styre brukernes forventninger er viktig for å gi brukere et så realistisk syn som mulig på hva chatboten kan hjelpe med og ikke. Foruten arbeidet eier av en chatbot gjør, påvirkes forventningene også av arbeid andre organisasjoner gjør med sine chatboter. Chatbotens utseende, navn og velkomstmeldingen den presenterer seg med er alle med på å forme brukerens forventninger i møte med chatboten. Jo høyere disse forventningene er, jo vanskeligere blir det for chatboten å leve opp til dem. Dersom forventningene er lave, kan brukeren bli positivt overrasket.

5.4 Kompetanse

Kompetansen chatboten kan levere, eller gir inntrykk av at den kan levere, spiller inn på både forventningene brukerne skaper seg om systemet og ikke minst om de godtar svarene de får. I dette delkapitlet ser vi på hvordan den faglig kompetanse kommer til syne for brukerne, hvilke opplevelser respondentene på spørreundersøkelsen kan fortelle om og hva intervjuobjektene tenker er det riktige nivået for chatboten å være på.

En annen del av brukeropplevelsen er den faglige kompetansen den virtuelle agenten utviser i møte med brukere. Det er naturligvis viktig at chatboten ikke leverer informasjon som er feil, men den faglige kompetansen strekker seg lengre enn det. Det handler om å forstå hva brukeren er ute etter og gi gode svar. Den virtuelle agenten fra casebedriften viser sin kompetanse som følge av arbeidet chatbot-trenerne og andre gjør i bakgrunnen.

En måte å **demonstrere kompetansen** for brukerne er å være ærlige på chatbotens begrensninger og varsle brukeren om tema den ikke kan hjelpe med.

"Men vi er i hvert fall veldig tydelige på de områdene hvor vi ikke har noen svar i Frida enda, at det her ikke er noe som Frida kan svare på. [...] Og der får jo brukerne beskjed om med en gang, at her er Frida fortsatt under opplæring, og de kan finne informasjon på NAV.no eller chatte videre med en veileder [...] Sånn at hvis det er noe som Frida ikke er opplært på enda i hvert fall, så skal det komme veldig tydelig frem. Det er ikke noe som vi prøver å legge skjul på det, at det er noen områder som ikke er utviklet enda." (IS6, chatbot-trener).

Et annen måte chatboten viser sin kompetanse på, går på språkforståelse og ikke minst hvordan den svarer. For casebedriften er det et mål at chatboten skal opptre så menneskelig som mulig. Det vil si at man går lengre enn å kopiere informasjonen man finner på nettsiden samtidig som man formidler både fagområdets seriøsitet sammen med en menneske-lignende varme.

"NAV sin rolle da, at vi er på en måte en seriøst, offentlig aktør for å si det forsiktig, også samtidig så skal vi være en viss form for folkelig oppe i det hele. Så det har vi prøvd å jobbe mye med, rundt også Frida og også hvordan vi kommuniserer skriftlig på chat sånn generelt sett når de kommer videre. Så dette med tone, språk og stil [...] så får jo også chatbot-trenerne opplæring i tone, språk og stil, samme som chatveilederne våre gjør. Nettopp for at de skal ... både skrive dialogen på en ... ha en, ja en chat som på en måte ikke bare blir oppramsing av NAV.no, men som formuleres på en måte sånn at det oppleves som en chat, da. Så det har vært viktig for oss. Men jeg tror nok vi har fortsatt et potensial der, for å jobbe videre med disse tingene." (IS1, mellomleder).

Casebedriften har valgt en linje hvor chatboten ikke responderer på tullemeldinger, hvor man tidligere kunne få den til å vitse. Dette valget henger sammen med organisasjonens posisjon som en del av det offentlige, hvor det er viktig at man ivaretar profesjonaliteten i alle ledd av interaksjoner med kundegrupper

"I starten så hadde vi jo litt sånn standard chatbot ting, som fortell en vits. Svar på, snakker på forskjellige språk. Sånne ting, som standard med pakken når vi kjøpte den. Og det meste av det har vi fjernet, så nå er det kun fokus på fagområdene til NAV og veldig lite rundt-om-kring snakk. Om det er rett, det vet jeg ikke, men det er det vi har gått for. Men, sånn smalltalk, det er jo veldig begrenset hva en chatbot får til, enda. Så det er kanskje best å ikke ha det før det er perfeksjonert." (IS11, chatbot-koordinator).

Vi ser at casebedriften har valgt et nivå når det kommer til chatbotens kompetanse hvor man vil oppfattes som så folkelig som mulig i språket. Samtidig som de har kuttet småprat og fokuserer på faglige innhold, er chatboten tydelig dersom den får spørsmål den ikke kan svare på.

Blant svarene i spørreundersøkelsen er det flere som drar frem språkforståelsen som en faktor i **opplevd kompetanse** hos chatboter. For mange av respondentene handler dette

om den virtuelle agentens evne til å forstå norsk som inneholder skrivefeil eller fremmedspråk som for eksempel engelsk.

"Jeg er ikke norsk native taler og fant dem ikke i stand til å håndtere forespørsler på engelsk." (spørreundersøkelse).

En respondent innrømmer å ha lite erfaring med chatboter, men uttrykker at interaksjoner med chatboter oppleves som et spill, der man må treffe det riktige nøkkelordet for å få det svaret man er ute etter. Vedkommende drar, som flere andre, også en sammenligning til det å google etter svar.

"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"." (spørreundersøkelse).

En annen språklig utfordring for det offentlige i Norge kan følge av at nynorsk er et offentlig skriftspråk på lik linje med bokmål. En av respondentene opplever at chatboter ikke støtter dette språket, og reiser samtidig spørsmål om hvordan situasjonen er for brukere med lese- skrive vansker.

"Det har gått ein trend i å bruke chat-bot, når det framleis er veldig ny og unøyaktig teknologi. Eg stiller meg spørrande til språkleg mangfald, som t.d. at dei aldri forstår nynorsk. Men og korleis dei forstår dårleg rettskriving, som følge av dysleksi o.l." (spørreundersøkelse).

Språkforståelsen er svært viktig når det kommer til å gi brukerne en god opplevelse av chatbotens kompetanse. Samtidig ser vi at respondentene fra spørreundersøkelsen melder om utfordringer på dette punktet, for de som sliter med språkvansker eller ikke har norsk som morsmål. Inkludering av engelsk eller andre store språkgrupper kunne ha hjulpet på dette punktet. Samtidig er det overraskende at nynorsken tilsynelatende blir ignorert. Dersom dette er faktum for chatboter i tjeneste hos offentlig sektor, kan det være grunn for debatt om nynorskens posisjon. Chatboten i caset forstår nynorsk, men velger å svare på bokmål.

I intervjuene var det også et tema hvor det **optimale kompetansenivået** for chatboten til organisasjonen skal være. I den forbindelse ble både dagens kompetansenivå og hva man jobber mot diskutert. Informantene var enige i at chatboten per i dag scorer høyere på kompetanse enn på varme i svarene den gir. Samtidig er det en uro for misforståelser i interaksjonen som gjør at svaret chatboten gir blir upassende, og det er viktig for organisasjonen å finne en balanse mellom det faktabaserte og menneskelige aspektet.

"Så, vi valgte ikke å gjøre det da [om valg av menneskelig avatar fremfor robotaktig] og vi kunne sikkert ha gjort det, men 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 og 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." (IS1, mellomleder).

Det endelige målet for chatboten til casebedriften er at den skal være et fullverdig tilbud for brukerne. Den vil kunne oppnå dette når brukeren er fornøyd med svaret man har fått og

ikke går til andre kanaler for bekreftelse. Hvordan chatboten kan levere informasjonen på en så overbevisende måte at brukeren får denne bekreftelsen kan være det som avgjør om chatboten er "ferdig" utviklet eller ikke. Dette er også avhengig av at svaret chatboten gir er riktig. Hvis en bruker opplever å få feil svar og slår seg til ro med det, vil det ha konsekvenser for omdømmet til hele organisasjonen på sikt. Språkforståelse henger tett sammen med dette, og det er en forventning om at chatboten i fremtiden vil kunne ta seg av henvendelser på engelsk.

"Jeg tror den største utfordringen det er jo tillit [...]. Har jeg fått riktig svar. Og dette med at chatboten da kan gi riktig svar, men også kan bekrefte til deg, gi tryggheten på at ja, du har forstått dette riktig. Når vi er der, da tenker jeg at vi har et fullverdig tilbud på chatbot. Men foreløpig så er det nok det der med at, ja, jeg fikk et svar her, men kan ... altså stoler jeg på det, er jeg trygg nok på det, må jeg sjekke det ut en ny plass for å være helt sikker på det? Eller må jeg rett og slett ringe inn bare for å få en bekreftelse på at ja, du har forstått det riktig. Det tror jeg er en utfordring vi må jobbe med fremover." (IS10, mellomleder).

Chatbotens optimale kompetansenivå handler om å oppnå gode nok språklige ferdigheter, både med tanke på språkforståelse og svarene chatboten gir brukerne. Det er om å gjøre å svare på en måte som gir brukeren en trygghet på at det man får opplyst er riktig i den konteksten brukeren spør om.

5.5 Utfordringer

I dette delkapitlet tar vi for oss en del potensielle utfordringer casebedriften står ovenfor. En del av dette går utenfor rammene rundt chatboten, mens andre skyldes misforståelser mellom brukeren og chatboten, eller forventninger om mer enn det systemet kan levere. Enkelte problemer avslører forbedringspotensialet og henter om hvordan problemer kan løses eller unngås i fremtiden.

En del av brukerne hos casebedriften starter chat med chatboten uten å gjøre noe mer enn det. Denne brukergruppen blir ikke prioritert i særlig grad hos casebedriften. Dette begrunnes med at brukergruppen kan ha åpnet chatboten av nysgjerrighet og ikke har et behov for hjelpen den kan gi, eller at den blir startet ved at brukerne bare tilfeldigvis kommer borti knappen for å starte chat. Uavhengig av grunnen er disse potensielle brukerne vanskelig å komme i kontakt med og dermed **utenfor casebedriftens ansvarsområde**.

"Ja, vi vet jo at noen starter chatboten med en feil også, altså den ligger veldig lett tilgjengelig, og i hvert fall før vi fikk den plakaten med å bekrefte samtykke på personopplysninger så startet jo bare dialogen uansett. Så det at vi hadde jo en viss sann prosent som vi trakk ifra alt av resultater og alt, vi så at her har de bare fått velkomstmeldingen til Frida og så har det ikke skjedd noe mer. Da anså vi det egentlig som en feil, ikke at de bevisst hadde byttet kanal engang. [...] Så den typen henvendelser har vi jo ikke prioritert å gjøre så mye med, annet enn at vi har prøvd å forbedre brukerservice så vi ikke skal få den typen henvendelser." (IS10, mellomleder).

Ettersom det er flere aktører innenfor det offentlige som benytter chatboter, kan det stilles spørsmål om hvem som har det overordnede ansvaret for brukeropplæring. Blant eldre generasjoner, innvandrere og andre som ikke nødvendigvis er så teknologivante, kan behovet for opplæring være større enn den informasjonen velkomstmeldingen gir. For casebedriften er det også uaktuelt å reklamere for chatboten i for eksempel tv.

"Personlig mener jeg at det er jo noe som digitaliseringsdirektoratet egentlig burde ha laget. Sånn generelt. For å informere brukere." (IS11, chatbot-koordinator).

Blant brukerne av chatboten finnes det også de som tydelig viser at de ikke mestrer samtalen med den virtuelle agenten. Casebedriften ser dette gjennom chatlogger over brukernes interaksjoner med chatboten. Enkelte brukere blir fanget i en loop hvor chatboten ender opp med å gi samme svar flere ganger. En informant trekker frem brukernes alder som en mulig forklaring på dette fenomenet, at brukerne ikke er vant til teknologien. Å havne i en loop er det også flere av respondentene i spørreundersøkelsen som melder, uten at man kan si at dette gjelder casebedriftens chatbot. Respondentene det gjelder er heller ikke blant eldre generasjoner.

"Hvis du ser på chatloger, så er det jo av og til det kommer folk som, helt klart ikke vet hvem de snakker med. Og du kan sitte å se på noen ganger, når de kommer helt i en loop og det gjentar seg og de skjønner ikke. [...] 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." (IS11, chatbot-koordinator).

Enkelte brukere kommer også til systemet med negative holdninger som utgangspunkt. Disse kan være skeptiske til teknologien fordi de har hatt dårlige erfaringer ved tidligere møter med den digitale verdenen. For dem er andre kanaler som telefon eller personlig oppmøte mer relevante.

"[...] Mange er takknemlige for svarene de får og de takker og alt sånt her [...], men at vi skal tenke at ... ja, vi må prøve å få hjulpet majoriteten. 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 hvert fall." (IS5, chatbot-koordinator).

Fra spørreundersøkelsen blir også det totale tjenestetilbudet innenfor offentlig sektor tatt opp, og som respondenten skriver er det viktig at de offentlige tjenestene har mer enn en måte for befolkningen å komme i kontakt på.

"Jeg er også usikker på hvor godt tilbudet blir til folk som av diverse grunner ikke kan bruke disse tjenestene. Ingen offentlige tjenester burde hvile på å kun gi én type kontakt-tilbud til brukerne sine (hvis man f.eks. ikke kan bruke chatbot, må det finnes et annet alternativ eller flere)." (spørreundersøkelse).

Innføring av ny teknologi i det offentlige kan være utfordrende med tanke på hvem som har ansvar for opplæring av potensielle brukere. Casebedriften anser ikke de brukere som starter chatboten uten å si noe som et stort problem, og de bruker ikke store ressurser på

denne gruppen. Andre brukere kan ha negative opplevelser med chatboten, som for eksempel ender opp med å få samme svar flere ganger, dette kan tyde på at vedkommende ikke forstår hvordan man får mest mulig ut av samtalen med chatboten. Det er usikkert om chatbot og chat er det riktige valget for de brukerne dette gjelder.

En stor andel respondenter har hatt interaksjoner med chatboter som har ledet til **frustrasjon blant brukerne**. Veldig mange tilbakemeldinger handler om at svarene chatbotene gir er dårlige, at chatboter fungerer som en dårlig søke-funksjon (sammenlignet med google), eller at de uansett gir svar som allerede er på nettsiden og som brukeren da allerede har sett.

"Disse chatbotene er en tungvinn og tidkrevende måte å få tak i info som allerede står på websidene. For spørsmål boten har svar har jeg med 100% sikkerhet funnet svaret selv tidligere. Jeg tar kontakt direkte når svare IKKE ligger på nett. (spørreundersøkelse).

For enkelte handler interaksjonene med chatbotene om å komme seg videre så fort som mulig, og disse ser på chatboten mer som et irritasjonsmoment før man kan få hjelp fra mennesker.

"De har plassert enda en meningsløs sperre i veien for faktisk kundeservice." (spørreundersøkelse).

Vi ser her at det er flere som både er frustrerte over at de må møte chatboter og at de ikke får noe ut av interaksjonen. Denne frustrasjonen skyldes mer enn at man blir misforstått eller får feil svar. Det kan virke som at frustrasjonen skyldes at man opplever teknologien som et hinder, fremfor et hjelpemiddel.

For mange henger nok denne problematikken sammen med **høye forventninger** om hva en virtuell agent kan levere. Det går igjen at man vil ha svar innenfor egen sak, mens noen forklarer at deres tillit til en offentlig tjeneste har blitt svekket som følge av dårlig informasjon.

"Fordi tjenesten er så informasjonsmessig dårlig, og gjør kronglete tekst ut av noe som burde vært løst med bedre søkefunksjon." (spørreundersøkelse).

På samme måte som brukerne ble frustrert av interaksjonene over, er det enkelte som har høye forventninger til hva chatboter kan hjelpe med. I noen tilfeller ligger disse forventningene også over mulighetene de fleste kommersielle chatboter leverer. Svar i spesifikke saker for eksempel er avhengig av innlogging, de er ikke av generell karakter. Dermed er dette svar chatboten i casebedriften ikke kan hjelpe med.

Sammen med at utfordringene ble oppdaget, ble det også avdekket et **forbedringspotensial** i forbindelse med bruken av virtuelle agenter. Noe av dette jobbes det allerede med, og enkelte løsninger vil ha kommet på plass mens forskningen for denne oppgaven pågikk. Selv om det tidligere ble omtalt som noe man ikke investerer mye ressurser i, er casebedriften interessert i å få ned differansen mellom de som starter chatboten og de som bruker den.

"Vi må få ned den differansen mellom hvor mange som starter Frida og hvor mange som faktisk bruker Frida. Det er litt viktig for oss. Så det er jo ting som vi ser på." (IS1, mellomleder).

For å lykkes med dette er det viktig at case-bedriften får tak i informasjon om brukeropplevelsene alle deres brukere har, og dette måles mot de andre kanalene organisasjonen benytter, telefon og epost. En utfordring i sammenligningen av kanalene ligger i chatbotens rolle kontra chaten med veiledere. Brukere som får hjelp av menneskelige veiledere får en brukerundersøkelse som da handler om chatten med veileder, samtidig som denne brukeren kan ha hatt en dårlig opplevelse med chatboten i forkant. Disse brukerne kan derfor ha verdifull informasjon om hva som ikke fungerte.

"Når du bare snakker med chatbot så går brukerundersøkelsen på chatboten, men har du blitt satt over til en veileder så går den bare på veileder uavhengig om hva slags hjelp du har fått av Frida i forkant." (IS3, chatbot-trener).

Videre har casebedriften også tilgang til egne brukerutvalg og en testlab som kunne vært utnyttet i arbeidet med å forbedre chatboten.

"Så der har vi nok mer å hente i sånne etablerte brukerutvalg for NAV, hvor de mer har fått et oppdrag for å være med og påvirke eller å gi innspill til å forbedre tjenestene. Og de brukerutvalgene har vi i liten grad brukt. Så der har vi et forbedringspotensial." (IS10, mellomleder).

Når det kommer til forbedringspotensialet fra respondentene i spørreundersøkelsen er det spesielt to grupperinger av svar som står frem. Samtidig som det kan være flere årsaker til hvorfor nettopp disse problemene oppstår. Flere trekker frem at de ikke får svar på spørsmålene sine fra interaksjoner med chatboter. Dette kan henge sammen med både forventningene til brukerne og språkforståelsen til den aktuelle chatboten.

"Den klarte ikke finne spesifikke svar på mitt konkrete spørsmål. Den forsto ikke mine nøkkelord, uansett hvordan jeg formulerte spørsmålet." (spørreundersøkelse).

Den andre hovedgruppen handler om sammenligning med telefon. Her er det respondenter som både foretrekker å bruke chatbot fordi man slipper å ringe eller ikke vil ringe. Samtidig er det enkelte brukere som trekker frem chatboten som kanal fordi det ikke var praktisk gjennomførbart å ringe.

"Fordi telefon er slitsomt og er mer tidkrevende." og "Fordi det var umulig å få tak i de på telefon" (spørreundersøkelse).

Tilbakemeldingene organisasjonen innhenter fra chater hvor innbyggeren har vært i kontakt med både chatbot og menneskelig veileder handler bare om interaksjonen med veileder. Der henvendelsen omhandler generell informasjon chatboten er godt egnet til å svare på er interaksjonen mellom denne brukeren og chatboten viktig. For gruppen med respondenter som opplever at chatboten gir de feil svar, har vi ikke nok informasjon til å dra konklusjoner, annet enn stor andel av respondentene har hatt denne opplevelsen. Ettersom disse tilbakemeldingene kommer fra spørreundersøkelsen, er det også uvisst noen av de gjelder Frida. Forventningsstyring kan hjelpe med å gi brukerne et realistisk bilde av hva chatboter kan brukes til. At enkelte respondenter fra spørreundersøkelsen sammenligner

interaksjoner med chatboter fordelaktig kontra telefonsamtaler er informasjon som kan brukes som reklame.

5.6 Tillit

Hvorvidt en bruker bestemmer seg for å bruke et system handler om brukeren tror man får nytte av det. Det handler om tilliten man har til at systemet kan dekke behovet man har. Tillit i forbindelse med chatboten jobbes det aktivt med i casebedriften. Noen brukeropplevelser kan svekke tilliten, mens andre vil styrke den. I dette delkapitlet tar vi for oss informantenes syn på tillit generelt før vi kommer tilbake til brukeropplevelser og tiltak som svekker eller styrker tilliten til virtuelle agenter.

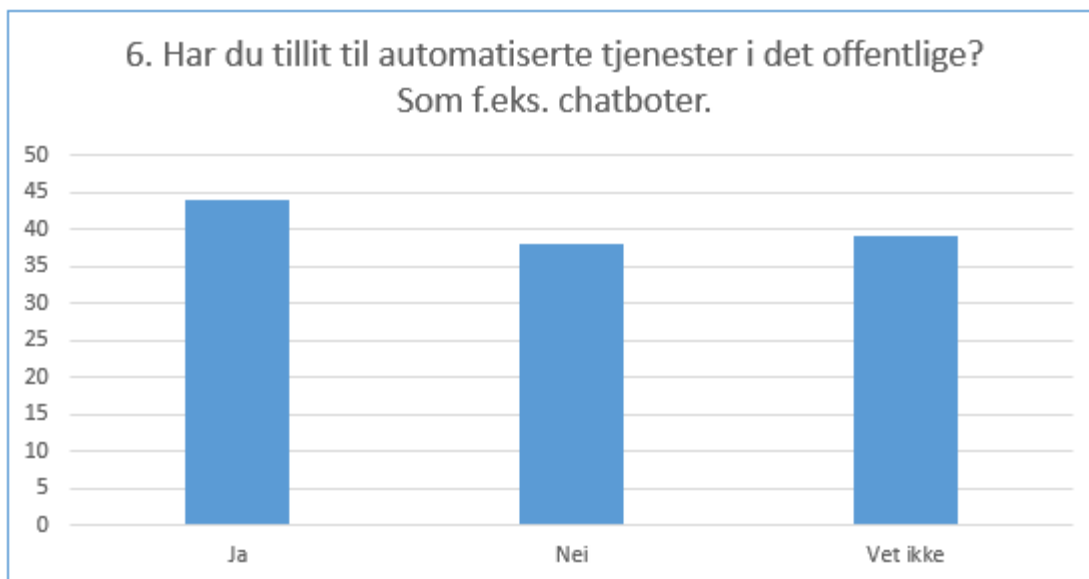
Arbeid med **tillit til en chatbot** tar tid, som en av informantene hos casebedriften beskriver det må man gjøre seg fortjent til det. Opplevelsen fra de ansatte er at brukerne har tillit til innholdet i det chatboten skriver, og dette er jo også noe de arbeider med.

"Det er jo, altså vi har jo snakket om en del av dette og også er det selvsagt også sånn relative bagateller som brukergrensesnitt og utseende på Frida og slike ting som man jo kontinuerlig jobber med og forbedrer. Men ut over det så tenker jeg at det viktigste rundt Frida er at hun gradvis bygger tillit. Tillit er jo ikke noe du kan kreve, du må jo gjøre deg fortjent til tillit. Og jeg opplever at hun gradvis har gjort seg mer og mer kjent og at. Ja, Frida kan du faktisk stole på." (IS12, leder).

Flere trekker frem dette med behovet for å bekrefte forståelsen av svar som en utfordring for chatboten. Det er ikke nødvendigvis sånn at brukeren ikke har tillit til svaret, men at brukeren bare vil være helt sikker på at man har forstått svaret riktig, og at det gjelder i akkurat den situasjonen brukeren står i.

"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." (IS4, chatbot-trener).

På direkte spørsmål om tillit til automatiserte systemer i offentlig sektor svarer 23 ja, 20 nei og de resterende 10 vet ikke. Alle som fikk dette spørsmålet, hadde tidligere bekreftet å ha vært i kontakt med en chatbot innenfor offentlig sektor det siste året. Forholdet mellom disse endrer seg i liten grad dersom man inkluderer respondentene som ikke har hatt kontakt med en chatbot det siste året (se figur 5).



Figur 5 Alle 121 respondentene i spørreundersøkelsen svarte på spørsmål om de har tillit til automatiserte tjenester i det offentlige.

Respondentene som svarte at de har tillit begrunnet svarene sine blant annet med at de har tillit til det offentlige, fordi det er det offentlige, fordi de regner med at leverandørene bak systemet kan levere godt nok eller fordi de har fått svar på henvendelsene sine. Av respondentene som svarer nei på spørsmål 6 svarer de fleste i retning av at de ikke har fått hjelp av denne typen tjenester, at det er begrenset hva slags hjelp man kan få eller at man foretrekker andre kanaler. Blant de som svarer vet ikke er det flere som trekker frem usikkerhet rundt om de kommer til å få den hjelpen de trenger gjennom bruk av systemet.

Det kan se ut som om de fleste brukerne som får riktig svar av chatboten slår seg til ro og har tillit til dette. Enkelte har et behov for å få bekreftet at svaret de har fått gjelder i akkurat deres tilfelle. Tillit utvikler seg over tid og kan være en medvirkende årsak til resultatet fra spørreundersøkelsen. I casebedriften er arbeidsområdene for chatboten blitt kraftig utvidet over det siste året og det vil ta tid å lære brukerne om mulighetene som finnes i chatkanalen.

En del interaksjoner med virtuelle agenter fører til at brukerens **tillit svekkes**. At man ikke får svar og at dette skaper frustrasjon trekkes frem av informantene. Dersom chatboten svarer etter en lengre redegjørelse fra brukeren, kan det føre til at chatboten gir svar som oppleves som upassende eller støtende. Det kan skje som følge av at chatboten feiltolker den informasjonen den har fått, og sjansen for dette øker med lengden og kompleksiteten i brukerens beskjeder.

"Ja, opplever at. Jeg opplever at vi har tillit i den grad at når de får informasjon så ser det ut som ... at de er fornøyde med den informasjonen da og avslutter chat. Men, hvis de ikke får informasjonen de ønsker med en gang. Da merker jeg at frustrasjonen skjer. Og det virker som tilliten da faller veldig fort." (IS9, chatbot-trener).

Til en viss grad bekreftes også dette utsagnet av respondentene i spørreundersøkelsen, der flere svarer at de har tillit til chatboter så lenge de får riktig svar eller gjør det brukeren ønsker. En del av respondentene påpeker også at spørsmålene de stiller må være av en enkel karakter for at de skal få svar

"Jeg har tillit i den forstand at jeg forventer at de kan svare på det som stort sett ellers er tilgjengelig av informasjon, men som de forhåpentligvis kan finne for meg raskere." (spørreundersøkelse).

En av de vanligste årsakene til at tilliten svekkes i interaksjonen med en chatbot er at brukerne ikke får et godt nok svar på det de er ute etter. Uavhengig av årsakene i de ulike interaksjonene ender resultatet ofte opp som en frustrert bruker som må se etter svar i andre kanaler.

I forbindelse med tillit ble det i intervjuene diskutert flere mulige **tiltak for å skape tillit** blant brukerne av chatboten. Enkelte av disse ble gjennomført i løpet av forskningsperioden, mens andre er en del av en pågående og kontinuerlig prosess. Det er komme på plass en erklæring om personvern og databehandling når man starter interaksjonen med chatboten. Dette har sammenheng med at man tidligere opplevde at brukerne lot være å lese velkomstmeldingen og heller begynte å prøve seg frem.

"Når du starter Frida, så får du opp en popup som sier noe om personvern. Hvis vi er tydelig på at du ikke skal skrive inn personlige opplysninger. Vi er tydelige på hvordan disse opplysningene eventuelt behandles og så videre. Og at du på en måte gir en aksept for at du har lest og forstått, før du da fortsetter [...]" (IS1, mellomleder).

Et av aspektene det jobbes kontinuerlig med, er det å være så personlig og folkelig som mulig i språket, samtidig som man er sikker på at man svarer riktig. Språklig presisjon, i den form at chatboten forklarer omstendighetene rundt svaret, kan også hjelpe til å styrke brukerens tillit. En informant forteller om eksempler på ulike scenarier som chatboten kan svare med dersom samme spørsmål blir stilt flere ganger. Dette kan også gi den bekreftelsen man ellers får ved å ta kontakt i andre kanaler og stille de samme spørsmålene. Det kontinuerlige arbeidet med å forbedre chatboten handler om å se mulighetene teknologien gir, fremfor begrensningene som ligger i den.

"Nei det går jo mer på at vi leser og prøver å få inntrykk har de forstått, spør de opp igjen. Et konkret grep vi har gjort på en del ting der vi ser at mange spør opp igjen er at vi en del steder rett og slett lager eksempler. Sånn at når de har lest det generelle svaret så kan du trykke eksempel og se. [...] Da får de litt den bekreftelsen [...]" (IS4, chatbot-trener).

Et annet område hvor man kanskje kan skape tillit til systemene er i de tilfellene der chatboten ikke kan hjelpe brukeren. Det går på prosessen med å bli satt over til en menneskelig veileder. Her går casebedriften inn for at det skal være så sømløst som mulig for brukeren. Som en del av dette er det viktig at brukeren oppfatter at han nå snakker med et menneske. I og med at det fortsatt er snakk om chat må det legges til rette for dette i brukerdessign og brukergrensesnitt.

"Så har vi også jobbet med dette med hvordan dette overføres fra [chatboten] til menneske. At vi får den søvnløse overføringen. Så vi slipper å stille spørsmålene på nytt igjen. Du skal få med chatdialogen. Og at det skal være en markering. Tydelig markering, på at nå snakker du med et menneske, nå snakker du med chatboten [...] Også er det jo en diskusjon som er pågående rundt dette med det å skape tillit til Norges befolkning. Hvis du bruker denne typen løsninger, det er jo at når Frida bestemmer seg for å sette samtalen videre, så er det ofte et spørsmål fra vår side hvorvidt vi.. [...] Hvilket nivå skal vi legge oss på? Hva forventer brukerne? Hvordan skal NAV opptre på en måte?" (IS1, mellomleder).

For casebedriften er noen av svarene IS1 stiller over svart på fordi organisasjonen er en del av noe større gjennom det å være en offentlig etat i Norge. Designmessig har de en overordnet profil de skal følge.

"De vurderingene sånn overordnet er jo lagt i NAV sin designprofil, og hvordan [...] tjenester skal se ut i stort. Så det er jo derfor på en måte de fargene er valgt og de ikonene er valgt, det er jo fordi dette ligger i NAV sin overordnede identitet og en designprofil som er utformet for at NAV skal fremstå helhetlig. Så noen rammer har man jo når man skal lage et chatvindu." (IS10, mellomleder).

For å bidra til at man når de overordnede målene for bruk av chatboten jobber organisasjonen med promoteringsarbeid i andre kanaler brukere tar kontakt. Dette er en del av den helhetlige kanal-strategien hos casebedriften og demonstrerer hvordan chatboten og dens oppgaver ikke kan sees på uten å ta med konteksten rundt organisasjonen som helhet.

"Også har vi jobbet med ulike strategier, nettopp for å forflytte brukerne og de som ikke trenger å stå i kø på telefonen i ti, femten, tjue minutter. Vi sender ut SMS med lenke til Frida og vi tester ulike måter å gjøre det på, for eksempel snike i køen, eller har du spørsmål av generell karakter, eller vet du at du kan chatte med oss, trykk her. Så får de lenke og så kan samtalen termineres, og så går du inn i en chat-kø i stedet" (IS1, mellomleder).

Noe potensielle brukere ikke nødvendigvis er klar over at svarene fra chatboten i caset er kvalitetssikret av mennesker og dermed skal svarene chatboten gir være korrekte. Å opplyse brukerne om dette, kanskje i sammenheng med noen av tiltakene man allerede gjør for å selge in chat og chatbot som tjeneste, kan være med å knytte chatboten til resten av organisasjonen. Denne promoteringen skjer i dag blant annet ved å opplyse noen av brukerne i telefonkø om muligheten for å få hjelp gjennom chat.

"Bare for å tenke høyt, burde man for eksempel ha skrevet noe om at chatboten ikke lever sitt eget liv og at det er mennesker som sitter bak og kvalitets-sikrer at svarene er helt riktige og at du får det nyeste regelverket og så videre. Hver eneste dag, nettopp for å skape denne tilliten og den ... ja ... tilfredsheten med å på en måte vite at dette er riktig, det er god hjelp jeg får hos denne chatboten" (IS1, mellomleder).

Dersom chatboten møter spørsmål den ikke kan svare på, opplyser den om dette. Det er en endring fra tiden da chatboten var ny og heller listet opp hva den kunne hjelpe med.

Informantene forteller at det ble vurdert om man skulle la chatboten fortelle hva den ikke kunne hjelpe med i forkant av samtalene, men man valgte bort dette fordi man var redd det ville gi inntrykk av at den ikke var så flink.

En annen kilde for å skape tillit til både organisasjonen og chatboten er gjennom gladsaker i media. Som eksempel fortalte en informant at en avis hadde sjekket svar i forbindelse med covid-19 pandemien, og fått bekreftet at chatboten hadde oppdaterte svar -i likhet med nettsidene til organisasjonen. Som en del av denne typen nyhetssaker vil det være mulig å utdype organisasjonens kanal-strategi og bruke muligheten til å forklare hvilke behov chatboten skal dekke.

"Frida hadde jo svart helt perfekt da, i forhold til de spørsmålene de hadde. For de var jo inne og kvalitetssikret de, og spurte da kommunikasjonsavdelingen i NAV [...] Så det er jo også på en måte en måte å forhåpentligvis tydeliggjøre ovenfor Norges befolkning at det virker. [...] Så sånne typer gladsaker er veldig bra for oss. Og kanskje er vi for dårlige på en måte på å gjøre de innsalgene ovenfor Norges befolkning på en måte, og mer reklamere for de tilbudene man har og være tydelig på hvordan den kanalstrategien, at det er viktig for NAV at vi kan bruke energien på de som trenger det mest, og at du som er digital og er selvbetjent, bruker for eksempel Frida som kan svare på, ja, flere hundre tusen spørsmål for eksempel da. Så det er jo en sånn type strategi som man egentlig ikke har diskutert så mye, det å drive et innsalg på den måten ute i media og så videre. Vi driver jo innsalg som, med å markedsføre dette på telefoni-siden som et eksempel." (IS12, leder).

Som vi ser, har casebedriften gjort mange tiltak med tanke på å skape en bedre brukeropplevelse og bygge tillit blant brukerne.

"Så ... det er ikke noen tvil om at det å skape en god brukeropplevelse handler om å skape tillit, og at det flyter godt, det fungerer sømløst. Og det har vi jo sånn sett jobbet ganske mye med, men vi har fortsatt et godt stykke igjen å gå der." (IS1, mellomleder).

Casebedriften jobber kontinuerlig med å skape en bedre brukeropplevelse for å bygge tillit gjennom dette. At chatboten kan komme med avklaringer i noen tilfeller er et eksempel på dette arbeidet. Et annet eksempel er arbeidet som legges ned i å gjøre overgangen fra chatbot til menneske så sømløst som mulig. Positiv omtale i media viser hvordan dette arbeidet er med på å skape en bedre opplevelse, og demonstrerer samtidig en kanal man kan informere brukerne i.

6. Diskusjon

I dette kapitlet diskuteres funnene sett i sammenheng med relevant teori for å gi en fyldig beskrivelse av tillit og brukeropplevelse vurdert mot forskningsspørsmålet. De to viktigste hovedtemaene for forskningsspørsmålet, tillit og brukeropplevelse har hver sin seksjon i kapitlet. Disse skaper et grunnlag for å diskutere forskningsspørsmålet i seg selv. Andre hovedtema fra funn, behov, forventninger, kompetanse og utfordringer flettes inn der de tematisk er koblet sammen med brukeropplevelser og/eller tillit. Kapitlet avsluttes med egne seksjoner om de praktiske implikasjonene av forskningen.

6.1 Tillit

Tilliten til chatboten i denne casen antas å være høy sett i lys av chatbotens hensikt, prosess og utførelse (Aoki, 2020). Det er få medie-saker som snakker om hvorfor NAV har tatt i bruk en chatbot, men hensikten bak det kan tolkes ved å se på den generelle strategien for digitalisering av offentlig sektor i Norge. Den overordnede hensikten med digitaliseringen er å skape effektive og bedre tjenester for brukeren. Chatboten bidrar til dette ved at den svarer på generelle spørsmål fra innbyggerne, spørsmål som de ellers måtte bruke tid på å lete seg frem til selv, eller få besvart ved hjelp av organisasjonens veiledere via chat eller andre kanaler. Chatboten frigjør de menneskelige ressursene slik at de kan fokusere på mer kompliserte arbeidsoppgaver.

Fra både casebeskrivelsen og svarene informantene gir, ser vi at arbeidsområdene hvor chatboten har forutsetningene for å gi disse svarene er kraftig utvidet det siste året. De fleste brukerne som får svar av chatboten på sine henvendelser, slår seg også til ro med disse svarene. En brukergruppe skiller seg ut ved at de har fått riktig svar av chatboten, men likevel ikke har fått dekket sitt behov for bekreftelse på at svaret chatboten har gitt gjelder i akkurat deres sak. Dette kan være knyttet til forventninger om å få svar som er mer tilpasset situasjonen den individuelle brukeren er i, og at den kognitive tilliten til systemet blir svekket når det ikke lever opp til disse (Glikson & Woolley, 2020). En måte å øke sjansen for at disse brukerne får tilfredsstillende svar, vil være å gjøre at chatboten kommer med avklarende eksempler dersom brukeren stiller det samme eller lignende spørsmål flere ganger på rad. Over tid bør det føre til økt tillit som følge av at chatbotens utførelse blir bedre ivarettatt. For brukerne som opplever at chatboten ikke kan svare på deres henvendelser, svekkes tilliten til chatboten.

En annen interessant brukergruppe er de som benytter seg av muligheten for å bli satt over til mennesker med en gang de møter en chatbot. Vi kan skille disse brukerne i to, de som vet at deres henvendelse gjelder et område chatboten ikke kan hjelpe med, og de som ikke er interessert i å gi chatboten en sjanse. Den første av disse er ikke interessant for denne forskningen. Sistnevnte brukergruppes oppførsel kan skyldes lav tillit som følge av arbeidsområdet chatboten opererer i, i tråd med et av hovedfunnene hos Aoki (2020). Den lave tilliten disse brukerne viser, kan også skyldes bevisste valg som følge av forventninger om at chatboten ikke kan hjelpe (Glikson & Woolley, 2020). Selv om svarene fra spørreundersøkelsen ikke handler om chatboten hos caseorganisasjonen, er dette en chatbot innenfor et område med mye ansvar og kompliserte saker hvor utfallene har direkte konsekvenser for brukernes liv.

Selv om forskningen fra Følstad et al. (2018) gjelder kommersielle chatboter, er det verd å merke seg at enkelte respondenter speiler resultatene. De oppgir å ha tillit til automatisering i offentlig sektor fordi det er offentlig sektor. Vi kan se dette som at merkevaren, "det offentlige", står sterkt i Norge. For å bygge forsterke denne tilliten kan det offentlige være transparente i forklaringer om hvordan algoritmene bak AI-en fungerer (Q. V. Liao et al., 2020). Samtidig er det usikkert om dette vil ha noen effekt, da denne informasjonen ikke er interessant for store deler av befolkningen (Aoki, 2020). Hvor godt chatboter det offentlige bruker klarer å svare på brukernes henvendelser vil da også påvirke tilliten til organisasjonene de representerer. Denne tilliten som følger av at det er en offentlig tjeneste det er snakk om er også i tråd med funn fra Alzahrani et al. (2017) som sier at etatens rykte, og brukertilfredsheten med etatens digitale tjenester er avgjørende for fremtidig bruk av offentlige digitale tjenester. I intervjuene ble det også avdekket at de fleste brukerne er fornøye med tjenesten. Respondentene i spørreundersøkelsen ga uttrykk til mer skepsis til automatiseringen av offentlige tjenester. Dette spørsmålet var mer generelt den konkrete chatboten som omtales i caset. At respondentene var splittet tilnærmet likt mellom å ha tillit, ikke ha tillit og svare vet ikke kan tyde på det ikke er et tema man har et bevisst forhold til.

Dersom vi ser på Frida med utgangspunkt i de konseptuelle metaforene om kompetanse og varme (Khadpe et al., 2020), scorer chatboten ifølge informantene høyt på faglig kompetanse, og lavere på varme. Det vil si at man ikke gir brukerne de beste forutsetningene for å ha en god opplevelse, ved at de kan forvente mer av chatboten enn den kan levere. Samtidig, dersom vi ser nærmere på hvordan brukerne kan få inntrykk av chatbotens kompetanse i caset, er det ikke mye som informerer om kompetansen den ene eller den andre veien. Ser vi på chatbotens måte å vise kompetanse på, med tanke på intelligens, ansvar og dyktighet, er det to måter disse kommer til uttrykk på i caset. For det første har vi evnen til å svare på brukernes henvendelser, og for det andre har vi det som handler om språkforståelse. Et uttrykk for språkforståelse ser vi gjennom chatbotens evne til å forstå nynorsk, selv om den svarer på bokmål. Dersom en henvendelse gjelder et fagområde chatboten ikke kan svare på, gir den beskjed om at den ikke kan hjelpe og tilbyr brukeren å bli satt over til et menneske. For brukere som opplever dette, kan det gi et inntrykk av lavere kompetanse enn chatboten faktisk har. Samtidig kan man være trygg på at informasjon man mottar fra chatboten er kvalitetssikret gjennom arbeid fagkoordinatorer gjør innen de ulike fagfeltene organisasjonen arbeider med.

For caseorganisasjonen er det også viktig at chatboten spiller på lag med resten av organisasjonen i den forstand at chatboten representerer organisasjonens tyngde i samfunnet. På grunn av dette har man valgt å fjerne muligheter for småprat, som ellers kunne vært en måte å vise noe varme på. Når det gjelder hvordan chatboten formidler varme i samtalen med brukerne, er chatboten i caset forsiktig med dette fordi man er redd for at det skal oppstå misforståelser der innbyggeren føler seg krenket. Informantene peker også på begrensninger i teknologien, og forventer at ettersom denne utvikler seg og blir mer avansert, vil det også åpne for at chatboten kan få en mer menneskelig tilnærming til å vise empati. Det er også et viktig poeng at organisasjonen ikke benytter seg av metaforer i stor grad når de beskriver chatboten, det at det er en chatbot skal styre forventningene i seg selv (Khadpe et al., 2020).

Chatboten i casen har også andre arbeidsoppgaver enn å hjelpe brukerne. Det vil si, chatboten hjelper alle den kan, dette frigjør kapasitet hos de menneskelige veilederne sånn at de kan konsentrere seg om de mer kompliserte sakene. Ved å ta unna henvendelser å frigjøre menneskelig kapasitet, fungerer chatboten som et filter. Denne arbeidsoppgaven er ikke direkte knyttet til folks tillit. Samtidig er innbyggernes tillit til systemet, i form av villighet til å benytte det avgjørende, for at man kan realisere gevinstene chatboten skaper.

6.2 Brukeropplevelser

Respondenter som har høyere forventninger enn chatboten de interagerer med kan levere, har dårlige brukeropplevelser. For oppgaveorienterte chatboter, som denne casen handler om, er det avgjørende for brukeropplevelsen at chatboten kan hjelpe med brukerens problem. I den sammenheng er de pragmatiske brukeropplevelsene sentrale for om brukerne har gode opplevelser (Følstad & Brandtzaeg, 2020). Som vi også får bekreftet fra informantene, fører feil svar og misforståelser til frustrasjon hos brukeren. De høye forventningene har en mulig forklaring ved at enkelte respondenter mener chatboten erstatter et menneske, og derfor må være i stand til å utføre jobben som om den var et menneske. En annen mulig kilde kan være chatbotens utseende, da mer menneskelignende utseende skaper forventninger om menneskelignende egenskaper (Glikson & Woolley, 2020). Chatboten i caset har en avatar og navn som ser og høres menneskelig ut. For chatboten kan de menneskelignende trekkene også være en kilde til tillit, spesielt blant eldre brukere (ibid.). Den eldre brukergruppen kan også ha problemer med å nyttiggjøre seg tjenestene som følge av at de mangler den digitale kompetansen disse krever (Bratteteig & Verne, 2012). Uoverensstemmelsen som da oppstår mellom forventninger av menneskelig ytelse og mangel på brukerkompetanse for å utnytte mulighetene i systemet, kan være uheldig for denne gruppen. Brukeropplæring rettet spesielt mot eldre for å heve deres digitale kompetanse, kan muligens være en løsning for denne gruppen. Innen det offentlige i Norge er det også allment akseptert at vi skal ta vare på de svakeste i samfunnet. Samtidig kan det i denne sammenhengen være fornuftig ressursbruk at denne brukergruppen benytter andre kanaler til kontakt med offentlig forvaltning. For enkelte respondenter henger frustrasjonen tett sammen med at de ser på teknologien som et hinder, fremfor en kilde til hjelp. Dette er en gruppe hvor chatboter kanskje ikke er den beste løsningen for å få hjelp, og for dem vil det være mer nyttig å søke hjelp i andre kanaler.

For å bedre forstå innbyggerne i møte med chatboten, kan man se på interaksjonene deres i sammenheng med modellen for innhold (Hassenzahl, 2018). Brukeropplevelsen starter med et hvordan: interaksjon med "chat med oss knappen" på organisasjonens hjemmeside. Selve interaksjonen med chatboten, og om man trenger mer hjelp enn den kan gi, interaksjon med veileder, svarer på hva. Svarene på spørsmålet om hvorfor er sentrale i denne konteksten. Hvorfor tar brukerne kontakt gjennom chat kanalen? Respondentenes svar er varierte, men alle har utgangspunkt i ønske om å få dekket et konkret behov og en forventning om at handlingene deres vil lede til dette. Behovet chatboten i caset dekker, er svar på generelle spørsmål. De ulike egenskapene respondentene beskriver samtaler med chatboter med, svarer på hvorfor de benytter den. Positive beskrivelser inkluderer at det er en effektiv måte å kommunisere med en organisasjon på, det er raskt og enkelt. Det er et alternativ utenfor vanlig åpningstid. For enkelte er det også positivt at man slipper å snakke

med andre mennesker. Når interaksjonen lykkes, har man hatt en god brukeropplevelse og sitter ifølge teorien igjen med følelsen av velvære (ibid.).

Gjennom arbeid med å styre brukernes forventninger til systemet øker, man sjansen for at brukerne unngår å ha høyere forventninger enn det systemet kan levere. For chatboten i caset finnes det tre muligheter til å påvirke brukeren før de interagerer med chatboten. Designet rundt chatboten, velkomstmeldingen som forteller om interaksjonen og omgivelsene rundt organisasjonen sammen med samfunnet for øvrig. Designet omfatter chatbotens avatar, navn, hvordan chatvinduet er utformet og den øvrige design-profilen hos organisasjonen. Velkomstmeldingen forteller brukere, som leser den, om chatboten. Informantene gir uttrykk for at det utfordrende å fortelle brukerne nok, uten å presentere dem for en "vegg av tekst" som ikke blir lest. Forventningene som stammer fra omgivelsene rundt, og samfunnet for øvrig er vanskeligere å påvirke. Her finner vi forventningene som skapes av oppslag i media, av andre aktører som benytter lignende eller tilsvarende teknologi, og ikke minst i samtaler mellom innbyggerne.

I design og utviklingsarbeidet av chatboten fra case-organisasjonen er sluttbrukerne, innbyggerne i Norge, indirekte involvert. Dette skiller seg fra anbefalt praksis innen menneske-maskin interaksjon (Svanæs & Gulliksen, 2008). Brukerne representeres ved at man tar utgangspunkt i interaksjoner mellom sluttbruker og organisasjonen når man designer og videreutvikler chatboten. Disse interaksjonene er ikke knyttet direkte til utviklingsarbeidet, og sluttbrukerne er ikke mer involvert enn de uansett ville vært for å få dekket sine behov. Brukerinvolveringene kommer til syne gjennom chat med veiledere, hvor veilederen kan representere brukeren videre. Det samles også inn tilbakemeldinger etter at en samtale med chatboten er ferdig dersom brukeren tar seg tid til å gjøre dette. Årlige brukerundersøkelser som omfatter hele organisasjonen og interne oversikter som gir informasjon om hva slags type henvendelser de bruker å motta til ulike tider av året.

Sammen danner disse et bilde av behovene en typisk bruker har, og er avgjørende for at denne måten å drive utviklingsarbeid på skal være vellykket. Det vil si at en forutsetning for at indirekte brukerinvolvering i systemutvikling skal fungere, er god kunnskap om brukernes behov og forventninger. Indirekte brukerinvolvering kan også være ressursbesparende dersom man allerede sitter på tilstrekkelig kunnskap om brukernes behov innad i organisasjonen. Gjennom tilbakemeldingene casebedriften innhenter gjennom de nevnte kanalene, samt brukerundersøkelser og medieomtale ser det ut som organisasjonen har tilstrekkelig kunnskap og har lykket med indirekte brukerinvolvering i design og utviklingsarbeidet av chatboten.

6.3 Forholdet mellom tillit og brukeropplevelser

Basert på teori og funnene fra intervju og spørreundersøkelse kan vi slå fast at tillit og brukeropplevelser henger tett sammen. Ulike forhold rundt brukeren og omgivelsene rundt organisasjonen som benytter chatboten, danner grunnlaget for brukerens forventninger og holdninger til systemet. Sammen danner disse basis for brukerens grunnleggende tillit til chatboten (Aoki, 2020). I interaksjoner mellom innbyggeren og chatboten blir denne tilliten testet. Dersom systemet gir en god brukeropplevelse, styrkes tilliten. En dårlig brukeropplevelse vil redusere tilliten til at man kan få løst behovet man var ute etter gjennom bruken av chatboten (Følstad & Brandtzaeg, 2020). Svekket tillit som følge av

dårlige brukeropplevelser, kan også påvirke forventningene til chatboten for innbyggeren som hadde den dårlige brukeropplevelsen, men også medvirke til lavere forventninger ved at innbyggeren påvirker andre potensielle brukere i omgivelsene rundt seg selv.



Figur 6 Illustrasjon av forhold mellom forventninger, tillit og brukeropplevelser.

Forventningene til systemet blir på denne måten svært avgjørende for forholdet mellom tillit og brukeropplevelse (se figur 6). Forsøk på å styre brukernes forventninger, slik at man i møte med chatboten ikke opplever å bli skuffet gjennom en dårlig brukeropplevelse er avgjørende for om en chatbot over tid vil være suksessfull med tanke på mål om effektivisering og forbedringer av digitale tjenester i offentlig sektor.

6.4 Praktiske implikasjoner

Tilliten innbyggere har til chatboter i offentlig sektor kan styrkes med tanke på brukergruppen som stiller de samme spørsmålene til menneskelige veiledere etter å ha fått riktig svar av chatboten. Dette kan løses ved at chatboten kommer med et eller flere eksempler som er relevante for situasjonen brukeren opplever. Eksemplifisering er en god måte for virtuelle AI-er å overtale mennesker på, sammenlignet med eksempel sammen med resonnement (Kraus, 2012).

Tilliten innbyggerne har fordi chatboten representerer noe større, i denne casen representert gjennom offentlig forvaltning, har også muligheter for forbedringer. Innbyggerne som velger kanalen kan gjøres oppmerksom på at dersom chatboten ikke kan hjelpe, vil man få hjelp av et menneske. Dersom dette er godt kjent blant brukerne, kan det knyttes til emosjonell tillit (Glikson & Woolley, 2020).

For å skape en enda bedre brukeropplevelse kan NAV jobbe mer med hvordan de innhenter tilbakemeldinger fra brukere av chatboten. Tilbakemeldings-skjema brukerne får etter endt chat, når brukeren har vært i kontakt med både chatbot og veileder, kan ha fordel av å inkludere tilbakemeldinger på begge samtalepartnerne. I tilfeller der innbyggeren har generelle spørsmål, som chatboten har gode forutsetninger for å svare på, er det spesielt interessant å få belyst hvorfor noen ikke ga chatboten en sjanse.

I arbeidet med å forme forventningene til chatboter i offentlig sektor, kan det også være en fordel å forklare brukerne 1) hvordan man bruker chatboten på best mulig måte. 2) Hvordan chatboten bidrar positivt i kanalmiksen hos organisasjonen ved at den frigjør ressurser for de innbyggerne som har behov for dem. Sistnevnte burde veie tungt i Norge hvor vi som et sosialdemokrati er enige om å bruke ressurser på de som har behov for hjelp.

7. Konklusjon

Denne oppgaven har sett nærmere på og beskrevet forholdet mellom tillit og brukeropplevelser når det er snakk om bruken av virtuelle agenter innen offentlig sektor i Norge (se figur 6). Brukernes tillit til en chatbot og brukeropplevelsene av chatboten henger svært tett sammen. Tillit til at chatboten kan hjelpe fører til at brukeren er villig til å gi systemet en sjanse. Gode brukeropplevelser øker tilliten til at man kan få hjelp. Har man lav tillit til at chatboten kan hjelpe, er det ikke sikkert man velger å gi den en sjanse. Brukeren vil da forsøke å finne alternative kanaler for å få dekket behovet sitt.

For oppgave-orienterte chatboter, slik som i caset, er det viktigste for innbyggeren at den kan løse oppgaven den er satt til. Gode brukeropplevelser er tett knyttet mot nettopp å få løst dette behovet. Feil eller manglende svar, sammen med utfordringer knyttet til chatbotens evne til å forstå brukeren er de viktigste kildene til dårlige brukeropplevelser. Enkelte respondenter drar at chatboten er enkel og rask å bruke som en fordel. Dette krever at man får dekket behovet man hadde. Blant de gode og dårlige brukeropplevelsene respondentene beskriver er de fleste pragmatiske av karakter (Følstad & Brandtzaeg, 2020; Hassenzahl, 2018). Disse er også de viktigste å løse for brukeren tillit til oppgaveorienterte chatboter.

Misforståelser som oppstår i brukernes møte med chatboter, fører ofte til frustrasjoner for brukerne. Disse har ofte opphav i forventninger som viser seg å ikke stemme. Funn fra spørreundersøkelsen viser at forventningene som fører til frustrerte brukere, er ofte knyttet til forventninger om menneskenære opplevelser. Det vil si at brukeren forventer at chatboten skal ha tilsvarende språkforståelse, og kunnskaper som om de snakket med et menneske. Inkludert i dette har man blant annet ønske om spesifikke svar, knyttet til brukernes egen saksbehandling. Chatboten i caset kan svare på spørsmål av generell karakter. For chatboten i caset er generell informasjon all informasjon som vil være likt, uavhengig av hvem som spør om den. Det vil si at for enkelte brukere kan dette oppleves som ganske spesifikk informasjon.

En del brukere som får hjelp fra chatboten søker likevel svar på de samme spørsmålene fra menneskelige veiledere. Det kan henge sammen med et behov om være helt sikker på at informasjonen man har mottatt, gjelder i egen sak. Chatboten kan muligens hjelpe disse gjennom bruk av eksempler som illustrerer hvordan informasjonen som er delt skal tolkes (Kraus, 2012).

Caseorganisasjonens bruk av chatbot er et godt eksempel på hvordan offentlig sektor kan gi brukerne en god opplevelse, selv om man ikke følger normer for tradisjonell brukerinvolvering i systemutvikling. God kunnskap om brukernes behov er en forutsetning for å lykkes med den indirekte brukerinvolveringen. I caseorganisasjonen lykkes dette ved at man involverer veiledere og deres kompetanse om brukerbehovene i videreutviklingen av chatboten. I forbindelse med denne oppgaven er videreutviklingen det er snakk om, knyttet opp mot nye fagområder for chatboten, ikke videreutvikling av AI-teknologien som ligger bak chatbotens kommunikasjonsferdigheter.

7.1 Begrensninger og fremtidig arbeid

Den største begrensningen i denne masteroppgaven var tidsbegrensningen på et semester. Tidsbegrensningen la grunnlag for tiden som ble satt av til å studere temaene chatbot, tillit og bruk av datasystemer i offentlig sektor. Det kan ha hatt betydning for teorien som er brukt i oppgaven uten at forfatteren er klar over det.

Frida@work-prosjektet skapte også noen begrensninger med tanke på tilgang til informanter. Samtidig er det viktig å trekke frem at oppgaven hadde vært mye mer begrenset enn den ble, dersom den ikke hadde vært en del av samarbeidsprosjektet.

Konklusjonene som er nådd i oppgaven medfører noe usikkerhet med tanke på i hvor stor grad de faktisk gjelder for Frida. Dette henger sammen med at spørreundersøkelsen ikke ble gjennomført blant brukere av den konkrete chatboten.

Respondentene som tok spørreundersøkelsen reflekterer også en relativt homogen gruppe, ved at en stor overvekt av de har universitetsutdannelse. Dette gjør at man kan stille spørsmål knyttet til i hvilken grad resultatet gjelder resten av befolkningen. Man kan anta at denne gruppen er mer teknologivant enn andre deler av befolkningen. Dette er også et mulig område å forske videre på. Er akademikere representative for befolkningen som helhet på dette område?

En annen mulighet for fremtidig forskning er å følge opp de to brukergruppene som nevnes spesielt i diskusjonen. De som stiller samme spørsmål til menneskelige veiledere, og de som ikke er villige til å gi chatboten en sjanse, selv om de har generelle spørsmål den vil kunne svare på. Bruk av spørreundersøkelser som den Følstad og Brandtzaeg (2020) benyttet kan være en mulig fremgangsmåte for å finne svar for disse brukergruppene.

I Norge drar det offentlige nytte av høy tillit blant befolkningen, også når det kommer til digitale tjenester. Hvordan er dette utenfor Norge? Fungerer oppgaveorienterte chatboter like god for offentlig sektor i land der myndighetene har lavere tillit blant innbyggerne?

8. Referanser

- Ala-Pietilä, P., Bauer, W., Bergmann, U., Bieliková, M., Bonefeld-Dahl, C., Bonnet, Y., ... Yeung, K. (2019). *Ethics guidelines for trustworthy AI*. European Commission,. Hentet fra <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>
- Alzahrani, L., Al-Karaghoul, W. & Weerakkody, V. (2017). Analysing the critical factors influencing trust in e-government adoption from citizens' perspective: A systematic review and a conceptual framework. *International business review*, 26(1), 164-175.
- Anskaffelsesloven. (2016). Lov om offentlige anskaffelser (LOV-2016-06-17-73). Hentet fra <https://lovdata.no/dokument/NL/lov/2016-06-17-73>
- Aoki, N. (2020). An experimental study of public trust in AI chatbots in the public sector. *Government Information Quarterly*, 37(4), 101490. <https://doi.org/10.1016/j.giq.2020.101490>
- Arbeids- og velferdsetaten. (2021, 25.05.2021). Samhandling med brukere. Hentet 14.07.2021 fra <https://www.nav.no/no/nav-og-samfunn/kunnskap/fou-midler/pagaende-fou-prosjekter2/navs-tiltak-og-virkemidler>
- Arbeids- og velferdsetaten. (U.Å., 07.05.2020). Kva er NAV? Hentet 03.12.2020 fra <https://www.nav.no/no/nav-og-samfunn/om-nav/fakta-om-nav/kva-er-nav>
- Baxter, P. & Jack, S. (2008). Qualitative case study methodology: Study design and implementation for novice researchers. *The qualitative report*, 13(4), 544-559.
- Bratteteig, T., Saplacan, D., Soma, R. & Svanes Oskarsen, J. (2020). Strengthening human autonomy in the era of autonomous technology. Contemporary perspectives on interaction with 'autonomous things'. *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society* (s. 1-3).
- Bratteteig, T. & Verne, G. B. (2012). Conditions for Autonomy in the Information Society: Disentangling as a public service. *Scand. J. Inf. Syst.*, 24(2), 3.
- Chattaraman, V., Kwon, W.-S., Gilbert, J. E. & Li, Y. (2014). Virtual shopping agents. *Journal of Research in Interactive Marketing*.
- Ciechanowski, L., Przegalinska, A., Magnuski, M. & Gloor, P. (2019). In the shades of the uncanny valley: An experimental study of human-chatbot interaction. *Future generation computer systems*, 92, 539-548. <https://doi.org/10.1016/j.future.2018.01.055>
- Cotter, K. & Reisdorf, B. C. (2020). Algorithmic Knowledge Gaps: A New Horizon of (Digital) Inequality. *International Journal of Communication*, 14, 21.
- Cuddy, A. J., Fiske, S. T. & Glick, P. (2008). Warmth and competence as universal dimensions of social perception: The stereotype content model and the BIAS map. *Advances in experimental social psychology*, 40, 61-149.
- Én digital offentlig sektor. (2019). Hentet fra <https://www.regjeringen.no/no/dokumenter/en-digital-offentlig-sektor/id2653874/>
- Fiske, S. T., Cuddy, A. J., Glick, P. & Xu, J. (2002). A model of (often mixed) stereotype content: competence and warmth respectively follow from perceived status and competition. *Journal of personality and social psychology*, 82(6), 878.
- Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., ... Rossi, F. (2018). AI4People—an ethical framework for a good AI society: opportunities, risks, principles, and recommendations. *Minds and Machines*, 28(4), 689-707. Hentet fra <https://link.springer.com/article/10.1007/s11023-018-9482-5>
- Følstad, A. & Brandtzaeg, P. B. (2020). Users' experiences with chatbots: findings from a questionnaire study. *Quality and User Experience*, 5(1), 1-14.
- 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* (s. 194-208): Springer.

- Glikson, E. & Woolley, A. W. (2020). Human trust in artificial intelligence: Review of empirical research. *Academy of Management Annals*, 14(2), 627-660.
- Goyal, P., Pandey, S. & Jain, K. (2018). Deep learning for natural language processing. *New York: Apress*.
- Hassenzahl, M. (2018). The thing and I: understanding the relationship between user and product. I *Funology 2* (s. 301-313). Springer.
- 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.
- International Organization for Standardization. (2019). *Ergonomics of human-system interaction – Part 210: Human-centred design for interactive systems* (ISO 9241-210:2019). Hentet fra <https://www.iso.org/obp/ui/#iso:std:iso:9241:-210:ed-2:v1:en>
- Intervju Subjekt 1. (2021, 14.01.2021). Innledende intervju. I.
- 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.
- Knowles, B. & Richards, J. T. (2021). The Sanction of Authority: Promoting Public Trust in AI. *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (s. 262-271).
- Kraus, M. F. a. I. Z. a. S. (2012). Guiding user choice during discussion by silence, examples and justifications. *ECAI 2012: 20th European Conference on Artificial Intelligence* (s. 330): IOS Press.
- Law, E. L.-C., Roto, V., Hassenzahl, M., Vermeeren, A. P. & Kort, J. (2009). Understanding, scoping and defining user experience: a survey approach. *Proceedings of the SIGCHI conference on human factors in computing systems* (s. 719-728).
- Levitan, D. M. (1946). The responsibility of administrative officials in a democratic society. *Political Science Quarterly*, 61(4), 562-598.
- Liao, Q. V., Gruen, D. & Miller, S. (2020). Questioning the AI: informing design practices for explainable AI user experiences. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (s. 1-15).
<https://doi.org/https://doi.org/10.1145/3313831.3376590>
- Liao, R. (2020, 13.07.2020). Microsoft spins out 5-year-old Chinese chatbot Xiaoice. Hentet fra <https://news.microsoft.com/apac/features/much-more-than-a-chatbot-chinas-xiaoice-mixes-ai-with-emotions-and-wins-over-millions-of-fans/>
- Mayer, R. C., Davis, J. H. & Schoorman, F. D. (1995). An Integrative Model of Organizational Trust. *The Academy of Management review*, 20(3), 709-734.
<https://doi.org/10.2307/258792>
- Meld. St. 27. ((2015–2016)). *Digital agenda for Norge – IKT for en enklere hverdag og økt produktivitet* Kommunal- og moderniseringsdepartementet.
- Mikalef, P., Fjørtoft, S. O. & Torvatn, H. Y. (2019). Artificial Intelligence in the public sector: a study of challenges and opportunities for Norwegian municipalities. *Conference on e-Business, e-Services and e-Society* (s. 267-277): Springer.
- Mori, M., MacDorman, K. F. & Kageki, N. (2012). The uncanny valley [from the field]. *IEEE Robotics & Automation Magazine*, 19(2), 98-100.
- Nasjonal strategi for kunstig intelligens. (2020). Hentet fra <https://www.regjeringen.no/no/dokumenter/nasjonal-strategi-for-kunstig-intelligens/id2685594/>
- NAV. (U.Å). NAV. Hentet fra <https://www.nav.no/person/kontakt-oss/nb>
- Oates, B. J. (2006). *Researching information systems and computing*. London: Sage Publications.
- Olson, G. M. & Olson, J. S. (2000). Distance matters. *Human-computer interaction*, 15(2-3), 139-178.

- Quiñones, D., Rusu, C. & Rusu, V. (2018). A methodology to develop usability/user experience heuristics. *Computer standards & interfaces*, 59, 109-129.
- Ringnes, I. F. (2018). Satser stort på digitale tjenester til foreldre Hentet 10.01.2021 fra <https://memu.no/artikler/satser-stort-pa-digitale-tjenester-til-foreldre/>
- Ringnes, I. F. (2020). Frida jobber døgnet rundt Hentet 25.11.2020 fra <https://memu.no/artikler/frida-jobber-dognet-rundt/>
- Sem & Stenersen Prokom. (U.Å.). Kommune-Kari. Hentet fra <https://prokom.no/produkter-og-tjenester/chatbots/kari/>
- Simonsen, L. (2019). *Når brukerdialogen automatiseres--hva blir vanskelig?* (Masteroppgave). Universitetet i Oslo. Hentet fra <https://www.duo.uio.no/handle/10852/80708>
- Simonsen, T. (2021, 10.02.2021). Resultater 2020. I. Skatteetaten. (U.Å.). Kontakt oss. Hentet fra <https://www.skatteetaten.no/kontakt/>
- Standard Norge. (2012, 12.12.2013). Regjeringens digitaliseringsprogram. Hentet fra <https://www.standard.no/nyheter/nyhetsarkiv/ikt/2012/regjeringens-digitaliseringsprogram/>
- Svanæs, D. & Gulliksen, J. (2008). Understanding the context of design: towards tactical user centered design. *Proceedings of the 5th Nordic conference on Human-computer interaction: building bridges* (s. 353-362).
- Thiebes, S., Lins, S. & Sunyaev, A. (2020). Trustworthy artificial intelligence. *Electronic Markets*, 1-18.
- Tjora, A. (2018). *Kvalitative forskningsmetoder i praksis*. Oslo: Gyldendal Norsk Forlag AS.
- Valevatn, J. (2017). Teknologien som lærer deg å kjenne. Hentet 13.11.2020 fra <https://teknologiradet.no/teknologien-som-laerer-deg-a-kjenne/>
- Walsham, G. (1995). Interpretive case studies in IS research: nature and method. *European Journal of Information Systems*, 4(2), 74-81. <https://doi.org/https://doi.org/10.1057/ejis.1995.9>
- Ågerfalk, P. J. (2020). Artificial intelligence as digital agency. *European Journal of Information Systems*, 29(1), 1-8.

9. Vedlegg

Vedlegg A: Intervjuguide innledende intervju

Introduksjon om oppgavene

Vi er to studenter som forsker på tillit til chatboter. I den forbindelse samarbeider vi om datainnsamling, men skriver to forskjellige oppgaver med litt forskjellig vinkling. Knut skriver om tillit og brukeropplevelse, mens Maria fokuserer på opplevd kompetansenivå og tillit.

- A. Hvordan forholder NAV seg til brukernes behov og brukerinvolvering i utviklingen av chatboten? Konkret hva gjør man for å forstå brukerbehov og involvere brukerne?
- B. Hvordan arbeider NAV med videreutviklingen av chatboten, spesielt med tanke på samtaler som ikke blir avsluttet på "korrekt" måte?
- C. Frida ber deg omformulere hvis hun ikke forstår spørsmålet ditt. Hvilke tanker ligger bak hvordan NAV har valgt å presentere Fridas kompetansenivå og begrensninger?
- D. Spesifikt om tillit, er det gjort noe med tanke på brukere som avbryter samtalen når de finner ut at de snakker med en robot? (hvis ja, hva er gjort og hvordan har det hjulpet?) (hvis nei, er det et stort problem? Hvorfor er det ikke gjort noe med det?)
- E. I samtaler mellom Frida og brukere - oppfatter dere at brukere har tillit til Frida? På hvilken måte?
- F. Hvilke tiltak har dere gjort for å få Frida til å oppfattes mer tillitsfull av brukere? Hvilke fordeler tror dere dette vil ha?
- G. Identifiserer Frida brukerne på noe vis og gir ulike responser basert på det? F.eks. om noen skriver dårlig norsk - får de raskere spørsmål om de vil bli satt over til menneskelige agenter?

Takk for svar.

Vedlegg B: Intervjuguide ansatt

Introduksjon om oppgavene

Vi er to studenter som forsker på tillit til chatboter. I den forbindelse samarbeider vi om datainnsamling, men skriver to forskjellige oppgaver med litt forskjellig vinkling. Knut skriver om tillit og brukeropplevelse, mens Maria fokuserer på opplevd kompetansenivå og tillit.

Om utvikling (ikke innhold) & brukerinvolvering

1. Med tanke på Fridas utseende og hvordan hun presenterer seg selv. Hvilke valg er tatt? Hvilke vurderinger ligger bak avgjørelsene?
2. Konkrete eksempler på tiltak som er gjort for å skape bedre tillit blant brukerne?
3. 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?
4. 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

5. Hvordan selger dere inn Frida til brukere?
6. I samtale mellom Frida og brukere - oppfatter dere at brukere har tillit til Frida? På hvilken måte?
7. Hvilke tiltak har dere gjort for å få Frida til å oppfattes mer tillitsfull av brukere? Hvilke fordeler tror dere dette vil ha?
8. Om kompetanse
9. Frida sier ikke noe om hvordan brukeren burde stille spørsmål eller hva hun kan svare på i velkomsthilsenen. Hvilke vurderinger er gjort rundt dette?
10. Hvordan behandler/styrer dere brukernes forventninger til Frida med tanke på hennes kompetansenivå?

Om empati

11. 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?

Takk for svar.

Vedlegg C: Intervjuguide leder

Introduksjon om oppgavene

Vi er to studenter som forsker på tillit til chatboter. I den forbindelse samarbeider vi om datainnsamling, men skriver to forskjellige oppgaver med litt forskjellig vinkling. Knut skriver om tillit og brukeropplevelse, mens Maria fokuserer på opplevd kompetansenivå og tillit.

Om utvikling (ikke innhold) & brukerinvolvering

- i. Er det en overordnet strategi for Fridas utseende og måten hun presenterer seg selv?
- ii. Er det måter dere ønsker å involvere brukere på i forbindelse med videreutvikling av Frida?

Om tillit

- iii. Hvordan ser NAV på sammenhengen mellom tillit og brukeropplevelse?
- iv. Tillit til AI kan uttrykkes konseptuelt langs akser for varme og kompetanse. Hvor befinner Frida seg langs disse aksene?

Om kompetanse

- v. Hvordan behandler/styrer dere brukernes forventninger til?
- vi. Hvilke forventninger ønsker dere at brukerne skal ha til Frida?

Om empati

- vii. Vi vet at enkelte tema, f.eks. død, gjør at du blir satt over til mennesker med en gang. Er det et mål at Frida skal kunne svar på alt dette i fremtiden? Hva må i så fall på plass før dette blir en realitet?

Takk for svar.

Bruk av chatbot i offentlig sektor

Obligatoriske felter er merket med denne stjernen*

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? *

i 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? *

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?»

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? *

 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? *

 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? *


 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? *


 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? *


 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? *

 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? *

 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? *

 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? *

 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? *

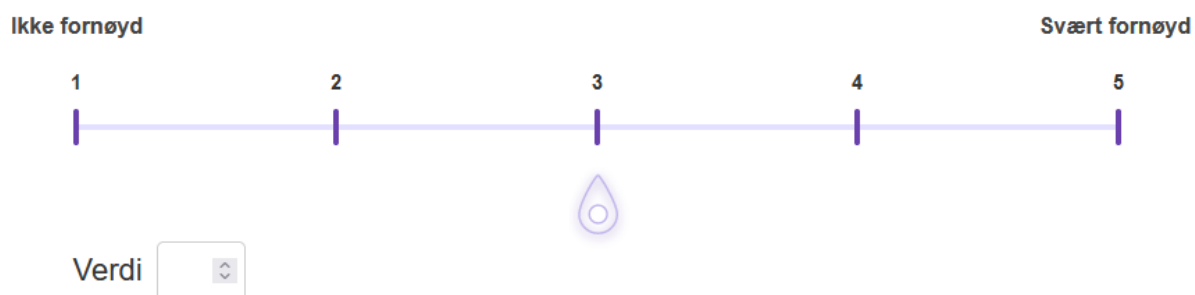
 Dette elementet vises kun dersom alternativet «Jeg ba om å bli satt over til et menneske» er valgt i spørsmålet «9b. Hvorfor ble du satt over til et menneske?»

Husk å ikke oppgi personlig informasjon.

10. Hvor fornøyd var du med interaksjonen? *


 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)



Kompetanse, brukeropplevelse og tillit

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

 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.

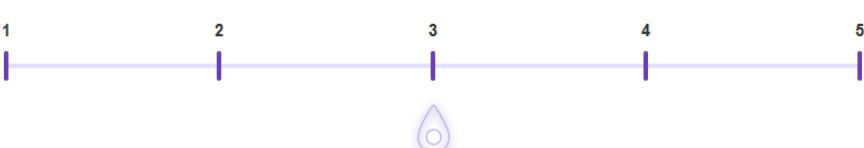
12. Hvilket inntrykk fikk du av chatbotens faglige kompetansenivå under samtalen? *

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

1 2 3 4 5



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? *


 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.

 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? *

 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? *

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

Takk for din deltakelse! Husk å trykk "Send" før du avslutter.

